

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE BIOCÊNCIAS  
PROGRAMA DE PÓS GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

**TMPRSS2 E OUTRAS SERINO PROTEASES TRANSMEMBRANA  
TIPO II (TTSPS): SINAL DE UMA CORRIDA ARMAMENTISTA DE  
LONGO PERÍODO ENTRE VÍRUS E MAMÍFEROS  
PLACENTÁRIOS**

Dissertação para a obtenção de título de mestre em Genética e Biologia Molecular

Matheus de Castro Nóbrega

Orientação: Maria Cátira Bortolini

Brasília/Porto Alegre, Junho de 2022

Matheus de Castro Nóbrega

**TMPRSS2 E OUTRAS SERINO PROTEASES TRANSMEMBRANA  
TIPO II (TTSPS): SINAL DE UMA CORRIDA ARMAMENTISTA DE  
LONGO PERÍODO ENTRE VÍRUS E MAMÍFEROS  
PLACENTÁRIOS**

Dissertação apresentada ao Programa de Pós-Graduação *Strictu Sensu* em Genética e Biologia Molecular da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de mestre em Genética e Biologia molecular.

Brasília/Porto Alegre, Junho de 2022

**Brasília/Porto Alegre, Junho de 2022**

Este trabalho foi realizado no Laboratório de Evolução Humana e Molecular do Departamento de Genética da Universidade Federal do Rio Grande do Sul com o fomento do Conselho Nacional de Pesquisa e Desenvolvimento (CNPq) e pela CAPES.

## **Agradecimentos**

Agradeço a minha família em especial aos meus pais João e Elânia por todo o apoio e suporte durante essa etapa da minha vida em particularmente durante a pandemia que me desanimou muito a princípio e também ao meu irmão Gustavo por todo o interesse quando decidia explicar ou mostrar alguma coisa do trabalho, seu entusiasmo com cada nova figura me deixou mais animado com meu próprio desenvolvimento. Agradeço também a todos os colegas do Laboratório de Evolução e Humana e Molecular pelo acolhimento, mesmo a distância, em particular á Bibiana Fam, pela disponibilidade, ao Yuri Yépes, por toda a paciência e pelas diversas aulas em que me ensinou sobre novas metodologias com doses iguais de coleguismo e experiência, e principalmente a minha orientadora Maria Cátira, pela atenção, interesse e por sua visão peculiar sobre o projeto, sem a colaboração dos três esse trabalho não teria sido realizado. Também agradeço às instituições de fomento Capes e CNPq pela bolsa, e a UFRGS e ao PPGBM pela oportunidade de realizar esse mestrado.

“O que sabemos é uma gota,  
O que ignoramos é um oceano”

Sir Isaac Newton

## Contents

<b>Resumo .....</b>	<b>7</b>
<b>Abstract .....</b>	<b>9</b>
<b>1. Introdução .....</b>	<b>10</b>
<b>2. Objetivo .....</b>	<b>21</b>
<b>3. Resultados- Manuscrito .....</b>	<b>22</b>
<b>4. Considerações finais .....</b>	<b>22</b>
<b>5. Referências .....</b>	<b>25</b>

## Resumo

O gene da Serina Protease Transmembrana 2 (*TMPRSS2*), localizado no cromossomo 21q22.3, codifica uma proteína com o mesmo nome, sendo membro da família das Serino Proteases Transmembrana Tipo II (TTSPs). A *TMPRSS2* humana é normalmente relacionada a resposta a andrógenos, mas foi cooptada pelo SARS-CoV-2 para clivar a glicoproteína viral Spike, e assim permitir a infecção da célula hospedeira. Apesar da existência de outras proteases cooptadas pelo SARS-CoV-2 para realizar essa função, recentes estudos funcionais mostram uma ativação e penetração mais velozes do SARS-CoV-2 em células expressando *TMPRSS2* que naquelas dependentes de outras proteases. De forma a avaliar os padrões evolutivos que moldaram a relação entre a *TMPRSS2* do *Homo sapiens* e a Spike do SARS-CoV-2, buscamos identificar todas as TTSPs presentes no genoma do *Homo sapiens*, visto que o último estudo sobre o tema foi publicado em 2009. Foram encontradas 18 TTSPs pertencentes a 4 subfamílias, resgatando a relação filogenética original das subfamílias TTSPs. Porém quando foram analisados somente os 30 sítios que interagem com a Spike do SARS-CoV-2 um padrão filogenético distinto foi encontrado. Também investigamos a região codificante dos ortólogos do *TMPRSS2* de 182 espécies de mamíferos placentários. A variabilidade interespecífica em 33 sítios pode ser explicada por seleção positiva de acordo com a análise no pacote MEME, sendo que seis desses sítios (299, 340, 389, 413, 431, 438), ou seja, 15%, são reconhecidos como importantes para a interação com o vírus. Esses resultados podem indicar um sinal de uma corrida armamentista entre os coronavírus e os seus potenciais hospedeiros mamíferos. Em outras palavras, esse padrão de variação sugere que a maior

parte da variação entre espécies seja resultados de pressões seletivas que potencialmente vêm moldando a função normal da TMPRSS2 humana e de seus ortólogos nas células das espécies correspondentes. Por outro lado, a Spike viral estaria sendo moldada evolutivamente para se ligar em posições nas proteases dos hospedeiros mamíferos com menos propensão a terem variação promovida por ação de seleção positiva, o que conferiria vantagem ao vírus, pois ele teria menos chance de perder afinidade com o hospedeiro ao mesmo tempo que daria mais chances para saltos zoonóticos.

## Abstract

The Transmembrane Serine Protease 2 (TMPRSS2) gene, located at human chromosome 21q22.3, encodes a protein with the same name member from the type II transmembrane serine proteases (TTSPs). TMPRSS2 is usually related to the response to androgens but is co-opted by the SARS-CoV2 to cleave the viral Spike glycoprotein to infect the host cell. Despite the existence of other proteases co-opted by SARS-CoV2 to that function, recent functional studies show a more rapid activation and penetration of SARS-CoV2 in cells expressing TMPRSS2 than with those in which infection depends on other proteases. In order to assess the evolutionary patterns that shaped the relationship between the *Homo sapiens* TMPRSS2 and SARS-CoV2 Spike, we aimed to identify all TTSPs present in the *Homo sapiens* genome since the previous study with this purpose was published in 2009. One of our goals is to understand better why TMPRSS2 has been evolutionarily co-opted and is preferentially used to cleave Spike. Eighteen canonic *Homo sapiens* TTSPs, grouped in 4 clades were found, rescuing the original phylogenetic relationship of the TTSPs subfamilies. However, when only 30 sites that interact with the Spike of SARS-CoV-2 were used, a distinct phylogenetic pattern was found. We also investigated the coding region of TMPRSS2 orthologs of the 182 species of placental mammals. Using the MEME package, our evolutionary analysis shows that the interspecific variability in 33 sites can be explained by positive selection, six of them (299, 340, 389, 413, 431, 438), that is, 15%, with importance for the interaction with SARS-CoV-2. These results may be a sign of the biological arms race between coronaviruses and their potential mammalian hosts. In other words, this pattern of variation suggests that most of the variation between species is the result of selective pressures that have been shaping the normal function of human TMPRSS2 and its orthologs in the cells of the corresponding species. On the other hand, the viral Spike would be evolutionarily shaped to bind at positions in the proteases of mammalian hosts that are less likely to have variation promoted by positive selection action, which would give the virus an advantage, as it would have less chance of losing affinity with the host while giving more chances for zoonotic jumps.

# 1. Introdução

## 1.1 Contexto geral

No contexto atual de pandemia da COVID-19 um esforço científico global foi empreendido para compreender os diversos aspectos que levaram um vírus a promover uma infecção tão ampla e com um número tão grande de mortes. Como exemplo, uma rápida busca no Google Acadêmico com a palavra chave “COVID-19” resgata ~466.000 publicações (busca no dia 15 de maio de 2022). Visando ajudar nesse esforço, esse trabalho teve como objetivo principal elucidar a história evolutiva da Serino Protease Transmembrana 2 (do inglês Transmembrane Serine Protease 2 or TMPRSS2) uma das proteínas envolvidas na infecção pelo  $\beta$ -coronavírus SARS-CoV-2 causador da COVID-19, e que está envolvida também em outras infecções virais.

A TMPRSS2 é uma proteína codificada por um gene de mesmo nome e faz parte da família das Serino proteases transmembrana tipo 2 (*Type II Transmembrane Serine Protease* - TTSP) com três domínios funcionais: um receptor LDL classe A, um domínio scavenger (*Scavenger Receptor Cysteine-Rich* - SRCR) e um domínio peptidase extracelular S1 (Tripsina) (Hussain *et al.*, 2020).

As TTSPs compõem uma família de proteínas identificada recentemente, em 2001 por Hooper *et al.* Os autores identificaram, à época, 17 membros em mamíferos placentários (Hooper *et al.*, 2001). Em 2009, Bugge *et al.* aumentou o número de proteínas identificadas em humanos e as classificou em 4 subfamílias usando como critério os domínios que cada proteína possui: HAT (human airway trypsin – Tripsina humana de vias aéreas) /DESC (7 genes/proteínas), Hepsina/TMPRSS (7), Matriptase (4), and Corina (1). *TMPRSS2* é membro da subfamília Hepsin/TMPRSS junto com *TMPRSS3*, *TMPRSS4*, *TMPRSS5/Spinesin*, MSP (Mosaic serine protease) e Enteropeptidase.

No entanto, nenhum trabalho mais específico sobre a história evolutiva da TMPRSS2 e sua família pode ser encontrado. Desse modo, o presente estudo busca contribuir para preencher essa lacuna de conhecimento, visando, dentre outras coisas,

avaliar uma potencial e quem sabe antiga guerra armamentista biológica entre vírus e seus potenciais hospedeiros mamíferos.

Nos itens abaixo poderá ser encontrado uma revisão mais detalhada sobre o tema SARS-CoV-2 e COVID-19. Vale destacar, contudo, que uma introdução abrangente, porém focada mais na TMPRSS2, e em proteínas parálogas no *Homo sapiens*, bem como aquelas ortólogas em outros mamíferos placentários, também poderá ser encontrada no item **Resultados- Manuscrito** (página 21), de modo que alguma repetição é inevitável.

## 1.2 A pandemia da COVID-19

A COVID-19 (Coronavirus Disease 2019) é uma doença causada pela infecção do  $\beta$ -coronavírus SARS-CoV-2 sendo caracterizada por um amplo espectro de sintomas. Os indivíduos infectados podem permanecer assintomáticos, apresentar sintomas brandos, apresentar pneumonia de graus leve a grave, bem como apresentar a síndrome respiratória aguda grave, caracterizada por um comprometimento respiratório severo que pode levar a óbito (WHO, 2022). Dados no primeiro ano na pandemia indicavam que os quadros mais severos e as taxas de mortalidade mais altas dessa doença estavam associados a idosos (PADHAN & PRABHEESH, 2021), pacientes imunossuprimidos (ISER *et al.* 2020) e a pacientes com condições preexistentes como hipertensão e problemas cardíacos (ROBINSON *et al.*, 2020, CHEN *et al.*, 2020). Como outros vírus respiratórios, entre os sintomas mais comuns encontravam-se a febre, tosse, dificuldade respiratória, dor muscular e fadiga, bem como uma pouco comum perda severa de olfato e paladar. No entanto, já era perceptível no início da pandemia que um ou mais desses sintomas poderiam estar ausentes (ISER *et al.* 2020). Além do pulmão, logo foi possível detectar que o SARS-CoV-2 também podia afetar coração, sistema digestivo, rins e cérebro (CHEN *et al.*, 2020, NERSISYAN *et al.*, 2020). Devido ao amplo espectro de sintomas e complicações a COVID-19 tem sido reportada desde sua emergência como uma doença multi-sistêmica (Temgoua *et al.*, 2020; Mir *et al.*, 2021). Ainda, por ter sintomas e quadros clínicos semelhantes aos da gripe comum e outros resfriados, juntamente com a presença de portadores assintomáticos, a COVID-19 apresentava um quadro difícil de ser diagnosticado de forma clínica imediata, o que ajudou, inicialmente, sua rápida transmissão pelo mundo.

Em 31 de dezembro de 2019, a Organização Mundial da Saúde (OMS) foi alertada sobre vários casos incomuns de pneumonia grave na cidade de Wuhan, na China. Tratava-se de uma nova cepa de coronavírus que não havia sido identificada antes em seres humanos. Em 30 de janeiro de 2020, depois de um surto que começava a assustar a Itália e a Europa inteira, a OMS decretou que o caso constituía uma “Emergência de Saúde Pública de Importância Internacional”. Desde então a pandemia da COVID-19 tem movimentado tanto os poderes políticos dos vários países, quanto a sociedade de modo geral, causando um impacto negativo nos sistemas de saúde globais, forçando os governos a fechar as fronteiras, restringir viagens, promover “lockdowns” e tomar precauções contra colapsos econômicos (BIEBER, 2020; NICOLA *et al.*, 2020; PADHAN & PRABHEESH, 2021; MEYER *et al.*, 2022). Para se ter uma ideia, o setor primário foi impactado principalmente pela falta de pessoal e pela insegurança em relação aos empregos. Além disso, a deficiência de pessoal para o transporte e verificação dos produtos trouxe grandes consequências para o setor (NICOLA *et al.*, 2020). O setor secundário foi igualmente afetado pela falta de pessoal e o fechamento das fronteiras. A quebra na cadeia de suprimentos, em uma escala nunca vista antes (ROZHKOVA *et al.*, 2022; MEYER *et al.*, 2022), e o afastamento de pessoal levou a um declínio da indústria (NICOLA *et al.*, 2020) assim como a redução na demanda por produtos do setor (MEYER *et al.*, 2022). O setor terciário, por sua vez, foi o setor que provavelmente teve mais áreas afetadas. Os setores de turismo, hoteleiro, aviação, tiveram profunda queda na demanda (MEYER *et al.*, 2022) enquanto os de saúde, farmacêutico e alimentício tiveram um aumento na demanda, seguido de baixa nas vagas para pacientes ou no estoque de produtos e adaptações significativas seja na infraestrutura ou na entrega dos produtos e serviços (NICOLA *et al.*, 2020). De forma ampla os três setores foram afetados pela redução no número de empregos, que conseqüentemente afeta a distribuição de capital, pelo aumento nos custos de transações internacionais, pelo declínio do turismo e dos serviços de modo amplo e geral (PADHAN & PRABHEESH, 2021). Mais recentemente, de acordo com o relatório econômico da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) de setembro de 2021 (OCDE, 2021), o PIB global atual é superior ao nível anterior à pandemia. No entanto, a produção em meados de 2021 ainda era 3,5% menor do que o esperado antes da disseminação do COVID-19, resultando em perdas de empregos e renda (MARTINEZ *et al.*, 2022).

Do ponto de vista social, as medidas de distanciamento, como o fechamento das escolas, elevaram a preocupação com o aumento nos níveis de violência doméstica, que inclui abuso físico, emocional e sexual, inclusive de crianças (NICOLA *et al.*, 2020; ABRAMS *et al.*, 2022). Para os especialistas, um período maior de confinamento significa que pessoas vulneráveis estão mais expostas ao abuso e tem mais dificuldade para buscar ajuda (NICOLA *et al.*, 2020; ABRAMS *et al.*, 2022). O prolongamento da pandemia também levou ao aumento da insegurança alimentar (ABRAMS *et al.*, 2022) e problemas mentais como depressão e ansiedade, advenços que tornaram a situação de vida ainda pior para aqueles em condições sociais de vulnerabilidade (ABRAMS *et al.*, 2022; WIRKNER *et al.*, 2022). Soma-se a esse cenário a discussão intensa e disseminação em escala global de teorias da conspiração, caracterizadas por idéias pseudocientíficos e fantasiosas, que levaram a efeitos sociais bastante negativos como diminuição da confiança nas instituições públicas e ao apoio das medidas de contenção como o distanciamento social e aplicação de vacinas (PUMMERER *et al.*, 2022).

Por outro lado, a emergência mundial também mobilizou a comunidade científico-acadêmica de um modo nunca antes visto. Os esforços têm buscado avançar no conhecimento sobre o coronavírus e como combatê-lo. Para se ter uma idéia da amplitude desse esforço coletivo da comunidade científico-acadêmica, uma rápida busca no PubMed (dia 25/09) com as palavras chaves “SARS-CoV2 e COVID-19” identifica 35.338 artigos publicados em 2020. A mesma pesquisa feita em 2022 (04/05) revelou que no período de 2019 a 2022 o número se elevou para 209.982 artigos científicos. Já no Google Acadêmico, como já comentado, o número salta para ~466.000 publicações. Dentre esses milhares de estudos muitos, incluindo de nosso grupo de pesquisa (FAM *et al.*, 2020; YÉPEZ *et al.*, 2022), focaram na compreensão de como o SARS-CoV-2 era tão eficiente evolutivamente, ou seja, quais eram os meios moleculares que este organismo se apropriou para infectar seres humanos (ver outros exemplos em SONG *et al.*, 2019; ANDERSEN *et al.*, 2020; CAO *et al.*, 2020; LI *et al.*, 2020; MOUSAVIZADEH & GHASEMI, 2020; GHEWARE *et al.*, 2022; HOROWITZ *et al.*, 2022).

## 1.2 Origem do SARS-CoV-2

Até 2003 os coronavírus (família Coronaviridae) eram conhecidos por causar apenas, de modo geral, doenças mais leves em humanos, como o resfriado comum sazonal (vírus HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1). No entanto, desde então, três coronavírus com capacidade para causar impactos graves na saúde humana foram identificados. No ano de 2003 foi identificado o  $\beta$ -coronavírus SARS-CoV, responsável pela epidemia da Síndrome Respiratória Aguda Grave (SARS “Severe Acute Respiratory Syndrome”). O SARS-CoV estava sendo transmitido, inicialmente, entre pacientes de hospitais na China, e foi posteriormente identificado como um coronavírus originário da espécie de morcego, *Rhinolophus sinicus*. No entanto, teria sido transmitido ao homem por espécie intermediária a civeta da palmeira (*Paguma larvata*), um pequeno carnívoro asiático (SONG *et al.*, 2019). Passado cerca de 10 anos, outro coronavírus da mesma família foi isolado inicialmente na Arábia Saudita. O  $\beta$ -coronavírus chamado de MERS-CoV foi o responsável pela epidemia da Síndrome Respiratória do Oriente Médio (MERS, “Middle East Respiratory Syndrome”), que acometeu diferentes países, sendo transmitido para humanos inicialmente por dromedários (*Camelus dromedarius*) (AZHAR *et al.* 2014, SONG *et al.* 2019; LAMMERS & HAAGMANS *et al.*, 2022). Estudos demonstram que o reservatório natural do MERS-CoV também seriam morcegos da família Rhinolophidae, tendo os dromedários como hospedeiros intermediários, antes do transbordamento zoonótico para o *Homo sapiens* (SONG *et al.* 2019).

Vale destacar que semelhante ao SARS-CoV, o SARS-CoV-2 utiliza a Enzima Conversora de Angiotensina 2 (ACE2 da sigla do inglês para Angiotensin-Converting Enzyme 2) como portal de entrada para as células hospedeiras (ver revisão em FAM *et al.*, 2020). Ainda, de acordo com a OMS e vários pesquisadores (ver revisão em YÈREZ *et al.*, 2022), a taxa de mortalidade média estimada, considerando casos detectáveis/notificados, para COVID-19 é menor (2,72%) do que a doença causada por MERS-CoV (34,4%) e SARS-CoV (9,6%). Esse número permanece baixo mesmo considerando que o número de mortes causadas pela COVID-19 pode estar subestimado. Apesar dessa taxa de mortalidade relativamente baixa, em 15 de maio de 2022 foi estimado que a infecção por SARS-CoV-2 já tinha levado a mais de 518 milhões de casos confirmados e mais de seis milhões de mortes de acordo com OMS (<https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---18-may-2022>). Comparativamente, SARS-CoV e

MERS-CoV infectaram 8.098 e 2.566 pessoas e mataram 774 e 866 pessoas, respectivamente (YÈPEZ *et al.*, 2022).

Desse modo é bem conhecido que os genomas virais mais semelhante ao SARS-CoV-2 são oriundos de linhagens de coronavírus de morcego da espécie *Rhinolophus affinis*, tais como “RaTG13” e “bat-SL-CoVZC45” (ZHOU *et al.*, 2020, SINGH & YI, 2021). Inicialmente, SONG *et al.* (2019) postularam que esses vírus de morcegos não possuíam domínios de ligação iguais aos do SARS-CoV-2, o que sugeria que provavelmente não se conectariam bem com o seu receptor celular em humanos (ACE2), indicando a possibilidade de existir hospedeiro intermediário na origem do SARS-CoV-2. O genoma com o domínio de ligação mais semelhante ao SARS-CoV-2 foi encontrado em um coronavírus do pangolim malaio (*Manis javanica*) denominado Pangolin-CoV (ZHANG *et al.*, 2020). No entanto, tanto o bat-SL-CoVZC45 quanto o Pangolin-CoV encontrados nesses animais não possuem os sítios de clivagens presentes no SARS-CoV-2, o que sugere que nenhum dos dois é o progenitor direto do SARS-CoV-2 (SONG *et al.*, 2019). A recombinação entre linhagens virais também tem sido proposta há algum tempo (SALLARD *et al.*, 2020, SINGH & YI, 2021). Porém, é importante notar que nenhuma dessas espécies de morcegos e pangolins foi amostrada o suficiente para abarcar toda a variedade de coronavírus potencialmente presente nelas. Para adquirir os sítios de clivagem e os domínios de ligação necessários para infectar seres humanos o animal que serve de reservatório para o vírus precisaria ter uma população suficientemente grande onde esses sítios/domínios poderiam aparecer por meio de mutações aleatórias com posterior ação da seleção natural para um efetivo salto zoonótico bem sucedido (ANDERSEN *et al.*, 2020).

Mais recentemente, TEMMAN *et al.* (2022) salientam que o reservatório animal do SARS-CoV-2 ainda é controverso, apesar de relatos de similaridade com vírus de morcegos *Rhinolophus affinis*, como já comentado acima. Os autores reforçaram a proposta de que o SARS-CoV-2 teria um genoma em mosaico, sugerindo a contribuição de diferentes progenitores. TEMMAN *et al.* (2022) sustentam que o SARS-CoV-2 seria resultado de uma recombinação de sequências pré-existentes em espécies de morcegos *Rhinolophus* que vivem nos extensos sistemas de cavernas de calcário do Sudeste Asiático e do sul da China, incluindo *R. malayanus* e *R. pusillus*. Além disso, as distribuições de *R. marshalli*, *R. malayanus* e *R. pusillus* se sobrepõem na sub-região da Indochina, o que significa que eles podem compartilhar cavernas como abrigos e habitats de forrageamento.

Suas descobertas, portanto, indicariam que vírus semelhantes ao SARS-CoV-2 transmitidos por morcegos potencialmente infecciosos para humanos circulam em *Rhinolophus spp.* na península da Indochina (TEMMAN *et al.*, 2022). Os últimos autores comentam, no entanto, que nenhum desses coronavírus de morcegos apresentavam o sítio com quatro aminoácidos (PRRA) na junção S1/S2 da proteína Spike que é clivado pela Furina. Esse sítio de clivagem por Furina é importante na determinação da infectividade viral do SARS-CoV-2. Com base na comparação das sequências em torno do local de clivagem entre os sítios S1 e S2 da proteína Spike dos morcegos foi sugerido que o local de clivagem pela Furina no SARS-CoV-2 poderia ter se originado por eventos de recombinação entre os coronavírus BANAL-116, BANAL-247, bat RmYN02 e bat RacCS203. Alternativamente, o sítio de clivagem pela Furina no SARS-CoV-2 poderia ter emergido através de passagens do vírus em um hospedeiro alternativo ou mesmo durante uma circulação precoce pouco sintomática e não relatada em humanos (TEMMAN *et al.*, 2022).

É oportuno salientar que assim como outros Coronaviridae, SARS-CoV-2 é um vírus envelopado com genoma de fita simples de RNA (~ 30 kb), de sentido positivo, o que significa que seu RNA genômico serve tanto como um modelo direto para tradução quanto para a replicação. SARS-CoV-2 apresenta duas “fases de leitura abertas” (em inglês *open reading frame* ou *ORF*) sobrepostas, *ORF1a* e *ORF1b* que geram polipeptídeos contínuos que são clivados em um total de 16 proteínas não estruturais. A tradução de *ORF1b* é mediada por mudança na matriz de leitura que permite que a tradução continue além do códon de parada de *ORF1a*. Além de *ORF1a* e *ORF1b*, todos os outros ORFs virais canônicos são traduzidos de RNAs sub genômicos. Esses RNAs subgenômico codificam as 4 proteínas estruturais conservadas - Spike (S), membranares (M), de envelope (E) e nucleocapsídeo (N), - e vários polipeptídeos acessórios, tais como as denominadas *ORF3a*, *ORF6*, *ORF7a*, *ORF7b*, *ORF8* e *ORF10* que desempenhariam papéis importantes no ciclo de vida viral e poderiam contribuir para sua patogênese e virulência (HASSAN *et al.*, 2021). Ainda em 2020 FINKEL *et al.* (2020) identificaram um número maior de *ORFs* do que aquelas propostas inicialmente, e sugeriram que essas *ORFs* adicionais, quando traduzidas, resultariam em polipeptídeos com papel regulatório. Atualmente, são reconhecidos pelo menos 13 *ORFs* (MALONE *et al.*, 2022).

Finalmente, é sabido que cinco dos seis aminoácidos críticos nos respectivos domínios de ligação ao receptor (RBD) da Spike viral, são diferentes entre SARS-CoV-2 e SARS-CoV (L455Y, F486L, Q493N, S494D, N501T, respectivamente; TANG *et al.*, 2020).

### 1.3 Infecção celular por SARS-CoV-2

Como já mencionado, o SARS-CoV-2 utiliza a peptidase de membrana ACE2 como receptor celular para infectar as células humanas (GHEWARE *et al.*, 2022; HOROWITZ *et al.*, 2022). Essa enzima foi anteriormente descrita como receptor de entrada para outros dois coronavírus humanos: o  $\beta$ -coronavírus SARS-CoV e o  $\alpha$ -coronavírus NL63 (HOFMANN *et al.*, 2005, HOFMANN & PÖHLMANN 2004). E está associada à severidade da doença no caso da COVID 19 (GHEWARE *et al.*, 2022; HOROWITZ *et al.*, 2022).

A ACE2 é uma enzima transmembrana responsável pela clivagem de angiotensina tipos 1 e 2 (Omim número de referência \*300335) e foi identificada como sítio de ligação para a proteína Spike do coronavírus SARS-CoV-2 (LI *et al.*, 2020; LAMMERS & HAAGMANS *et al.*, 2022).

Além de ACE2, há outros elementos genéticos chaves no sucesso da infecção. Dentre estes, o receptor Serino Protease Transmembrana 2 (no inglês Transmembrane Serine Protease 2 ou TMPRSS2), cujo papel é fundamental para uma bem sucedida infecção pelo SARS-CoV-2 (HOFFMANN *et al.*, 2020; LAMMERS & HAAGMANS *et al.*, 2022). O papel da TMPRSS2 (Omim número de referência \* 602060) no hospedeiro normalmente está relacionado à resposta à sinalização de andrógenos, mas foi evolutivamente recrutada pelo SARS-CoV-2 para processar a proteína Spike (HOFFMANN *et al.*, 2020; LAMMERS & HAAGMANS *et al.*, 2022).

A proteína Spike do SARS-CoV-2 possui 2 subunidades: a subunidade S1 que se liga ao ACE2 e a subunidade S2 que media a fusão de membranas (JACKSON *et al.*, 2022; LAMMERS & HAAGMANS *et al.*, 2022). As duas subunidades são separadas pelo sítio S1/S2 que contém um motivo de clivagem para Furina (JACKSON *et al.*, 2022; LAMMERS & HAAGMANS *et al.*, 2022). Para a infecção ocorrer com sucesso, ambas as proteínas são necessárias, após se ligar a ACE2 da célula alvo, a proteína Spike é clivada

pela TMPRSS2 no sítio S2'. A clivagem ativa a subunidade S2 promovendo a fusão entre a célula hospedeira e o vírus (JACKSON *et al.*, 2022; LAMMERS & HAAGMANS *et al.*, 2022). Outra entrada que pode ser utilizada pelo vírus é através dos endossomos. Nessa via a Catepsina pode clivar a proteína Spike, mas estudos recentes mostram que essa via não é eficiente considerando células epiteliais primárias (JACKSON *et al.*, 2022; LAMMERS & HAAGMANS *et al.*, 2022). Outros co-receptores como a Neuropilina 1 e proteases como Catepsina L, os parálogos de T TMPRSS2, MPRSS11d e TMPRSS13 podem também estar envolvidas com a infecção por SARS-CoV-2, mas não está claro ainda como essas vias moleculares alternativas operam (JACKSON *et al.*, 2022; LAMMERS & HAAGMANS *et al.*, 2022).

#### 1.4 O Gene *TMPRSS2* e seu produto como mediador de infecção viral

O gene *TMPRSS2*, também conhecido como *PRSS10*, codifica a proteína TMPRSS2 pertencente à família das Serina-Proteases (“Type II Transmembrane Serine Protease” – TTSP). O gene possui 43.854 nucleotídeos, 15 éxons e está localizado no cromossomo 21 humano, na região 21q22.3 (PAOLINI-GIACOBINO *et al.*, 1997). TMPRSS2 contém três domínios funcionais: um receptor LDL classe A, um domínio scavenger (“Scavenger Receptor Cysteine-Rich” - SRCR) e um domínio peptidase extracelular S1 (Tripsina) (HUSSAIN *et al.*, 2020).

Baseando-se na sua estrutura, foi especulado que a proteína deveria funcionar como mediadora de sinais entre o ambiente extracelular e a célula (VAARALA *et al.*, 2001). Análises posteriores *in vitro* relataram que essa proteína regulava a atividade de canais de sódio, estando ela associada a processos fisiológicos e patológicos como a digestão, remodelamento tecidual, invasão celular tumoral, apoptose e dor (THUNDERS & DELAHUNT, 2020).

A expressão do *TMPRSS2* também é observada durante o desenvolvimento e aumenta com o envelhecimento. *TMPRSS* tem alta transcrição no cérebro fetal, mas baixa no cérebro adulto e uma baixa transcrição também no pulmão fetal quando comparado ao pulmão adulto (THUNDERS & DELAHUNT, 2020). O gene *TMPRSS2* possui expressão diferenciada em outros órgãos, tais como fígado, coração, e trato gastrointestinal, órgãos que podem ser afetados pelo SARS-CoV-2 durante a infecção. Esse gene também é

altamente expresso no tecido epitelial do lúmen da próstata (VAARALA *et al.*, 2001; THUNDERS & DELAHUNT, 2020). Na próstata ele contribui para uma cascata proteolítica que resulta na ativação do antígeno próstata-específico, a protease presente no fluido seminal com atividade enzimática análoga à coagulação do sangue (THUNDERS & DELAHUNT, 2020). *TMPRSS2* possui elementos responsivos a andrógenos na sua região 5' UTR, portanto, dentro do seu promotor. A testosterona e a di-hidrotestosterona regulam a transcrição desse gene através do estímulo do receptor andrógeno presente nas células epiteliais do lúmen da próstata, um dos locais onde o gene é expresso (THUNDERS & DELAHUNT, 2020).

Devido a sua atividade proteolítica, *TMPRSS2* foi ao longo da evolução cooptada por vírus para facilitar suas entradas nas células hospedeiras. Porém, não se sabe desde quando isso vem acontecendo, pois estudos com outros animais e seus respectivos vírus são praticamente inexistentes. Inicialmente foi descrito o papel de *TMPRSS2* para o sucesso das infecções por cepas do vírus influenza, considerando populações humanas. A replicação do vírus influenza humano é iniciada pela glicoproteína de superfície hemaglutinina (HA) que media a ligação aos receptores de superfície celulares contendo ácido salicílico e fusão do envelope viral com a membrana celular (BÖTTCHER-FRIEBERTSHÄUSER *et al.*, 2010). HA é sintetizada como uma molécula precursora que precisa ser clivada pela célula hospedeira nas subunidades HA1 e HA2 para adquirir sua capacidade de fusão com a membrana da célula hospedeira (BÖTTCHER-FRIEBERTSHÄUSER *et al.*, 2010). A clivagem induz mudanças conformacionais em pH baixo que expõe o peptídeo de fusão hidrofóbico N-terminal da subunidade HA2 e inicia a fusão (BÖTTCHER-FRIEBERTSHÄUSER *et al.*, 2010). Em seres humanos a clivagem pode ser feita por duas serino-proteases: a tripsina humana de vias aéreas – (human air way trypsin-like protease, HAT também conhecida como *TMPRSS11D*) ou pela *TMPRSS2* (BÖTTCHER-FRIEBERTSHÄUSER *et al.*, 2010). Estudos subsequentes mostraram que *TMPRSS2* e a protease relacionada *TMPRSS4* clivam a HA do vírus influenza H1N1 (família Orthomyxoviridae), responsável pela pandemia de 1918. Além disso, *TMPRSS2* também ativaria a proteína de fusão do Metapneumovirus (MPV; família Paramyxoviridae) humano identificado em 2001 (BÖTTCHER-FRIEBERTSHÄUSER *et al.*, 2010).

Análises *in vitro* com plasmídeos realizadas por GLOWACKA *et al.* (2011) demonstraram que as proteínas Spike do SARS-CoV eram clivadas de duas formas pela

TMPRSS2: pela clivagem em cis e pela clivagem em trans. Na clivagem em cis, TMPRSS2 e a proteína Spike do SARS-CoV eram expressos juntos na mesma célula e a clivagem resultava na quebra da proteína Spike e na liberação dos fragmentos desta proteína no sobrenadante. Os fragmentos funcionam como chamarizes de anticorpos e provocam a inibição da resposta imune mediada por anticorpos. Ainda, os mesmos testes *in vitro* indicaram que a clivagem em trans ocorria quando uma célula expressava a proteína Spike viral na membrana, enquanto outra célula expressava TMPRSS2. Quando ocorria o encontro de ambas moléculas ocorria a clivagem da proteína Spike e a posterior fusão das células em questão. Isso indica que a proteína Spike precisa de um processamento para permitir a entrada do vírus na célula, e em seres humanos esse processamento ocorre através do recrutamento da TMPRSS2 (GLOWACKA *et al.*, 2011). Posteriormente, BERTRAM *et al.* (2013) demonstraram que TMPRSS2 estava envolvida com a infecção pelo  $\alpha$ -coronavírus humano CoV-229E, enquanto IWATA-YOSHIKAWA *et al.* (2019) corroboraram que TMPRSS2 ativava SARS-CoV, sendo também indispensável para a infecção bem sucedida do  $\beta$ -coronavírus MERS-CoV. Com o advento da COVID-19, os estudos se voltaram para identificar se TMPRSS2 também exercia um papel crucial na infecção por SARS-CoV-2, o que de fato foi demonstrado por vários autores (HOFFMANN *et al.*, 2020; MATSUYAMA *et al.*, 2020; JACKSON *et al.*, 2022; LAMMERS & HAAGMANS *et al.*, 2022). Por outro lado, como já comentado, o papel de ortólogos de TMPRSS2 nas infecções de vírus presentes em outras espécies de mamíferos ainda precisa ser investigado.

Como já comentado, a TMPRSS2 não é a única capaz de processar Spike, sendo a Catepsina L pode realizar esse processo. No entanto, a rota da TMPRSS2 aparenta ser mais veloz, e portanto mais eficiente (KOCH *et al.*, 2021; JACKSON *et al.*, 2022; LAMMERS & HAAGMANS *et al.*, 2022). Além disso, outras proteínas da mesma família da TMPRSS2 podem ser utilizadas. Esses achados indicam que a proteína TMPRSS2 além de outras da mesma família vem sendo cooptadas muito antes da interação entre SARS-CoV-2 e o *Homo sapiens*.

1.5 As Serino Proteases Transmembrana Tipo 2 (“Type II Transmembrane Serine Protease” - TTSP).

As TTSPs são uma família identificada recentemente em 2001 por HOOPER et al, consistindo à época de 17 membros em mamíferos placentários. Em 2009 BUGGE et al aumentou o número de proteínas identificadas em humanos e as classificou em 4 subfamílias usando como critério os domínios que cada proteína possui: HAT/DESC (7 genes/proteínas), Hepsina/TMPRSS (7), Matriptase (4), and Corina (1). *TMPRSS2* é membro da subfamília Hepsin/TMPRSS junto com *TMPRSS3*, *TMPRSS4*, *TMPRSS5/Spinesin*, MSPL e Enteropeptidase.

Desde então, estudos sobre esta família gênica não têm sido publicados, pelo menos considerando parálogos humanos e ortólogos presentes em outras espécies de mamíferos.

## 2. Objetivo

Tendo em vista que a protease *TMPRSS2* é a molécula preferencial para a clivagem da Spike do SARS-CoV-2, buscamos neste trabalho atingir o seguinte objetivo de identificar eventuais padrões de diversidade na região codificadora dessa molécula e seus parálogos no *Homo sapiens*, bem como de ortólogos em mamíferos, com intuito de responder, basicamente, duas principais perguntas: 1) O padrão de diversidade genética encontrado poderia ser resultado, mesmo que parcialmente, de pressão seletiva exercida por uma corrida armamentista biológica entre vírus-hospedeiro? e 2) Porque *TMPRSS2* foi preferencialmente cooptada evolutivamente pelo SARS-CoV-2 para o processamento da Spike, enquanto outras TTSPs não o foram?

### 3. Resultados- Manuscrito

Os resultados podem ser encontrados no manuscrito a seguir:

Castro-Nóbrega M, Yépez Y, Fam B, and Bortolini MC. TMPRSS2 AND OTHER TYPE II TRANSMEMBRANE SERINE PROTEASES (TTSPs): SIGNATURE OF A LONG-TERM ARMS RACE BETWEEN VIRUSES AND PLACENTAL MAMMALS.

### 4. Considerações finais

O presente trabalho contribuiu para atualizar a história das TTSPs, uma família de proteínas cujo último trabalho em relação a sua história de evolução molecular foi publicado em 2009. A oportunidade de um estudo dessa natureza se faz premente já que um ou mais membros dessa família (exemplo TMPRSS2) foi (foram) cooptado (s) evolutivamente pelo vírus SARS-CoV-2 para seu processamento e uma bem sucedida entrada na célula do hospedeiro *Homo sapiens*. SARS-CoV-2 é o agente causador da pandemia da COVID-19.

Dezoito parálogos TTSPs foram identificados no genoma do *Homo sapiens*, sendo que fomos capazes de alocar todos eles em quatro subfamílias, incluindo TMPRSS12 no clado das Matriptase.

Um alinhamento filogenético dessas 18 TTSP do *Homo sapiens* resgata o esperado considerando a história evolutiva das moléculas. Interessantemente quando a mesma análise é feita considerando somente os 30 sítios de TMPRSS2 que interagem com o vírus, a topologia da árvore se altera, sugerindo que há uma certa identidade em pontos chaves, não obstante a história filogenética das TTSPs, considerando suas sequências codificadoras

completas. Em outras palavras, as proteínas dividiram-se em dois clados distintos, sendo que um deles continha todas as proteínas que já foram documentadas como capazes de realizar a clivagem da proteína Spike do SARS-CoV-2. No outro clado encontravam-se as TTSPs cuja interação com SARS-CoV-2 tem sido controversa ou descartada.

Finalmente, respondendo as questões principais formuladas:

1) O padrão de diversidade genética encontrado poderia ser resultado, mesmo que parcialmente, de pressão seletiva exercida por uma corrida armamentista biológica entre vírus-hospedeiro?

Sim. Em relação a evolução da TMPRSS2 humana e seus ortólogos em outros mamíferos placentários, nosso trabalho identificou 33 sítios onde a variação encontrada entre pode ser explicada por seleção positiva, cinco deles em posições que interagem com o vírus, ou seja, somente 15%. Esse padrão de variação sugere que, talvez, a maior parte da variação entre espécies seja resultados de pressões seletivas que vêm moldando a função normal da TMPRSS2 e de seus ortólogos nas células das espécies correspondentes. Por outro lado, também indica que o vírus, através de sua proteína Spike, que necessita de clivagem, estaria sendo moldado evolutivamente para se ligar em posições nas proteases dos potenciais hospedeiros mamíferos com menos propensão a terem variação promovida por ação de seleção positiva devido a demandas do próprio hospedeiro. Isso daria vantagem ao vírus, pois ele teria, por exemplo, menos chance de perder afinidade com o hospedeiro ao mesmo tempo que daria mais chances para saltos zoonóticos. Nesse contexto, o padrão de variabilidade nos mamíferos placentários indica profundidade de

tempo evolutivo, algo esperado já que TMPRSS2 e outras TTSPs vêm sendo cooptadas por outras famílias virais, tais como *Orthomyxoviridae*. Considerando que tanto as famílias de vírus como as de mamíferos placentários convivem a milhões de anos, esses achados indicam uma corrida armamentista evolucionária de longo tempo.

2) Porque TMPRSS2 foi preferencialmente cooptada evolutivamente pelo SARS-CoV-2 para o processamento da Spike, enquanto outras TTSPs não o foram?

Há um provável conjunto de cerca de 30 aminoácidos em sítios pouco variáveis devido à seleção positiva, o que garante uma identidade chave entre TMPRSS2 e outras TTSPs, incluindo de membros de diferentes subfamílias. É o conjunto de aminoácidos nesses sítios que potencialmente garante a reiterada cooptação evolutiva dessas proteases pelo SARS-CoV-2 e outros vírus para o processamento de suas Spikes. Por outro lado, TTSPs que não compartilham esse conjunto de aminoácidos não tem sido preferencialmente cooptada.

## 5. Referências

- Abrams, E. M., Greenhawt, M., Shaker, M., Pinto, A. D., Sinha, I., & Singer, A. (2022). The COVID-19 Pandemic: Adverse Effects on the Social Determinants of Health in Children and Families. *Annals of Allergy, Asthma & Immunology*.  
<https://doi.org/10.1016/j.anai.2021.10.022>
- Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., & Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nature Medicine*, 26(450–452).  
<https://doi.org/10.1038/s41591-020-0820-9>
- Azhar, E. I., El-Kafrawy, S. A., Farraj, S. A., Hassan, A. M., Al-Saeed, M. S., Hashem, A. M., & Madani, T. A. (2014). Evidence for Camel-to-Human Transmission of MERS Coronavirus. *New England Journal of Medicine*, 370(26), 2499–2505.  
<https://doi.org/10.1056/nejmoa1401505>
- Bertram, S., Dijkman, R., Habjan, M., Heurich, A., Gierer, S., Glowacka, I., Welsch, K., Winkler, M., Schneider, H., Hofmann-Winkler, H., Thiel, V., & Pohlmann, S. (2013). TMPRSS2 Activates the Human Coronavirus 229E for Cathepsin-Independent Host Cell Entry and Is Expressed in Viral Target Cells in the Respiratory Epithelium. *Journal of Virology*, 87(11), 6150–6160.  
<https://doi.org/10.1128/jvi.03372-12>
- Bieber, F. (2020). Global Nationalism in Times of the COVID-19 Pandemic. *Nationalities Papers*, 50(1), 1–13. <https://doi.org/10.1017/nps.2020.35>
- Böttcher-Friebertshäuser E., Freuer, C., Sielaff, F., Schmidt, S., Eickmann, M., Uhlendorff, J., Steinmetzer, T., Klenk, H.-D., & Garten, W. (2010). Cleavage of Influenza Virus Hemagglutinin by Airway Proteases TMPRSS2 and HAT Differs in Subcellular Localization and Susceptibility to Protease Inhibitors. *Journal of Virology*, 84(11), 5605–5614. <https://doi.org/10.1128/jvi.00140-10>

- Bugge, T. H., Antalis, T. M., & Wu, Q. (2009). Type II Transmembrane Serine Proteases. *Journal of Biological Chemistry*, 284(35), 23177–23181.  
<https://doi.org/10.1074/jbc.r109.021006>
- Cao, Y., Li, L., Feng, Z., Wan, S., Huang, P., Sun, X., Wen, F., Huang, X., Ning, G., & Wang, W. (2020). Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discovery*, 6(1).  
<https://doi.org/10.1038/s41421-020-0147-1>
- Chen, W., Yuan, P., Yang, M., Yan, Z., Kong, S., Yan, J., Liu, X., Chen, Y., Qiao, J., & Yan, L. (2020). SARS-CoV-2 Entry Factors: ACE2 and TMPRSS2 Are Expressed in Peri-Implantation Embryos and the Maternal–Fetal Interface. *Engineering*, 6(10), 1162–1169. <https://doi.org/10.1016/j.eng.2020.07.013>
- Fam, B. S. O., Vargas-Pinilla, P., Amorim, C. E. G., Sortica, V. A., & Bortolini, M. C. (2020). ACE2 diversity in placental mammals reveals the evolutionary strategy of SARS-CoV-2. *Genetics and Molecular Biology*, 43(2).  
<https://doi.org/10.1590/1678-4685-gmb-2020-0104>
- Finkel, Y., Mizrahi, O., Nachshon, A., Weingarten-Gabbay, S., Morgenstern, D., Yahalom-Ronen, Y., Tamir, H., Achdout, H., Stein, D., Israeli, O., Beth-Din, A., Melamed, S., Weiss, S., Israely, T., Paran, N., Schwartz, M., & Stern-Ginossar, N. (2020). The coding capacity of SARS-CoV-2. *Nature*, 1–6.  
<https://doi.org/10.1038/s41586-020-2739-1>
- Gheware, A., Ray, A., Rana, D., Bajpai, P., Nambirajan, A., Arulselvi, S., Mathur, P., Trikha, A., Arava, S., Das, P., Mridha, A. R., Singh, G., Soneja, M., Nischal, N., Lalwani, S., Wig, N., Sarkar, C., & Jain, D. (2022). ACE2 protein expression in lung tissues of severe COVID-19 infection. *Scientific Reports*, 12(1).  
<https://doi.org/10.1038/s41598-022-07918-6>
- Glowacka, I., Bertram, S., Müller, M. A., Allen, P., Soilleux, E., Pfefferle, S., Steffen, I., Tsegaye, T. S., He, Y., Gnirss, K., Niemeyer, D., Schneider, H., Drosten, C., & Pöhlmann, S. (2011). Evidence that TMPRSS2 Activates the Severe Acute Respiratory Syndrome Coronavirus Spike Protein for Membrane Fusion and Reduces Viral Control by the Humoral Immune Response. *Journal of Virology*, 85(9), 4122–4134. <https://doi.org/10.1128/JVI.02232-10>

- Guaita Martínez, J. M., Carracedo, P., Gorgues Comas, D., & Siemens, C. H. (2022). An analysis of the blockchain and COVID-19 research landscape using a bibliometric study. *Sustainable Technology and Entrepreneurship*, *1*(1), 100006. <https://doi.org/10.1016/j.stae.2022.100006>
- Hassan, Sk. S., Choudhury, P. P., & Roy, B. (2021). Rare mutations in the accessory proteins ORF6, ORF7b, and ORF10 of the SARS-CoV-2 genomes. *Meta Gene*, *28*, 100873. <https://doi.org/10.1016/j.mgene.2021.100873>
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N.-H., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*, *181*(2). <https://doi.org/10.1016/j.cell.2020.02.052>
- Hofmann, H., & Pöhlmann, S. (2004). Cellular entry of the SARS coronavirus. *Trends in Microbiology*, *12*(10), 466–472. <https://doi.org/10.1016/j.tim.2004.08.008>
- Hofmann, H., Pyrc, K., van der Hoek, L., Geier, M., Berkhout, B., & Pöhlmann, S. (2005). Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proceedings of the National Academy of Sciences*, *102*(22), 7988–7993. <https://doi.org/10.1073/pnas.0409465102>
- Hooper, J. D., Clements, J. A., Quigley, J. P., & Antalis, T. M. (2001). Type II Transmembrane Serine Proteases. *Journal of Biological Chemistry*, *276*(2), 857–860. <https://doi.org/10.1074/jbc.r000020200>
- Horowitz, J. E., Kosmicki, J. A., Damask, A., Sharma, D., Roberts, G. H. L., Justice, A. E., Banerjee, N., Coignet, M. V., Yadav, A., Leader, J. B., Marcketta, A., Park, D. S., Lanche, R., Maxwell, E., Knight, S. C., Bai, X., Guturu, H., Sun, D., Baltzell, A., & Kury, F. S. P. (2022). Genome-wide analysis provides genetic evidence that ACE2 influences COVID-19 risk and yields risk scores associated with severe disease. *Nature Genetics*, *54*(4), 382–392. <https://doi.org/10.1038/s41588-021-01006-7>
- Hussain, M., Jabeen, N., Amanullah, A., Ashraf Baig, A., Aziz, B., Shabbir, S., Raza, F., & Uddin, N. (2020). Molecular docking between human TMPRSS2 and SARS-CoV-2 spike protein: conformation and intermolecular interactions. *AIMS Microbiology*, *6*(3), 350–360. <https://doi.org/10.3934/microbiol.2020021>

- Iser, B. P. M., Sliva, I., Raymundo, V. T., Poletto, M. B., Schuelter-Trevisol, F., & Bobinski, F. (2020). Definição de caso suspeito da COVID-19: uma revisão narrativa dos sinais e sintomas mais frequentes entre os casos confirmados. *Epidemiologia E Serviços de Saúde*, 29(3). <https://doi.org/10.5123/s1679-49742020000300018>
- Iwata-Yoshikawa, N., Okamura, T., Shimizu, Y., Hasegawa, H., Takeda, M., & Nagata, N. (2019). TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. *Journal of Virology*, 93(6). <https://doi.org/10.1128/jvi.01815-18>
- Jackson, C. B., Farzan, M., Chen, B., & Choe, H. (2022). Mechanisms of SARS-CoV-2 entry into cells. *Nature Reviews Molecular Cell Biology*, 23, 1–18. <https://doi.org/10.1038/s41580-021-00418-x>
- Koch, J., Uckeley, Z. M., Doldan, P., Stanifer, M., Boulant, S., & Lozach, P. (2021). TMPRSS2 expression dictates the entry route used by SARS-CoV-2 to infect host cells. *The EMBO Journal*, 40(16). <https://doi.org/10.15252/emj.2021107821>
- Lamers, M. M., & Haagmans, B. L. (2022). SARS-CoV-2 pathogenesis. *Nature Reviews Microbiology*, 1–15. <https://doi.org/10.1038/s41579-022-00713-0>
- Landau, L. J. B., Fam, B. S. de O., Yépez, Y., Caldas-Garcia, G. B., Pissinatti, A., Falótico, T., Reales, G., Schüler-Faccini, L., Sortica, V. A., & Bortolini, M. C. (2021). Evolutionary analysis of the anti-viral STAT2 gene of primates and rodents: Signature of different stages of an arms race. *Infection, Genetics and Evolution*, 95, 105030. <https://doi.org/10.1016/j.meegid.2021.105030>
- Li, X., Zai, J., Zhao, Q., Nie, Q., Li, Y., Foley, B. T., & Chaillon, A. (2020). Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *Journal of Medical Virology*, 92(6), 602–611. <https://doi.org/10.1002/jmv.25731>
- Malone, B., Urakova, N., Snijder, E. J., & Campbell, E. A. (2022). Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design. *Nature Reviews Molecular Cell Biology*, 23. <https://doi.org/10.1038/s41580-021-00432-z>

- Matsuyama, S., Nao, N., Shirato, K., Kawase, M., Saito, S., Takayama, I., Nagata, N., Sekizuka, T., Katoh, H., Kato, F., Sakata, M., Tahara, M., Kutsuna, S., Ohmagari, N., Kuroda, M., Suzuki, T., Kageyama, T., & Takeda, M. (2020). Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proceedings of the National Academy of Sciences*, *117*(13), 7001–7003.  
<https://doi.org/10.1073/pnas.2002589117>
- Meyer, B. H., Prescott, B., & Sheng, X. S. (2022). The impact of the COVID-19 pandemic on business expectations. *International Journal of Forecasting*, *38*(2).  
<https://doi.org/10.1016/j.ijforecast.2021.02.009>
- Mir, T., Almas, T., Kaur, J., Faisaluddin, M., Song, D., Ullah, W., Mamtani, S., Rauf, H., Yadav, S., Latchana, S., Michaelson, N. M., Connerney, M., & Sattar, Y. (2021). Coronavirus disease 2019 (COVID-19): Multisystem review of pathophysiology. *Annals of Medicine and Surgery*, *69*, 102745.  
<https://doi.org/10.1016/j.amsu.2021.102745>
- Mousavizadeh, L., & Ghasemi, S. (2020). Genotype and phenotype of COVID-19: Their roles in pathogenesis. *Journal of Microbiology, Immunology and Infection*, *54*(2).  
<https://doi.org/10.1016/j.jmii.2020.03.022>
- Nersisyan, S., Shkurnikov, M., Turchinovich, A., Knyazev, E., & Tonevitsky, A. (2020). Integrative analysis of miRNA and mRNA sequencing data reveals potential regulatory mechanisms of ACE2 and TMPRSS2. *PLOS ONE*, *15*(7), e0235987.  
<https://doi.org/10.1371/journal.pone.0235987>
- Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., & Agha, R. (2020). The Socio-Economic Implications of the Coronavirus and COVID-19 Pandemic: A Review. *International Journal of Surgery*, *78*(78), 185–193. NCBI. <https://doi.org/10.1016/j.ijisu.2020.04.018>
- Paoloni-Giacobino, A., Chen, H., Peitsch, M. C., Rossier, C., & Antonarakis, S. E. (1997). Cloning of the TMPRSS2 Gene, Which Encodes a Novel Serine Protease with Transmembrane, LDLRA, and SRCR Domains and Maps to 21q22.3. *Genomics*, *44*(3), 309–320. <https://doi.org/10.1006/geno.1997.4845>
- Pummerer, L., Böhm, R., Lilleholt, L., Winter, K., Zettler, I., & Sassenberg, K. (2022). Conspiracy Theories and Their Societal Effects During the COVID-19 Pandemic.

- Social Psychological and Personality Science*, 13(1), 194855062110002.  
<https://doi.org/10.1177/19485506211000217>
- Robinson, E. L., Alkass, K., Bergmann, O., Maguire, J. J., Roderick, H. L., & Davenport, A. P. (2020). Genes encoding ACE2, TMPRSS2 and related proteins mediating SARS-CoV-2 viral entry are upregulated with age in human cardiomyocytes. *J Mol Cell Cardiol*, 147, 88–91.  
<https://doi.org/https://dx.doi.org/10.1016%2Fj.yjmcc.2020.08.009>
- Rozhkov, M., Ivanov, D., Blackhurst, J., & Nair, A. (2022). Adapting supply chain operations in anticipation of and during the COVID-19 Pandemic. *Omega*, 102635.  
<https://doi.org/10.1016/j.omega.2022.102635>
- Sallard, E., Halloy, J., Casane, D., van Helden, J., & Decroly, É. (2020). Retrouver les origines du SARS-CoV-2 dans les phylogénies de coronavirus. *Médecine/Sciences*, 36(8-9), 783–796. <https://doi.org/10.1051/medsci/2020123>
- Singh, D., & Yi, S. V. (2021). On the origin and evolution of SARS-CoV-2. *Experimental & Molecular Medicine*, 53. <https://doi.org/10.1038/s12276-021-00604-z>
- Song, Z., Xu, Y., Bao, L., Zhang, L., Yu, P., Qu, Y., Zhu, H., Zhao, W., Han, Y., & Qin, C. (2019). From SARS to MERS, Thrusting Coronaviruses into the Spotlight. *Viruses*, 11(1), 59. <https://doi.org/10.3390/v11010059>
- Tang, T., Bidon, M., Jaimes, J. A., Whittaker, G. R., & Daniel, S. (2020). Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antiviral Research*, 178, 104792. <https://doi.org/10.1016/j.antiviral.2020.104792>
- Temgoua, M. N., Endomba, F. T., Nkeck, J. R., Kenfack, G. U., Tochie, J. N., & Essouma, M. (2020). Coronavirus Disease 2019 (COVID-19) as a Multi-Systemic Disease and its Impact in Low- and Middle-Income Countries (LMICs). *SN Comprehensive Clinical Medicine*, 2(9), 1377–1387. <https://doi.org/10.1007/s42399-020-00417-7>
- Temmam, S., Vongphayloth, K., Salazar, E. B., Munier, S., Bonomi, M., Regnault, B., Douangboubpha, B., Karami, Y., Chrétien, D., Sanamxay, D., Xayaphet, V., Paphaphanh, P., Lacoste, V., Somlor, S., Lakeomany, K., Phommavanh, N., Pérot, P., Dehan, O., Amara, F., & Donati, F. (2022). Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. *Nature*. <https://doi.org/10.1038/s41586-022-04532-4>

- Thunders, M., & Delahunt, B. (2020). Gene of the month: TMPRSS2 (transmembrane serine protease 2). *Journal of Clinical Pathology*, *jclinpath-2020-206987*.  
<https://doi.org/10.1136/jclinpath-2020-206987>
- Vaarala, M. H., Porvari, K. S., Kellokumpu, S., Kyllönen, A. P., & Vihko, P. T. (2001). Expression of transmembrane serine protease TMPRSS2 in mouse and human tissues. *The Journal of Pathology*, *193*(1), 134–140. [https://doi.org/10.1002/1096-9896\(2000\)9999:9999::aid-path743>3.0.co;2-t](https://doi.org/10.1002/1096-9896(2000)9999:9999::aid-path743>3.0.co;2-t)
- Wirkner, J., Christiansen, H., Knaevelsrud, C., Lüken, U., Wurm, S., Schneider, S., & Brakemeier, E.-L. (2022). Mental Health in Times of the COVID-19 Pandemic. *European Psychologist*, *26*(4), 310–322. <https://doi.org/10.1027/1016-9040/a000465>
- Yépez, Y., Marcano-Ruiz, M., Bezerra, R. S., Fam, B., Ximenez, J. P., Silva Jr, W. A., & Bortolini, M. C. (2022). Evolutionary history of the SARS-CoV-2 Gamma variant of concern (P.1): a perfect storm. *Genetics and Molecular Biology*, *45*(1).  
<https://doi.org/10.1590/1678-4685-gmb-2021-0309>
- Zhang, C., Zhang, Y., Zhang, S., Wang, Z., Sun, S., Liu, M., Chen, Y., Dong, N., & Wu, Q. (2020). Intracellular autoactivation of TMPRSS11A, an airway epithelial transmembrane serine protease. *Journal of Biological Chemistry*, *295*(36), 12686–12696. <https://doi.org/10.1074/jbc.ra120.014525>
- Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., Chen, H.-D., Chen, J., Luo, Y., Guo, H., Jiang, R.-D., Liu, M.-Q., Chen, Y., Shen, X.-R., Wang, X., & Zheng, X.-S. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*(7798).  
<https://doi.org/10.1038/s41586-020-2012-7>