


Cotton flower metabolites inhibit SARS-CoV-2 main protease

Yufang Zhang¹, Wenkang Li^{1,2}, Yiming Hu³, Tianze Ding¹, Muhammad Mubashar Zafar¹, Xue Jia¹, Liya Zhang¹ , Maozhi Ren^{1,2,4,5,6}, Fuguang Li^{1,2,4,6} and Wenjing Wang^{1,2,4,6}

1 Zhengzhou Research Base, State Key Laboratory of Cotton Biology, School of Agricultural Sciences, Zhengzhou University, China

2 Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, China

3 Zhengzhou Technology and Business University, China

4 Hainan Yazhou Bay Seed Laboratory, Sanya, China

5 Laboratory of Space Biology, Institute of Urban Agriculture, Chinese Academy of Agricultural Sciences, Chengdu, China

6 Sanya Institute of Zhengzhou University, China

Keywords

biochemical assay; cotton flower; main protease; SARS-CoV-2; variants; virtual screening

Correspondence

Maozhi Ren, Fuguang Li and Wenjing Wang, Zhengzhou Research Base, State Key Laboratory of Cotton Biology, School of Agricultural Sciences, Zhengzhou University, No. 157, Science Avenue, High-Tech Development Zone, 450001, Zhengzhou, China

E-mails: renmaozhi01@caas.cn (MR); aylifug@caas.cn (FL); wangwj@zzu.edu.cn (WW)

Yufang Zhang and Wenkang Li contributed equally to this work

(Received 4 May 2022, revised 10 August 2022, accepted 24 August 2022)

doi:10.1002/2211-5463.13477

Edited by Cláudio Soares

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been spreading globally for over 2 years, causing serious contagious disease and incalculable damage. The introduction of vaccines has slowed the spread of SARS-CoV-2 to some extent, but there remains a need for specific and effective treatment. The high chemical diversity and safety profiles of natural products make them a potential source of effective anti-SARS-CoV-2 drugs. Cotton plant is one of the most important economic and medical crops and is the source of a large number of antiviral phytochemicals. In this work, we used SARS-CoV-2 main protein (M^{pro}) as the target to identify potential anti-SARS-CoV-2 natural products in cotton. An *in vitro* assay showed that of all cotton tissues examined, cotton flower extracts (CFs) exhibited optimal inhibitory effects against M^{pro} . We proceeded to use the CF metabolite database to screen natural M^{pro} inhibitors by combining virtual screening and biochemical assays. We identified that several CF natural products, including astragalin, myricitrin, and astilbin, significantly inhibited M^{pro} with half-maximal inhibitory concentrations (IC_{50} s) of 0.13, 10.73, and 7.92 μ M, respectively. These findings may serve as a basis for further studies into the suitability of cotton as a source of potential therapeutics for SARS-CoV-2.

SARS-CoV-2, a highly contagious and mutable virus that is responsible for the COVID-19 (coronavirus disease 2019) pandemic, has caused a global catastrophe since it was first noticed in late 2019 [1]. To date, very few drugs are known to inhibit SARS-CoV-2 effectively. The introduction of vaccines has slowed the spread of

SARS-CoV-2 to some extent. However, billions of persons who are awaiting vaccination and millions of immunocompromised persons who are unlikely to respond robustly to vaccination are still calling for specific and effective treatment [2]. Moreover, many thousands of SARS-CoV-2 variants have been identified [3],

Abbreviations

CF, cotton flowers; DMF, *N,N*-dimethylformamide; DMSO, dimethyl sulfoxide; EDTA, ethylenediaminetetraacetic acid; FRET, fluorescence resonance energy transfer; HRV 3C, human rhinovirus 3C protease; IC_{50} , half-maximal inhibitory concentration; M^{pro} , main protein; NSP, nonstructural protein; PCR, polymerase chain reaction; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SDS/PAGE, sodium dodecyl sulfate – polyacrylamide gel electrophoresis; VOC, variant of concern.

in which, five variants of concern (VOCs: Alpha, Beta, Gamma, Delta and Omicron) are spreading among global populations. These SARS-CoV-2 variants bring enormous challenges for vaccines [4]. Therefore, it remains important to discover broad-spectrum inhibitors against various SARS-CoV-2 variants.

SARS-CoV-2 genome encodes a total of 29 proteins: four structural proteins (including envelope, membrane, spike and nucleocapsid), 16 nonstructural proteins (NSPs), and nine accessory proteins [5]. Structural proteins are mainly to form virus particles and infect host cells, while NSPs mostly play important roles in viral replication and other functions. Among the SARS-CoV-2 variants, pernicious mutations occur mainly on the spike protein [6]. The NSPs, especially those closely related to novel coronavirus replication, such as PL^{Pro} and M^{Pro} (proteins essential for SARS-CoV-2 replication [7–11]), are less likely to mutate, which may provide conserved targets for drug discovery [12–14]. The characterization of SARS-CoV-2 M^{Pro} crystal structure provides a model for drug discovery [8]. Some studies have obtained several candidate M^{Pro} inhibitor through virtual screening [15–17], which have made great contribution to the screening of anti-SARS-CoV-2 drugs. Nevertheless, it is necessary to verify the real inhibitory effect of the screened M^{Pro} inhibitor through biochemical experiments.

Plant-based natural products have shown excellent antiviral effect against SARS-CoV-2 [18,19]. Cotton (*Gossypium* spp) is a wonderful plant grown for basic needs, such as clothing, food, and shelter. According to Chinese Materia Medica, cotton plant also possesses high medicinal value and health care functions. For example, the extract of cotton flowers (CFs) has been used as an ethical herb by the Uygur people in Xinjiang, China, for a long time to treat mental retardation [20]. Oral CF extract has been reported to be beneficial for age-related dementia and Alzheimer's disease [21]. In addition, cotton plant can also produce many secondary metabolites with antiviral activity [22,23]. Gossypin from cotton plant can inhibit herpes simplex viruses [22]. Gossypetin shows a high inhibitory effect on influenza virus *in vitro* [23]. Other metabolites, such as ferulic acid, astragalin, myricitrin, astilbin, kaempferol, and kaempferitrin, also possess certain antiviral properties [24–29]. Given that cotton plant contains many antiviral products, it is necessary to test whether any of them can inhibit SARS-CoV-2.

The current study was designed to identify potential inhibitors against SARS-CoV-2 M^{Pro} from cotton plant using computational simulation and biochemical assays. The findings of our study may provide some new options for researchers to fight against COVID-19.

Materials and methods

Sequence alignment of M^{Pro} in different SARS-CoV-2 variants

The M^{Pro} protein sequences of original SARS-CoV-2 strain (Wuhan-Hu-1) as well as the five VOCs, including alpha (B.1.1.7), beta (B.1.351), gamma (P.1), delta (B.1.617.2), and omicron (B.1.1.529), were downloaded from GISAID (<https://www.gisaid.org/>) and GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>). The GenBank accession numbers of M^{Pro} for Wuhan-Hu-1, B.1.1.7, B.1.351, P.1, B.1.617.2 and B.1.1.529 are [YP_009725301.1](#), [QUS74317.1](#), [QWW93434.1](#), [QYK37388.1](#), [QRC42887.1](#) and [UKO09917.1](#), respectively. All five sequences were aligned by DNAMAN version 8 software (LynnonBiosoft, San Ramon, CA, USA).

Structural comparison of M^{Pro} in the SARS-CoV-2 original strain and variants

Sequence alignment revealed that only the M^{Pro}s in beta (βM^{Pro}) and omicron variants (oM^{Pro}) were mutated, we therefore compared the structures of the two mutant M^{Pro}s with that of the original M^{Pro}. The M^{Pro} structures of original strain and omicron variant were downloaded from Protein Data Bank (PDB) with PDB ID of [6LU7](#) and [7TVS](#). The βM^{Pro} structure was predicted using ALPHAFOLD2 [30] via COLABFOLD (version 1.3) [31] with template of original M^{Pro} (PDB: [6LU7](#)) [8], MSA mode of “MMSEQS2 (Uniref + Environmental)” and pair_mode of “unpaired+paired”. ALPHAFOLD2 with MMSEQS2 was accessible on Google Colab (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/v1.3.0/AlphaFold2.ipynb>).

Other parameters were set as default. The oM^{Pro} and predicted βM^{Pro} structures were compared with the original M^{Pro} and then visualized by the graphics tool PYMOL version 2.3.2 (Schrödinger, New York, NY, USA).

Expression and purification of M^{Pro}

To generate an authentic N-terminus and C-terminus of M^{Pro} (without additional residues at the termini), our study used a previously reported strategy [8]. The full-length original SARS-CoV-2 M^{Pro} gene (ORF1ab polyprotein residues 3264–3569, GenBank code: [MN908947.3](#)) was chemically synthesized with codon optimization. Meanwhile, 12 nucleotides coding for the four amino acids AVLQ before the first Ser1 residue were added to form an autocleavage site (AVLQ↓SGFRK) recognized by M^{Pro}, and 24 nucleotides coding for the eight amino acids GPHis6 at the C-terminus were added to form a cleavage site (SGVTFFQ↓GPHis6) that can be removed by human rhinovirus 3C protease (HRV 3C). Then, the sequence was amplified by polymerase chain reaction (PCR) and inserted into the BamHI and XhoI sites of the pGEX-6p-1 plasmid (GE Healthcare, Chicago, IL, USA).

The recombinant plasmid was transformed into *Escherichia coli* BL21 (DE3) cells. First, bacteria were cultured at $2.5 \times g$ and 37°C until the OD value reached 1.0. Second, 0.5 mM IPTG was added to the cell culture to induce the expression at 18°C overnight. Third, the cells were collected by centrifugation at $3000 \times g$ for 15 min. Fourth, the cell pellets were resuspended in a lysis buffer containing 25 mM Tris (pH 8.0) and 150 mM NaCl, lysed by high-pressure homogenization, and then centrifuged at $25000 \times g$ for 40 min. Fifth, the supernatant was hung on a Ni-NTA affinity column (Qiagen, Shanghai, China) and washed twice with resuspension buffer containing 20 mM imidazole. Sixth, the collected protein was eluted by cleavage buffer (50 mM Tris-HCl pH 7.0, 150 mM NaCl) including 300 mM imidazole. Finally, the target protein was digested with Human rhinovirus 3C protease overnight to remove the C-terminal His-tag. The concentrated target protein was stored in using 10 mM Tris-HCl (pH 7.5) buffer.

Enzymatic activity assays

The activity of recombinant M^{Pro} was evaluated by using the fluorescence resonance energy transfer (FRET) substrate MCA-AVLQSGFR-Lys(Dnp)-Lys-NH₂ [32]. The amino acid sequence of the FRET substrate was derived from the N-terminal self-cleavage sequence of M^{Pro}. The fluorescence intensity was recorded by a Fluoraskan Ascent Fluorometer (Thermo, Waltham, MA, USA) with excitation and emission wavelengths of 320 and 405 nm, respectively. The reaction buffer contained 50 mM Tris-HCl (pH 7.3), 1 mM ethylenediaminetetraacetic acid (EDTA), and 30 nM M^{Pro}. For enzyme kinetics studies, the fluorescence intensity was recorded as soon as the FRET substrates (different concentrations ranging from 0 to 50 μM) were added to the reaction buffer. The initial velocity of each reaction was calculated by linear regression for the first 10 min of the kinetic progress curves. Finally, the kinetic parameters k_m and V_{max} were obtained by plotting the initial velocity against the substrate concentration using the Michaelis-Menten equation in GRAPHPAD PRISM 8 software (GraphPad Software, San Diego, CA, USA).

Preparation of extracts from various tissues of cotton

Zhongmian-24 (ZM24) from the Institute of Cotton Research, the Chinese Academy of Agricultural Sciences (CAAS), was used in this work. The cotton seeds were planted in the experimental field at the School of Agricultural Sciences, Zhengzhou University, China. After 90 days, different tissues, including flowers, leaves, roots, stems, and cotton bolls, were collected from the cotton plants at the same growth states. A total of 0.3 g fresh weight of each tissue was ground thoroughly to powder in liquid nitrogen and then ultrasound-treated for 3 h in

1 mL *N,N*-dimethylformamide (DMF; Aladdin) [33]. The extracts were centrifuged at $10\,000 \times g$ for 5 min. The supernatants were collected and stored at -20°C until further use.

Enzyme inhibition assay of cotton extracts on SARS-CoV-2 M^{Pro}

For the enzyme inhibition assay, each cotton extract (final concentration of $15\text{ mg}\cdot\text{mL}^{-1}$) was incubated with 0.2 μM SARS-CoV-2 M^{Pro} at 37°C for 30 min in reaction buffer containing 50 mM Tris-HCl (pH 7.3) and 1 mM EDTA. When the FRET substrate (final concentration of 20 μM) was added to the reaction, the fluorescence intensity was immediately monitored every 30 s for 1 h. DMSO was used as the negative control for the whole study. The inhibition rate was calculated by the ratio of the initial velocity.

Enzyme inhibition assay of CF metabolites on SARS-CoV-2 M^{Pro}

The enzyme inhibition assay of CF metabolites against SARS-CoV-2 M^{Pro} was performed by treating recombinant M^{Pro} with different CF metabolites (astragalin, myricitrin, astilbin, kaempferitrin, and kaempferol) followed by the addition of fluorescence resonance energy transfer (FRET) substrate. For astragalin, myricitrin, and astilbin, the final concentrations were varied from 0 to 20, 0 to 5000, and 0 to 10 000 μM , respectively. The experimental conditions were the same as that for cotton extracts on SARS-CoV-2 M^{Pro}. The IC₅₀ was calculated by plotting the initial velocity against various concentrations of inhibitors by GRAPHPAD PRISM 8 software. In addition, 10 μM kaempferitrin and kaempferol were used to compare their inhibitory effects against M^{Pro}.

Virtual screening of SARS-CoV-2 M^{Pro} inhibitors

The in-house CF metabolome in our laboratory was used to identify natural SARS-CoV-2 M^{Pro} inhibitors. First, the CF metabolite database was established by obtaining the 3D structures downloading from the Chemicalbook website or drawing with CHEM3D software (PerkinElmer, Waltham, MA, USA). Then, virtual screening was performed with Autodock vina in PYRX 0.8 [34] by using pre-processed SARS-CoV-2 M^{Pro} (6LU7) as the receptor and the small molecules in the CF metabolite database as the ligands. The grid box centered for M^{Pro} was $X = -11.82$, $Y = 14.94$, and $Z = 67.92$ and the dimensions of the grid box were set as $65 \times 65 \times 65\text{ \AA}$. The spacing value was the default value of 0.375 \AA . Other docking parameters were kept at their default. The docking result was analyzed and visualized in PYMOL and CHIMERAX 1.2.5 software (University of California, Los Angeles, CA, USA). The 2D interactions of small molecules and M^{Pro} were analyzed by the online tool PROTEINS PLUS (<https://proteins.plus>).

Results

Sequence alignment and structural comparison of M^{Pro} in different SARS-CoV-2 variants

SARS-CoV-2 has mutated thousands of times and several dominant variants have emerged, including alpha (B.1.1.7), beta (B.1.351), gamma (P.1), delta (B.1.617.2), and omicron (B.1.1.529), so it is necessary to identify a drug target that is conserved in different variants. The sequences of the five dominant variants were aligned by DNAMAN. There was almost no mutation in these variants except for one mutation of K90R in beta variants and one mutation of P132H in omicron variants (Fig. 1A). In addition, these mutations are far from the active center of M^{Pro} (Fig. 1B), which has no effect on the structure of the active center. Therefore, M^{Pro} of SARS-CoV-2 can be used as a reliable target for drug discovery against all the dominant variants.

Enzymatic activity for recombinant SARS-CoV-2 M^{Pro}

Recombinant SARS-CoV-2 M^{Pro} with native N and C termini was obtained according to a method reported previously [8]. The molecular weight of M^{Pro} was approximately 33 kDa according to sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS/PAGE; Fig. 2A). To characterize its enzymatic activity, a FRET assay was performed with a fluorescently labeled substrate, MCA-AVLQ↓SGFR-Lys(Dnp)-Lys-NH₂ [32]. In the presence of M^{Pro}, the fluorescence intensity increased rapidly at first and then gradually reached a plateau, whereas little change was observed in the absence of protein (Fig. 2B). To study the enzyme kinetics, a standard curve was generated to convert the relative fluorescence unit (RFU) to the amount of the cleaved substrate (Fig. 2C). Next, the enzymatic activity of this SARS-CoV-2 M^{Pro} was characterized by measuring the parameters *km*, *Kcat*, and *Vmax*. When 200 nm

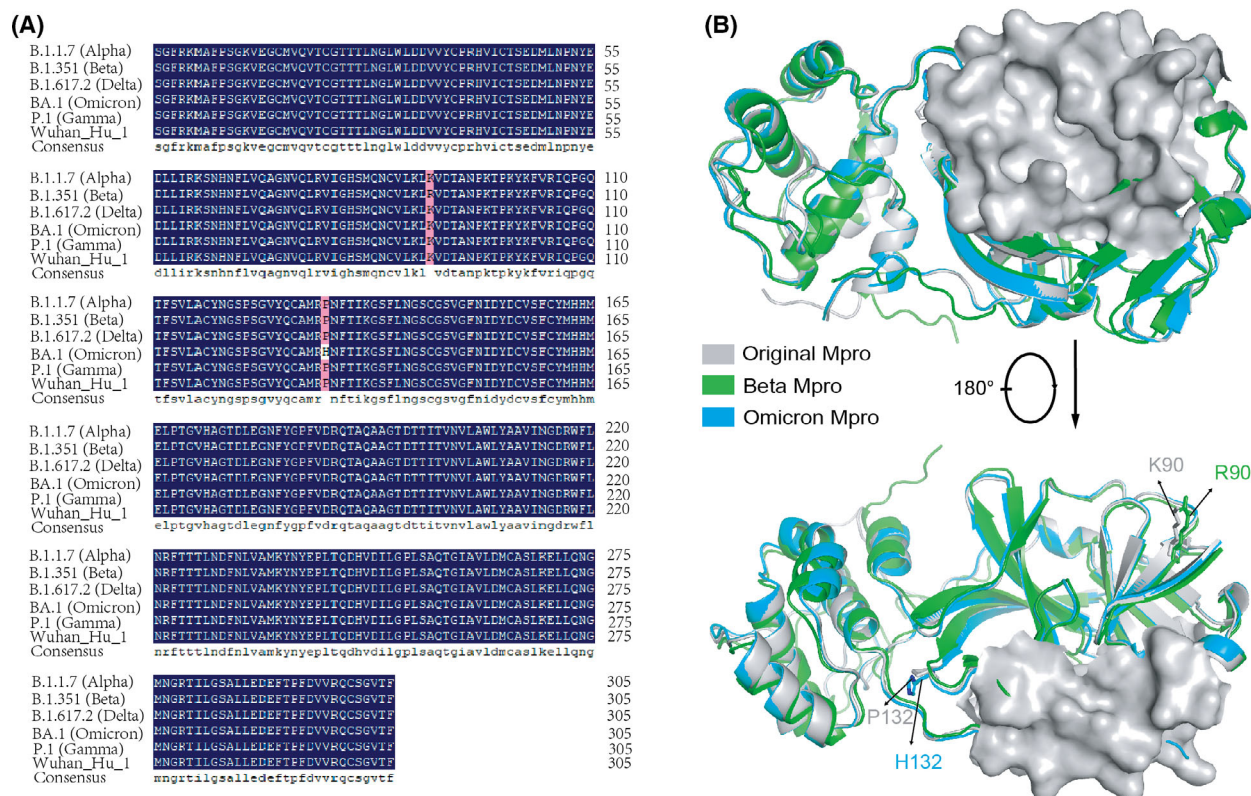


Fig. 1. Conservation analysis of M^{Pro} in five dominant SARS-CoV-2 variants. (A) Sequence alignment of SARS-CoV-2 M^{Pro} in the original strain of Wuhan-Hu-1 and five dominant variants of alpha, beta, gamma, delta, and omicron. The amino acid color in dark blue indicates 100% conservation. (B) Structural comparison of the original M^{Pro} with variant M^{Pro}s in the beta variant (βM^{Pro}) and omicron variant (oM^{Pro}). The surface shown in gray represents the active pocket of M^{Pro}. The amino acids of K90 and P132 in the original M^{Pro}, R90 in βM^{Pro}, and H132 in oM^{Pro} are shown with sticks.

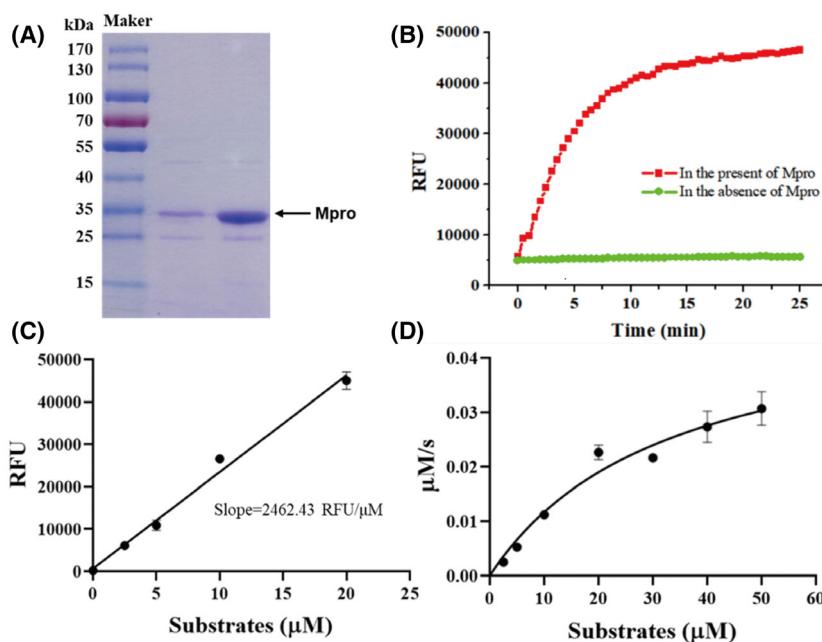


Fig. 2. SARS-CoV-2 M^{pro} expression and characterization. (A) SDS/PAGE of recombinant M^{pro} (lane 1–2). (B) The relative fluorescence unit (RFU) of FRET substrates with (red) or without (green) recombinant M^{pro} . (C) A standard curve converts RFU to the amount of the cleaved substrates. Error bars are the mean \pm SD ($n = 3$). (D) Michaelis–Menten plot of recombinant M^{pro} with various concentrations of FRET substrates. Error bars are the mean \pm SD ($n = 3$).

M^{pro} was mixed with different concentrations of FRET substrates (0–50 μM), the initial velocity was measured and plotted against substrate concentration using the Michaelis–Menten equation in GRAPH PAD PRISM 8 (Fig. 2D). The best-fit values of k_m and V_{max} were 33.12 μM and 0.05 $\mu\text{M}\cdot\text{s}^{-1}$, respectively. The calculated $k_{\text{cat}}\cdot k_m^{-1}$ was 7608.69 $\text{s}^{-1}\cdot\text{M}^{-1}$, suggesting that the obtained M^{pro} are active enough for further research.

The cotton flower extract strongly inhibits SARS-CoV-2 M^{pro}

To test whether the cotton plant extracts possess inhibition potential against SARS-CoV-2, we prepared extracts of different cotton tissues, such as flowers, bolls, leaves, roots, and stems. The inhibitory ratios of the extracts at the same concentrations against SARS-CoV-2 M^{pro} are shown in Fig. S1. The crude extract of the cotton flower exhibits a significant inhibitory effect against M^{pro} , suggesting that cotton flower contains some candidate inhibitory ingredients that can inhibit SARS-CoV-2 M^{pro} .

Virtual screening of SARS-CoV-2 M^{pro} inhibitors based on the metabolite database of cotton flower

To identify the natural products that inhibit SARS-CoV-2 M^{pro} in cotton flowers, we established a

metabolite database according to the in-house CF metabolome data in our laboratory (data not yet published). A total of 228 metabolites were identified (Fig. S2A, Table S1), including 89 flavonols, 60 flavonoids, 33 isoflavones, 11 anthocyanins, etc., of which 131 metabolites have well-defined three-dimensional structures and can be used for virtual screening. Virtual screening was carried out with CF metabolites as ligands and SARS-CoV-2 M^{pro} (PDB: 6LU7) as the receptor. As shown in Fig. S2B, all the ligands can bind to the active pocket of M^{pro} with docking scores less than $-5.5 \text{ kcal}\cdot\text{mol}^{-1}$. Moreover, most of the metabolites exhibited docking scores less than $-7.0 \text{ kcal}\cdot\text{mol}^{-1}$, indicating their great inhibitory potentials against SARS-CoV-2 M^{pro} .

Predicted binding poses of CF metabolites to SARS-CoV-2 M^{pro}

As reported, the CF metabolites astragalin, myricitrin, astilbin, kaempferitrin, and kaempferol have broad-spectrum antiviral activity [24–29], so these five metabolites were researched further. Astragalin could most strongly bind to M^{pro} with the lowest docking score of $-9.4 \text{ kcal}\cdot\text{mol}^{-1}$ through many favorable interactions, such as hydrogen bonds (H-bonds) and hydrophobic interactions (Fig. 3A). Specifically, astragalin forms H-bonds with the main chains of L141,

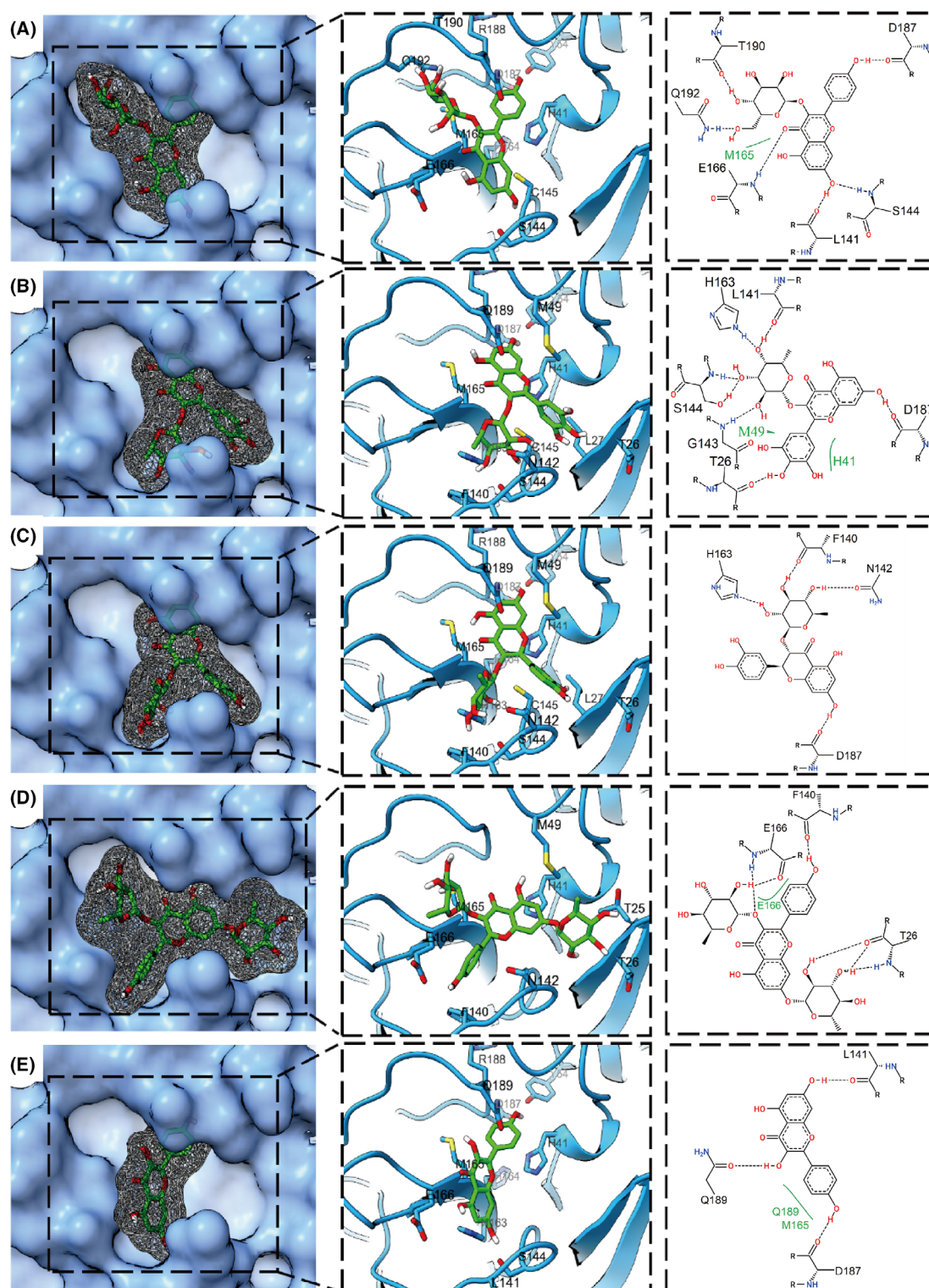


Fig. 3. Binding poses of six CF metabolites with SARS-CoV-2 M^{pro}. (A–E) Three views of astragalalin (A), myricitrin (B), astilbin (C), kaempferitrin (D), and kaempferol (E) binding with SARS-CoV-2 M^{pro} (PDB: 6LU7). SARS-CoV-2 M^{pro} is represented by the blue surface (left), blue illustration (middle), and 2D amino acid residues (right). CF metabolites are represented by green sticks (left), green sticks with gray mesh (middle), and 2D structures (right). The dashed black lines and solid green lines in the right figures represent hydrogen bonds and hydrophobic interactions, respectively.

S144, C145, E166, D187, T190, etc., and the side chains of S144, Q192. In addition, hydrophobic interactions also exist between astragaloside and residues H41, M165, and Q189 of M^{pro}. The docking scores of myricitrin and astilbin against M^{pro} are -9.2 and -9.1 kcal·mol⁻¹, respectively. Myricitrin binds to M^{pro} by 7 H-bonds with L141, H163, S144, G143, T26, and D187, as well as two hydrophobic interactions with M49 and H41 (Fig. 3B). For astilbin, the interactions against M^{pro} are mainly H-bonds formed with H163, F140, N142, and D187 (Fig. 3C). Kaempferitrin, another classic antiviral molecule, can also bind tightly to M^{pro} with a binding score of -8.5 kcal·mol⁻¹ by several H-bonds and hydrophobic interactions against F140, T26, and E166 (Fig. 3D). As a comparison, kaempferol, the aglycon of kaempferitrin, was also used to evaluate the binding ability against M^{pro}. Slightly weaker interactions between kaempferol and

M^{pro} are observed with a binding score of -7.8 kcal·mol⁻¹ by 3 H-bonds and 2 hydrophobic interactions (Fig. 3E). In total, the binding of the above molecules can surpassingly block the active pocket and cover the catalytic residues H41 and C145 of M^{pro}, indicating their excellent inhibition potentials against M^{pro}.

Strong inhibition of CF metabolites against SARS-CoV-2 by *in vitro* assay

We tested the inhibitory activity of the above five major ingredients from cotton flower: astragaloside, myricitrin, astilbin, kaempferitrin, and kaempferol. Among the metabolites, the strongest inhibition of SARS-CoV-2 M^{pro} was observed for astragaloside, with a lowest half-maximal inhibitory concentration (IC₅₀) of 0.13 μ M (Fig. 4A). The IC₅₀ values were 10.73 μ M for

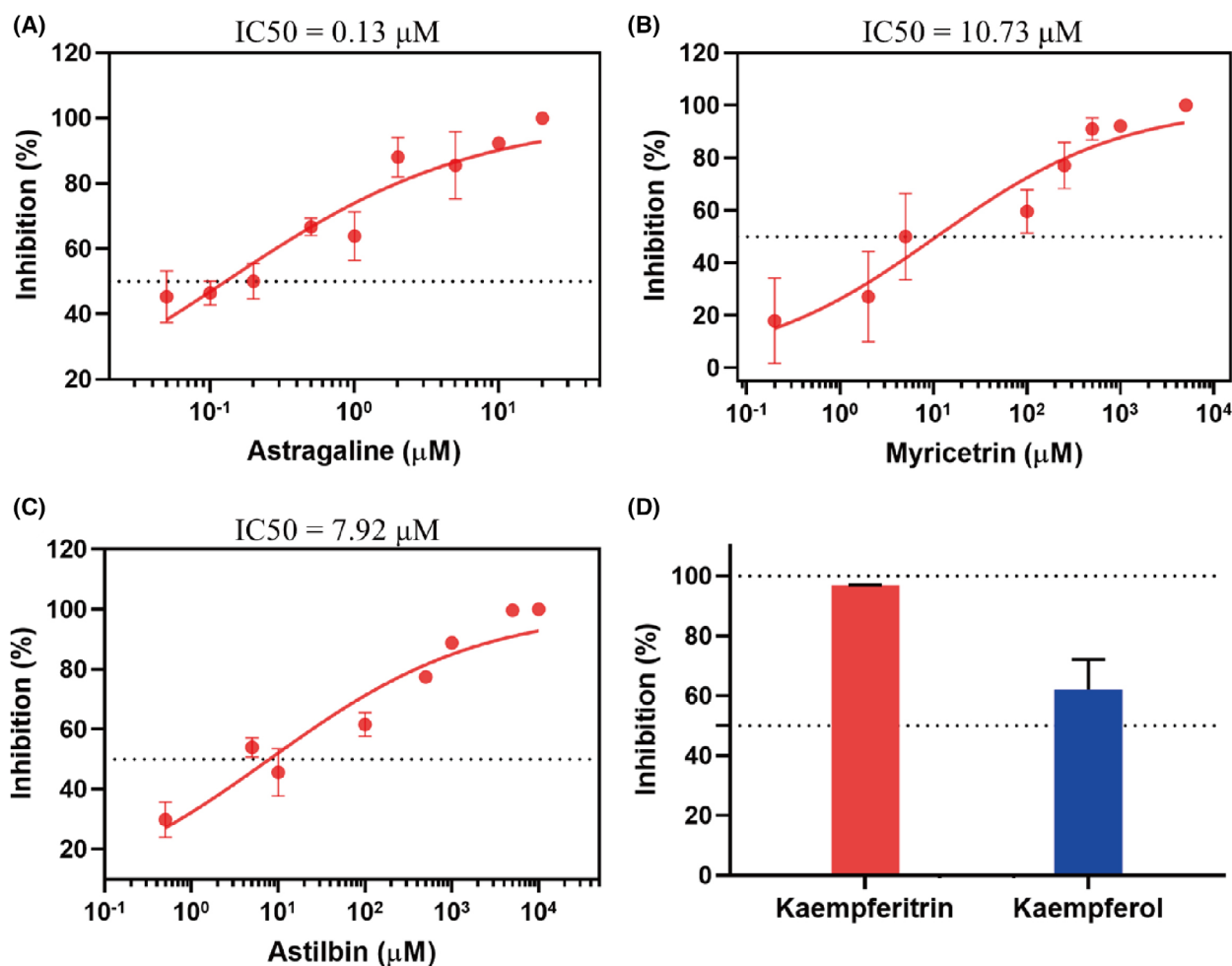


Fig. 4. Inhibition of SARS-CoV-2 M^{pro} by six CF metabolites. (A–C) the *in vitro* inhibition curves of SARS-CoV-2 M^{pro} by astragaloside (A), myricitrin (B), and astilbin (C). (D) The inhibition rates of SARS-CoV-2 M^{pro} by 10 μ M kaempferitrin (red) and kaempferol (blue). The data are the means \pm standard deviation (SD) of three repeats.

myricitrin and 7.92 μM for astilbin (Fig. 4B,C), which were slightly higher than the IC_{50} value of astragalín. The inhibitory activity of kaempferol against M^{pro} was slightly less than the inhibitory activity of kaempferitrin (Fig. 4D). At a final concentration of 10 μM , kaempferol inhibited M^{pro} by 96.81%, whereas the inhibition ratio of kaempferol was only 62.05%. In total, the above compounds showed high inhibitory activities against SARS-CoV-2 M^{pro} , especially astragalín, myricitrin and astilbin, indicating their excellent inhibitory potentials on the replication of SARS-CoV-2.

Discussion

SARS-CoV-2 is still spreading and mutating and is likely to coexist with humans for a long time. Improving the daily diet and consuming more functional foods containing natural SARS-CoV-2 inhibitors may be an effective means for preventing SARS-CoV-2. In this work, SARS-CoV-2 M^{pro} is selected as the target for natural inhibitor screening for its high conservation in different mutant strains. Through theoretical screening and enzymological verification, we found that astragalín, myricitrin, astilbin, kaempferitrin, and kaempferol in cotton flowers showed significant inhibition against SARS-CoV-2 M^{pro} , indicating that cotton flower may be an effective natural material for inhibiting SARS-CoV-2.

Generally, studies on cotton plant have mainly focused on its agronomic properties, such as fiber and oil, and have often ignored its medicinal and edible value. Cotton flowers are rich in flavonoids, such as quercetin, isoquercitrin, and quercimeritrin. Interestingly, most of these flavonoids have a broad spectrum of antiviral activity. Astragalín, for instance, shows great antihepatitis C virus [25] and anti-influenza activity [35]. Myricitrin, in addition to inhibiting influenza virus, can also inhibit coxsackie A16 viruses [26]. Astilbin also possesses antiviral properties against human immunodeficiency virus-1 [27] as well as bovine and equine herpesviruses [36]. In addition, kaempferitrin and its aglycone kaempferol can inhibit a series of viruses, such as Dengue virus-2 [37], pseudorabies virus [28], and bovine herpesvirus 1 [29]. Since the SARS-CoV-2 outbreak, although many studies have indicated the inhibitory potential of some substances against M^{pro} by virtual screening, few of them have been validated by enzymic methods. In this work, based on the great inhibitory activity of cotton flowers against SARS-CoV-2 M^{pro} , we selected and verified several natural inhibitors of M^{pro} , that is, astragalín, myricitrin, astilbin, kaempferitrin, and kaempferol.

However, it remains to further study whether these M^{pro} inhibitors can surpassingly prevent the replication of SARS-CoV-2.

Continuous variation in SARS-CoV-2 brings great difficulties to epidemic prevention. Considering that cotton flowers are rich in natural products that effectively inhibit highly conserved M^{pro} , which might be processed as a functional food for daily use to prevent various SARS-CoV-2 variants.

Acknowledgements

The authors are grateful to Xinpei Han for providing the cotton seeds. This research was funded by Sponsored by State Key Laboratory of Cotton Biology Open Fund (CB2021A08) and the Key Science and Technology Research of Henan Province (No. 222102310043).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

MR, FL and WW conceived and designed the project; WL, YZ, TD and WW acquired the data; XJ provided the resources; WL, TD, YH and WW analyzed and interpreted the data; LZ, MZ, MR, FL and WW wrote and reviewed the paper; WL, TD and WW revised the paper. All authors have read and agreed to the published version of the manuscript.

Data availability statement

All data generated or analyzed during this study are included in this published article (and its supporting information files). Samples of the compounds can be available from the authors.

References

- 1 Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, et al. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol.* 2020;5:536–44.
- 2 Tao K, Tzou PL, Nouhin J, Bonilla H, Jagannathan P, Shafer RW. SARS-CoV-2 antiviral therapy. *Clin Microbiol Rev.* 2021;34:e0010921.
- 3 Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol.* 2020;5:1403–7.

- 4 Wang R, Zhang Q, Ge J, Ren W, Zhang R, Lan J, et al. Analysis of SARS-CoV-2 variant mutations reveals neutralization escape mechanisms and the ability to use ACE2 receptors from additional species. *Immunity*. 2021;**54**:1611–21.e5.
- 5 Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*. 2020;**583**:459–68.
- 6 Khateeb J, Li Y, Zhang H. Emerging SARS-CoV-2 variants of concern and potential intervention approaches. *Crit Care*. 2021;**25**:244.
- 7 Shin D, Mukherjee R, Grewe D, Bojkova D, Baek K, Bhattacharya A, et al. Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature*. 2020;**587**:657–62.
- 8 Jin Z, Du X, Xu Y, Deng Y, Yang H. Structure of Mpro from COVID-19 virus and discovery of its inhibitors. *Nature*. 2020;**582**:1–9.
- 9 Yao YS, Luo Z, Zhang XW. In silico evaluation of marine fish proteins as nutritional supplements for COVID-19 patients. *Food Funct*. 2020;**11**:5565–72.
- 10 Baig MH, Sharma T, Ahmad I, Abohashrh M, Alam MM, Dong J-J. Is PF-00835231 a pan-SARS-CoV-2 Mpro inhibitor? A comparative study. *Molecules*. 2021;**26**:1678.
- 11 Sur VP, Sen MK, Komrskova K. In silico identification and validation of organic triazole based ligands as potential inhibitory drug compounds of SARS-CoV-2 main protease. *Molecules*. 2021;**26**:6199.
- 12 Chen J, Zhang Y, Zeng D, Zhang B, Ye X, Zeng Z, et al. Merbromin is a mixed-type inhibitor of 3-chymotrypsin like protease of SARS-CoV-2. *Biochem Biophys Res Commun*. 2022;**591**:118–23.
- 13 Vandyck K, Abdelnabi R, Gupta K, Jochmans D, Jekle A, Deval J, et al. ALG-097111, a potent and selective SARS-CoV-2 3-chymotrypsin-like cysteine protease inhibitor exhibits in vivo efficacy in a Syrian hamster model. *Biochem Biophys Res Commun*. 2021;**555**:134–9.
- 14 Chiou W-C, Chen J-C, Chen Y-T, Yang J-M, Hwang L-H, Lyu Y-S, et al. The inhibitory effects of PGF and EGCG against the SARS-CoV-2 3C-like protease. *Biochem Biophys Res Commun*. 2022;**591**:130–6.
- 15 Ibrahim MAA, Abdeljawaad KAA, Abdelrahman AHM, Hegazy M-EF. Natural-like products as potential SARS-CoV-2 Mpro inhibitors: in-silico drug discovery. *J Biomol Struct Dyn*. 2021;**39**:5722–34.
- 16 Ibrahim MAA, Mohamed EAR, Abdelrahman AHM, Allemailem KS, Moustafa MF, Shawky AM, et al. Rutin and flavone analogs as prospective SARS-CoV-2 main protease inhibitors: in silico drug discovery study. *J Mol Graph Model*. 2021;**105**:107904.
- 17 Ibrahim MAA, Abdelrahman AHM, Hussien TA, Badr EAA, Mohamed TA, El-Seedi HR, et al. In silico drug discovery of major metabolites from spices as SARS-CoV-2 main protease inhibitors. *Comput Biol Med*. 2020;**126**:104046.
- 18 Song JK, Zhang L, Xu YF, Yang DZ, Zhang L, Yang SY, et al. The comprehensive study on the therapeutic effects of baicalin for the treatment of COVID-19 in vivo and in vitro. *Biochem Pharmacol*. 2021;**183**:114302.
- 19 Zhang Y, Gao H, Hu X, Wang Q, Zhong F, Zhou X, et al. Structure-based discovery and structural basis of a novel broad-spectrum natural product against the main protease of coronavirus. *J Virol*. 2022;**96**:e0125321.
- 20 Zhao Y, Dou J, Wu T, Aisa HA. Investigating the antioxidant and acetylcholinesterase inhibition activities of *Gossypium herbaceum*. *Molecules*. 2013;**18**:951–62.
- 21 Liu Y, Aisa HA, Ji C, Yang N, Zhu H, Zuo P. Effects of *Gossypium herbaceum* extract administration on the learning and memory function in the naturally aged rats: neuronal niche improvement. *J Alzheimers Dis*. 2012;**31**:101–11.
- 22 Lee J, Kim Y, Lee C, Lee H, Han S. Antiviral activity of some flavonoids on herpes simplex viruses. *Korean J Pharmacogn*. 1999;**30**:34–9.
- 23 Jeong HJ, Ryu YB, Park SJ, Kim JH, Kwon HJ, Kim JH, et al. Neuraminidase inhibitory activities of flavonols isolated from *Rhodiola rosea* roots and their in vitro anti-influenza viral activities. *Bioorg Med Chem*. 2009;**17**:6816–23.
- 24 Cui MY, Xiao MW, Xu LJ, Chen Y, Liu AL, Ye J, et al. Bioassay of ferulic acid derivatives as influenza neuraminidase inhibitors. *Arch Pharm*. 2020;**353**:e1900174.
- 25 Luo H, Li Y. Inhibitory effect of astragaloside on hepatitis C virus. *Chin J Biol*. 2017;**30**:623–7.
- 26 Van NTH, Vien TA, Nhiem NX, Van Kiem P, Van Minh C, Long PQ, et al. Chemical components of *Ardisia splendens* leaves and their activity against coxsackie A16 viruses. *Nat Prod Commun*. 2014;**9**:643–5.
- 27 Liangsakul P, Kuhakarn C, Hongthong S, Jariyawat S, Suksen K, Akkarawongsapat R, et al. Anti-HIV 1 activity of xanthones from the bark of *Mammea harmandii*. *Nat Prod Commun*. 2018;**13**:53–6.
- 28 Chen X, Chen YQ, Yin ZQ, Wang R, Hu HY, Liang XX, et al. Kaempferol inhibits pseudorabies virus replication in vitro through regulation of MAPKs and NF-kappa B signaling pathways. *J Integr Agric*. 2021;**20**:2227–39.
- 29 Zhu L, Wang P, Yuan W, Zhu G. Kaempferol inhibited bovine herpesvirus 1 replication and LPS-induced inflammatory response. *Acta Virol*. 2018;**62**:220–5.
- 30 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;**596**:583–9.

- 31 Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods*. 2022;**19**:679–82.
- 32 Ni W-J, Chen X-X, Wei S-Y, Lan L-L, Qiu R-J, Ling Y-P, et al. Study on the mechanism of active components of Liupao tea on 3CLpro based on HPLC-DAD fingerprint and molecular docking technique. *J Food Biochem*. 2021;**45**:e13707.
- 33 Wang WJ, Wu Y, Xu HH, Shang Y, Chen YC, Yan ME, et al. Accumulation mechanism of indigo and indirubin in *Polygonum tinctorium* revealed by metabolite and transcriptome analysis. *Ind Crops Prod*. 2019;**141**:111783.
- 34 Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. In: Hempel JE, Williams CH, Hong CC, editors. *Chemical biology: methods and protocols*. New York, NY: Springer New York; 2015. p. 243–50.
- 35 Ding F, Liu J, Du R, Yu Q, Gong L, Jiang H, et al. Qualitative and quantitative analysis for the chemical constituents of *Tetragium hemsleyanum* Diels et Gilg using ultra-high performance liquid chromatography/hybrid quadrupole-orbitrap mass spectrometry and preliminary screening for anti-influenza virus components. *Evid Based Complement Alternat Med*. 2019;**2019**:9414926.
- 36 Parreira RM, Simoni IC, Favero OA, Lago JHG, Mecchi MC, Barrosa KH, et al. Phytochemical profile and in vitro evaluation of extracts from leaves of *Drimys brasiliensis* (Winteraceae) against bovine and equine herpesviruses. *J Appl Pharm Sci*. 2017;**7**:122–7.
- 37 Mello CD, Valente LMM, Wolff T, Lima-Junior RS, Fialho LG, Marinho CF, et al. Decrease in dengue virus-2 infection and reduction of cytokine/chemokine production by *Uncaria guianensis* in human hepatocyte cell line Huh-7. *Mem Inst Oswaldo Cruz*. 2017;**112**: 458–68.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. The *in vitro* anti-SARS-CoV-2 M^{pro} activity of cotton extracts. Error bars are the mean \pm SD ($n = 3$).

Fig. S2. Virtual screening of M^{pro} inhibitors based on CF metabolites. (A) The statistics and classification for all identified metabolites (blue) or that with well-defined three-dimensional structures (red) in CF tissue. (B) The statistics for binding energies of the CF metabolites to the SARS-CoV-2 M^{pro}. The list of CF metabolites and corresponding binding energies are shown in Table S1.

Table S1. The list of CF Compounds and their binding scores with SARS-CoV-2 M^{pro}.