

# A computationally designed antigen eliciting broad humoral responses against SARS-CoV-2 and related sarbecoviruses

Received: 5 November 2021

Accepted: 23 August 2023

Published online: 25 September 2023

 Check for updates

Sneha Vishwanath<sup>1</sup>, George William Carnell<sup>1</sup>, Matteo Ferrari<sup>2</sup>, Benedikt Asbach<sup>3</sup>, Martina Billmeier<sup>3</sup>, Charlotte George<sup>1</sup>, Maria Suau Sans<sup>1</sup>, Angalee Nadesalingam<sup>1</sup>, Chloe Qingzhou Huang<sup>1</sup>, Minna Paloniemi<sup>1</sup>, Hazel Stewart<sup>4</sup>, Andrew Chan<sup>1</sup>, David Arthur Wells<sup>2</sup>, Patrick Neckermann<sup>3</sup>, David Peterhoff<sup>3,5</sup>, Sebastian Einhauser<sup>3</sup>, Diego Cantoni<sup>6</sup>, Martin Mayora Neto<sup>6</sup>, Ingo Jordan<sup>7</sup>, Volker Sandig<sup>7</sup>, Paul Tonks<sup>1</sup>, Nigel Temperton<sup>6</sup>, Simon Frost<sup>2,8,9</sup>, Katharina Sohr<sup>10</sup>, Maria Teresa Lluesma Ballesteros<sup>10</sup>, Farzad Arbabi<sup>10</sup>, Johannes Geiger<sup>10</sup>, Christian Dohmen<sup>10</sup>, Christian Plank<sup>10</sup>, Rebecca Kinsley<sup>1,2</sup>, Ralf Wagner<sup>2,3,5</sup> & Jonathan Luke Heeney<sup>1,2</sup>✉

The threat of spillovers of coronaviruses associated with the severe acute respiratory syndrome (SARS) from animals to humans necessitates vaccines that offer broader protection from sarbecoviruses. By leveraging a viral-genome-informed computational method for selecting immune-optimized and structurally engineered antigens, here we show that a single antigen based on the receptor binding domain of the spike protein of sarbecoviruses elicits broad humoral responses against SARS-CoV-1, SARS-CoV-2, WIV16 and RaTG13 in mice, rabbits and guinea pigs. When administered as a DNA immunogen or by a vector based on a modified vaccinia virus Ankara, the optimized antigen induced vaccine protection from the Delta variant of SARS-CoV-2 in mice genetically engineered to express angiotensin-converting enzyme 2 and primed by a viral-vector vaccine (AZD1222) against SARS-CoV-2. A vaccine formulation incorporating mRNA coding for the optimized antigen further validated its broad immunogenicity. Vaccines that elicit broad immune responses across subgroups of coronaviruses may counteract the threat of zoonotic spillovers of betacoronaviruses.

Among the coronaviruses of the greatest pandemic risk are the viruses of the *Betacoronavirus* genus that bind to angiotensin-converting enzyme 2 (ACE-2)<sup>1,2</sup>. Over the past two decades, two ACE-2-binding sarbecoviruses have spilled over into human populations, causing the severe acute respiratory syndrome (SARS) epidemic in 2002–2003 and the current SARS coronavirus 2 (SARS-CoV-2) pandemic. Bats are a

reservoir of a large number of SARS-CoV-like ACE-2-binding sarbecoviruses which pose a constant threat for future spillover into humans, with the potential to cause new epidemics<sup>3,4</sup>. In addition to the emergence of new ACE-2-binding viruses from zoonotic reservoirs, another concern is the emergence of mutations in variants of these viruses that are capable of escaping vaccine-induced immunity—a constant observation and

A full list of affiliations appears at the end of the paper. ✉e-mail: [jlh66@cam.ac.uk](mailto:jlh66@cam.ac.uk)

concern in the current pandemic. As human infections increase globally during the pandemic, the virus has continued to accrue mutations, most notably in the spike protein<sup>5</sup>. An accumulating number of variants of concern (VOCs) have implications for increased transmission and escape from natural and vaccine-induced immunity<sup>6–9</sup>. The N501Y (asparagine to tyrosine) substitution in the receptor binding domain (RBD) of the spike protein is a common feature of VOCs and is associated with increased affinity of the viral spike protein to the ACE-2 receptor and with a subsequent increase in transmission<sup>10</sup>. Notably, the majority of these mutations reported in VOCs are in the receptor binding motif in the RBD (or around it), which interacts with ACE-2, as well as in the regions that induce highly potent neutralizing antibodies<sup>11,12</sup>. In the key RBD epitopes, the Delta VOC<sup>13</sup> has L452R and T478K mutations, whereas the Omicron lineage VOCs have multiple mutations<sup>14</sup>. The continued emergence of these VOCs during the ongoing coronavirus disease 2019 (COVID-19) pandemic and the constant threat of new zoonotic spillovers of coronaviruses from animals to humans highlight the need for next generation vaccines with broader protection from ACE-2-binding sarbecoviruses as well as from the emerging VOCs.

In this study, to increase the coverage to all the viruses of the *Sarbecovirus* subgenus of betacoronaviruses, we used a digitally immune-optimized synthetic vaccine (DIOSynVax) technology to design antigens. These computationally immune-optimized and structurally engineered antigens are selected in vivo to induce immune responses across a group of related viruses. First, we generated a phylogenetically informed RBD-based antigen by comparing all the known human and animal reservoir *Sarbecovirus* sequences. This antigen design was further used as a backbone for designing both epitope-optimized and immune-refocused designs using available structural data for the spike protein in complex with RBD-binding monoclonal antibodies, in this case specifically those that bound both SARS-CoV and SARS-CoV-2, such as S309 (ref. 15) and CR3022 (ref. 16). The nucleic acid sequences of these in-silico-designed antigens were optimized for expression in human cells, and synthetic genes expressing each unique antigen were shuttled in an expression cassette for consecutive in vitro and in vivo screens in Bagg albino laboratory (BALB/c) mice. The best-in-class immunologically optimal antigen, which we designated as 'T2\_17', was further validated by DNA immunization screens in guinea pigs and rabbits. To further validate the utility of this antigen to boost specific responses on the background of spike-specific immune responses to pre-existing early Wuhan isolates (used by most licensed vaccines), the T2\_17 antigen was administered as a heterologous boost using either DNA or MVA immunogens to transgenic mice expressing human ACE-2 (K18-hACE-2) previously primed with the AZD1222 vaccine for SARS-CoV-2. RBD-specific immune responses were observed in groups immunized with the T2\_17 antigen. Further immunogenicity of the T2\_17 antigen was confirmed in mice and guinea pigs as a messenger (m)RNA-delivered immunogen based on chemically modified mRNA<sup>17</sup> in a lipidoid nanoparticle formulation (LNP)<sup>18</sup>. These studies confirmed that these computationally derived antigens can induce broad humoral responses, using a single RBD-based antigen covering SARS-CoV, SARS-CoV-2 (including VOCs) and related bat sarbecoviruses.

## Results

### In silico design of antigens

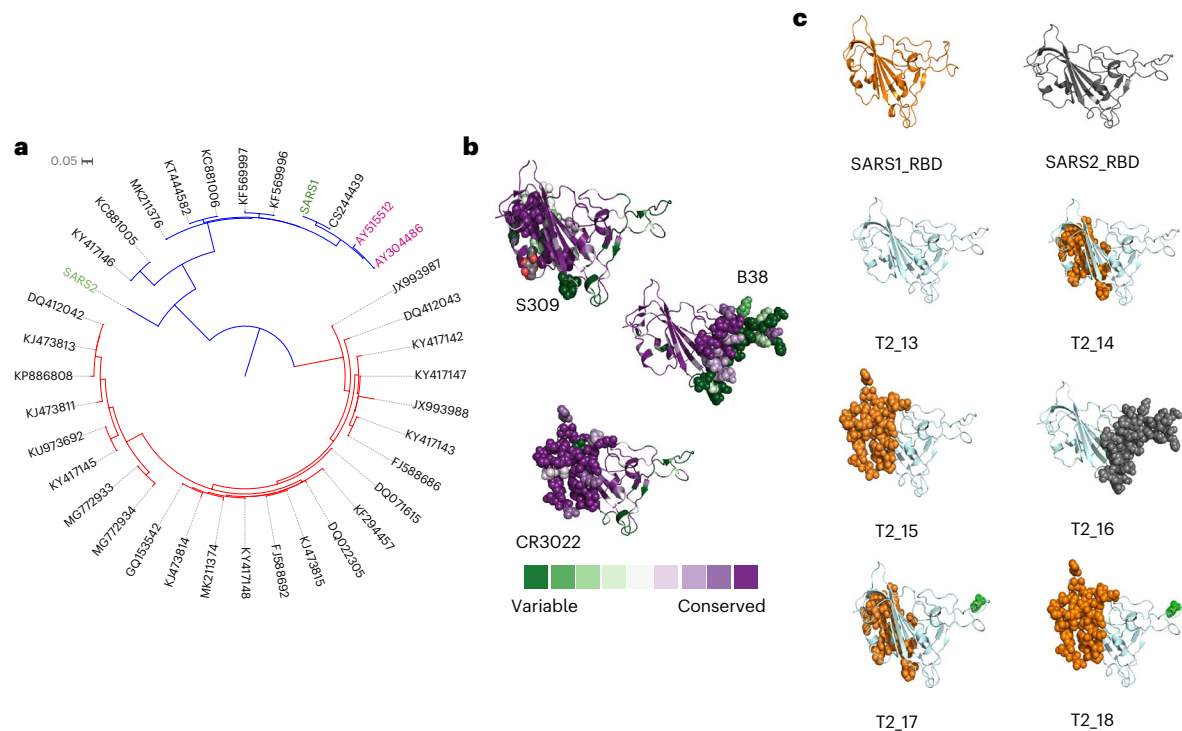
Sequences of spike protein of viruses belonging to the *Sarbecovirus* subgenus were compiled from the National Centre for Biotechnology Information (NCBI) virus database<sup>19</sup> and further pruned to remove poor-quality and redundant sequences. The hCoV-19/Wuhan/IVDC-HB-01/2019 strain of SARS-CoV-2 was used for the analyses. The phylogenetic tree of these sequences is presented in Fig. 1a. Two distinct clades are observed in the tree: clade 1 viruses, which do not interact with the ACE-2 receptor<sup>1,20</sup>, and clade 2 viruses, which do. Clade 1 viruses share many of the sequence features of the members of clade 2 but

possess deletions within the ACE-2 binding region (Supplementary Fig. 1). An optimized core sequence (T2\_13) was designed, such that the antigen was phylogenetically the closest to all the sarbecoviruses represented in the phylogenetic tree shown in Fig. 1a. Due to the evolutionary relatedness between the input sequences and the evolution model used in the pipeline, the design captures both the conserved as well as distinct epitopes of the input sequences. To further understand the importance of the amino acid composition of epitopes in generating antibody responses, we modified T2\_13 to display the exact amino acid sequences of epitopes of SARS-CoV for monoclonal antibodies S309 (ref. 15) (T2\_14) and CR3022 (ref. 16) (T2\_15) and of SARS-CoV-2 for monoclonal antibody B38 (ref. 11) (T2\_16). The sequences of epitopes for monoclonal antibodies S309 (ref. 15) and CR3022 (ref. 16) are highly conserved across the sequences considered in this study, while the sequence of epitopes for monoclonal antibody B38 (ref. 11) is highly divergent (Fig. 1b). These antibodies were chosen for the analyses, as these were the only biochemically and structurally characterized antibodies available at the time of generating the antigen designs. We further modified the epitope region for monoclonal antibody B38 (ref. 11) by introducing a glycosylation site on the backbone of T2\_14 (T2\_17) and T2\_15 (T2\_18). This was done to mask the divergent epitope region and enhance the presentation of the conserved epitopes to the immune system. The masking of epitopes by introducing glycans has been exploited by many viruses such as hepatitis C virus<sup>21</sup>, Lassa virus<sup>22</sup> and influenza virus<sup>23</sup> to escape natural immunity and we used this strategy to train the immune system to elicit preferential immune response towards the conserved but subdominant epitopes. To compare the immunogenicity of soluble and membrane-anchored RBD-based vaccines, membrane-anchored forms of T2\_13 and T2\_17 (T2\_13\_TM and T2\_17\_TM, respectively) were generated. The structural stability of these designs was evaluated in silico using the BUILD module of the FoldX<sup>24</sup> algorithm, with T2\_13 as the reference model. Structural models of these vaccine antigens are presented in Fig. 1c.

### Antigen selection and immunogenicity confirmation in BALB/c mice

In vivo screening in BALB/c mice was performed by immunizing mice with the in silico designed antigens and SARS-CoV-2 RBD (hCoV-19/Wuhan/IVDC-HB-01/2019) as a DNA immunogen (Fig. 2a). The sera from immunized mice were assessed for cross-reactive antibodies against spike proteins in a flow cytometry-based cell-surface display assay. Binding against four spike proteins, namely SARS-CoV (SARS-Tor2), SARS-CoV-2 (hCoV-19/Wuhan/IVDC-HB-01/2019), WIV16 and RaTG13, were tested. Sera taken 2 weeks following the second immunization with the antigen designs demonstrated the binding profile of sera induced by the vaccine candidates for different spike proteins (Fig. 2b). Sera from all antigen-immunized mice showed higher binding than those of PBS-immunized mice across the four spike proteins, suggesting seroconversion of the mice on immunization with the antigens.

Across the four spike proteins, no significant differences in binding were observed for sera from mice immunized with T2\_13 and sera from mice immunized with SARS-CoV-2 RBD (all  $P > 0.05$ , Mann-Whitney  $U$  (MWU) test), suggesting that epitopes in this design are biased towards SARS-CoV-2 RBD. For the T2\_16 design, in which the epitope region for mAb B38 was mutated to the epitope region on SARS-CoV-2, binding to SARS-CoV, WIV16 and RaTG13 declined in comparison with T2\_13 ( $P < 0.05$ , MWU test) without significant changes in binding to SARS-CoV-2. Matching of the epitopes of S309 and CR3022 to SARS-CoV (T2\_14 and T2\_15) enhanced the binding to SARS-CoV spike ( $P < 0.05$ , MWU test) but not to other spike proteins. Introduction of a glycosylation site in design T2\_17 significantly enhanced the binding of elicited antibodies to SARS-CoV and RaTG13 ( $P < 0.01$ , MWU test) in comparison with T2\_14, but no difference was observed in T2\_18 in comparison with T2\_15. There was no statistical difference between transmembrane-anchored and soluble designs when delivered as a DNA



**Fig. 1** **In silico design of antigen candidates.** **a**, Phylogenetic tree generated for sarbecoviruses using the protein sequence of the RBD of the spike proteins. The tree was generated using IQ-TREE<sup>37</sup>. Human viruses are represented in green, palm civet viruses in pink and bat viruses in dark grey. The two distinct clades are coloured in red (non-ACE-2 binding) and blue (ACE-2 binding). **b**, Structural models of RBD with epitope regions highlighted as spheres. The backbone of RBD is coloured according to the conservation of amino acid as calculated

by the CONSURF algorithm<sup>39</sup> using the alignment utilized for construction of the phylogenetic tree. The figure was generated and rendered using PyMOL<sup>60</sup> utilizing PDB<sup>39</sup> IDs 6WPS (ref. 15), 6W41 (ref. 16) and 7BZ5 (ref. 11). **c**, Structural representation of the different antigen designs used in the study. The epitopes that were modified to match the wild-type SARS-CoV (coloured orange) and wild-type SARS-CoV-2 (coloured grey) are represented as spheres. The further glycosylation site modification is represented in green.

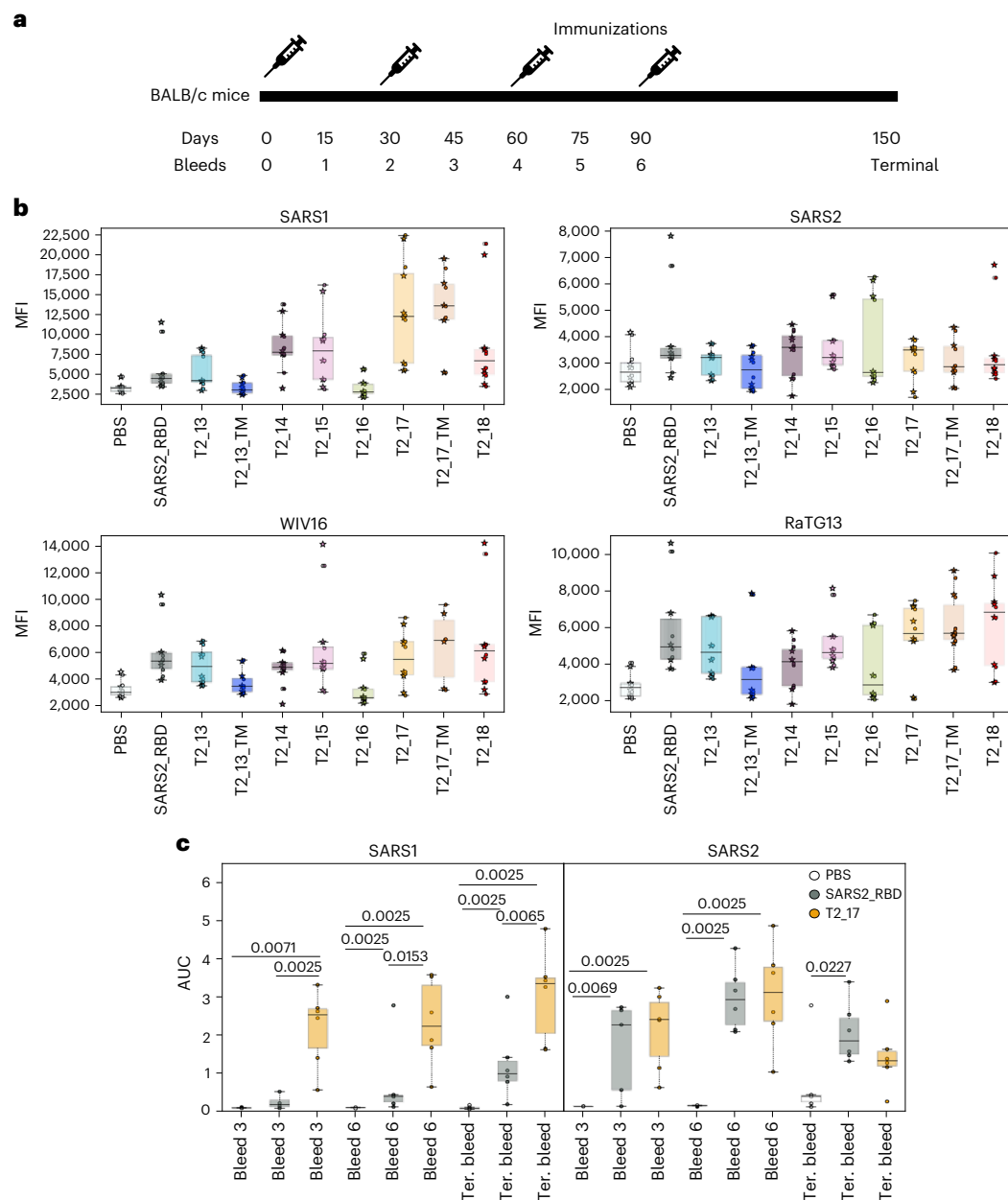
immunogen. As T2\_17 has either the best (or second best) median binding to the four spike proteins, we choose T2\_17 as the lead candidate for further immunological assays.

Elicitation of cross-binding antibodies by T2\_17 was further confirmed by enzyme-linked immunosorbent assay (ELISA) with SARS-CoV RBD and SARS-CoV-2 RBD (Fig. 2c), revealing robust binding antibody responses to both SARS-CoV and SARS-CoV-2 within 2 weeks of the second immunization. T2\_17 elicited stronger responses against SARS-CoV in comparison with SARS-CoV-2 RBD. Against SARS-CoV-2, the two antigens SARS-CoV-2 RBD and T2\_17 generated similar binding antibody responses.

### Immunogenicity confirmation of T2\_17 in outbred animals

To determine the breadth of antibody response and neutralization potency of T2\_17 as a DNA immunogen in outbred animals, guinea pigs were immunized using the European conformity approved and clinically validated Pharmajet Tropis needleless, intradermal delivery device to ensure standardized intradermal delivery (Fig. 3a). As a control, we used a C-terminal glycosylation modified SARS-CoV-2 RBD (SARS2\_RBD\_P521N) (Fig. 3b) which we had previously evaluated in BALB/c mice<sup>25</sup> (Supplementary Fig. 2). Generation of neutralizing antibodies to both SARS-CoV and SARS-CoV-2 was confirmed using pseudoviruses expressing full-length spike proteins of SARS-CoV and SARS-CoV-2. While both T2\_17 and SARS2\_RBD\_P521N generated binding antibodies against both SARS-CoV and SARS-CoV-2 (Extended Data Fig. 1) after one immunization, T2\_17 elicited significantly higher antibodies than SARS2\_RBD\_P521N to SARS-CoV and comparable antibodies against SARS-CoV-2. Higher titres of binding antibodies were detected for T2\_17 to SARS-CoV in comparison with SARS2\_RBD\_P521N after two immunizations, while the responses were comparable for SARS-CoV-2. After

three immunizations, SARS2\_RBD\_P521N induced a higher response to SARS-CoV-2, while T2\_17 had higher responses to SARS-CoV (Extended Data Fig. 1). Neutralizing antibodies were detected for SARS-CoV-2 after the first immunization, while significant neutralizing responses to SARS-CoV developed after two immunizations, although these were more potent from T2\_17 than from SARS2\_RBD\_P521N (Fig. 3c). Higher binding and neutralizing responses by SARS2\_RBD\_P521N to SARS-CoV-2 were expected as it differs from SARS-CoV-2 by only one amino acid. To further confirm whether the T2\_17 antigen generates broader responses than SARS2\_RBD\_P521N, we compared sera induced by these two antigens 28 d post third immunization for neutralization against SARS-CoV (SARS-Tor2), SARS-CoV-2 (hCoV-19/Wuhan/IVDC-HB-01/2019), WIV16 and RaTG13 pseudoviruses. Statistically significant higher neutralizing antibody titres were generated by T2\_17 against SARS-CoV, WIV16 and RaTG13 (Fig. 3d). To further confirm that anti-sera to T2\_17 could abrogate hACE-2 receptor binding, we performed an ELISA-based competition assay (Fig. 3e), demonstrating that T2\_17 and SARS2\_RBD\_P521N anti-sera abrogated binding to the hACE-2 receptor and are comparable to the WHO standard of pooled convalescent COVID-19 patient sera (NIBSC standard 20/162). These findings demonstrated important proof-of-concept of T2\_17 as a single gene-delivered, structurally engineered antigen capable of eliciting broad pan-sarbecovirus neutralizing antibodies. Before clinical trials in humans, a good manufacturing practice lot of T2\_17 DNA was manufactured and evaluated for safety and immunogenicity in rabbits using the same gene delivery device to ensure uniform intradermal administration (Fig. 3f). After one immunization, binding antibodies to SARS-CoV and SARS-CoV-2 were elicited (Extended Data Fig. 2), increasing on subsequent immunizations until a plateau was reached by the fourth immunization. Robust neutralizing antibodies



**Fig. 2 | In vitro selection and in vivo immunogenicity of antigens in BALB/c mice.** **a**, Immunization and bleed schedule of BALB/c mice. Mice were immunized at intervals of 30 d and bled every 15 d. **b**, FACS binding data for the antigens. Sera from mice immunized with designed antigens were screened for binding to SARS-CoV, SARS-CoV-2, WIV16 and RaTG13 spike proteins. The x axis represents all the vaccine designs considered for screening, and the y axis represents the mean fluorescence intensity (MFI). For each mouse serum, two replicates of MFI (represented as filled circles and filled stars) are reported. **c**, Elicitation of

binding antibodies against SARS-CoV and SARS-CoV-2 by T2\_17 was confirmed using ELISA, with SARS-CoV-2RBD as control vaccine design. T2\_17 generated cross-binding antibodies. The x axis represents the bleeds, and the y axis represents the area under the curve (AUC) from ELISA binding curves. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution, the whiskers represent the minimum and maximum (excluding outliers) and the fliers represented as filled circles denote outliers. Two-tailed Mann-Whitney *U* tests demonstrated statistical significance.  $n = 6$  for **b** and **c**.

were observed 2 weeks following the third immunization (Fig. 3g). Sera post 14 d after four immunizations (bleed 4) showed broad neutralizing antibody responses against the SARS-CoV, SARS-CoV-2, Beta, Gamma, Delta and Omicron (BA.1) VOCs and bat sarbecoviruses, WIV16 and RaTG13 (Fig. 3h).

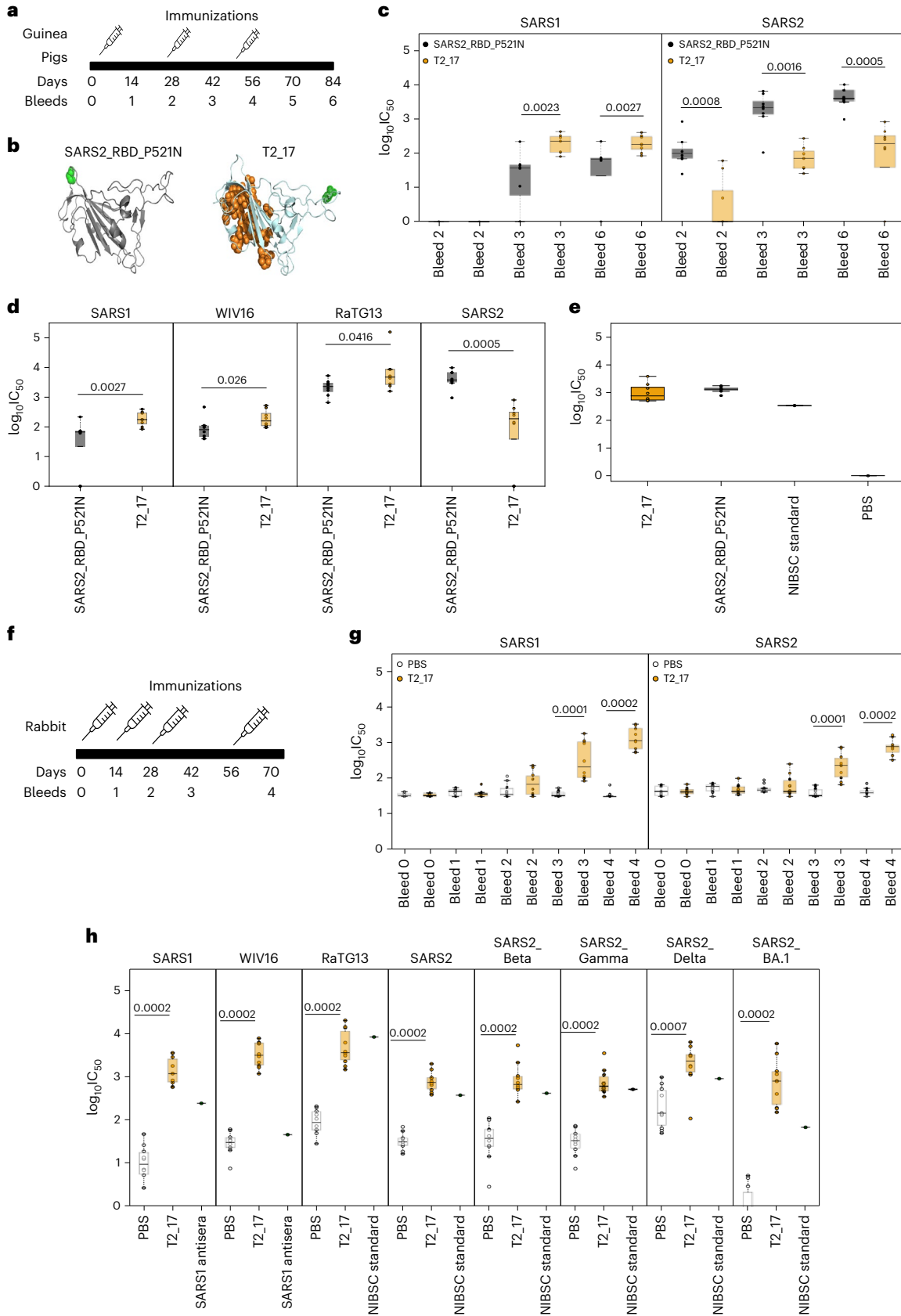
### Challenge studies in mice expressing K18-hACE-2

As almost all the human population is seroconverted either due to natural infection, vaccination or both, we tested the efficacy of the T2\_17 antigen when given as a booster following AZD1222 (ChAdOx1

nCoV-19) as prime vaccine. Although the present immune landscape of the human population is complex due to multiple immunizations and infections by different SARS-CoV-2 variants, we aimed to study the baseline immunogenicity of T2\_17 in the background of vaccination. To address this, homozygous K18-hACE-2 transgenic mice were immunized with  $1.4 \times 10^9$  viral particles of AZD1222 and 4 weeks later boosted with either T2\_17 or the licensed AZD1222 vaccine (Fig. 4a), while the control group received only PBS with each immunization. We administered T2\_17 either as a DNA immunogen or a modified vaccinia virus Ankara (MVA) immunogen. The ChadOx-MVA prime-boost

regime has been shown to be effective in Ebola<sup>26,27</sup>. Eight weeks post boost, all groups of mice were challenged with either a January 2020 isolate of SARS-CoV-2 (Victoria) or the Delta variant of SARS-CoV-2 (Table 1).

Increased binding antibody titres to both SARS-CoV and SARS-CoV-2 after boosting with either AZD1222 or T2\_17 were observed (Extended Data Fig. 3). Statistically significant differences in antibody titres to SARS-CoV-2 were observed 4 weeks after boosting with T2\_17 as



**Fig. 3 | Immunogenicity studies in guinea pigs and rabbits.** **a**, Immunization and bleed schedule of guinea pigs. Guinea pigs were immunized with DNA delivered intradermally using the Tropis PharmaJet device at 28-d intervals and bled every 14 d. **b**, Structure models of the antigen designs used for the study in guinea pigs. The glycosylation site and the modified epitope are represented as green and orange spheres, respectively. **c**, Neutralization of SARS-CoV and SARS-CoV-2 pseudoviruses by guinea pig sera immunized with T2\_17 and SARS2\_RBD\_P521N. The x axis represents the bleed number, and the y axis represents the  $\log_{10}IC_{50}$  values for neutralization curves. **d**, Broad neutralization of SARS-CoV, WIV16, RaTG13 and SARS-CoV-2 pseudoviruses by T2\_17 in comparison with SARS2\_RBD\_P521N. Sera post 28 d after three immunizations (bleed 6) were used for comparison. **e**, ACE-2 competition ELISA. Sera from guinea pigs immunized with T2\_17 and SARS2\_RBD\_P521N. The NIBSC standard (20/162) was used as

control. **f**, Immunization and bleed schedule of rabbits. Rabbits were immunized at intervals of 14 d and bled every 14 d. **g**, Neutralization of SARS-CoV and SARS-CoV-2 pseudoviruses by rabbit sera immunized with T2\_17. The x axis represents the bleed number, and the y axis represents the  $\log_{10}IC_{50}$  values for neutralization curves. **h**, Broad neutralization of SARS-CoV, WIV16, RaTG13, SARS-CoV-2, SARS-CoV-2 Beta, SARS-CoV-2 Gamma, SARS-CoV-2 Delta and SARS-CoV-2 BA.1 pseudoviruses by T2\_17. Sera taken 14 d after the fourth immunization (bleed 4) were used for comparison. NIBSC standards for SARS-CoV-2 and SARS-CoV antiserum were used as reference. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution, the whiskers represent the minimum and maximum (excluding outliers) and the fliers represented as filled circles denote outliers. Two-tailed Mann–Whitney *U* tests demonstrated statistical significance.  $n = 8$  for **c**, **d** and **e**;  $10$  for **g** and **h**.

DNA or MVA immunogen in comparison with boosting with AZD1222, while a statistically significant increase in binding antibody titres to SARS-CoV was observed on boosting with T2\_17 as MVA immunogen (Extended Data Fig. 3). Generation of neutralizing antibodies to SARS-CoV, SARS-CoV-2 and the Delta VOC was confirmed using pseudoviruses expressing full-length spike proteins of SARS-CoV, SARS-CoV-2 and the Delta VOC. Neutralizing antibodies for SARS-CoV-2 and the Delta VOC were detected for all of the groups, except the control group before challenge, while neutralizing antibodies against SARS-CoV were detected only for the T2\_17 MVA-boosted group (Fig. 4b). Two weeks post boost, the antibodies elicited by T2\_17 both as DNA and MVA immunogens neutralized the Delta variant significantly better than the sera from mice boosted with AZD1222 (Fig. 4b). Mice from all the groups, except the controls, survived and continued to gain weight following challenge with either the Victoria strain or the Delta variant (Fig. 4c).

#### Longitudinal serology study in K18-hACE-2 mice

As all the groups were protected in the challenge studies and similar levels of neutralizing antibodies were observed for SARS-CoV-2, we explored whether this could be due to the short interval between the prime and the boost in the challenge study. For this, we primed another group of K18-hACE-2 mice with AZD1222 vaccine and boosted the mice 20 weeks afterwards (Fig. 4d). The mice were boosted with either AZD1222, T2\_17(DNA), T2\_17(MVA) or PBS (Table 2). A group of PBS primed mice was immunized with T2\_17(MVA) at week 20 as a control. The AZD1222/PBS group was included to monitor the antibody titres over time in the absence of the boost.

Only neutralizing antibody titres against SARS-CoV-2 were measured for this longitudinal analysis. In this study as well, significantly higher titres were again observed for the T2\_17(MVA)-boosted group (Fig. 4e) at 4 weeks post boost. No antibody titres were observed in the T2\_17(MVA) primed group, suggesting that MVA is a weaker immunogen when delivered only as prime. The antibody levels were maintained up to 44 weeks post prime. As T2\_17 is an RBD-based antigen, we further explored whether higher RBD-specific antibodies were generated

on boosting with T2\_17 in comparison with boosting with AZD1222. Terminal bleed sera from 4 mice with the highest neutralizing antibody titres for the vaccine groups PBS/PBS, PBS/T2\_17(MVA), AZD1222/AZD1222 and AZD1222/T2\_17(MVA) were tested against 15-mer peptides with overlap of 14 from SARS-CoV RBD, SARS-CoV-2 RBD and T2\_17 using PEPperPRINT microarray technology. The intensities of the vaccination groups were normalized to that of the PBS/PBS mice group. The microarray data are shown in Fig. 4f. A higher number of peptide hits was observed for the T2\_17(MVA)-boosted group in comparison with the AZD1222-boosted group, suggesting that the T2\_17-boosted group induced a greater number of RBD-specific antibodies.

#### Immunogenicity of the vaccine candidate as mRNA

To further validate the immunogenicity of the T2\_17 antigen in the mRNA technology, we immunized BALB/c mice with T2\_17 as an mRNA immunogen. A previous study on Middle East Respiratory Syndrome (MERS)-based vaccine has shown that membrane-anchored, prefusion-stabilized, full-length MERS spike antigen elicited more potent pseudovirus-neutralizing antibody responses than the soluble form, as mRNA immunogen<sup>28</sup>. To test this, T2\_17 was also delivered as a membrane-anchored form (T2\_17\_TM) mRNA immunogen to mice. The mRNA immunogen (T2\_17 and T2\_17\_TM) was delivered in a prime–boost regime at 4-week intervals in BALB/c mice (Extended Data Fig. 4a) at different doses (5  $\mu$ g and 10  $\mu$ g). Full-length spike protein with a double proline mutation in the lipid formulation, similar to the one used for T2\_17 and T2\_17\_TM, was used as a control. In addition, BNT162b2 vaccine was used as a further control. All the antigen-immunized mice generated binding antibodies against SARS-CoV-2. The transmembrane-anchored T2\_17 (T2\_17\_TM) generated higher binding antibody titres at the 5- $\mu$ g dose in comparison with soluble T2\_17 (Extended Data Fig. 4b). No significant difference was observed for T2\_17\_TM at the two test doses. T2\_17 at the higher dose of 10  $\mu$ g generated binding antibody titres equivalent to those of T2\_17\_TM (Extended Data Fig. 4b). As higher antibody titres were observed for T2\_17\_TM at lower doses, we further evaluated immunogenicity of T2\_17\_TM in guinea pigs. Guinea pigs were

#### Fig. 4 | Immunogenicity and challenge studies in K18-hACE-2 mice.

**a**, Immunization, bleed and challenge schedule of K18-hACE-2 mice. K18-hACE-2 mice were primed with AZD1222 vaccine and then boosted with either AZD1222 or T2\_17 after 4 weeks. The mice were challenged after 8 weeks with either the Victoria strain of SARS-CoV-2 or the Delta variant. **b**, Neutralization of SARS-CoV, SARS-CoV-2 and the Delta variant of SARS-CoV-2 pseudoviruses by K18-hACE-2 mice sera. Sera of mice boosted with T2\_17(DNA) and T2\_17(MVA) significantly neutralized the Delta variant (B.1.617.2) in comparison with those boosted by AZD1222 at bleed 4. The x axis represents the bleed number, and the y axis represents the  $\log_{10}IC_{50}$  values for neutralization curves. **c**, Weight loss profile of K18-hACE-2 mice following challenge with the Victoria strain and the Delta variant. All mice except the naive were protected. **d**, Immunization and bleed schedule of K18-hACE-2 mice for longitudinal analysis. **e**, Neutralization of

SARS-CoV-2 pseudoviruses by K18-hACE-2 mice sera. Neutralization by sera of mice boosted with T2\_17(MVA) is statistically higher than that by sera of mice boosted with AZD1222 at bleed 2. The x axis represents the bleed number, and the y axis represents the  $\log_{10}IC_{50}$  values for neutralization curves. **f**, Peptide microarray analyses of the ACE-2 longitudinal analysis. The x axis represents the mice sera, and the y axis represents the different linear peptides. The last column represents the conservation of the corresponding peptide in SARS-CoV, SARS-CoV-2 and T2\_17. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution, the whiskers represent the minimum and maximum (excluding outliers) and the fliers represented as filled circles denote outliers. Two-tailed Mann–Whitney *U* tests demonstrated statistical significance.  $n = 12$  for **b**;  $6$  for **c** and **e**.



**Table 1 | Prime–boost regime of the K18-hACE-2-mice challenge**

Prime	Boost	Challenge
PBS	PBS	Victoria
AZD1222	AZD1222	Victoria
AZD1222	T2_17(DNA)	Victoria
AZD1222	T2_17(MVA)	Victoria
PBS	PBS	Delta
AZD1222	AZD1222	Delta
AZD1222	T2_17(DNA)	Delta
AZD1222	T2_17(MVA)	Delta

**Table 2 | Prime–boost regime of the longitudinal serology study in K18-hACE-2 mice**

Prime	Boost
PBS	PBS
PBS	T2_17(MVA)
AZD1222	PBS
AZD1222	AZD1222
AZD1222	T2_17(DNA)
AZD1222	T2_17(MVA)

immunized either with T2\_17\_TM at a dose of 3.15 µg or full-length spike with double proline mutations at a dose of 15 µg at 3-week intervals (Fig. 5a). Three weeks post prime, T2\_17\_TM induced binding antibodies against SARS-CoV as well as SARS-CoV-2, while the full-length spike antigen did not induce binding antibodies against SARS-CoV but induced binding antibodies against SARS-CoV-2 (Extended Data Fig. 5). T2\_17\_TM induced significantly higher binding antibody titres against SARS-CoV-2 in comparison with full-length spike 3 weeks post boost (Extended Data Fig. 5). Three weeks post boost, higher neutralizing antibody titres against SARS-CoV were observed for T2\_17\_TM. Three of the guinea pigs immunized with full-length spike induced neutralizing titres against SARS-CoV 3 weeks post boost. Neutralizing titres declined to low levels afterwards in this group but remained high in animals vaccinated with T2\_17\_TM, although neutralizing antibody titres against SARS-CoV-2 were lower after vaccination with T2\_17\_TM in comparison with full-length spike. The breadth of antibody immune responses elicited by the T2\_17\_TM antigen is demonstrated by significantly higher neutralizing titres than those induced by full-length spike against WIV16, SARS-CoV and the SARS-CoV-2 Omicron variant (BA.1) pseudoviruses 6 weeks post boost (Fig. 5c). The neutralizing titres induced by the full-length spike were almost negligible for the SARS-CoV-2 Omicron (BA.1) variant. We further tested for neutralization titre against one of the recent prominent variants, XBB.1.5. No neutralizing titres were observed for both T2\_17\_TM at 3.15-µg dose and full-length spike vaccine at 15-µg dose. To check whether the low neutralizing titre of T2\_17\_TM was due to the low dose of 3.15 µg, we tested the sera of guinea pigs immunized with 15 µg of T2\_17\_TM and compared them with sera of guinea pigs immunized with 15 µg of full-length spike. The sera from guinea pigs immunized with 15 µg of T2\_17\_TM neutralized pseudoviruses expressing the XBB.1.5 variant (Fig. 5d).

## Discussion

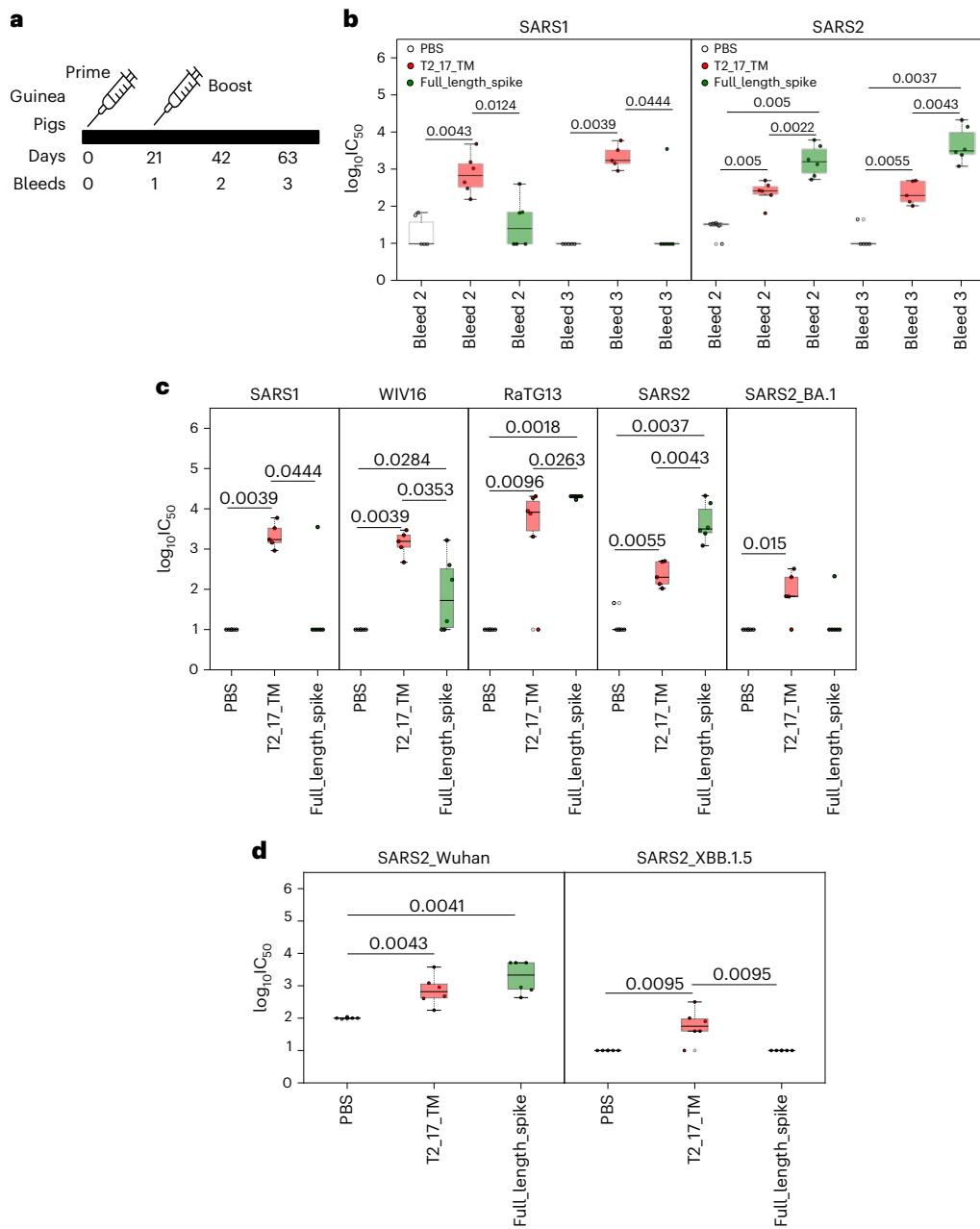
Two human SARS epidemics were caused in the past two decades by zoonotic sarbecoviruses that infect through the ACE-2 receptor.

Vaccines that can provide broad protection against such lineages are urgently needed. Strategies to achieve pan-sarbecovirus and pan-betacoronavirus protection include synthetic mRNAs expressing chimaeric versions of different spike proteins<sup>29</sup>, and mosaic or cocktail nanoparticles expressing RBDs derived from different coronaviruses<sup>30</sup>. These strategies have been reported to be effective in generating immune responses to pan-sarbecoviruses and pan-betacoronaviruses but require the synthesis, manufacture and formulation of multiple gene constructs, which adds regulatory and technical complexity to large-scale manufacturing. In addition to the zoonotic spillover from the related bat or other mammal sarbecoviruses, another cause of concern is the rapid accumulation of immune-escape mutations in circulating SARS-CoV-2. Since late 2020, many mutations leading to immune escape or to increased transmissibility (or to both) have been reported, with the latest circulating Omicron lineage reporting the greatest number of mutations in the SARS-CoV-2 viral genome<sup>14</sup>. An ideal vaccine candidate targeting circulating and emerging VOCs would be a single antigen providing protection against the diverse group of sarbecoviruses.

In this study we report preclinical data for a single-antigen RBD-subunit-based vaccine that induces immune responses against SARS-CoV, SARS-CoV-2, RaTG13, WIV16 and VOCs. The core backbone of the antigen was designed using DIOSynVax technology. It integrates phylogenetic relationships between the input sequences and structural bioinformatics to generate a core antigen sequence that ideally should generate immune responses against a diverse group of phylogenetically related viruses. We further modified the core antigen sequence by mutating some of the known epitopes within the RBD or by introducing a glycosylation site (or both) to enhance the immunogenicity of the antigen. This resulted in a panel of antigens, referred to as T2\_13 to T2\_17. The designs were generated in June 2020 and the epitope information was derived from the structural data available in June 2020. Epitope regions of CR3022 (ref. 16), S309 (ref. 15) and B38 (ref. 11) were considered for these designs. Antibodies binding to the RBD were later classified into four classes (classes 1–4)<sup>31</sup>. Out of the four classes, three classes (namely, class 1, class 2 and class 3) are potent neutralizers. The antibody B38 belongs to class 1, S309 to class 3 and CR3022 to class 4. By introducing a glycosylation site at an epitope of B38, we masked the epitope region corresponding to class-1 antibodies, while the antibodies targeting class 2 and class 3 are still available in the constructs and would induce neutralizing antibodies. The introduction of glycosylation was aimed to mask one of the divergent epitopes and to re-rank the order of the epitope presentation to the immune system, thus generating a pool of polyclonal sera enriched in both neutralizing and variant-resilient antibodies.

The strength and breadth of immune responses against different spike proteins were confirmed in BALB/c mice. From the binding profile of the sera of mice immunized with the designed antigens, we selected T2\_17 for further preclinical studies. Mice immunized with T2\_17 as a DNA immunogen produced significant binding antibody titres against both SARS-CoV and SARS-CoV-2. Neutralizing antibodies were also measured in outbred guinea pigs and outbred rabbits against both SARS-CoV and SARS-CoV-2. Notably, rabbit sera neutralized a wide panel of SARS-CoV-2 VOCs, namely Alpha, Beta, Gamma, Delta and Omicron BA.1. These broad humoral responses validate the use of the DIOSynVax technology for the generation of a pan-sarbecovirus RBD-based antigen. The breadth of the antibody immune responses elicited by the antigen T2\_17 against VOCs up to BA.1 is particularly encouraging, as the antigen was designed using the hCoV-19/Wuhan/IVDC-HB-01/2019 strain of SARS-CoV-2, suggesting the applicability of the technology to capture some of the future variants, if not all of them.

To assess the suitability of a T2\_17 antigen as a booster within the background of the non-naive population, K18-hACE-2 mice were primed with the AZD1222 vaccine and boosted with AZD1222 or T2\_17



**Fig. 5 | Immunogenicity of mRNA in guinea pigs.** **a**, Immunization and bleed schedule of guinea pigs. The guinea pigs were immunized with mRNA at 3-week intervals. **b**, Neutralization of SARS-CoV and SARS-CoV-2 pseudoviruses by guinea pig sera. The x axis represents the bleed number, and the y axis represents the  $\log_{10}IC_{50}$  values for neutralization curves. **c**, Broad neutralization of SARS-CoV, WIV16, RaTG13, SARS-CoV-2 and SARS-CoV-2 Omicron by T2\_17 at a dose of 3.15  $\mu$ g. Sera taken 6 weeks post boost (bleed 3) were used for comparison.

**d**, Neutralization of SARS-CoV-2 and SARS-CoV-2 XBB.1.5 by T2\_17 at a dose of 15  $\mu$ g. Sera taken 6 weeks post boost (bleed 3) were used for comparison. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution, the whiskers represent the minimum and maximum (excluding outliers) and the fliers represented as filled circles denote outliers. Two-tailed Mann-Whitney *U* tests demonstrated statistical significance. *n* = 6 for **b**, **c** and **d**.

as a DNA or MVA-vectored immunogen at 4-week intervals and challenged with either the Victoria or Delta strains of SARS-CoV-2. All the antigen-immunized mice were protected against the challenge, with increases in neutralizing antibody titres (4 weeks after boost) against Delta in the T2\_17-boosted group. Neutralizing antibodies against SARS-CoV were observed in the T2\_17(MVA) group. T2\_17(DNA) did not induce neutralizing antibodies against SARS-CoV. We believe that this is due to the difference in vaccine vector between AZD1222 and DNA, but this has not been addressed in the present study. Responses against SARS1 were substantially better in the group that was boosted with T2\_17(MVA). A longitudinal serology study was carried out to

understand the influence of boosting K18-hACE-2 mice at 20 weeks post prime. The antibody titres remained high for 12 weeks post prime. Four weeks after boosting, the T2\_17(MVA)-boosted group showed higher neutralizing antibody titres than other groups. Equivalent neutralizing antibody titres were observed for all the boosted groups at terminal bleed. Peptide microarrays were used on the terminal sera of the K18-hACE-2 mice from the longitudinal study to check for differential immune responses in different boost groups. A higher number of peptide hits for the RBD regions were observed for the T2\_17(MVA)-boosted group, suggesting a greater elicitation of the RBD-specific humoral response in the regime of adenovirus prime and MVA boost.

Considering the important successes of mRNA vaccines in the current COVID-19 pandemic<sup>32</sup>, T2\_17 was tested as an mRNA immunogen in mice and guinea pigs using chemically modified mRNA<sup>17</sup> in an LNP formulation<sup>18</sup>. A previous report on an mRNA vaccine has shown that a membrane-anchored and prefusion-stabilized full-length MERS spike antigen elicited more potent pseudovirus-neutralizing antibody responses than the soluble form<sup>28</sup>. In our study, in addition, T2\_17 was also delivered as a transmembrane-anchored form (T2\_17\_TM) mRNA immunogen to mice at different doses of 5 µg and 10 µg. T2\_17\_TM showed significantly higher binding antibody titres than T2\_17 at the lower dose of 5 µg, but both had comparable binding antibody titres at the 10 µg dose. On the basis of these observations, T2\_17\_TM was further validated as an mRNA immunogen in guinea pigs. Both binding and neutralizing antibodies were induced by T2\_17\_TM. Six weeks post boost, only guinea pigs immunized with T2\_17\_TM showed neutralizing antibodies against SARS-CoV and the SARS-CoV-2 Omicron (BA.1) variant at the lower dose of 3.15 µg and against XBB.1.5 at the higher dose of 15 µg. The group immunized with the full-length SARS-CoV-2 spike did not show a robust neutralizing immune response against SARS-CoV or the SARS-CoV-2 Omicron variants at the 15-µg dose. Higher antibody titres against SARS-CoV-2 and RaTG13 were observed for the full-length spike, but it must be noted that the full-length homotrimer spike presents three RBD subunits that are homologous to the SARS-CoV-2 spike tested here and would always induce higher titres in comparison with any other heterologous antigens. Moreover, the high similarity between the S2 regions of SARS-CoV-2 and RaTG13 would also induce higher cross-neutralizing antibodies between RaTG13 and SARS-CoV-2 and would lead to higher antibody titres for the full-length spike in comparison with T2\_17.

At the time of the design of T2\_17, none of the SARS-CoV-2 variants had yet been observed. Although T2\_17 still generates neutralizing antibodies against VOCs including the recent XBB.1.5, the titres are lower than observed against the Wuhan strain. With the extraordinary variation owing to the global distribution of SARS-CoV-2 in animals and humans, future updates may be needed for T2\_17, such as including the sequence information of the VOCs as well as combining it with other conserved structural and non-structural antigens. Furthermore, to understand the immunogenicity of T2\_17 in the background of the current complex immunity observed in the human population, phase-1 clinical trials have now been initiated.

All the studies combined indicate that T2\_17 is an efficacious single antigen for targeting multiple sarbecoviruses and support its applicability across different vaccine technologies. Immunization with T2\_17 generated a robust humoral immune response against SARS-CoV, SARS-CoV-2, RaTG13, WIV16 and the SARS-CoV-2 variants Alpha, Beta, Gamma, Delta and Omicron (BA.1, XBB.1.5). That the design predated the emergence of these VOCs and none of the sequences were included in the initial design is strong validation of the efficacy of the DIOSynVax technology. Given the continuous emergence of new variants, new vaccine antigens should be substantially different from the Wuhan strain or from other variants to surpass the boosting of the immunodominant epitopes conserved in these strains<sup>33,34</sup>. All the current vaccines use the full-length spike as the antigen and only 16% of the antibodies generated against the spike antigen are RBD-directed<sup>35</sup>. As T2\_17 is an RBD-based antigen substantially differing from the Wuhan-Hu-1 strain of SARS-CoV-2, it could be used as a booster vaccine candidate for overcoming immune imprinting by vaccines that use the full-length spike.

## Methods

### Phylogenetic analysis

Protein sequences of spike proteins were downloaded from the NCBI virus database for all the known sarbecoviruses (June 2020). A multiple sequence alignment was generated using MUSCLE<sup>36</sup>. The resulting multiple sequence alignment was pruned to the RBD region, filtered at 95% sequence identity and used as input for phylogenetic tree

reconstruction. The phylogenetic tree was generated with IQ-TREE<sup>37</sup> using the protein model with the best Bayesian information criterion score. The resultant tree was used for generating the phylogenetically optimized design using HyPhy<sup>38</sup>.

### Epitope identification

Available structural data (June 2020) for spike protein-antibody complexes for SARS-CoV and SARS-CoV-2 were downloaded from the Protein Data Bank (PDB)<sup>39</sup>. Structural data were then pruned for antigen-antibody complexes where the epitopes were on the RBD. Amino acid residues of antigen that have at least one atom within 5 Å radii of at least one atom of an amino acid of the antibody were defined as epitope residues, with epitope regions defined as contiguous stretches of at least 5 amino acids.

### Glycosylation site modification

The position of the glycosylation site was determined by in silico mutation of triplets of amino acids in the epitopes to the glycosylation sequon, N-X-T/S<sup>40</sup>, using the FoldX algorithm<sup>24</sup>. Briefly, residues succeeding the N-X motif, where X can be any amino acid except proline, were mutated to either threonine or serine, or residues preceding X-T/S, where X can be any amino acid except proline, were mutated to asparagine to generate N-X-T/S motifs. The mutations with the least energy cost as calculated by the Build module of FoldX<sup>24</sup> were selected.

### Molecular modelling

Structural models were generated for T2\_13 using MODELLER<sup>41,42</sup>, with both SARS-CoV and SARS-CoV-2 structures as templates. The structural model with the highest discrete optimized potential energy score<sup>43</sup> was chosen as the working model for further molecular modelling. The side chains for the model were further optimized using SCWRL<sup>44</sup> and energy minimized using GROMACS<sup>45</sup>. For T2\_14 to T2\_18, mutations were introduced using T2\_13 as the reference structure, utilizing the BUILD module of the FoldX algorithm<sup>24</sup>, and checked for structural stability using the FoldX forcefield<sup>24</sup>.

### Production and transformation of plasmids

Sequences of antigens were codon optimized for expression in humans via the GeneOptimizer algorithm<sup>46</sup>. These genes were cloned into pEVAC plasmid (GeneArt) via restriction digestion. Plasmids were transformed via heat shock in chemically induced competent *E. coli* DH5α cells (Invitrogen, 18265-017). Plasmid DNA was extracted from transformed bacterial cultures via the Plasmid Mini Kit (Qiagen, 12125). All plasmids were subsequently quantified using UV spectrophotometry (NanoDrop, Thermo Scientific)<sup>46</sup>.

### Vaccination experiments in mice

Ten groups of six 8–10-week-old female BALB/c mice were purchased from Charles River Laboratories. Mice were immunized a total of 4 times at 30-d intervals. A total volume of 50 µl of PBS containing 50 µg of plasmid DNA was administered via subcutaneous route in the rear flank. Blood was sampled from the saphenous vein at 15-d intervals and animals were terminally bled by cardiac puncture under non-recovery anaesthesia at day 150.

### Fluorescence-activated cell sorting (FACS) assay

HEK293T cells were transfected with an expression plasmid expressing wild-type spike glycoprotein of each of the four ACE-2 binding sarbecoviruses including SARS-CoV (SARS-Tor2), SARS-CoV-2 (hCoV-19/Wuhan/IVDC-HB-01/2019), WIV16 (accession ID: [ALK02457](#)) and RaTG13 (accession ID: [QHR63300](#)). At 48 h after transfection, cells were transferred into V-bottom 96-well plates (50,000 cells per well). Cells were incubated with sera (diluted at 1:50 in PBS) or anti-mouse IgG isotype negative control (Invitrogen, 10400C, diluted to 20 µg ml<sup>-1</sup> in PBS) for 30 min, washed with FACS buffer (PBS, 1% FBS, 0.02% Tween 20) and stained with goat anti-mouse IgG (H+L) Alexa Fluor 647 secondary

antibody (Invitrogen, A32728, diluted at  $20 \mu\text{g ml}^{-1}$  in FACS buffer) for 30 min in the dark. Cells were washed with FACS buffer and samples were run on an Attune NxT flow cytometer (Invitrogen) with a high-throughput auto sampler. Dead cells were excluded from the analysis by staining cells with 7-aminoactinomycin D (7-AAD) and gating 7-AAD-negative live cells (Supplementary Fig. 3).

### ELISA

The assays were adapted from those originally described in ref. 47. Briefly, Nunc MaxiSorp flat-bottom plates were coated with  $50 \mu\text{l}$  per well of  $1 \mu\text{g ml}^{-1}$  RBD from SARS-CoV (SARS1) or SARS-CoV-2 (SARS2) in DPBS ( $-\text{Ca}^{2+}/-\text{Mg}^{2+}$ ) and incubated overnight at  $4^\circ\text{C}$ . The next day, the plates were blocked with 3% milk in PBST (0.1% w/v Tween 20 in PBS) for 1 h. After removing the blocking buffer,  $50 \mu\text{l}$  per well of serum samples diluted in PBST-NFM (1% w/w non-fat milk in PBST) was added to the plates and incubated on a plate shaker for 2 h at  $20^\circ\text{C}$ . The plates were washed 3 times with  $200 \mu\text{l}$  of PBST, and then  $50 \mu\text{l}$  of HRP-conjugated goat anti-mouse IgG (H and L chains) (Jackson ImmunoResearch) was added to each well and left to incubate on a plate shaker for 1 h. Plates were washed 3 times with  $200 \mu\text{l}$  of PBST before  $50 \mu\text{l}$  per well of 1-Step Ultra TMB chromogenic substrate (Sigma) was added to the plates and the chemical reaction was stopped 3 min later with  $50 \mu\text{l}$   $2 \text{ N H}_2\text{SO}_4$ . The optical density at a wavelength of  $450 \text{ nm}$  ( $\text{OD}_{450}$ ) was measured using a BioRad microplate reader. Values from the dilution curve were used to determine the area under the curve.

### Intradermal nucleic acid immunization in guinea pigs

Two groups of eight 7-week-old female Dunkin Hartley guinea pigs (Envigo RMS) were immunized a total of 3 times at 28-d intervals. A total volume of  $200 \mu\text{l}$  of PBS containing  $400 \mu\text{g}$  of plasmid DNA was administered using the PharmaJet Tropis intradermal device, split over each hind leg. Blood was sampled from the saphenous vein at 14-d intervals.

### Intradermal nucleic acid immunization in rabbits

Ten mature (5 male, 5 female) rabbits were immunized with a good manufacturing practice lot of pEVAC\_T2\_17 (clinical pEVAC\_PS) intradermally using PharmaJet Tropis needleless delivery to the upper left and right hind limbs ( $300 \mu\text{l}$  at  $2 \text{ mg ml}^{-1}$ ). For the control group, 10 mature (5 male, 5 female) rabbits were injected with PBS. Arterial blood was sampled at 14-d intervals.

### Production of lentiviral pseudotypes

Lentiviral pseudotypes were produced by transient transfection of HEK293T/17 cells with packaging plasmids p8.91 (refs. 48,49) and pCS-FLW<sup>50</sup> and different SARS-CoV-2 VOC spike-bearing expression plasmids using the FuGENE HD Transfection Reagent<sup>51,52</sup>. Supernatants were collected after 48 h, passed through a  $0.45 \mu\text{m}$  cellulose acetate filter and titrated on HEK293T/17 cells transiently expressing human ACE-2 and TMPRSS2. Target HEK293T/17 cells were transfected 24 h earlier with  $2 \mu\text{g}$  pCAGGS-huACE-2 and  $75 \text{ ng}$  pCAGGS-TMPRSS2 (refs. 53,54).

### Pseudotype-based microneutralization assay

Pseudotype-based microneutralization assays (pMN) were performed as described previously<sup>55</sup>. Briefly, serial dilutions of serum were incubated with SARS-CoV-2, RaTG13, SARS-CoV, WIV16 and SARS-CoV-2 variant spike bearing lentiviral pseudotypes for 1 h at  $37^\circ\text{C}$  and 5%  $\text{CO}_2$  in 96-well white cell culture plates. HEK293T/17 cells ( $1.5 \times 10^4$ ) transiently expressing human ACE-2 and TMPRSS2 were then added per well and plates incubated for 48 h at  $37^\circ\text{C}$  and 5%  $\text{CO}_2$  in a humidified incubator. Bright-Glo (Promega) was then added to each well and luminescence read after a 5-min incubation period. Experimental data points were normalized to 100% and 0% neutralization controls, and nonlinear regression analysis performed in GraphPad Prism 9 to produce neutralization curves and half-maximal inhibitory concentration ( $\text{IC}_{50}$ ) values.

### ACE-2 competition assay

The SARS-CoV-2 surrogate virus neutralization test (SVNT, Genscript) was carried out following manufacturer instructions. Briefly, sera were diluted in PBS across an 8 point 1:2 dilution series from a starting concentration of 1:50. Samples were further diluted in the provided sample buffer at a 1:9 ratio and then mixed with HRP conjugated to SARS-CoV-2 RBD protein, incubated at  $37^\circ\text{C}$  for 30 min and added to human ACE-2 protein coated wells in 96-well plate format. The reaction was incubated at  $37^\circ\text{C}$  for 15 min and then washed 4 times with the provided wash buffer. TMB solution was then added and the plates were incubated for 15 min in the dark at room temperature to allow the reaction to develop. The reaction was then quenched using the provided stop solution and absorbance was read at  $450 \text{ nm}$ .

### MVA production

The MVA strain used in this study was MVA-CR19. Recombinant MVA that expresses SARS-CoV-2 RBD T2-17 was generated as described previously<sup>56</sup>. In brief, for in vivo recombination, adherent AGE1.CR.pIX were infected with parental MVA-CR19 TK-GFP at different multiplicity of infection (MOIs) ranging from 0.5 to 0.006 particle forming unit. After 2 h, the cells were transfected with  $0.4 \mu\text{g}$  of the shuttle vector pMVA\_RBD T2\_17 using Effectene (Qiagen) according to manufacturer instructions. After 48 h, the cells were collected, lysed using three freeze-thaw cycles and sonicated. Pure recombinant viruses were obtained by sequential plaque purification under agarose overlays and confirmed to be free of contaminating parental MVA-CR19 TK-GFP by PCR screening. This recombinant MVA encoding SARS-CoV-2 RBD T2-17 was plaque purified for an additional three rounds. The resulting recombinant MVA-CR19 RBD-T2\_17 (MVA T2\_17) virus stock was produced in suspension AGE1.CR.pIX cells, purified via two ultracentrifugation rounds over a 35% sucrose cushion and titrated on DF-1 cells using crystal violet staining. The sequence of recombinant MVA and the absence of revertant MVA was confirmed by PCR amplification and Sanger sequencing. The expression of RBD T2\_17 was confirmed by western blot analysis with monoclonal antibody CR3022, with cell lysates from HEK293 cells collected 24 h after infection (MOI 2) with MVA T2\_17.

### Vaccine boost efficacy studies in K18-hACE-2 mice

Eight groups of six 8–15-week-old homozygous female K18-hACE-2 mice (Jax) were primed with  $1.4 \times 10^9$  viral particles of AZD1222 or PBS by intramuscular route, in a total volume of  $100 \mu\text{l}$  split over the two rear legs. After 28 d, 2 groups of 6 mice were boosted with either PBS, AZD1222, T2\_17(DNA) or T2\_17(MVA). Mice were bled at 2-week intervals and challenged at day 84 with either Victoria/1/2020 (B-lineage) or Delta SARS-CoV-2 by intranasal route, in a total volume of  $40 \mu\text{l}$  over both nares. Mice were weighed daily and monitored for clinical signs for a period of 14 d before being humanely culled by terminal bleed.

### Longitudinal serology studies in K18-hACE-2 mice

Six groups of six 8–15-week-old homozygous female K18-hACE-2 mice (Jax) were primed with  $1.4 \times 10^9$  viral particles of AZD1222 or PBS by intramuscular route, in a total volume of  $100 \mu\text{l}$  split over the two rear legs. After 20 weeks, groups of 6 mice were boosted either with PBS, AZD1222, T2\_17(DNA) or T2\_17(MVA). Mice were bled at 12 and 24 weeks post prime, and terminally bled at week 44 post prime.

### Peptide microarray

Four samples were selected from the terminal bleeds of the K18-hACE-2 mice used for the longitudinal study (ACE-2/AZD1222 vaccine prime, T2\_17 boost samples). The samples were selected on the basis of the quality of serum and performance in the pMN assay against SARS-CoV-2 Wuhan. Aliquots ( $30 \mu\text{l}$ ) of each sample were sent to PEPperPRINT for peptide microarray analysis. Briefly, 15-mer peptides spanning the SARS-CoV RBD (213 amino acid (aa)), SARS-CoV-2 RBD (214 aa) and

T2\_17 RBD (214 aa) with a 14-aa overlap were printed in duplicate per array copy for a total of five array copies. HA and c-Myc control peptides were included in each array copy. The protein sequences were elongated by neutral GSGSGG linkers to avoid truncated peptides, and identical peptides were removed. In total, 1,310 peptide sequences were synthesized and spotted in duplicate onto the PEPperCHIP microarray platform. The corrected raw intensities were log transformed for all the serum samples. For each vaccine group, namely PBS/T2\_17MVA, AZD1222/AZD1222 and AZD1222/T2\_17MVA, the peptides with raw intensities twofold higher than the maximum intensity observed in the PBS/PBS group were considered as antibody epitope hits.

### mRNA vaccine production

mRNA sequences encoding the SARS-CoV-2 spike protein with 2 proline mutations, T2\_17 and T2\_17\_TM, were synthesized by in vitro transcription from linearized plasmid DNA templates using modified nucleotides to generate partial modified mRNAs. After in vitro transcription, mRNAs were dephosphorylated and enzymatically polyadenylated. Purification steps were performed by precipitation and mRNA subsequently formulated in water for injection at a concentration of 1 mg ml<sup>-1</sup>. mRNAs were stored at -80 °C until LNP encapsulation. Each mRNA was LNP encapsulated via nanoprecipitation by microfluidic mixing of mRNA in citrate buffer (pH 4.5) with ionizable, structural, helper and polyethylene glycol lipids in ethanol, followed by buffer exchange and concentration via tangential flow filtration. mRNA LNPs were filtered through a 0.2-µm membrane and stored at -20 °C until use. The drug product was analytically characterized and evaluated as acceptable for in vivo use.

### Immunization of BALB/c mice with mRNA

Seven groups of six 8–10-week-old female BALB/c mice were purchased from Charles River Laboratories. Mice were immunized twice at a 21-d interval. A total volume of 100 µl of vehicle containing various amounts of mRNA was administered via intramuscular route in each hind leg. Blood was sampled from the saphenous vein at 21-d intervals and animals were terminally bled by cardiac puncture under non-recovery anaesthesia at day 63.

### Immunization of guinea pigs with mRNA

Three groups of six 8–10-week-old female Hartley guinea pigs were purchased from Envigo. Guinea pigs were immunized twice at a 21-d interval. A total volume of 200 µl of vehicle containing various amounts of mRNA (3.15 µg or 15 µg) was administered via intramuscular route in each hind leg. Blood was sampled from the saphenous vein at 21-d intervals and animals were humanely euthanized after the final bleed on day 63.

### Statistical analyses

Two-tailed Mann–Whitney *U* tests were performed for all the comparisons using the Python sklearn package<sup>57</sup>. All plots were generated using the Python Matplotlib package and statannotat package<sup>58</sup>.

### Animal-work ethics

Animal studies were approved by the Animal Welfare and Ethical Review Body, University of Cambridge, and experiments were carried out under an approved UK home office licence (P8143424B).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The main data supporting the results in this study are available within the paper and its Supplementary Information. The sequences used for designing the vaccine antigens were retrieved from the

publicly available NCBI virus database. The structure coordinates of the antigen–antibody complexes used for the analyses are available in the Protein Data Bank. The sequences of the antigens have been patented under UK Patent Application No. 2303150.3, Coronavirus Vaccines.

### References

- Liu, K. et al. Cross-species recognition of SARS-CoV-2 to bat ACE2. *Proc. Natl Acad. Sci. USA* **118**, e2020216118 (2021).
- Olival, K. J. et al. Possibility for reverse zoonotic transmission of SARS-CoV-2 to free-ranging wildlife: a case study of bats. *PLoS Pathog.* **16**, e1008758 (2020).
- Hu, B. et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.* **13**, e1006698 (2017).
- Menachery, V. D. et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat. Med.* **21**, 1508–1513 (2015).
- Vilar, S. & Isom, D. G. One year of SARS-CoV-2: how much has the virus changed? *Biology* **10**, 91 (2021).
- Wong, T. Y. et al. Evaluating antibody mediated protection against Alpha, Beta, and Delta SARS-CoV-2 variants of concern in K18-hACE2 transgenic mice. *J. Virol.* **96**, e0218421 (2022).
- Planas, D. et al. Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nat. Med.* **27**, 917–924 (2021).
- Wang, P. et al. Increased resistance of SARS-CoV-2 variant P.1 to antibody neutralization. *Cell Host Microbe* **29**, 747–751.e4 (2021).
- Leung, K., Shum, M. H., Leung, G. M., Lam, T. T. & Wu, J. T. Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020. *Euro Surveill.* **26**, 2002106 (2021).
- Tian, F. et al. N501Y mutation of spike protein in SARS-CoV-2 strengthens its binding to receptor ACE2. *eLife* **10**, e69091 (2021).
- Wu, Y. et al. A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science* **368**, 1274–1278 (2020).
- Hwang, W. C. et al. Structural basis of neutralization by a human anti-severe acute respiratory syndrome spike protein antibody, 80R. *J. Biol. Chem.* **281**, 34610–34616 (2006).
- Campbell, F. et al. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Euro Surveill.* **26**, 2100509 (2021).
- Willett, B. J. et al. SARS-CoV-2 Omicron is an immune escape variant with an altered cell entry pathway. *Nat. Microbiol.* **7**, 1161–1179 (2022).
- Pinto, D. et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290–295 (2020).
- Yuan, M. et al. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **368**, 630–633 (2020).
- Kormann, M. S. D. et al. Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nat. Biotechnol.* **29**, 154–157 (2011).
- Jarzębińska, A. et al. A single methylene group in oligoalkylamine-based cationic polymers and lipids promotes enhanced mRNA delivery. *Angew. Chem. Int. Ed.* **55**, 9591–9595 (2016).
- Hatcher, E. L. et al. Virus Variation Resource—improved response to emergent viral outbreaks. *Nucleic Acids Res.* **45**, D482–D490 (2017).
- Li, W. et al. Animal origins of the severe acute respiratory syndrome coronavirus: insight from ACE2-S-protein interactions. *J. Virol.* **80**, 4211–4219 (2006).

21. Lavie, M., Hanoulle, X. & Dubuisson, J. Glycan shielding and modulation of hepatitis C virus neutralizing antibodies. *Front. Immunol.* **9**, 910 (2018).
22. Watanabe, Y. et al. Structure of the Lassa virus glycan shield provides a model for immunological resistance. *Proc. Natl Acad. Sci. USA* **115**, 7320–7325 (2018).
23. Eggink, D., Goff, P. H. & Palese, P. Guiding the immune response against influenza virus hemagglutinin toward the conserved stalk domain by hyperglycosylation of the globular head domain. *J. Virol.* **88**, 699–704 (2014).
24. Schymkowitz, J. et al. The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, W382–W388 (2005).
25. Carnell, G. W. et al. Glycan masking of a non-neutralising epitope enhances neutralising antibodies targeting the RBD of SARS-CoV-2 and its variants. *Front. Immunol.* **14**, 1118523 (2023).
26. Ewer, K. et al. A monovalent chimpanzee adenovirus Ebola vaccine boosted with MVA. *N. Engl. J. Med.* **374**, 1635–1646 (2016).
27. Zhou, Y. & Sullivan, N. J. Immunology and evolution of the adenovirus prime, MVA boost Ebola virus vaccine. *Curr. Opin. Immunol.* **35**, 131–136 (2015).
28. Corbett, K. S. et al. SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature* **586**, 567–571 (2020).
29. Martinez, D. R. et al. Chimeric spike mRNA vaccines protect against Sarbecovirus challenge in mice. *Science* **373**, 991–998 (2021).
30. Walls, A. C. et al. Elicitation of broadly protective sarbecovirus immunity by receptor-binding domain nanoparticle vaccines. *Cell* <https://doi.org/10.1016/j.cell.2021.09.015> (2021).
31. Barnes, C. O. et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).
32. Fang, E. et al. Advances in COVID-19 mRNA vaccine development. *Sig. Transduct. Target. Ther.* **7**, 94 (2022).
33. Aguilar-Bretones, M., Fouchier, R. A., Koopmans, M. P. & van Nierop, G. P. Impact of antigenic evolution and original antigenic sin on SARS-CoV-2 immunity. *J. Clin. Invest.* **133**, e162192 (2023).
34. Ju, B. et al. Antigenic sin of wild-type SARS-CoV-2 vaccine shapes poor cross-neutralization of BA.4/5/2.75 subvariants in BA.2 breakthrough infections. *Nat. Commun.* **13**, 7120 (2022).
35. Voss, W. N. et al. Prevalent, protective, and convergent IgG recognition of SARS-CoV-2 non-RBD spike epitopes. *Science* **372**, 1108–1112 (2021).
36. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).
37. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
38. Pond, S. L. K., Frost, S. D. W. & Muse, S. V. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679 (2005).
39. Berman, H. M. et al. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
40. Kornfeld, R. & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **54**, 631–664 (1985).
41. Eswar, N. et al. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* <https://doi.org/10.1002/0471140864.ps0209s50> (2007).
42. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
43. Shen, M. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524 (2006).
44. Krivov, G. G., Shapovalov, M. V. & Dunbrack, R. L. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **77**, 778–795 (2009).
45. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
46. Raab, D., Graf, M., Notka, F., Schödl, T. & Wagner, R. The GeneOptimizer Algorithm: using a sliding window approach to cope with the vast sequence space in multiparameter DNA sequence optimization. *Syst. Synth. Biol.* **4**, 215–225 (2010).
47. Amanat, F. et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**, 1033–1036 (2020).
48. Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L. & Trono, D. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat. Biotechnol.* **15**, 871–875 (1997).
49. Naldini, L. et al. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263–267 (1996).
50. Demaison, C. et al. High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of immunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. *Hum. Gene Ther.* **13**, 803–813 (2002).
51. Sampson, A. T. et al. Coronavirus pseudotypes for all circulating human coronaviruses for quantification of cross-neutralizing antibody responses. *Viruses* **13**, 1579 (2021).
52. Di Genova, C. et al. Production, titration, neutralisation, storage and lyophilisation of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) lentiviral pseudotypes. *Bio Protoc.* **11**, e4236 (2021).
53. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
54. Bertram, S. et al. Influenza and SARS-coronavirus activating proteases TMPRSS2 and HAT are expressed at multiple sites in human respiratory and gastrointestinal tracts. *PLoS ONE* **7**, e35876 (2012).
55. Carnell, G., Grehan, K., Ferrara, F., Molesti, E. & Temperton, N. J. An optimized method for the production using PEI, titration and neutralization of SARS-CoV spike luciferase pseudotypes. *Bio Protoc.* **7**, e2514 (2017).
56. Jordan, I. et al. A deleted deletion site in a new vector strain and exceptional genomic stability of plaque-purified modified vaccinia Ankara (MVA). *Virology* **35**, 212–226 (2020).
57. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
58. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
59. Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).
60. Schrödinger, L. & DeLano, W. The PyMOL Molecular Graphics System, version 2.5 (2021).

## Acknowledgements

J.L.H., R.K., G.W.C. and S.V. were funded to perform the DNA/MVA vaccine studies by Innovate UK DIOS-CoVax, award number 72845 (DIOSynVax). B.A., M.B., P.N., D.P. and S.E. were funded by COVID (Bavarian Ministry of Science and Arts) and EU H2020 programme Grant 899619 to R.W. The mRNA preparation at Ethis and the related preclinical studies carried out by DIOSynVax were supported by CEPI.

## Author contributions

S.V., D.A.W. and S.F. developed the computational pipeline. G.W.C., M.F., C.G., M.S.S., A.N., C.Q.H., M.P., H.S. and A.C. performed the in vitro assays. G.W.C. and P.T. performed the animal experiments. B.A., M.B., P.N., D.P. and S.E. performed DNA purification and preparation. M.B., I.J. and V.S. performed MVA production and purification. K.S., M.T.L.B., F.A., J.G., C.D. and C.P. performed mRNA production and

purification. D.C., M.M.N. and N.T. provided key reagents. S.V. and G.W.C. analysed and visualized the data. R.K., R.W. and J.L.H. acquired funding, conceptualized the investigation, reviewed data and administered the project. R.W. and J.L.H. supervised the project. S.V. wrote the original draft. S.V., G.W.C. and J.L.H. reviewed and edited the manuscript. All authors provided feedback on the paper.

### Competing interests

M.F., D.A.W., S.F., R.K., R.W. and J.L.H. are employees or shareholders of DIOSynVax Ltd. S.F. is an employee of Microsoft. I.J. and V.S. are employees of ProBioGenAG. K.S., M.T.L.B., F.A., J.G., C.D. and C.P. are employees of Ethris GmbH. The sequences of the antigens have been patented under UK Patent Application No. 2303150.3, Coronavirus Vaccines. All other authors have no competing interests.

### Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41551-023-01094-2>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41551-023-01094-2>.

**Correspondence and requests for materials** should be addressed to Jonathan Luke Heeney.

**Peer review information** *Nature Biomedical Engineering* thanks Cheng-Feng Qin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

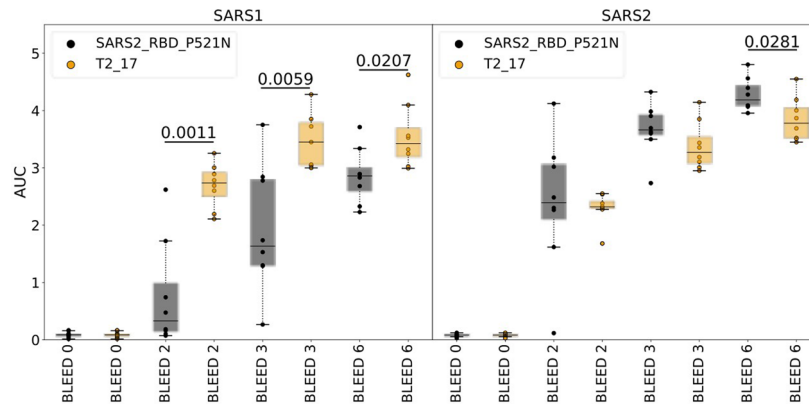
**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

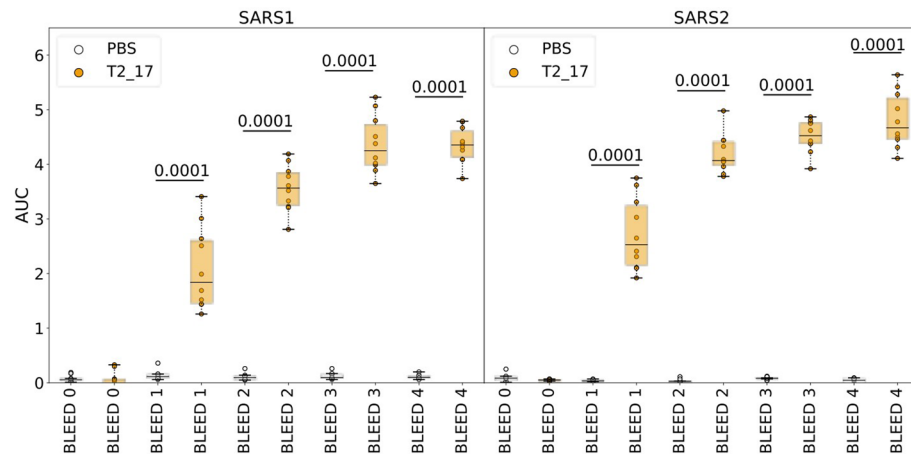
<sup>1</sup>Lab of Viral Zoonotics, Department of Veterinary Medicine, University of Cambridge, Cambridge, UK. <sup>2</sup>DIOSynVax Ltd, University of Cambridge, Cambridge, UK. <sup>3</sup>Institute of Medical Microbiology and Hygiene, University of Regensburg, Regensburg, Germany. <sup>4</sup>Department of Pathology, University of Cambridge, Cambridge, UK. <sup>5</sup>Institute of Clinical Microbiology and Hygiene, University Hospital Regensburg, Regensburg, Germany. <sup>6</sup>Viral Pseudotype Unit, Medway School of Pharmacy, The Universities of Kent and Greenwich at Medway, Chatham, UK. <sup>7</sup>ProBioGenAG, Berlin, Germany. <sup>8</sup>London School of Hygiene and Tropical Medicine, London, UK. <sup>9</sup>Microsoft Health Futures, Redmond, WA, USA. <sup>10</sup>Ethris GmbH, Planegg, Germany. ✉e-mail: [jlh66@cam.ac.uk](mailto:jlh66@cam.ac.uk)



**Extended Data Fig. 1 | Binding antibody data for T2\_17 in guinea pigs.**

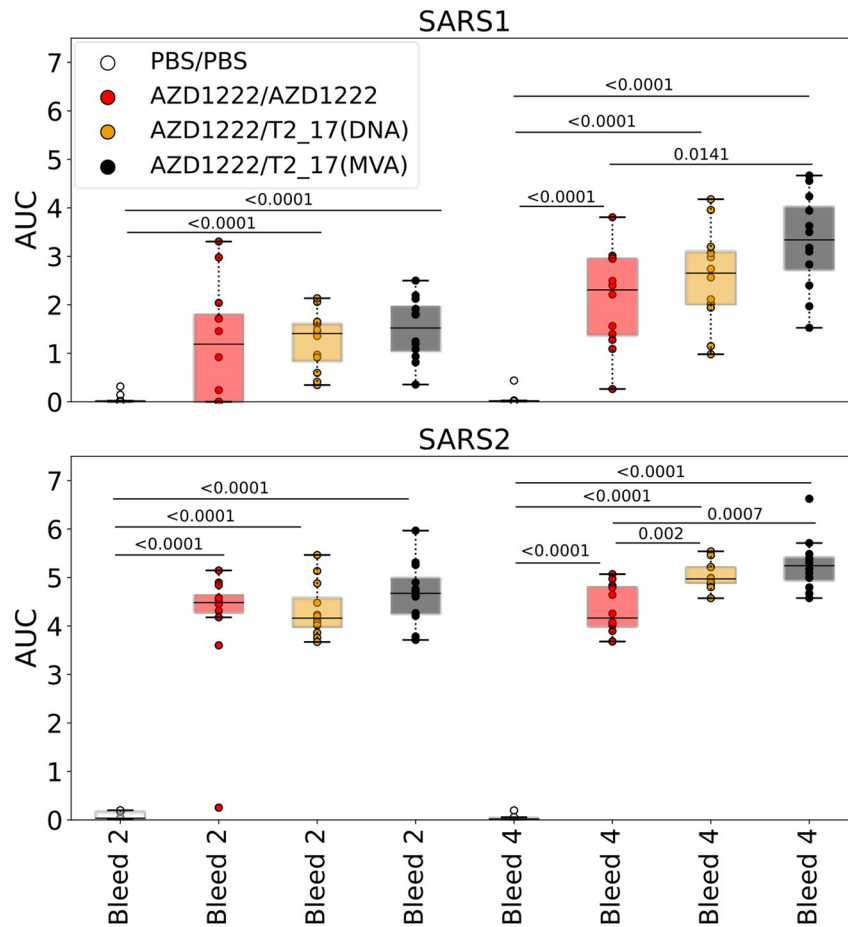
Elicitation of binding antibodies against SARS-CoV and SARS-CoV-2 by T2\_17 and SARS2\_RBD\_P521N was confirmed using ELISA. The pre-bleed (Bleed 0) is considered as the control for non-specific binding. The X-axis represents the bleed number, and the Y-axis represents the area under the curve (AUC) from

the ELISA binding curve. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution, and the whiskers represent the minimum and maximum of the distribution (excluding outliers) and the fliers represented as filled circle represent the outliers. Two-tailed Mann-Whitney U demonstrated statistical significance. Sample size is eight.



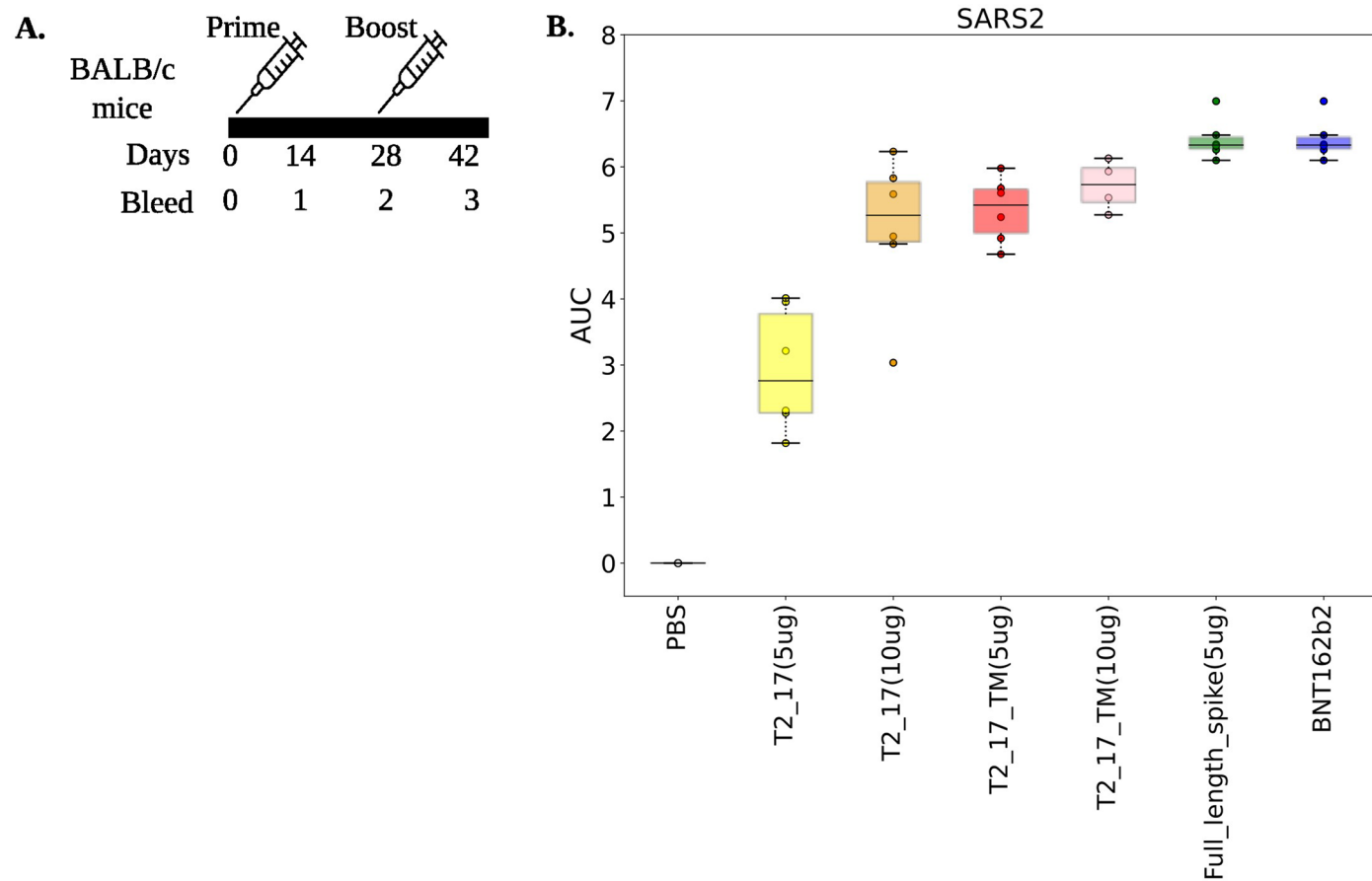
**Extended Data Fig. 2 | Binding antibody data for T2\_17 in rabbits.** Elicitation of binding antibodies against SARS-CoV and SARS-CoV-2 by T2\_17 was confirmed using ELISA. The X-axis represents the bleed number, and the Y-axis represents the area under the curve (AUC) from the ELISA binding curve. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution,

and the whiskers represent the minimum and maximum of the distribution (excluding outliers) and the fliers represented as filled circle represent the outliers. Two-tailed Mann-Whitney U demonstrated statistical significance. Sample size is ten.



**Extended Data Fig. 3 | ELISA binding data of K18-hACE2 sera.** Binding antibodies were observed 4 weeks post immunization with AZD1222 and 4 weeks post boosting. The X-axis represents the bleed number, and the Y-axis represents the area under the curve (AUC) from ELISA binding curves. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution, and the whiskers represent the minimum and maximum of the distribution

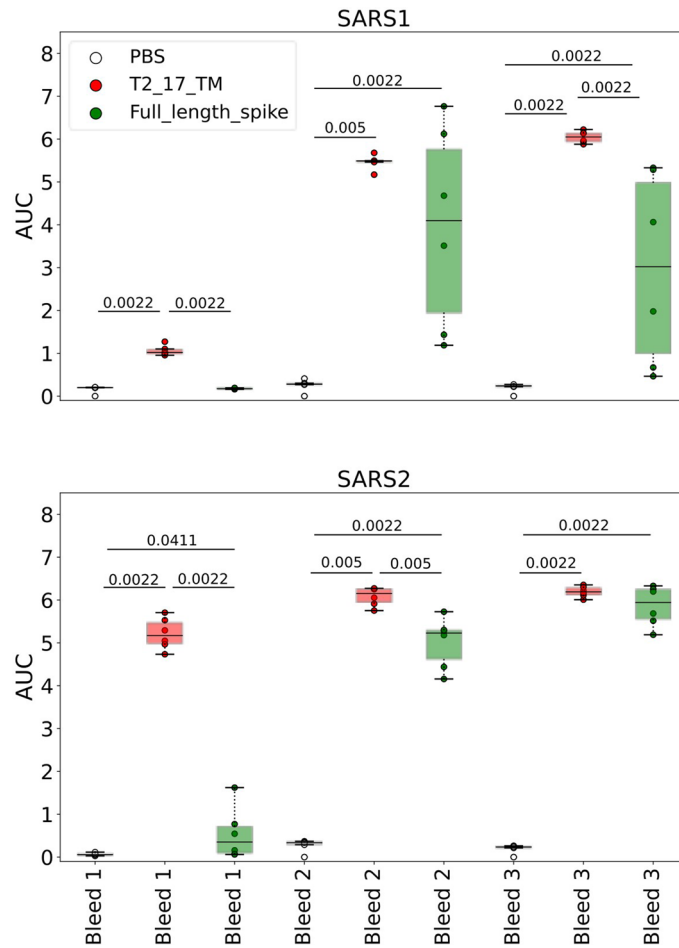
(excluding outliers) and the fliers represented as filled circle represent the outliers. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution (excluding outliers) and the fliers represented as filled circle represent the outliers. Two-tailed Mann-Whitney U demonstrated statistical significance, Sample size is twelve.



**Extended Data Fig. 4 | Immunogenicity of mRNA vaccine in BALB/c mice.**

**(a)** Immunization and bleed schedule of BALB/c mice. The mice were immunized with mRNA at 4-week intervals. **(b)** Elicitation of binding antibodies against SARS-CoV-2 was confirmed using ELISA for sera 2 weeks post boost (bleed 3). The X-axis represents the antigens, and the Y-axis represents the area under the

curve (AUC) from ELISA binding curves. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution, and the whiskers represent the minimum and maximum of the distribution (excluding outliers) and the fliers represented as filled circle represent the outliers. Sample size is six.



**Extended Data Fig. 5 | Immunogenicity of mRNA vaccine in guinea pigs.**

Elicitation of binding antibodies against SARS-CoV and SARS-CoV-2 in guinea pigs was confirmed using ELISA. The X-axis represents the bleeds, and the Y-axis represents the area under the curve (AUC) from ELISA binding curves. The boxes represent the quartiles (25th, 50th and 75th percentiles) of the distribution,

and the whiskers represent the minimum and maximum of the distribution (excluding outliers) and the fliers represented as filled circle represent the outliers. Two-tailed Mann-Whitney U demonstrated statistical significance. Sample size is six.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

Sequence data were downloaded from the NCBI virus database using the NCBI download tool.  
Structure data were downloaded from the PDB using the PDB download tool.

Data analysis

Sequence cleaning and pruning were done using custom codes written in python 3.  
Sequence alignments were generated using the MUSCLE algorithm version 3.8.  
The phylogenetic tree was built using the iQtree algorithm version 1.6.1 with the in-built statistical packages.  
Novel sequences were designed using the HyPhy package, version 2.5.  
Structure visualization and the rendering of images were done using PyMol ver 2.5.  
Stability calculations were done using the FOLDX 5 algorithm.  
Energy minimization was done using the Gromacs-2019.4 package with in-built force-fields and parameters.  
Plots were generated using custom codes written in python and using the matplotlib package.  
Statistical analyses were done using the numpy package in python 3.  
All the software was used with Academic license.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The main data supporting the results in this study are available within the paper and its Supplementary Information. The sequences used for designing the vaccine antigens were retrieved from the publicly available NCBI virus database. The structure coordinates of the antigen–antibody complexes used for the analyses are available in the Protein Data Bank. The sequences of the antigens have been patented under UK Patent Application No. 2303150.3, Coronavirus Vaccines.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	<input type="text" value="The study did not involve human participants."/>
Reporting on race, ethnicity, or other socially relevant groupings	<input type="text" value="—"/>
Population characteristics	<input type="text" value="—"/>
Recruitment	<input type="text" value="—"/>
Ethics oversight	<input type="text" value="—"/>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<input type="text" value="We used groups of 6 mice as the lowest number of mice suitable for performing the statistical analyses."/>
Data exclusions	<input type="text" value="No data were excluded from the analyses."/>
Replication	<input type="text" value="All the reported experiments were done in duplicate."/>
Randomization	<input type="text" value="The animals in all the animals studies were randomly assigned to groups."/>
Blinding	<input type="text" value="The investigators were blinded to the assigned groups."/>

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials &amp; experimental systems

n/a	Involvement
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

## Methods

n/a	Involvement
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used	Clone CR3022 (absolute antibody Ab01680-10.0 and Ab01680-3.0). IgG Isotype negative control (Invitrogen 10400C). Alexa Fluor 647 Secondary Antibody (Invitrogen A32728). HRP-conjugated goat anti Ig (H and L chains) (Jackson ImmunoResearch, 109-035-088 (anti-human), and 706-035-148 (anti-guinea pig)). HRP-conjugated goat anti Ig (H and L chain) (Thermo Fischer Scientific, 31460 (anti-rabbit)).
Validation	All antibodies used were from commercial sources and validated by them.

## Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	HEK293T/17, from ATCC. HEK293T, from ATCC. Suspension of AGE1.CR.pIX cells, from ProBioGenAG.
Authentication	None of the cell lines used were authenticated.
Mycoplasma contamination	All cells lines were tested, and confirmed negative for mycoplasma.
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No commonly misidentified cell lines were used.

## Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	BALB/c mice (BALB/cAnNCrI Charles River), 8-to-10-weeks old, female (Charles River Laboratories). Homozygous K18-hACE2 mice (B6.Cg-Tg(K18-ACE2)2PrImn/J, Jax), 8-to-15-weeks old, female (Charles River Laboratories). All mice used in this study were female, 8–15 weeks old and weighing 20–30g at the time of use. Dunkin Hartley Guinea pigs, 7-week-old female (Envigo). Rabbits, 5 female and 5 male (Charles River Laboratories).
Wild animals	The study did not involve wild animals.
Reporting on sex	Only females were used in this study, with the exception of rabbits (50% male, 50% female).
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	UK Home Office License P8143424B. Characterization of vaccine candidates against viral diseases. Each individual animal study was submitted to a portal to the local AWERB University of Cambridge.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

## Methodology

Sample preparation	Transfected HEK293T (from ATCC).
Instrument	Attune NxT Flow Cytometer
Software	Attune Cytometric Software
Cell population abundance	Not applicable for binding studies.
Gating strategy	Preliminary FSC/SSC gates were set on the starting cell population. Singlets were then gated by plotting FSC-H versus FSC-A, followed by gating live cells as 7-AAD negative. The live-cell population was visualized as a histogram in the RL-1 channel. The PMT of the negative cell population was set between $10^2$ and $10^3$ to reduce background noise. The MFI values of the negative population, which consisted of untransfected cells stained with primary and secondary AF647 antibodies, were subtracted from the MFI of the test samples to correct for non-specific binding.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.