

Evolução em ação: lições de uma pandemia*

Rhewter Nunes¹, José Alexandre Felizola Diniz Filho², Mariana Pires de Campos Telles^{1,3}

¹ Laboratório de Genética & Biodiversidade, ICB, Universidade Federal de Goiás, Goiânia

² Laboratório de Ecologia Teórica & Síntese, Instituto de Ciências Biológicas, Universidade Federal de Goiás

³ Escola de Ciências Médicas e da Vida, Pontifícia Universidade Católica de Goiás

Autor para correspondência – tellesmpc@gmail.com

Palavras-chave: coronavírus, filogenia molecular, genômica, modelagem epidemiológica, SARS-COV-2, variação adaptativa

* Material produzido no contexto do Instituto Nacional de Ciência e Tecnologia em Ecologia, Evolução e Conservação da Biodiversidade (INCT_EECBio) e do Projeto Mapeamento das variações genéticas do Coronavírus (SARS-COV-2) em Goiás (FAPEG).



O surgimento de novas mutações em linhagens do vírus SARS-CoV-2 ganhou grande repercussão na mídia e nas redes sociais no segundo ano da pandemia. O ponto central da questão é que algumas mutações em determinados genes dessas linhagens podem conferir a elas maior capacidade de infecção e, talvez, maior patogenicidade. Tendo em vista que o vírus tem um tempo de geração extremamente curto, o surgimento e a manutenção dessas mutações permite observar a evolução em ação. Nesse artigo é discutido como os diferentes processos evolutivos podem ser percebidos ao acompanhar o surgimento de novas linhagens de SARS-CoV-2 durante a pandemia. Além disso, será apresentado um exemplo de aplicação de um modelo epidemiológico chamado de SIR (Suscetível-Infetado-Recuperado). Ele servirá para exemplificar uma situação hipotética em que novas mutações aumentam a capacidade de transmissão do vírus, ilustrando um exemplo de evolução adaptativa.

Evolução em diferentes escalas

Uma das maiores dificuldades para ensinar evolução é que se torna necessário pensar em processos que acontecem em grandes escalas de tempo. Pode-se comparar os organismos e entender as relações de ancestralidade a partir da identificação de padrões de compartilhamento de características em diferentes momentos no tempo (o que forma uma “filogenia”). Também é possível mapear certas características sobre a sequência histórica de diferenciação e correlacionar as características com variações climáticas ou biogeográficas. Tudo isso permite ajudar no entendimento dos processos, ou mecanismos, subjacentes aos padrões evolutivos que se pretende estudar. De modo geral, o maior interesse está em entender os processos adaptativos, ou seja, aqueles que são resultado da ação de **seleção natural**, que implicam aptidão diferencial, em termos de sobrevivência ou fecundidade, entre linhagens ou populações e que explicam o ajuste do organismo ao ambiente. Mas é possível também tentar avaliar de forma integrada o efeito de outros processos que explicam de forma mais ampla a evolução dos organismos, em termos de **mutações** (dando origem à variabilidade), **migração** (criando padrões geográficos de variação) e **deriva genética** (evolução neutra).

Quando se lida com padrões e processos em grandes escalas de tempo, normalmente os biólogos imaginam eventos que ocorreram há milhares ou milhões de anos, tempo suficiente ou necessário para que certas mudanças aconteçam nas populações e nas espécies. Entretanto, pensando de forma mais ampla, o tempo é relativo ao tipo de organismo que está sendo analisado. A medida comum e mais geral seria o que se pode chamar de “tempo de geração”, ou seja, qual o intervalo entre duas gerações sucessivas? Já que os mecanismos de mudança evolutiva ocorrem durante a reprodução, essa é a medida apropriada de tempo, ou seja, intervalo entre dois

eventos reprodutivos na população. Há muitos detalhes técnicos para estimar o tempo de gerações, mas de qualquer modo é possível avaliar que o tempo de geração da espécie humana (digamos, 15-20 anos) é bem maior do que o tempo de geração de algumas espécies de pequenos animais ou plantas, que podem ser anuais (ou seja, a cada ano temos uma nova geração), de um mosquito (alguns poucos meses) ou de uma bactéria ou vírus (alguns minutos). Ou seja, o tempo de uma geração entre espécies diferentes pode ser considerado um tempo relativo e que o tempo absoluto (em horas, dias, meses, etc) dessa mesma uma geração varia dentro desse conjunto de organismos. Sendo assim, em alguns casos será possível observar a evolução ocorrendo em um curto espaço de tempo.

Como os coronavírus replicam-se de forma assexuada dentro de uma célula, criando mais e mais cópias do seu genoma (diferente dos animais e outros organismos celulares em que o material genético é composto de DNA, em alguns vírus, como os coronavírus, o genoma é composto de RNA de fita simples +), na realidade, quando o SARS-COV-2 infecta uma pessoa observa-se ali milhões de eventos reprodutivos acontecendo simultaneamente e sequencialmente, de modo que o tempo de geração é realmente muito curto (minutos). Isso significa que eventos evolutivos que levariam milhares de anos para acontecer em animais ou plantas, acontecem em questão de dias, no máximo semanas, para os vírus. Realmente é isso que os pesquisadores estão observando desde o início da pandemia, ou seja, a evolução em ação.

O objetivo deste ensaio é apresentar dados e informações acumuladas durante a pandemia para discutir, em um contexto mais amplo, diferentes processos evolutivos, vinculado a um contexto atual e cotidiano dos estudantes, com o intuito de fornecer material para ser utilizado em trilhas de aprendizagem interdisciplinar (ecologia, evolução e genética), contribuindo para o processo de ensino e aprendizagem.

Seleção Natural - processo em que ocorre sobrevivência e reprodução diferencial de determinados indivíduos devido às diferenças no fenótipo.

Mutações - alterações que ocorrem no material genético dos organismos. É um mecanismo evolutivo que promove variabilidade genética.

Migração - movimento de alguns indivíduos, ou de seus gametas, entre diferentes subpopulações de uma espécie. O movimento resulta no que chamamos de fluxo gênico e tende a diminuir as diferenças genéticas entre as subpopulações.

Deriva genética - mudanças nas frequências dos alelos de um determinado locus devido à amostragem aleatória. A amostragem aleatória pode fazer com que determinados alelos desapareçam da população ao longo do tempo ou mesmo que alelos, inicialmente raros, passem a ser mais frequentes.



Sequenciamento de genomas e monitoramento da evolução de SARS-CoV-2

O evento inicial que marcou o transbordamento (*spillover*) do SARS-CoV-2 na China ocorreu em algum momento no final de 2019, e isso envolveu certamente uma adaptação para um tipo diferente de hospedeiro, a espécie humana.

Comparando as sequências dos genomas de SARS-Cov-2 ao redor do mundo, é possível reconstruir uma **filogenia**, relações evolutivas ou de ancestralidade entre as linhagens e, além disso, mapear as frequências dessas linhagens para avaliar sua origem em termos geográficos (Figura 1). No mundo todo, os genomas de SARS-CoV-2 que foram e estão sendo sequenciados são normalmente depositados em um banco de dados público de genomas chamado GISAID e, dessa forma, essas informações podem estar disponíveis para outros pesquisadores ao redor do mundo, auxiliando na compreensão da importância de alguns processos evolutivos. Na parte “A” da figura 1, obtida da plataforma GISAID, observa-se uma filogenia gerada a partir das sequências de genomas de SARS-CoV-2, amostradas em diferentes datas. A partir da filogenia é possível perceber grupos e subconjuntos de genomas, que formam **clados**. Atualmente, são reconhecidos 12 principais clados ao redor do mundo e sua frequência, por país, pode ser observada na parte B da figura 1. Além disso, na parte C da figura 1 é possível observar a proporção de cada clado ao longo do tempo.

Os primeiros clados que surgiram em 2019, no começo da pandemia, quase não aparecem na porção mais à direita da figura, indicando que, efetivamente, o vírus evoluiu desde o seu surgimento e que as frequências das linhagens foram mudando ao longo do tempo. Em síntese, portanto, essas linhagens são caracterizadas por **MUTAÇÕES** diferentes que ocorreram ao longo do tempo e que se espalham

na população como consequência dos outros processos evolutivos (seleção natural, deriva genética e migração). O primeiro passo para as análises, nesse sentido, é entender o genoma do SARS-Cov2 para avaliar como e onde essas mutações acontecem.

Uma primeira versão da sequência do genoma de SARS-CoV-2 foi publicada em março de 2020 na revista científica *Nature* e teve a sua importância discutida em uma matéria na revista *Superinteressante* – ver “Para saber mais” ao final desse texto). Essa sequência foi obtida a partir de um conjunto de genomas de RNA presentes no fluido broncoalveolar de um paciente que apresentava síndrome respiratória grave em dezembro de 2019. A abordagem de sequenciar uma coleção de genomas presentes em uma determinada amostra é chamada de metagenômica e permitiu identificar a presença de um novo genoma de vírus da família Coronaviridae, 89,1% similar ao genoma de SARS-CoV (vírus causador da doença SARS que ocorreu na China em 2002), permitindo, dessa maneira, definir aquele que foi denominado SARS-CoV-2, o novo coronavírus.

Com a primeira sequência do genoma, obtida para SARS-CoV-2, foi possível identificar que o seu genoma é composto de 29.903 bases nitrogenadas e apresenta dois conjuntos principais de genes: os complexos de poliproteínas chamados de pp1a e pp1b, e as proteínas constituintes da partícula viral, conforme pode ser observado na Figura 2. Os complexos de poliproteínas são quebrados em proteínas menores que participam dos processos de transcrição de genes e da replicação do genoma, enquanto os genes de constituição da partícula viral codificam proteínas que formarão glicoproteínas de membrana (M), proteínas do capsídeo (N), proteínas do envelope (E) e proteínas *Spike* (S), como pode ser observado na Figura 2. Dentre essas proteínas, a *Spike* é particularmente importante por causa de sua função biológica, pois está presente na parte mais externa da partícula viral e confere aos coronavírus o aspecto de coroa que se popularizou nas representações gráficas do vírus na mídia. A proteína *Spike* é responsável por mediar a ligação entre a partícula do vírus e a célula do hospedeiro, consequentemente, promover a infecção.

Filogenia – também chamada de árvore filogenética, é um diagrama que apresenta as relações evolutivas entre um grupo de organismos com base em uma inferência obtida a partir de métodos específicos.

Clado – grupo de linhagens ou organismos que possivelmente tenham evoluído de um mesmo ancestral comum.



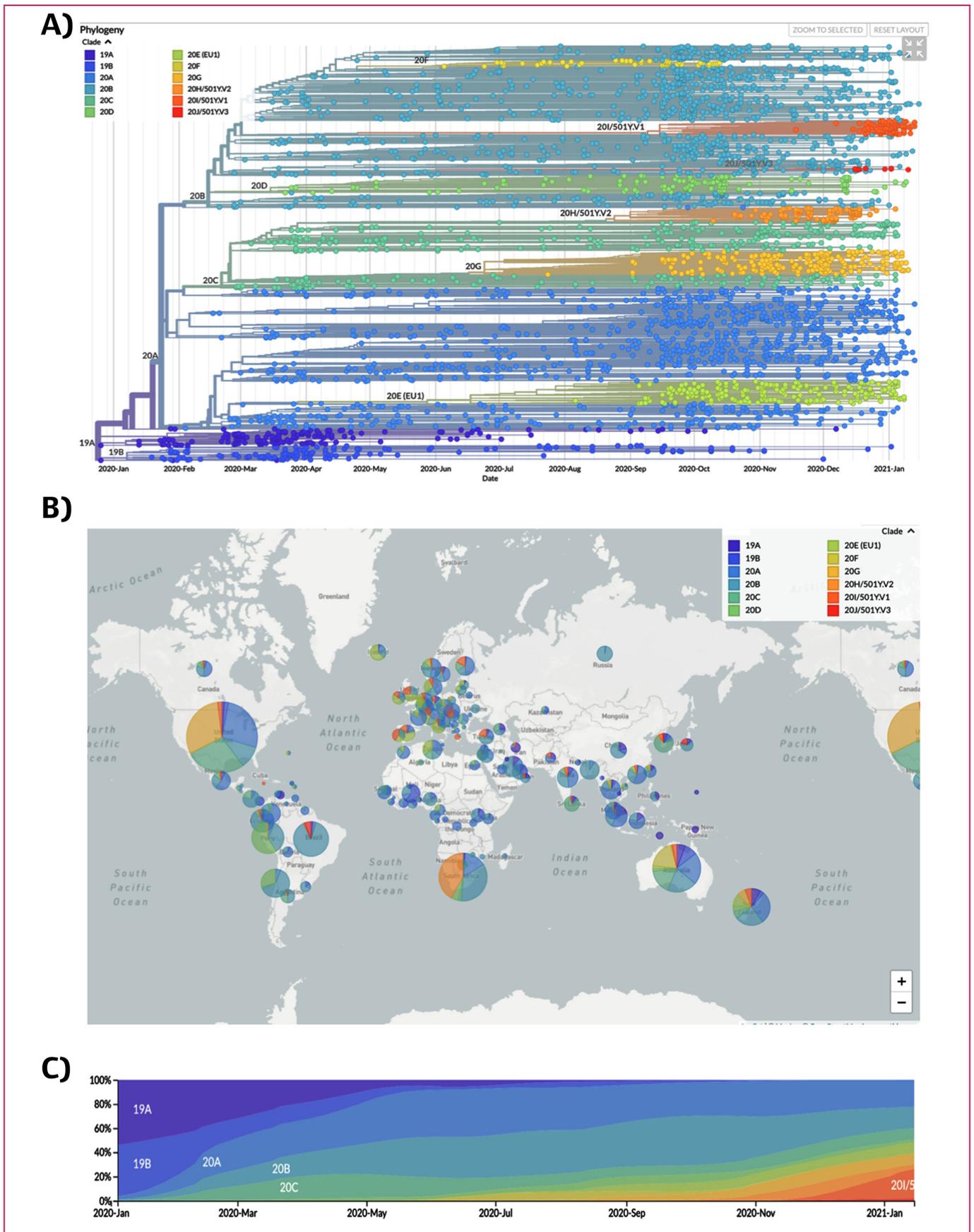
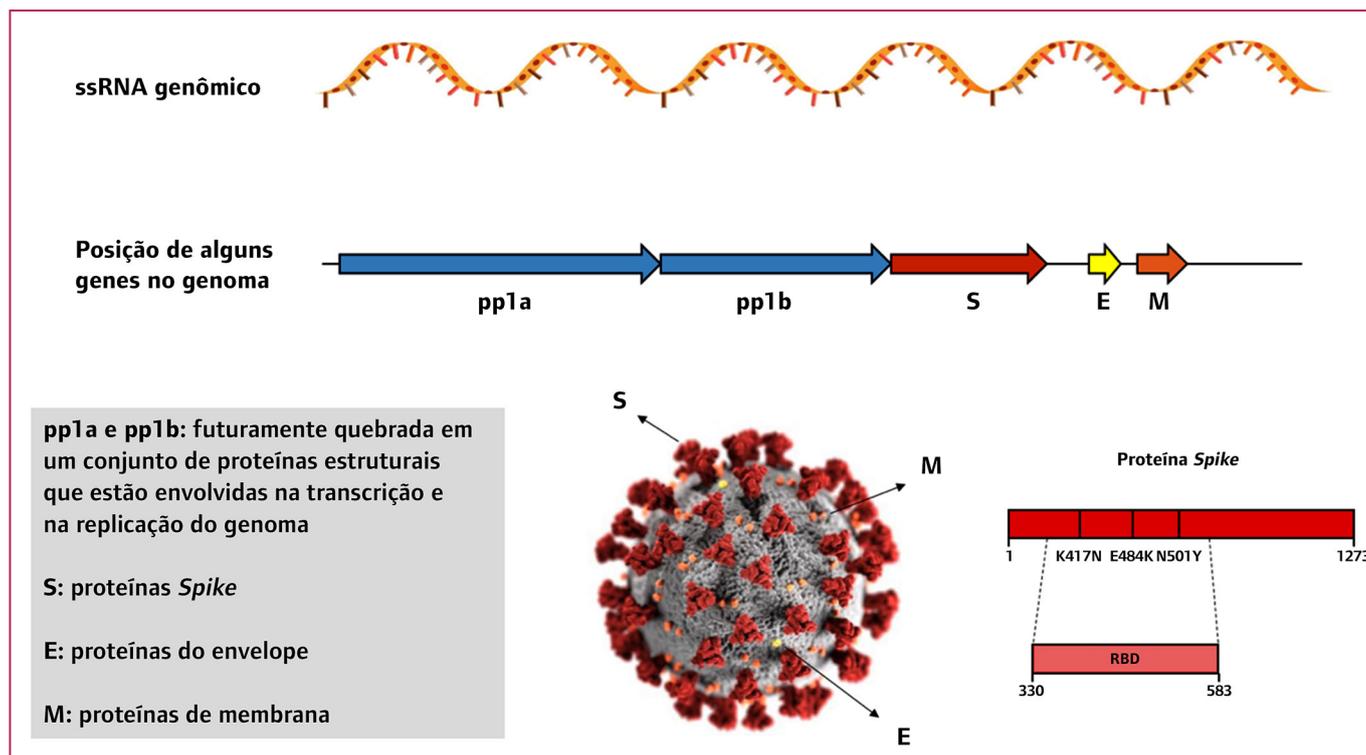


Figura 1. Principais clados formados na análise filogenética de seqüências genômicas de SARS-CoV-2 entre janeiro de 2020 e janeiro de 2021. **A)** Amostras sequenciadas plotadas na filogenia seguindo a data de amostragem e coloridas conforme o clado a que pertencem; **B)** Distribuição da proporção de cada clado por país em 2021 e; **C)** Proporção de cada clado ao longo do tempo. Figuras extraídas da plataforma Nextstrain em 15/01/2021.



**Figura 2.**

Representação da posição de alguns genes que compõem o genoma de SARS-CoV-2. ssRNA: *single strand* RNA (RNA de fita única, em português) e foco na proteína Spike, constituída de 1273 aminoácidos, sendo os aminoácidos 330 a 583 os que formam a região chamada de **RBD**, e onde as mutações K417N, E484K e N501Y se encontram.

RBD - do inglês, *receptor-binding domain*, ou domínio de ligação ao receptor. É uma região da proteína Spike de SARS-CoV-2 que se liga fortemente aos receptores da enzima conversora de angiotensina 2 (ACE2) do hospedeiro. O vírus utiliza da ligação entre a proteína Spike e o receptor ACE2 para promover a infecção da célula do hospedeiro.

Refletindo sobre a função da proteína Spike, não é difícil concluir que é um dos principais alvos para estudos no desenvolvimento de vacinas, terapias utilizando anticorpos e testes para diagnóstico baseados em antígenos. Mas, além das questões relacionadas com prevenção, diagnóstico e tratamento, é importante avaliar o gene envolvido (S) no que pode ser chamado de monitoramento genético da epidemia, que nada mais é que avaliar a evolução da estrutura do genoma do vírus ao longo do tempo. Basicamente, o monitoramento é feito pelo sequenciamento e análise de genomas de SARS-CoV-2 oriundos de pacientes de diferentes localidades e em diferentes datas.

No Brasil, um primeiro esforço para sequenciar genomas de SARS-CoV-2 de pacientes daqui foi publicado na revista científica *Science* em setembro de 2020. Neste estudo, foi possível detectar quais eram os grupos de genomas que circulavam no país nos primeiros meses da pandemia, além de avaliar o efeito de intervenções não farmacológicas (fechamento de escolas, por exemplo) no avanço e disseminação do vírus no país. O primeiro estudo foi extremamente importante para traçar estratégias de enfrentamento ao coronavírus. No entanto, é preciso ressaltar que o esforço de sequenciamento do país como um todo no ano de 2020

foi muito baixo, pelo menos quando comparado com outros países emergentes, como a África do Sul, que depositou 2.882 sequências no banco de dados GISAID no final de 2020. Essa quantidade de genomas depositados era quase o dobro do valor brasileiro, embora o Brasil já apresentasse cerca de sete vezes mais casos confirmados que a África do Sul.

As mutações que podem ser observadas nos genomas do SARS-Cov2 ao longo do tempo e que caracterizam as diferentes linhagens implicam na substituição de uma determinada base nitrogenada do RNA por outra, ou pela inserção ou deleção de uma ou mais bases em uma determinada posição. De acordo com o código genético, três bases nitrogenadas (códon) codificam para um único aminoácido. As mudanças nesses códons (trincas de bases nitrogenadas), por efeito de mutação pontual durante a replicação, podem refletir em troca de aminoácidos, alterando, dessa forma, a estrutura primária da proteína. Em outras palavras, uma mutação pode gerar uma nova versão da proteína, diferente da conhecida anteriormente e que, por ter uma modificação em sua sequência de aminoácidos, pode vir a ter uma modificação também em seu funcionamento, tornando-se, por isso, uma mutação que pode alterar o valor adaptativo daquela linhagem.

De modo geral, as variações que podem ser observadas na filogenia do SARS-COV-2 parecem ocorrer no contexto do que se chama evolução neutra, como proposto pelo cientista japonês Motoo Kimura, ainda nos anos 1970, como no caso do DNA ou RNA, considerando o coronavírus. A maior parte das mutações não causa mudanças na aptidão, de modo que o aumento ou a diminuição de frequência dessas mutações na população ocorre por acaso e se fixa por processos estocásticos relacionados à deriva genética (a deriva é mais rápida em pequenas populações, ou pode explicar por que certas variantes ocorrem em um lugar e não em outro, pelo chamado “efeito fundador”, ou seja, simplesmente porque alguém contaminado com aquela linhagem chegou naquele lugar antes, por migração). Na realidade, existe uma discussão se há mesmo variações neutras, no sentido estrito, e a literatura traz o termo evolução **quase-neutra**, no sentido de que as diferenças de aptidão seriam muito pequenas, e nesse caso, processos de deriva genética podem ser mais importantes do que essa diferença no sentido de aumentar ou diminuir a frequência das mutações em diferentes lugares.

Assim, em princípio, as muitas mudanças nas linhagens do coronavírus, que podem ser observadas na Figura 1, ocorrem de forma neutra, e aparentemente não tiveram impacto muito grande sobre a sua transmissão ou patogenicidade. Entretanto, não é fácil de avaliar tal questão devido aos efeitos simultâneos de muitas variáveis que podem criar fatores que geram dúvidas e tornam as análises difíceis como a localização de diferentes linhagens dominantes que podem variar em termos de comportamento da população, política governamental, fatores sócio-econômicos, estrutura etária e, até mesmo, sistema de registro dos dados.

Variações Adaptativas

Entretanto, algumas mutações que começaram a aparecer já no final de 2020, quase 1 ano após o início da pandemia, parecem ser diferentes em termos de transmissão. Por exemplo, em Manaus foi detectada uma nova linhagem chamada de P.1 (Gamma), que é descendente da linhagem B.1.1.28, descrita primeiramente por

Nuno Farias e colaboradores, mas também detectada e descrita de forma independente por Felipe Naveca e colaboradores. O que chamou a atenção na P.1. foi a presença de mutações no gene codificador da proteína *Spike*, que receberam os nomes de S: K417N, S: E484K e S: N501Y. O “S” no nome delas se refere ao gene em que as mutações ocorrem, o gene S, que codifica a proteína *Spike*. Os demais caracteres referem-se à posição da mutação na proteína e a que tipo de alteração de aminoácidos é observada em relação ao genoma de referência. Por exemplo, no caso na N501Y, o aminoácido que ocupa a posição 501 da proteína, segundo o genoma de referência, é uma *Asparagina* (simbolizada pela letra N) e no genoma recém-sequenciado esse aminoácido foi substituído por uma *Tirosina* (simbolizado pela letra Y). Tais mutações apareceram em 13 dos 31 genomas sequenciados de amostras coletadas entre 15 e 23 de dezembro de 2020, mas não estavam presentes em 26 genomas de amostras coletadas entre março e novembro, indicando que o surgimento delas é recente no referido local. Outras mutações e linhagens têm aparecido desde então, demonstrando o potencial evolutivo do vírus.

A mutação N501Y já havia sido observada em outras linhagens, como a B.1.1.7 (Alpha) na Inglaterra. A mutação N501Y é bastante importante pois ocorre em uma porção superior da proteína *Spike*, em uma região de interação com um receptor nas células humanas. É como se essa região de interação, chamada de RBD (Domínio de ligação ao receptor), fosse um quebra-cabeça em que a forma típica do SARS-CoV-2 formasse um encaixe meio frouxo, enquanto a versão com a mutação N501Y já fosse mais ajustada e, portanto, aumentasse as chances de uma infecção bem sucedida. A mutação observada no gene S, além de outras como a E484K, parece ter evoluído de forma independente em diferentes linhagens em outros países do mundo como Austrália, Dinamarca, Japão, Holanda, África do Sul, País de Gales, Estados Unidos e, como já mencionado, na linhagem B.1.1.28 aqui no Brasil. A mutação E484K pode ser particularmente importante porque parece estar ligada também a uma maior probabilidade de reinfeção, o que apresenta muitas implicações epidemiológicas (ver próxima seção).

Evolução *quasi-neutra* -

teoria que propõe que algumas mutações não são estritamente neutras (que não estão sob ação da seleção natural), mas que podem ser levemente deletérias. Em populações grandes o suficiente, mutações levemente deletérias tendem a serem eliminadas por seleção purificadora enquanto que, em populações pequenas, elas são mais toleradas e se comportam de maneira neutra. Essa teoria foi desenvolvida por Tomoko Ohta a partir da teoria neutra proposta por Motoo Kimura.



Assim, quando se avaliam as mutações N501Y, E484K e as outras no gene S, é possível definir mais claramente que elas envolvem processos de adaptação do vírus, que se torna, evolutivamente falando, mais eficiente em invadir as células e, com isso, aumenta o seu poder de disseminação. Como o genoma viral é formado por uma sequência de RNA relativamente simples – mas altamente eficiente e especializada – dificilmente pode-se dizer que é improvável que as mesmas mutações evoluam de forma independente, em diferentes lugares.

Dado bastante tempo (no sentido do vírus) e uma enorme população de infectados, chega a ser quase inevitável que essas convergências adaptativas apareçam (tecnicamente, em análise filogenética, essas mutações N501Y e E484K, por exemplo, são **homoplasias**).

Homoplasia é um determinado evento em que uma característica ou mutação foi ganha ou perdida de forma independente em diferentes linhagens ao longo da evolução. Ela pode surgir tanto por pressão de seleção atuando na adaptação das linhagens quanto por efeitos da deriva genética.

Aparentemente, é isso que é possível detectar nas várias linhagens que estão aparecendo no Brasil (Manaus), na Inglaterra, na África do Sul e em outros países. E muitas outras coincidências devem acontecer, mesmo nas mutações neutras, mas sabe-se pouco sobre tal fato exatamente porque o número de genomas sequenciados é relativamente pequeno e não está distribuído uniformemente pelo mundo. Sendo assim, embora existam linhagens diferentes do coronavírus, que podem ser reconhecidas como "clados" (grupos formados por todos os descendentes de um mesmo evento evolutivo, no caso uma ou mais mutações), tais linhagens podem ter uma ou mais mutações que aparecem de forma independente, e isso é especialmente plausível se essas mutações forem adaptativas.

Um ponto importante, mais conceitual: é comum que as pessoas pensem nas adaptações biológicas quase que de forma intencional, como já discutido antes, mas de fato o que temos é simplesmente um processo de seleção natural como pensado por Darwin e Wallace em meados do século XIX. No caso do coronavírus, processo acontece dentro do hospedeiro. Se há pacientes com diferentes respostas imunes, em algum momento uma mutação que aconteceu ao acaso, modificando os aminoácidos gerados (como explicado anteriormente), que se mostre mais funcional no sentido de entrar na célula, esses pacientes particularmente vão transmitir mais e a nova mutação começará a aumentar de frequência na população.

Observa-se, desse modo, um processo de seleção natural dentro do organismo (população viral dentro do hospedeiro humano) e que se propaga para o nível populacional do hospedeiro (população humana mundial). Então, é importante entender que a mutação N501Y, por exemplo, não foi direcionada para infectar mais, em termos de intencionalidade, ou seja, o vírus não está evoluindo para se tornar mais infeccioso, isso é uma consequência que aparece após a reprodução diferencial da linhagem que carrega a nova mutação. Vale ressaltar que esse é um dos aspectos mais difíceis de ser ensinado e entendido conceitualmente em biologia evolutiva. O que acontece, de fato, é que qualquer mutação que apareça ao acaso e afete certas características do vírus, como as que acontecem na porção RBD, podem fazer com que ele seja mais eficiente para entrar na célula humana. Se ele infecta mais as pessoas, a carga viral pode ser maior e, consequentemente, a chance de haver transmissão e infecção de outra pessoa é mais alta. Como a sequência do vírus SARs Cov-2 é curta e relativamente simples, eventualmente mudanças ao acaso no genoma dele podem ser adaptativas no sentido positivo, para o vírus.

Outra preocupação tem sido descobrir se as mutações têm potencial para driblar a imunização gerada pelas vacinas ou por infecções anteriores. Um estudo preliminar com experimentos *in vitro* apresentou resultados alarmantes de que as mudanças naturais, que têm surgido em SARS-CoV-2, teriam o potencial para escapar de uma resposta imune eficaz e que, portanto, deve-se buscar por vacinas que consigam controlar as formas emergentes do vírus. Se esse for o caso, haverá um esforço contínuo no sentido de aprimorar as vacinas, como vemos no caso da gripe (influenza). Por outro lado, as empresas que desenvolveram algumas das vacinas (no caso a *Pfizer*) insistem que é pouco provável que essas mutações afetem a eficácia das mesmas, considerando que elas foram experimentalmente testadas com formas diferentes do coronavírus. Mais investigações sobre esse tema ainda são necessárias e, de qualquer modo, é preciso focar na imunização do maior número de pessoas possíveis com as vacinas disponíveis, mesmo que a eficiência delas possa ser parcialmente reduzida (se é que isso acontecerá) para novas linhagens.



Modelagem epidemiológica e o aumento da capacidade de transmissão do vírus

A expansão de uma epidemia normalmente é descrita por modelos epidemiológicos tipo SIR (de Suscetíveis, Infectados e Recuperados), que descrevem a trajetória desses três conjuntos na população de hospedeiros. A lógica do modelo mais simples é a seguinte: um patógeno (no nosso caso, o coronavírus) chega na população com um número de suscetíveis S inicial (S_0) igual ao tamanho da população (N), assumindo que não existe uma imunidade natural. À medida que o patógeno começa a infectar os hospedeiros, o número de suscetíveis diminui, assumindo que não existe reinfeção (ou seja, os recuperados – ou mortos – não voltam ao conjunto de suscetíveis). Em termos de modelo, a transferência dos indivíduos do compartimento de suscetíveis para os infectados é dada pela taxa de infecção (α), enquanto a taxa γ determina a taxa de recuperação desses infectados. A razão entre as duas taxas (α/γ) é o chamado número reprodutivo R_0 , que pode ser pensado também como número médio de novos infectados que surgem a partir de um infectado inicial. Quando o R_0 é maior do que 1, significa que as infecções ocorrem mais rápido do que as recuperações, de modo que a epidemia está, de fato, acelerando.

Na prática, o R_0 é a condição inicial das transmissões, mas à medida que o número de suscetíveis diminui, a probabilidade de novas infecções começa a diminuir, de modo que há uma relação inversa entre a proporção de suscetíveis e o número de transmissões. Nesse modelo, portanto, à medida que o tempo passa o R efetivo (R_e) é dado por $R_0 * S/N$. Com isso, cria-se um comportamento logístico no número de infectados e o número de infectados por tempo forma uma onda, um crescimento acelerado. A epidemia cresce de forma acelerada, chega em um pico e depois começa

a diminuir (pois o número de pessoas na população que ainda são suscetíveis é pequeno). No caso do coronavírus, o R_0 aparentemente estaria em torno de 2,5 ou 3,0, mas certamente esse número se reduz a partir do momento que se implementam medidas de distanciamento e isolamento social ou quarentena e depende de uma série de fatores.

A importância epidemiológica das mutações na proteína *Spike* pode ser entendida exatamente no contexto do R_0 , pois na Inglaterra e em outros países da Europa estima-se que essa nova linhagem com a mutação N501Y possua um R_0 entre 50% e 70% maior do que a linhagem originalmente mais comum até então. A redução do R_0 para R_e pode ser ainda menor se houver uma possibilidade maior de reinfeção por causa da mutação E484K, pois a nova linhagem pode ser mais eficiente em termos de se contrapor às defesas imunológicas do hospedeiro. Desse modo, ela se expande mais rápido e requer uma atitude mais enérgica das autoridades na tentativa de controlá-la. É importante destacar que embora o R_0 seja sempre entendido como uma **medida ecológica** de crescimento da população de infectados, na realidade ele termina sendo exatamente a medida de **aptidão darwiniana** do SARS-CoV-2, assumindo uma correlação entre a multiplicação no nível celular e uma maior transmissão entre os hospedeiros.

Para ilustrar o processo adaptativo é possível utilizar um modelo SIR simples, começando com uma população humana vivendo em uma cidade de 1 milhão de habitantes. Nessa cidade chegam 10 indivíduos infectados pelo coronavírus e começam a transmitir o vírus para outras pessoas. Para simplificar, é possível assumir que rapidamente as autoridades de saúde detectam a chegada da pandemia na cidade, pois existe de fato um evento de migração a partir de outros locais e implementam diferentes medidas de isolamento social que reduzem rapidamente o R_e , logo no início, para algo em torno de 1,25 (ou seja, cada 100 pessoas infectadas transmitem o vírus para outras 125 pessoas), uma redução em torno de 50% em relação ao R_0 . Assumindo também, para fins de simplificação, que o intervalo serial (o tempo entre as gerações) é entre a infecção e o máximo de transmissões, que ocorre em uma semana. Na verdade, o intervalo é

Modelagem epidemiológica

– é a construção de modelos matemáticos que descrevem o comportamento de uma doença na população durante uma epidemia. É uma ferramenta extremamente importante para direcionar práticas de contenção da doença.

Medidas ecológicas de crescimento da população

métricas definidas a partir das taxas de natalidade e mortalidade e sua variação ao longo do tempo.

Aptidão darwiniana

– considera que os organismos melhor adaptados ao meio em que vivem possuem maiores chances de deixar descendentes férteis que os organismos menos adaptados.

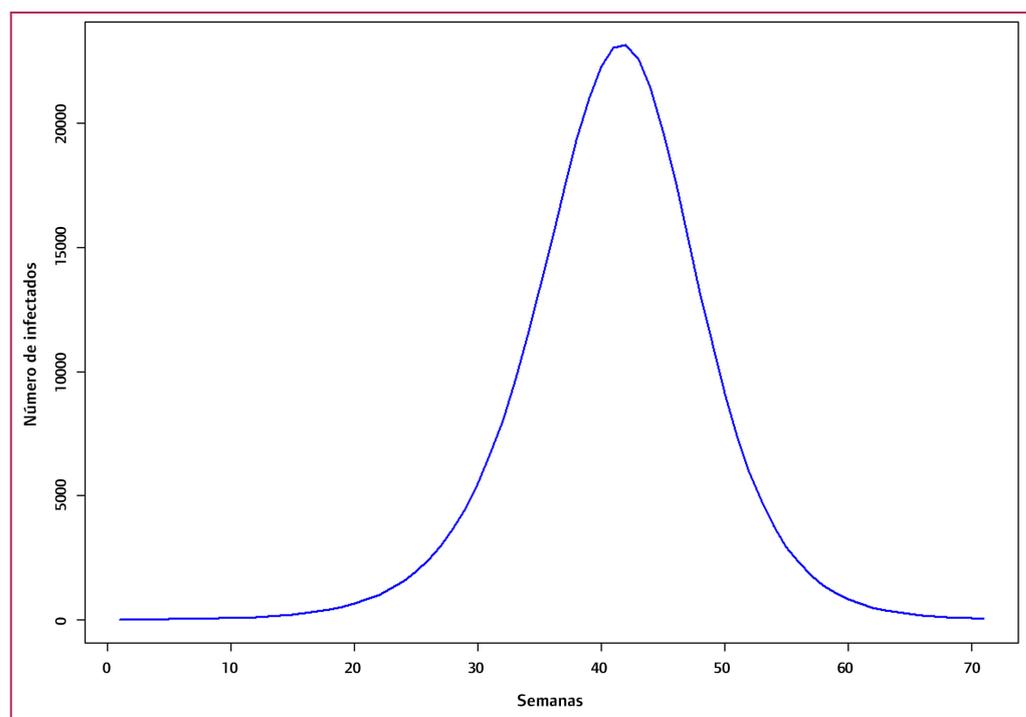


um pouco menor, pois é possível observar um pico de infectados pouco depois de 40 semanas após o início da epidemia, que é finalizado após cerca de 70 semanas, com cerca de 38% da população já tendo sido infectada. Note-se que, nessa simulação simples, está sendo demonstrado apenas o número de infectados, ou seja, todas as pessoas que tiveram contato com

o vírus, incluindo aquelas pessoas assintomáticas. Se houver interesse em avaliar o número de casos graves, hospitalizações ou óbitos, é possível utilizar as frequências desses eventos e estimá-los a partir do número de infectados. Por exemplo, para pensar em letalidade, utiliza-se o IFR (do inglês *Infected Fatality Rate*) que para a COVID-19 está em torno de 0,5%.

Figura 3.

Distribuição do número de infecções por SARS-CoV-2 ao longo das semanas considerando um modelo epidemiológico do tipo SIR (Suscetíveis, Infectados e Recuperados).



Este modelo inicial, entretanto, não considera o fato de que o SARS-CoV-2 está em constante evolução, como discutido anteriormente. Entretanto, se a linhagem com mutações como a N501Y possui uma maior capacidade de transmissão, significa que as duas variantes possuirão dinâmicas diferentes ao longo do tempo e, na perspectiva do vírus, a linhagem com a mutação N501Y é mais adaptada do que a linhagem ancestral. Assim, é possível refazer a simulação apresentada anteriormente e imaginar que uma linhagem com a mutação N501Y chega na cidade na semana 40 e consegue uma capacidade de transmissão 70% maior do que a linhagem original. Neste caso, o que apareceria sob um ponto de vista puramente epidemiológico seria uma 2ª onda de contaminações e casos aparecendo na população (Figura 4A), mas que, na verdade, isso seria em função da diminuição da linhagem representada em azul

e crescimento da linhagem com a mutação N501Y (curva vermelha da Figura 4B), que passou a ter uma trajetória independente da linhagem ancestral. Note que, como a linhagem com a mutação N501Y possui uma maior capacidade de transmissão, o pico da epidemia é mais alto.

É importante ressaltar também que, no modelo simples, o número de suscetíveis diminui com o total das infecções da Figura 4A, pois a simulação assume que alguém contaminado com a linhagem ancestral não se infectaria com a linhagem que apresenta a mutação N501Y. Mas, é possível que uma nova mutação seja capaz de infectar pessoas que tenham sido infectadas pela linhagem ancestral, o que parece ocorrer da linhagem com a mutação E484K. Se este realmente for o caso, a curva da linhagem mutante crescerá de forma ainda mais rápida, requerendo medidas mais rígidas de isolamento social e

quarentena, a fim de conter o avanço da pandemia. Não há evidência ainda de que isso acontece, mas nada impede que novas mutações gerem linhagens com maior patogenicidade, e nesse caso, nos modelos, o número de hospitalizações e de óbitos também seria diferente para essa linhagem. Ainda é possível determinar se realmente a 2ª. onda que

passou a ocorrer no Brasil a partir do final de 2020 é função da chegada de linhagens com novas mutações, ou se ela é principalmente fruto do aumento das transmissões por uma diminuição drástica do isolamento social a partir desse período. Em Manaus, o recrudescimento da epidemia, que causou um forte colapso no sistema de saúde no início de

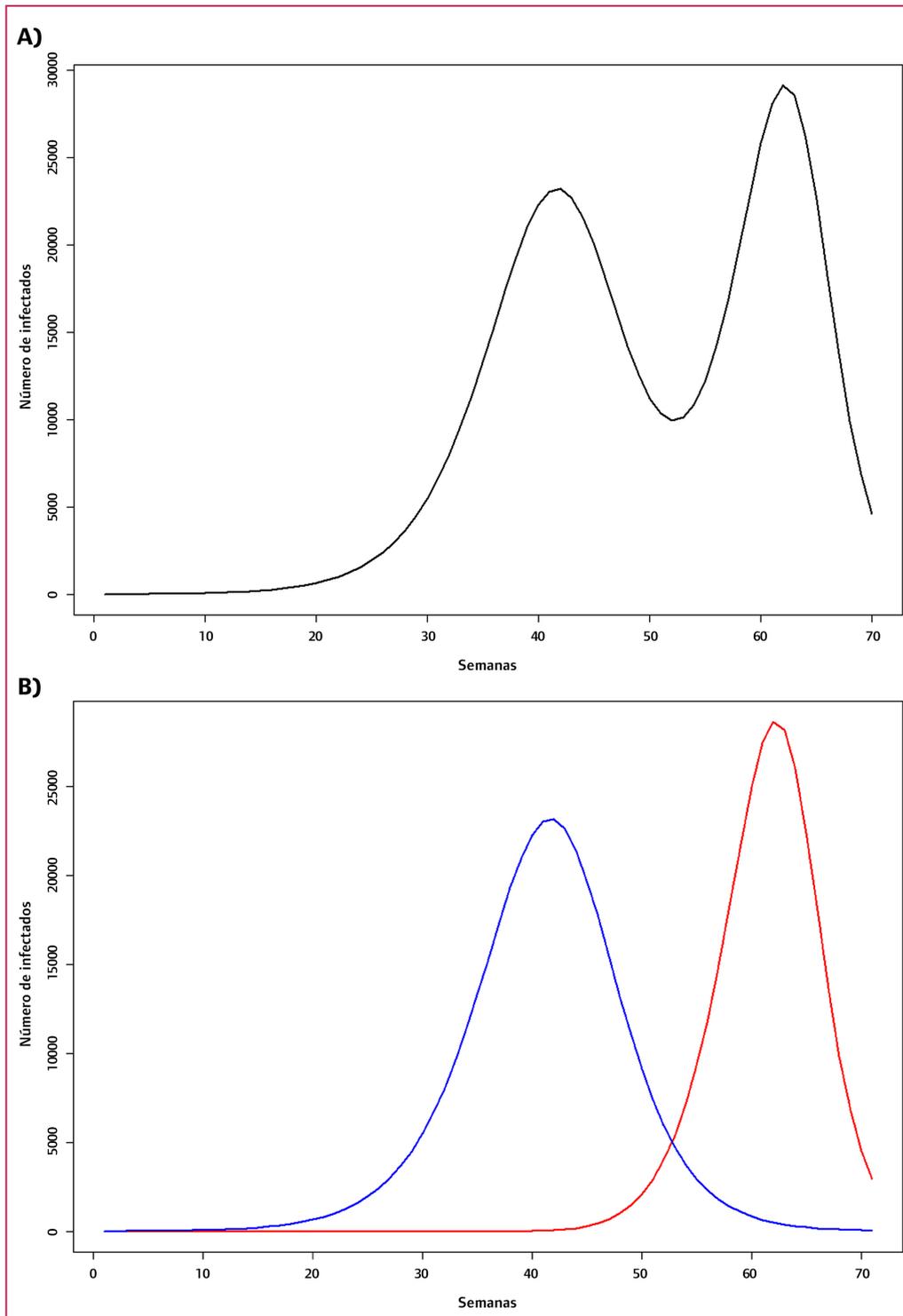


Figura 4. Distribuição do número de infectados por SARS-CoV-2 ao longo do tempo considerando um modelo epidemiológico SIR simples. **A)** O surgimento de uma mutação que confere maior capacidade de transmissão gerando uma segunda onda de infecções; e **B)** Linha azul é a distribuição da primeira onda de infecções e, a vermelha, a da segunda gerada pela nova variante.

2021, pode estar associado à linhagem P1 de forma mais clara. Na Inglaterra e em outros países da Europa a associação de um maior crescimento da pandemia com um aumento de frequência da linhagem com a mutação N501Y é mais clara, mesmo após o início da vacinação em dezembro de 2020.

Considerações finais

Em um sentido mais amplo, é importante entender que, para além da própria dinâmica ecológica e epidemiológica envolvida na pandemia, que é mais amplamente discutida nos meios de comunicação, existe também, na perspectiva do vírus uma dinâmica evolutiva subjacente que pode ter consequências aplicadas importantes. Assim, é preciso continuar estudando os genomas do SARS-CoV-2 e entender suas mudanças, principalmente em busca de linhagens que possam apresentar variações adaptativas com implicações epidemiológicas. Essa é uma importante contribuição científica que a Genética e a Biologia Molecular têm dado para mitigar os efeitos da pandemia, algo que também precisa ser ressaltado em um contexto de ensino.

Certamente é possível também tirar muitas lições importantes da pandemia em termos de ensino de evolução. Em síntese, um professor pode mostrar que as mutações que acontecem no coronavírus em geral são neutras e permitem rastrear a expansão da pandemia a partir da sua fixação em diferentes populações (deriva genética e processos relacionados), que são alcançadas por processos contínuos de migração. Entretanto, em algumas situações essas mutações podem ser adaptativas, estando sujeitas, portanto, à ação da seleção natural. Esse efeito se reflete em uma mudança de parâmetros epidemiológicos relacionados principalmente a uma maior capacidade de transmissão e um aumento no número de casos e óbitos, e isso já foi observado em alguns lugares do mundo. Espera-se que a ilustração da dinâmica evolutiva a partir do que se conhece sobre

a origem e expansão da pandemia da COVID-19, melhore a compreensão geral sobre a evolução e seus processos, pois é possível demonstrar a existência dessa dinâmica em um contexto mais próximo do cotidiano das pessoas, em relação a um assunto que tem sido amplamente discutido no caso da atual pandemia que gerou um grande impacto social.

Para saber mais:

- BADAIN et al., (2021). *Observatório Covid-19 BR*. Disponível em: <<https://covid19br.github.io/analises.html?aba=aba15>> Acessado em: 15/01/2021.
- CANDIDO, D. S. et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science*, v. 369, n. 6508, p. 1255-1260, 2020.
- CARBINATTO, B. (2020). *Por que sequenciar o genoma do novo coronavírus é importante*. Disponível em: < <https://super.abril.com.br/ciencia/por-que-sequenciar-o-genoma-do-novo-coronavirus-e-importante/>> Acessado em:
- CORUM, J.; ZIMMER, C. (2021). *Inside the B.1.1.7 Coronavirus Variant*. Disponível em: <<https://www.nytimes.com/interactive/2021/health/coronavirus-mutations-B117-variant.html>>, Acessado em: 15/01/2021.
- FARIA, N. R. et al. (2021). Genomic characterization of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings. Disponível em: < <https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586> >, acessado em: 15/01/2021.
- GISAID (2021). Genomic epidemiology of hCoV-19. Disponível em: <<https://www.gisaid.org/phylogenetics/global/nextstrain/>> Acessado em: 15/01/2021.
- NAVECA, F. et al. (2021). Phylogenetic relationship of SARS-CoV-2 sequences from Amazonas with emerging Brazilian variants harboring mutations E484K and N501Y in the Spike protein. Disponível em: < <https://virological.org/t/phylogenetic-relationship-of-sars-cov-2-sequences-from-amazonas-with-emerging-brazilian-variants-harboring-mutations-e484k-and-n501y-in-the-spike-protein/585>>, acessado em: 15/01/2021.
- WU, F. et al. A new coronavirus associated with human respiratory disease in China. *Nature*, v. 579, n. 7798, p. 265-269, 2020.

