Pre-existing antibodies targeting a dominant linear antibody epitope on SARS-CoV-2 S2 cross-reacted with commensal gut bacteria and shaped immune responses elicited by a candidate vaccine

Liqiu Jia[1, §], Shufeng Weng[2, §], Jing Wu[1, §], Xiangxiang Tian[3, 4], Yifan Zhang[3, 4], Xuyang Wang[1], Jing Wang[3, 5], Dongmei Yan[5], Wanhai Wang[4], Zhaoqin Zhu[3, #], Chao Qiu[6, #], Wenhong Zhang[1, 2, 7, 8, #], Ying Xu[2, #], Yanmin Wan[1, 2, 9, #]

[1] Department of Infectious Diseases, National Medical Center for Infectious Diseases, Shanghai Key Laboratory of Infectious Diseases and Biosafety Emergency Response, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai, China;
[2] State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Science, Fudan University, Shanghai, China;
[3] Department of laboratory medicine, Shanghai Public Health Clinical Center, Shanghai, China;
[4] Clinical Laboratory, The First Affiliated Hospital of Zhengzhou University, Key Laboratory of Laboratory Medicine of Henan Province, Zhengzhou, Henan, P.R. China;
[5] Department of Immunology, School of Basic Medical, Jiamusi University, Jiamusi, Heilongjiang Province, China;
[6] Institutes of biomedical sciences & Shanghai Key Laboratory of Medical Epigenetics, Fudan University, Shanghai, China;
[7] National Clinical Research Center for Aging and Medicine, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai, China;
[8] Key Laboratory of Medical Molecular Virology (MOE/MOH) and Institutes of Biomedical Sciences, Shanghai Medical College, Fudan University, Shanghai, China.
[9] Department of radiology, Shanghai Public Health Clinical Center, Shanghai, China
[§] These authors contribute equally to this work.
[#] Correspondence should be addressed to: Zhaoqin Zhu, zhaoqinzhu@163.com; Chao Qiu, qiuchao@fudan.edu.cn; Wenhong Zhang, zhangwenhong@fudan.edu.cn; Ying Xu, yingxu2520@fudan.edu.cn; Yanmin Wan, yanmin_wan@fudan.edu.cn

Running title: Antibody cross-reactivity between SARS-CoV-2 and gut microbiota

Word count: Abstract, 203; main text, 5746

Abstract

Pre-existing SARS-CoV-2 cross-reactive antibodies have been detected in both unexposed human and animals. However, the origins of these cross-reactive antibodies and their potential impacts on vaccine efficacy have not been completely clarified. In this study, we demonstrated that the S2 subunit was the predominant target of the pre-existing SARS-CoV-2 spike protein cross-reactive antibodies in both healthy human and naïve SPF mice. Through linear epitope mapping, we identified a dominant antibody epitope on the connector domain of S2 (aa1145-aa1162), which could be recognized by antibodies pre-existed in unexposed human and mice. Six monoclonal antibodies against this linear epitope were isolated from naïve SPF mice and were proved to cross-react with commensal gut bacteria collected from both human and mouse. Via immunizing mice with a candidate DNA vaccine encoding the full length of SARS-CoV-2 spike protein, we further demonstrated that high levels of pre-existing S2 cross-reactive antibodies did not impair the immunogenicity of the DNA vaccine. On the contrary, mice with high levels of pre-existing antibodies mounted stronger S2 specific binding antibody responses compared to mice with low levels of pre-existing antibodies. In addition, S1 specific T cell and binding antibody responses also tended to be enhanced in mice with high levels of pre-existing antibodies.

Key words: Cross-reactive antibody, SARS-CoV-2, Spike protein, commensal gut bacteria, vaccine immunogenicity

Introduction

Antibodies are vital components of the immune system that mediate protection against infections (1). When confronting infections, the actual role of pre-existing antibody depends on the following features (2): High titers of broadly neutralizing antibodies can protect the host against infection. While, when the pre-existing antibodies are non-neutralizing or with only a narrow neutralizing spectrum, hosts may not be sterilely protected or only be protected against specific serotypes of viruses. In addition to defending hosts against infections, pre-existing antibodies can also impact host immune responses upon infection or vaccination (3-5), which is best exemplified by the observations showing that pre-existing antibodies shaped the recall immune responses against influenza (6, 7).

For most occasions, pre-existing antibodies in adults derive from previous infection or vaccination except some “naturally” produced, poly-reactive antibodies (2, 8). When encountering a newly emerged or mutated virus, cross-reactive antibodies induced by previously occurred, phylogenetically closely related viruses constitute the main body of the pre-existing cross-reactive antibodies. The effect of this kind of pre-existing antibodies has been extensively investigated especially for infections of influenza (3, 7, 9) and flaviviruses (10-12). Of note, previous infection by phylogenetically similar viruses is not the sole source of pre-existing cross-reactive antibodies, as it has been clearly clarified that pre-existing antibodies against HIV-1 gp41 may stem from exposures to certain commensal gut bacteria (13-15). Besides, autoimmune diseases caused by cross-reactivities between microbial and self-antigens also implied that commensal gut bacteria represent important sources of cross-reactive antibodies (16-19).

Pre-existing antibodies against SARS-CoV-2 have also been observed in uninfected healthy individuals, which are speculated to be engendered by previous exposures to human common cold coronaviruses (20-26) or SARS-CoV (27-29). Meanwhile, sequence analyses (30) and a clinical observation (31) suggest that pre-existing SARS-CoV-2 antibodies might be engendered by common human pathogens and childhood vaccination. Although these two explanations are not mutually exclusive, they both need more experimental evidence to support.

In this study, we found that higher levels of SARS-CoV-2 S2 protein specific

antibodies existed in both healthy human and naïve SPF mice. To track the potential origins of these pre-existing cross-reactive antibodies, we mapped and located a dominant linear antibody epitope on S2, which could be recognized by pre-existing antibodies from both healthy human and naïve SPF mice. Monoclonal antibodies against this linear epitope were isolated from naïve SPF mice and proved to cross-react with commensal gut bacteria collected from both healthy human and naïve SPF mouse. Moreover, despite having been discussed iteratively (32, 33), the influences of pre-existing cross-reactive immunities on COVID-19 responses have not been clarified. Here we showed that high levels of pre-existing antibodies did not impair the immunogenicity of a candidate DNA vaccine encoding SARS-CoV-2 spike protein. On the contrary, mice with high levels of pre-existing antibodies mounted stronger S2 specific binding antibody responses compared with mice with low levels of pre-existing antibodies after immunization with a candidate DNA vaccine. Meanwhile, S1 specific T cell and binding antibody responses also tended to be stronger in mice with high levels of pre-existing antibodies, although no statistical significance was reached.

Materials and methods

Ethics statement

All experiments and methods were performed in accordance with relevant guidelines and regulations. Experiments using mice and samples of healthy human were approved by the Research Ethics Review Committee of the Shanghai Public Health Clinical Center Affiliated to Fudan University.

Plasma samples of healthy human

Two batches of plasma samples were collected from healthy individuals at the health screening clinic of Shanghai Public Health Clinical Center. A concurrent batch was collected in December 2020. All the 95 individuals enrolled in this batch reported no epidemiological link with confirmed COVID-19 patients and were confirmed to be free from any chronic or acute disease. Viral RNA tests confirmed that all individuals in this batch were free from SARS-CoV-2 infection. In addition, a historical batch of 78 plasma samples from healthy individual cohort (collected in 2016) were also measured for their cross reactivities with SARS-CoV-2 S protein. Demographical information about

these two cohorts was described in Table 1.

Detection of SARS-CoV-2 S1 and S2 specific binding antibodies

In-house enzyme-linked immunosorbent assays (ELISA) were developed to measure SARS-CoV-2 S1 and S2 specific binding antibodies. High-binding 96-well EIA plates (Cat# 9018, Corning, USA) were coated with purified SARS-CoV-2 S1 (Cat# 40591-V08H, Sino Biological, China) or S2 proteins (Cat# 40590-V08B, Sino Biological, China) at a final concentration of 1μg/ml in carbonate/bi-carbonate coating buffer (30mM $NaHCO_3$,10mM Na_2CO_3, pH 9.6). Subsequently, the plates were blocked with 1×PBS containing 5% milk for 1 hour at 37°C. Next, 100μl of diluted human plasma or mouse serum was added to each well. After 1-hour incubation at 37°C, the plates were washed with 1×PBS containing 0.05% Tween20 for 5 times. Then, 100μl of a HRP labeled goat anti-mouse IgG antibody (Cat# ab6759, Abcam, UK) or goat anti-mouse IgG antibody (Cat# 115-035-003, Jackson Immuno Research, USA) diluted in 1×PBS containing 5% milk were added to each well and incubated for 1 hour at 37°C. After a second round of wash, 100μl of TMB substrate reagent (Cat# MG882, MESGEN, China) was added to each well. 15 minutes later, the color development was stopped by adding 100μl of 1M H_2SO_4 to each well and the values of optical density at $OD450_{nm}$ and $OD630_{nm}$ were measured using 800 TS microplate reader (Cat# 800TS, Biotek, USA).

Competitive ELISA

According to the reference sequence of SARS-CoV-2 (Genebank accession number: NC_045512), peptides (18-mer overlapping by 11 residues, purities > 95%) encompass the full length of S protein were synthesized by GL Biochem (Shanghai, China). The experiment procedure was generally similar with the afore mentioned in-house ELISA assays, except that the diluted mouse serum or human plasma were incubated with synthesized peptides (5μg/ml) for 1 hour at room temperature before adding into the coated EIA plates. 1×PBS containing 0.01% DMSO (the solvent used to dissolve peptides) were used as the negative control in this assay.

Antibody avidity assay

Avidity of Ag-specific Ab Avidity of Ag-specific Ab was determined by avidity

170 ELISA as reported (34) with minor modifications. Briefly, plates were coated as
171 the regular ELISA assay described above. Diluted (1:200) mouse sera were
172 added into each well. After 1-hour incubation, ELISA plates were washed with
173 washing buffer and incubated with 1.5M NaSCN or PBS for 15 mins at room
174 temperature and then immediately washed with washing buffer. Ab avidity
175 index was defined as the ratio of the OD value of a sample with 1.5M NaSCN
176 treatment versus the OD value of the same sample with PBS treatment.
177

178 **FACS analysis of S2 specific B cells in mice**

179 Spleen and mesenteric lymph nodes were isolated from naïve SPF mice and
180 single-cell suspensions were freshly prepared. After counting, 1×10^6 single
181 cells were resuspended in 100μl R10 (RPMI1640 containing 10% fatal bovine
182 serum) and incubated with biotinylated S2 protein (Cat# 40590-V08B-B, Sino
183 Biological, China) for 30 minutes at room temperature. After incubation, the
184 cells were washed twice with 500μl R10. Then, the cells were incubated with
185 the mixture of PE-anti-mouse CD19 (Cat# 152408, Biolegend, USA, 1μl/test),
186 BV510-anti-mouse CD45 (Cat# 103137, Biolegend, USA, 1.25μl/test) and
187 Streptavidin-IF647 (Cat# 46006, AAT Bioquest, USA, 0.2μl/test) at room
188 temperature for 30 minutes. After washing, the stained cells were resuspended
189 in 200μl 1×PBS and analyzed using a BD LSRFortessa™ Flow Cytometer.
190 The data were analyzed using the FlowJo software (BD Biosciences, USA).
191

192 **Preparation of P144 specific monoclonal antibodies**

193 Monoclonal antibodies against P144 were prepared from one naïve BALB/c
194 mouse and one naïve C57BL/6 mouse respectively using the hybridoma
195 technique. Briefly, freshly isolated splenocytes were mixed and fused with
196 SP2/0 cells at a ratio of 1:10. Hybridoma cell clones secreting P144 specific
197 antibodies were screened by ELISA and monoclonal hybridoma cells were
198 selected by multiple rounds of limited dilution. Selected clones of hybridoma
199 cells were injected intraperitoneally into BALB/c×ICR hybrid mice. About 1-2
200 weeks later, peritoneal fluid was collected, and monoclonal IgG was purified
201 using Protein A resin. The purities of monoclonal antibodies were verified using
202 SDS-PAGE and the antibody concentrations were determined using a BCA kit
203 (Cat# P0012, Beyotime Biotechnology, China).
204

205 **Isolation of gut commensal bacteria and preparation of whole cell lysate**
206 **(WCL)**
207 About 2g of each fecal sample was suspended with 15ml sterile 1×PBS and
208 vortexed thoroughly to obtain uniform mixtures. After centrifugation at 200×g
209 for 5 min, the supernatants were collected, and the sediments were discarded.
210 This process was repeated twice. Next, all the supernatant samples were
211 centrifuged twice at 9000×g for 5 min and the supernatants were discarded.
212 The precipitated bacteria pellets were resuspended in 500μl of 1×PBS
213 (containing 1mM PMSF) and disrupted with an ultrasonic cell crusher (the
214 probe-type sonicator, Model JY92-II; Ningbo Scientz Biotechnology Co., Ltd,
215 China). After sonication, the samples were centrifuged at 10000rpm for 30
216 minutes to remove the cellular debris.
217
218 **Western blotting**
219 WCL containing 10μg of total protein was separated by SDS-PAGE (10%
220 acrylamide gels) and then transferred onto a PVDF membrane (Cat#
221 IPVH00010, Millipore, USA) or stained with Coomassie brilliant blue. After
222 blocking with 5% skim milk for 2h, the membrane was incubated with a P144
223 specific monoclonal antibody or a control mouse IgG at a concentration of 1
224 μg/ml. After washing, the membrane was incubated with HRP conjugated goat
225 anti-mouse IgG antibody (Cat # ab6759, Abcam, UK) diluted 1:5000 in TBST
226 (Tris-buffered saline, pH 8.0, 0.05% Tween 20) containing 5% skim milk. After
227 wash, the bands were developed with an ultra-sensitive ECL substrate (Cat#
228 K-12045-D10, Advansta, USA). The area corresponding to the specific WB
229 bands were excised from the gel stained with Coomassie blue and analyzed
230 using the mass spectrometry.
231
232 **Mass spectrometry analysis**
233 The FASP digestion was adapted for the following procedures in Microcon
234 PL-10 filters. After three-time buffer displacement with 8 M Urea and 100 mM
235 Tris-HCl, pH 8.5, proteins were reduced by 10 mM DTT at 37 °C for 30 min and
236 followed by alkylation with 30 mM iodoacetamide at 25°C for 45 min in dark.
237 Digestion was carried out with trypsin (enzyme/protein as 1:50) at 37°C for 12
238 h after a wash with 20% ACN and three-time buffer displacement with
239 digestion buffer (30 mM Tris-HCl, pH 8.0). After digestion, the solution was

240 filtrated out and the filter was washed twice with 15% ACN, and all filtrates
241 were pooled and vacuum-dried to reach a final concentration to 1 mg/ml.
242 LC-MS analysis was performed using a nanoflow EASYnLC 1200 system
243 (Thermo Fisher Scientific, Odense, Denmark) coupled to an Orbitrap Fusion
244 Lumos mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). A
245 one-column system was adopted for all analyses. Samples were analyzed on
246 a home-made C18 analytical column (75 μm i.d. × 25 cm, ReproSil-Pur 120
247 C18-AQ, 1.9 μm (Dr. Maisch GmbH, Germany)). The mobile phases consisted
248 of Solution A (0.1% formic acid) and Solution B (0.1% formic acid in 80% ACN).
249 The derivatized peptides were eluted using the following gradients: 2–5% B in
250 2 min, 5–35% B in 100 min, 35–44% B in 6 min, 44–100% B in 3 min, 100% B
251 for 10 min, at a flow rate of 200 nl/min. Data-dependent analysis was
252 employed in MS analysis: The time between master scan was 3s, and
253 fragmented in HCD mode, normalized collision energy was 30.
254
255 **Construction and preparation of a candidate DNA vaccine encoding**
256 **SARS-CoV-2 full length S protein**
257 The full-length *s* gene sequence of the reference SARS-CoV-2 strain was
258 optimized according to the preference of human codon usage and synthesized
259 by GENEWIZ life science company (Suchow, China). The codon optimized
260 spike gene was subcloned into a eukaryotic expression vector (pJW4303,
261 kindly gifted by Dr. Shan Lu's Laboratory at the University of Massachusetts)
262 (35). And the sequence of inserted gene was verified by Sanger sequencing
263 (Sangon Biotech Co., Ltd., Shanghai, China). An EndoFree Plasmid
264 Purification Kit (Cat#12391, Qiagen, Hilden, USA) was used to prepare the
265 recombinant plasmid for mouse vaccination.
266
267 **Mouse vaccination**
268 Peripheral blood samples were collected from female adult mice and
269 pre-existing S2 binding antibodies were measured using the previously
270 described in-house ELISA method. According to their pre-existing S2 binding
271 antibody levels (at 1:100 dilution of serum), the mice were divided into three
272 groups: low ($0.015 < OD_{450nm-630nm} \leqslant 0.130$, n=6), moderate ($0.130 <$
273 $OD_{450nm-630nm} \leqslant 0.750$, n=6) and high ($OD_{450nm-630nm} > 0.750$, n=6). All mice
274 were immunized intramuscularly with the candidate S protein DNA vaccine

275 (50μg/mouse) for three times at an interval of 2 weeks. Three weeks post the
276 third vaccination, the mice were euthanized. Peripheral blood, bronchial lavage
277 and spleen were collected for assays of S protein specific immune responses.
278

279 **Metagenomic analysis of mouse gut microbiota**

280 Stool (fecal) samples were self-collected and DNA was extracted from 250 mg
281 of each fecal sample in duplicate. The DNA samples were separated strictly
282 according to TIANamp Stool DNA Kit (Cat# DP328, TIANGEN, China). Then, a
283 total amount of 1μg DNA per sample was used as input material for the DNA
284 sample preparations. Sequencing libraries were generated using NEBNext®
285 Ultra™ DNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's
286 recommendations and index codes were added to attribute sequences to each
287 sample. Briefly, the DNA sample was fragmented by sonication to a size of
288 350bp, then DNA fragments were end-polished, A-tailed, and ligated with the
289 full-length adaptor for Illumina sequencing with further PCR amplification. At
290 last, PCR products were purified (AMPure XP system) and libraries were
291 analyzed for size distribution by Agilent 2100 Bioanalyzer and quantified using
292 real-time PCR. The clustering of the index-coded samples was performed on a
293 cBot Cluster Generation System according to the manufacturer's instructions.
294 After cluster generation, the library preparations were sequenced on an
295 Illumina Novaseq 6000 platform and paired-end reads were generated.
296

297 **SARS-CoV-2 pseudo-virus neutralization assay**

298 VSV-backboned SARS-CoV-2 pseudo-viruses were prepared according to a
299 reported method (36). The neutralization assay was conducted by following
300 the previously described procedure (36, 37). Briefly, 100μl of serially diluted
301 mice sera were added into 96-well cell culture plates. Then, 50μl of
302 pseudo-viruses with a titer of 13000 $TCID_{50}$/ml were added into each well and
303 the plates were incubated at 37°C for 1 hour. Next, Vero cells were added into
304 each well (2×10^4 cells/well) and the plates were incubated at 37°C in a
305 humidified incubator with 5% CO_2. 24 hours later, luminescence detection
306 reagent (Bright-Glo™ Luciferase Assay System, Promega, USA) was added to
307 each well following the manufacturer`s instruction. The luminescence was
308 measured using a luminescence microplate reader (GloMax®
309 Navigator Microplate Luminometer, Promega, USA) within 5 minutes. The

310 Reed-Muench method was used to calculate the virus neutralization titer.
311 Antibody neutralization titers were presented as 50% maximal inhibitory
312 concentration (IC_{50}).
313
314 **Detections of S protein specific cellular immune responses**
315 SARS-CoV-2 S protein specific IFN-γ releases were measured using the
316 method of enzyme-linked immunosorbent spot (ELISPOT) assays (Cat#
317 551083, BD Bioscience, USA) according to a previously described
318 procedure(38). Briefly, the 96-well ELISPOT plates were coated with purified
319 anti-mouse IFN-γ monoclonal antibody overnight at 4°C. Then, the plates were
320 blocked and 2×10^5 fresh splenocytes were added into each well and
321 incubated with peptide pools for 20 hours at 37°C in a humidified incubator
322 with 5% CO_2. The final concentration for each peptide was 1μg/ml. After
323 incubation, detecting antibody and Avidin-HRP were added sequentially.
324 Finally, the plates were developed using the BD™ ELISPOT AEC Substrate
325 Set (Cat#551951, BD Bioscience, USA) according to the manufacturer's
326 manual. Spots representing IFN-γ producing cells were enumerated using an
327 automated ELISPOT plate reader (ChampSpot III Elispot Reader, Saizhi,
328 Beijing, China).
329 At the same time, the supernatants in the wells of ELISPOT plates were also
330 collected for detecting secreted cytokines using a multiplexed cytokine beads
331 array kit (Cat# 741054, Biolegend, USA).
332
333 **Statistical analysis**
334 All statistical analyses were performed using GraphPad Prism 8
335 (GraphPad Software, Inc., La Jolla, CA, USA). Comparisons between two
336 groups were conducted by the method of *t*-test. Comparisons among three or
337 more group were done using one-way ANOVA. $P<0.05$ was considered as
338 statistically significant.
339
340 **Results**
341 **Pre-existing antibodies recognizing a dominant linear epitope on**
342 **SARS-CoV-2 S2 protein were detected in both human and mice**
343 Pre-existing antibodies cross-react with SARS-CoV-2 S protein have been
344 found in uninfected individuals by multiple previous studies (22, 25, 26, 39). It

345 was postulated that the pre-existing immunities against SARS-CoV-2 might be
346 induced by previous exposure to seasonal human coronaviruses (22, 30, 32,
347 33, 40). However, contradictive evidence suggested that human common cold
348 coronavirus infection did not necessarily induce antibodies cross-reactive with
349 SARS-CoV-2 spike protein (28, 41, 42). In addition to this hypothesis, an
350 alternative explanation suggested that the cross-reactive immunities to
351 SARS-CoV-2 might derive from other common human pathogens and
352 vaccines (43).
353 To track the origins of the pre-existing cross-reactive antibodies to
354 SARS-CoV-2 spike protein, in this study, we first measured the levels of
355 pre-existing S protein specific antibodies in healthy human individuals and
356 SPF mice. Our data showed that the cross-reactive antibody responses
357 against S2 were significantly stronger than those against S1 in plasma
358 samples of healthy human collected both pre (2016 cohort) and post (2020
359 cohort) the outbreak of COVID-19 pandemic (Fig.1A and 1B). More strikingly,
360 our data showed that binding antibodies targeting S2 could also be detected in
361 two strains of naïve SPF mice (Fig. 1C and 1D). And this finding was further
362 confirmed by Western-blotting (WB) assays, which showed that mouse sera
363 with high ELISA-detected OD values (Fig. 2A) bound specifically with purified
364 S2 while not S1 (Fig. 2B). Quite interestingly, the WB results indicated that
365 cross-reactive antibodies against S2 also existed in the serum of a mouse
366 (#487) with no detectable binding signal in the S2 ELISA assay (Fig. 2A and
367 2B). We next performed linear antibody epitope mapping using an in-house
368 developed method of peptide competition ELISA. Our data showed that a
369 single peptide (P144, aa1145-aa1162, 18 mer) accounted for most of the
370 observed pre-existing antibody responses towards S2 in mice (Supplementary
371 Figure 1). Via employing a series of truncated peptides based on P144, we
372 determined the minimal range of this epitope
373 (1147-SFKEELDKYFKNHT-1160), which locates on the connector domain
374 (adjacent to the N-terminal of HR2 domain) (Fig. 2C). We also proved that
375 antibodies recognizing this epitope widely existed in both healthy human and
376 naïve SPF mice using the competitive ELISA assay (Fig.3).
377
378 **The P144 specific antibody responses might be engendered by**
379 **exposures to certain commensal gut bacteria**

380 To explore the potential origins of the pre-existed P144 specific antibodies, we
381 first performed phylogenetic analyses among SARS-CoV-2 and other human
382 coronaviruses. The results showed that the aa sequence of P144 was highly
383 conserved among SARS-CoV-2 variants and SARS-CoV, while the similarities
384 between P144 and MERS-CoV or seasonal human coronaviruses were
385 relatively low, especially within the range of predicted antibody binding epitope
386 (boxed fragment) (Figure 4A). The possibility of coronavirus infection in our
387 SPF-mouse colonies was excluded by serum screening tests using
388 commercialized mouse hepatitis virus (MHV) antigen and antibody detection
389 kits (Data not shown).
390 Subsequently, to investigate whether environmental factors contribute to the
391 induction of these S2 cross-reactive antibodies, we compared the levels of
392 pre-existing S2 binding antibodies between mice housed in SPF condition and
393 mice maintained in a sterile isolation pack. Our data showed that the levels of
394 pre-existing S2 binding antibodies were significantly higher in SPF mice
395 (Figure 4B). Through metagenomic sequencing, we further demonstrated that
396 the compositions of commensal gut bacteria were significantly different
397 between mice housed in SPF condition and mice maintained in the sterile
398 isolation pack (Supplementary Figure 2A). The abundance of bacteroidaceae,
399 prevotellaceace and parabacteroides increased significantly in the commensal
400 gut bacteria of SPF mice (Figure 4C). Moreover, flowcytometry assays
401 indicated that the frequencies of S2 specific B cells ($CD3^{-}S2^{+}CD19^{+}$) in
402 mesenteric lymph nodes were significantly higher than those in spleen
403 (Supplementary Figure 2B). These results collectively implied that the S2
404 cross-reactive antibodies might be induced by exposures to certain microbial
405 antigens.
406
407 **P144 specific monoclonal antibodies reacted with commensal gut**
408 **bacteria from both human and mouse and showed limited neutralizing**
409 **activities**
410 To probe the potential antigens that might induce the P144 binding antibodies,
411 we isolated 6 mAbs from two naïve SPF mice (one C57BL/6J mouse and one
412 BALB/c mouse) with high levels of pre-existing S2 specific antibody responses.
413 Five of these mAbs recognized P144 solely, while one mAb (clone M3) bound
414 with P144 and P103 simultaneously (Supplementary Figure 3). Results of

415 competitive ELISA showed that the minimal epitopes varied slightly among the
416 five P144 specific mAbs, especially at the C-terminal of P144 (Supplementary
417 Figure 3A). The neutralizing potentials of these isolated monoclonal antibodies
418 were evaluated using a pseudo virus-based neutralization assay. Our results
419 showed that these monoclonal antibodies exhibited limited neutralizing activity
420 against 5 major SARS-CoV-2 variants (Supplementary Figure 4). Of note, G18
421 neutralized the Indian strain (B.1.617) by 31%; F5 neutralized the UK strain
422 (B.1.1.7) and the South African strain (B.1.351) by around 20%; G13
423 neutralized the South African strain (B.1.351) by nearly 20% (Supplementary
424 Figure 4).
425 To prove the cross-reactivities between S2 and commensal gut microbial
426 antigens, whole cell lysates (WCL) of mouse and human commensal gut
427 bacteria were prepared and used as antigens for Western blotting assays,
428 respectively. As shown in Figure 5A, specific bindings with the WCL of mixed
429 fecal bacteria prepared from mice either with low levels of pre-existing
430 antibodies (L) or with high levels of pre-existing antibodies (H) could be clearly
431 visualized for each isolated mAbs. It was noteworthy that all mAbs except E10
432 strongly recognized a band around 180KD in the sample from mice with high
433 pre-existing antibody responses. E10 predominantly recognized a band
434 around 55KD in both samples, while stronger binding with the sample from
435 mice with high pre-existing antibody responses could be visually observed
436 (Figure 5A). Among the six mAbs, F5 showed the most diverse binding
437 compacity. In addition to the band around 180KD, F5 bound with a band
438 around 55KD (similar with E10) and a band between 40KD-55KD (Figure 5A).
439 In comparison with the WB results of mouse samples, the recognized bands
440 were less consistent across different human fecal bacteria samples (Figure
441 5B), presumably due to the individual to individual variation of gut microbiota
442 composition. We found that a band around 70KD was recognized by most
443 mAbs in 4 (lanes 1, 5, 6, 7) out of 7 samples and a band between 50KD-70KD
444 was recognized by all mAbs in 3 (Lanes 2, 3, 7) out 7 samples. Besides, our
445 data showed that the recognition pattern of each sample was largely
446 consistent across different mAbs (Figure 5B).
447 Proteins corresponding to specifically recognized bands were excised from
448 Coomassie blue stained gels and analyzed by the mass spectrometry. For the
449 mouse fecal bacteria samples, protein bands with molecular weights around

450 180KD, 100KD, 55KD-70KD and 40KD-55KD (indicated by arrows in Figure
451 5A, panel F5) were selected. For human fecal bacteria samples, protein bands
452 with molecular weights around 50KD-70KD, 70KD and 70KD-100KD
453 (indicated by arrows in Figure 5B, panel F5) were selected. The lists of
454 proteins identified in mouse and human samples were shown in Table 2 and
455 Table 3, respectively. Proteins with molecular weights corresponding
456 approximately to the excised protein bands were identified for both human and
457 mouse fecal bacteria samples. Of note, multiple proteins within the theoretical
458 MW range of 58KD to 60KD were found to be identical between mouse and
459 human samples, which included Fumarate hydratase class I (Accession#
460 P14407), Formate-tetrahydrofolate ligase OS (Accession# Q189R2),
461 Phosphoenolpyruvate carboxykinase (ATP) OS (Accession# C4ZBL1 and
462 A6LFQ4) and 60 kDa chaperonin OS (Accession# A0Q2T1).
463
464 **Pre-existing cross-reactive antibodies impacted specific immunities**
465 **induced by a candidate DNA vaccine encoding SARS-CoV-2 S protein**
466 Pre-existing antibodies has been shown to be able to shape the recall immune
467 responses upon influenza infection and vaccination (6). And the concern about
468 how the pre-existing immunities may influence the effect of a SARS-CoV-2
469 vaccine has also attracted lots of attention (33). To investigate the impact of
470 the pre-existing P144 antibodies on the immunogenicity of a candidate DNA
471 vaccine, 18 BALB/c mice were divided into 3 groups according to their levels of
472 pre-existing S2 binding antibodies and immunized with a DNA vaccine
473 encoding the full length of SARS-CoV-2 S protein (Figure 6A and 6B). Our data
474 showed that mice with high levels of pre-existing S2 binding antibodies
475 mounted significantly higher S2 binding antibody responses after vaccination
476 compared to mice with low or moderate levels of pre-existing S2 binding
477 antibodies (Figure 6C). More specifically, the average level of P144 specific
478 antibody responses was also stronger in mice with high levels of pre-existing
479 S2 binding antibodies than mice with low levels of pre-existing S2 binding
480 antibodies (Figure 6D and 6E). By comparison, both the S1 binding antibody
481 and the neutralizing antibody titers did not significantly differ among all groups,
482 despite that mice with moderate or high levels of pre-existing antibodies
483 tended to mount higher average titer of S1 binding antibodies (Figure 6F and
484 6G). Mice without vaccination showed neither obvious S1 binding antibody

485 response nor neutralizing activity (Data not shown). We further investigated
486 the influence of pre-existing antibodies on humoral immune responses in
487 mouse respiratory tract after vaccination. And our data showed that the levels
488 of S1 specific IgG in BALF were similar among the three groups after DNA
489 vaccination (Figure 7A), while the average level of S2 specific IgG in BALF
490 from mice with high pre-existing S2 binding antibodies was significantly higher
491 than those from mice with low pre-existing antibodies (Figure 7B). S protein
492 specific IgA response did not increase significantly after vaccination as
493 compared with unvaccinated group (Figure 7C and 7D).
494 As most pre-existing antibodies in naïve SPF mice recognized P144 (Figure
495 2C and Supplementary Figure 1), we delineated the impact of pre-existing
496 antibody on the epitope recognition after vaccination. Our results showed that
497 the minimum epitope recognition pattern by the sera of mice with high levels of
498 pre-existing antibodies remained unchanged after vaccination (Figure 8).
499 Whereas the minimum epitope recognized by the sera of mice with moderate
500 and low levels of pre-existing antibody responses altered at either the
501 N-terminal or both terminals of P144 (Figure 8).
502 In addition to antibody measurement, we compared S protein specific T cell
503 responses among the three groups as well (Figure 9A). The results showed
504 that the candidate DNA vaccine elicited robust S protein specific T cell
505 responses in all groups (Figure 9). Although no statistical significance was
506 reached, interesting trends were observed: First, mice with high levels of
507 pre-existing S2 binding antibodies tended to mount relatively higher S1
508 specific cellular immune responses (Figure 9B, 9C and 9D); second, as
509 measured by the total responses of IL-6, IL-2 and TNF-a, mice with high levels
510 of pre-existing antibodies tended to mount stronger T cell responses against
511 S1 while lower responses against S2 (Figure 9C and 9D). Mice without
512 vaccination showed no S protein specific T cells responses (Data not shown).
513
514 **Discussion**
515 The origins of pre-existing cross-reactive immunities against SARS-CoV-2
516 have been investigated vigorously since the outbreak of the pandemic (44).
517 Accumulating data suggest that cross-reactive T cells (33, 45-48) in
518 SARS-CoV-2 unexposed human might be induced by previous infections of
519 other hCoVs. While the origins of pre-existing cross-reactive antibodies could

520 not be completely explained by previous infections of other coronaviruses, as
521 recent studies revealed that the magnitude of antibody responses to
522 SARS-CoV-2 S protein in the sera of patients with COVID-19 was not related
523 to HCoVs' S titers (49) and immunization with coronaviruses OC43 did not
524 induce significant SARS-CoV-2 S protein cross-reactive antibodies in mice.
525 Moreover, it has also been observed that SARS-CoV-2 S protein specific
526 binding antibody responses were weak in SARS-CoV-2 unexposed individuals
527 with obvious binding antibody responses against S proteins of common cold
528 hCoVs (50, 51).
529 To track the potential origins of the pre-existed cross-reactive antibodies
530 targeting SARS-CoV-2 spike protein, in this study, we first screened the
531 cross-reactive antibody responses in SARS-CoV-2 unexposed human plasma
532 collected in 2020 and 2016, respectively. In both cohorts, we found that the
533 magnitudes of S2 binding antibodies were significantly higher than those of S1
534 binding antibodies. This finding is consistent with previous studies showing
535 that pre-existing S2 cross-reactive antibody responses are stronger than S1
536 cross-reactive antibody responses in SARS-CoV-2 unexposed individuals (49,
537 52, 53). Since S2 cross-reactive antibody responses have also been observed
538 in unexposed animals (49), we continued to screen the cross-reactive antibody
539 responses in two strains of naïve SPF mice. Our data showed that the OD
540 values of S2 cross-reactive antibodies were significantly higher than those of
541 S1 cross-reactive antibodies in naïve BALB/c and C57BL/6 mice. Detections of
542 mouse sera collected from another two independent SPF animal facilities
543 confirmed this finding (Data not shown). We also tried to detect the
544 SARS-CoV-2 S protein specific T cells responses in mice with high
545 pre-existing S2 cross-reactive antibodies using the IFN-γ ELISPOT assay,
546 which showed that there was no pre-existing cross-reactive T cell response.
547 To facilitate the search of potential antigens that induced the cross-reactive
548 antibodies, we identified a dominant antibody epitope (P144) through a
549 method of competitive ELISA based linear antibody epitope mapping. P144 is
550 located within the connector domain of S2 (aa1147-aa1160, directly N-terminal
551 of the HR2 region). The same epitope has been predicted (43) and detected in
552 both SARS-CoV-2 unexposed and infected individuals by multiple previous
553 studies (22, 23, 25, 26, 32). In this study, we also detected P144 specific
554 antibody responses in plasma of healthy individuals collected in both 2020 and

555 2016. More interestingly, we found that the pre-existing S2 cross-reactive
556 antibodies in mice were predominantly against the single epitope. Of note,
557 although the aa sequence of P144 is highly conserved between SARS-CoV-2
558 and SARS-CoV, its similarity with four seasonal hCoVs is relatively low. A
559 recent study showed that this epitope was more frequently recognized than its
560 homologous peptides from common cold hCoVs by antibodies in plasma of
561 COVID-19 negative individuals (23). The above evidence implied again that
562 the pre-existing S2 specific antibodies might not be necessarily elicited by
563 previous common cold coronavirus infections.
564 As the pre-existing S2 binding antibodies in mice were predominantly against
565 P144, it provided us a good chance to unveil their origins. To do so, we first
566 labeled mouse B cells with purified S2 protein and found that the frequency of
567 S2 specific B cells was significantly higher in mesenteric LN than in spleen,
568 suggesting that the gastrointestinal tract might be the primary site where the
569 cross-reactive B cells were activated. Exposure to a certain gut microbial
570 antigen, which can promote B cell diversification and stimulate antibody
571 production in both T-dependent and -independent ways (54), might account for
572 the presence of the cross-reactive antibodies. To prove this hypothesis, we
573 next compared the levels of pre-existing S2 cross-reactive antibodies between
574 mice housed in a sterile isolation pack and mice maintained in SPF condition.
575 Our results showed that the levels of pre-existing S2 cross-reactive antibodies
576 were significantly higher in SPF mice. Through metagenomic sequencing, we
577 further demonstrated that the abundance of bacteroidaceae, prevotellaceace
578 and parabacteroides were also significantly higher in the commensal gut
579 bacteria of SPF mice. Therefore, we speculated that the S2 cross-reactive
580 antibodies might be induced by commensal gut bacteria.
581 To prove the above speculation directly, we isolated six P144 specific
582 monoclonal antibodies from a naïve BALB/c mouse and a naïve C57BL/6
583 mouse, respectively. All the six mAbs were confirmed to be able to bind with
584 P144 and showed weak neutralizing compacities against five SARS-CoV-2
585 variants. By leveraging these mAbs, we detected the potential cross-reactive
586 antigens in mouse and human fecal microbiota through WB assays. Compared
587 with a control mouse IgG, specific bands were observed for each mAb, which
588 proved antibody cross-reactivities between SARS-CoV-2 and commensal gut
589 bacteria. The strongly recognized protein bands were further analyzed by

590 LC-MS. Our data showed that antigens derived from bacteroides and
591 parabacteroides were frequently identified in fecal bacteria samples of both
592 human and mouse, which was in consistence with our metagenomic
593 sequencing data showing that the abundance of bacteroides and
594 parabacteroides was significantly higher in the commensal gut bacteria of mice
595 with high pre-existing S2 binding antibody levels. More intriguingly, five
596 potential cross-reactive microbial antigens were identified in mouse and
597 human samples simultaneously, which implied that the S2 cross-reactive
598 antibodies might naturally occur in different species of mammals. We analyzed
599 the sequence similarities between P144 and the proteins identified by LC-MS.
600 Our results showed that most identified proteins shared 40%-70% identities for
601 more than 8 residues with P144 (Data not shown). The cross-reactive epitopes
602 on the identified proteins could not specified based on current data. Besides,
603 the neutralizing mechanism(s) of P144 specific antibodies and their potential
604 influences on gut microbiota were not clarified in this study. We plan to look
605 into these issues in future.
606 In parallel with tracking the initial antigens that induced the S2 cross-reactive
607 antibodies, we investigated the impact of pre-existing antibodies on the
608 immunogenicity of a candidate DNA vaccine as well. According to previous
609 reports, pre-existing cross-reactive antibodies may influence the effects of
610 different vaccines differentially (6, 55). In this study, our data showed that the
611 pre-existing cross-reactive antibodies shaped the vaccine-induced immune
612 responses by promoting S2 binding antibodies in both peripheral blood and
613 bronchial lavage. Meanwhile, the production of SARS-CoV-2 neutralizing
614 antibodies was not impaired. Through epitope mapping, we further showed
615 that the pre-existing antibodies strongly imprinted the minimal epitope
616 recognition in mice with high levels of pre-existing antibodies, which suggested
617 that the influence of pre-existing cross-reactive antibodies on vaccine induced
618 antibody responses was primarily epitope specific. In addition to antibody
619 response, we also found that the high levels of pre-existed S2 binding
620 antibodies tended to induce higher S1 specific T cell responses while lower S2
621 specific T cell responses, which indicated that the pre-existing cross-reactive
622 antibodies might change the balance between humoral versus cellular immune
623 responses. Collectively, the above results manifested that the pre-existing
624 cross-reactive antibodies against SARS-CoV-2 S2 did not impair the

625 immunogenicity of the candidate DNA vaccine. Further experiments of live
626 virus challenge will help to clarify whether the increased T cell and S2 binding
627 antibody responses may mitigate the severity of disease caused by
628 SARS-CoV-2.
629 A deep understanding of pre-existing cross-reactive antibodies against
630 SARS-CoV-2 will enable better therapeutic, diagnostic and vaccine strategies.
631 In this study, we provided evidence showing that antibodies targeting a
632 conserved linear epitope on S2 cross-reacted with gut microbial antigens from
633 both human and mouse, manifesting that some of the pre-existing
634 cross-reactive antibodies might be induced by exposure to certain commensal
635 gut bacteria. We proved that the pre-existing S2 cross-reactive antibodies did
636 not impair the immunogenicity of a candidate DNA vaccine in a mouse model.
637 Further investigations into the role of P144 specific antibody after
638 SARS-CoV-2 infection and its impact on gut microbiota might provide clues to
639 elucidate the mechanisms underlying the gastrointestinal symptom of
640 COVID-19 (56-58).
641

642 **Acknowledgements**

643 We thank Miss Zhangyufan He from Huashan Hospital, Fudan University, for
644 her kind help with the language polishing. This work was funded by the
645 National Natural Science Foundation of China (Grant No. 81971559,
646 82041010, 81971900, 31872744), National Science and Technology Major
647 Project (Grant No. 2018ZX10731301-004, 2018ZX10302302-002 and
648 2018ZX10301-404-002-003) and the Science and Technology Commission of
649 Shanghai Municipality (Grant No. 20411950400).
650

651 **Author contributions**

652 Y.M.W., Y.X., C.Q., Z.Q.Z. and W.H.Z. designed the study. L.Q.J., S.F.W.,
653 Y.M.W., X.X.T., Y.F.Z., J.W., X.Y.W., and J.W. conducted the experiments.
654 Y.M.W., L.Q.J. and S.F.W. analyzed the data and drafted the manuscript.
655 Y.M.W., Y.X., C.Q., Z.Q.Z. and W.H.Z. revised the manuscript. D.M.Y. and
656 W.H.W. provided intellectual inputs in tackling technical challenges in tracking
657 the potential cross-reacting antigens.

Conflict of Interest

The authors declare that they have no relevant conflicts of interest.

References

1. Lu LL, Suscovich TJ, Fortune SM, Alter G. Beyond binding: antibody effector functions in infectious diseases. Nature reviews Immunology. 2018;18(1):46-61.
2. Warter L, Appanna R, Fink K. Human poly- and cross-reactive anti-viral antibodies and their impact on protection and pathology. Immunologic research. 2012;53(1-3):148-61.
3. Cobey S, Hensley SE. Immune history and influenza virus susceptibility. Current opinion in virology. 2017;22:105-11.
4. Mok DZL, Chan KR. The Effects of Pre-Existing Antibodies on Live-Attenuated Viral Vaccines. Viruses. 2020;12(5).
5. Zimmermann P, Curtis N. Factors That Influence the Immune Response to Vaccination. Clinical microbiology reviews. 2019;32(2).
6. Dugan HL, Guthmiller JJ, Arevalo P, Huang M, Chen YQ, Neu KE, et al. Preexisting immunity shapes distinct antibody landscapes after influenza virus infection and vaccination in humans. Science translational medicine. 2020;12(573).
7. Zhang A, Stacey HD, Mullarkey CE, Miller MS. Original Antigenic Sin: How First Exposure Shapes Lifelong Anti-Influenza Virus Immune Responses. Journal of immunology (Baltimore, Md : 1950). 2019;202(2):335-40.
8. Boes M. Role of natural and immune IgM antibodies in immune responses. Molecular immunology. 2000;37(18):1141-9.
9. Auladell M, Jia X, Hensen L, Chua B, Fox A, Nguyen THO, et al. Recalling the Future: Immunological Memory Toward Unpredictable Influenza Viruses. Frontiers in immunology. 2019;10:1400.
10. St. John AL, Rathore APS. Adaptive immune responses to primary and secondary dengue virus infections. Nature Reviews Immunology. 2019;19(4):218-30.
11. Izmirly AM, Alturki SO, Alturki SO, Connors J, Haddad EK. Challenges in Dengue Vaccines Development: Pre-existing Infections and Cross-Reactivity. 2020;11(1055).
12. Andrade P, Gimblet-Ochieng C, Modirian F, Collins M, Cárdenas M, Katzelnick LC, et al. Impact of pre-existing dengue immunity on human antibody and memory B cell responses to Zika. Nature Communications. 2019;10(1):938.
13. Liao HX, Chen X, Munshaw S, Zhang R, Marshall DJ, Vandergrift N, et al. Initial antibodies binding to HIV-1 gp41 in acutely infected subjects are polyreactive and highly mutated. The Journal of experimental medicine. 2011;208(11):2237-49.
14. Bonsignori M, Liao HX, Gao F, Williams WB, Alam SM, Montefiori DC, et al. Antibody-virus co-evolution in HIV infection: paths for HIV vaccine development. Immunological reviews. 2017;275(1):145-60.
15. Williams WB, Liao HX, Moody MA, Kepler TB, Alam SM, Gao F, et al. HIV-1 VACCINES. Diversion of HIV-1 vaccine-induced immunity by gp41-microbiota cross-reactive antibodies. Science (New York, NY). 2015;349(6249):aab1253.
16. Zitvogel L, Ayyoub M, Routy B, Kroemer G. Microbiome and Anticancer Immunosurveillance.

701 Cell. 2016;165(2):276-87.
702 17. Leng Q, Tarbe M, Long Q, Wang F. Pre-existing heterologous T-cell immunity and neoantigen
703 immunogenicity. Clinical & translational immunology. 2020;9(3):e01111.
704 18. Sioud M. T-cell cross-reactivity may explain the large variation in how cancer patients respond to
705 checkpoint inhibitors. Scandinavian journal of immunology. 2018;87(3).
706 19. Rose NR. Negative selection, epitope mimicry and autoimmunity. Current opinion in immunology.
707 2017;49:51-5.
708 20. Anderson EM, Goodwin EC, Verma A, Arevalo CP, Bolton MJ, Weirick ME, et al. Seasonal
709 human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with
710 protection. Cell. 2021.
711 21. Galipeau Y, Greig M, Liu G, Driedger M, Langlois MA. Humoral Responses and Serological
712 Assays in SARS-CoV-2 Infections. Frontiers in immunology. 2020;11:610688.
713 22. Klompus S, Leviatan S, Vogl T, Kalka IN, Godneva A, Shinar E, et al. Cross-reactive antibody
714 responses against SARS-CoV-2 and seasonal common cold coronaviruses. 2020:2020.09.01.20182220.
715 23. Shrock E, Fujimura E, Kula T, Timms RT, Lee IH, Leng Y, et al. Viral epitope profiling of
716 COVID-19 patients reveals cross-reactivity and correlates of severity. Science (New York, NY).
717 2020;370(6520).
718 24. Ortega N, Ribes M, Vidal M, Rubio R, Aguilar R, Williams S, et al. Seven-month kinetics of
719 SARS-CoV-2 antibodies and protective role of pre-existing antibodies to seasonal human
720 coronaviruses on COVID-19. 2021:2021.02.22.21252150.
721 25. Ladner JT, Henson SN, Boyle AS, Engelbrektson AL, Fink ZW, Rahee F, et al. Epitope-resolved
722 profiling of the SARS-CoV-2 antibody response identifies cross-reactivity with endemic human
723 coronaviruses. Cell reports Medicine. 2021;2(1):100189.
724 26. Ng KW, Faulkner N, Cornish GH, Rosa A, Harvey R, Hussain S, et al. Preexisting and de novo
725 humoral immunity to SARS-CoV-2 in humans. Science (New York, NY). 2020;370(6522):1339-43.
726 27. Yang R, Lan J, Huang B, A R, Lu M, Wang W, et al. Lack of antibody-mediated cross-protection
727 between SARS-CoV-2 and SARS-CoV infections. EBioMedicine. 2020;58:102890.
728 28. Lv H, Wu NC, Tsang OT, Yuan M, Perera R, Leung WS, et al. Cross-reactive Antibody Response
729 between SARS-CoV-2 and SARS-CoV Infections. Cell reports. 2020;31(9):107725.
730 29. Zhu Y, Yu D, Han Y, Yan H, Chong H, Ren L, et al. Cross-reactive neutralization of SARS-CoV-2
731 by serum antibodies from recovered SARS patients and immunized animals. Science advances.
732 2020;6(45).
733 30. Lipsitch M, Grad YH, Sette A, Crotty S. Cross-reactive memory T cells and herd immunity to
734 SARS-CoV-2. Nature reviews Immunology. 2020;20(11):709-13.
735 31. Sumbul B, Sumbul HE, Okyay RA, Gülümsek E, Şahin AR, Boral B, et al. Is there a link between
736 pre-existing antibodies acquired due to childhood vaccinations or past infections and COVID-19? A
737 case control study. PeerJ. 2021;9:e10910.
738 32. Majdoubi A, Michalski C, O'Connell SE, Dada S, Narpala SR, Gelinas JP, et al. A majority of
739 uninfected adults show pre-existing antibody reactivity against SARS-CoV-2. JCI insight. 2021.
740 33. Sette A, Crotty S. Pre-existing immunity to SARS-CoV-2: the knowns and unknowns. Nature
741 reviews Immunology. 2020;20(8):457-8.
742 34. Fan W, Wan Y, Li Q. Interleukin-21 enhances the antibody avidity elicited by DNA prime and
743 MVA boost vaccine. Cytokine. 2020;125:154814.
744 35. Lu S, Manning S, Arthos J. Antigen engineering in DNA immunization. Methods in molecular

745 medicine. 2000;29:355-74.
746 36. Nie J, Li Q, Wu J, Zhao C, Hao H, Liu H, et al. Establishment and validation of a pseudovirus
747 neutralization assay for SARS-CoV-2. Emerging microbes & infections. 2020;9(1):680-6.
748 37. Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, et al. The Impact of Mutations in SARS-CoV-2 Spike
749 on Viral Infectivity and Antigenicity. Cell. 2020;182(5):1284-94.e9.
750 38. Ren Y, Wang N, Hu W, Zhang X, Xu J, Wan Y. Successive site translocating inoculation
751 potentiates DNA/recombinant vaccinia vaccination. Scientific reports. 2015;5:18099.
752 39. Tso FY, Lidenge SJ, Peña PB, Clegg AA, Ngowi JR, Mwaiselage J, et al. High prevalence of
753 pre-existing serological cross-reactivity against severe acute respiratory syndrome coronavirus-2
754 (SARS-CoV-2) in sub-Saharan Africa. International journal of infectious diseases : IJID : official
755 publication of the International Society for Infectious Diseases. 2021;102:577-83.
756 40. Lee CH, Pinho MP, Buckley PR, Woodhouse IB, Ogg G, Simmons A, et al. Potential CD8+ T Cell
757 Cross-Reactivity Against SARS-CoV-2 Conferred by Other Coronavirus Strains. Frontiers in
758 immunology. 2020;11:579480.
759 41. Freeman B, Lester S, Mills L, Rasheed MAU, Moye S, Abiona O, et al. Validation of a
760 SARS-CoV-2 spike protein ELISA for use in contact investigations and sero-surveillance.
761 2020:2020.04.24.057323.
762 42. Post N, Eddy D, Huntley C, van Schalkwyk MCI, Shrotri M, Leeman D, et al. Antibody response
763 to SARS-CoV-2 infection in humans: A systematic review. PloS one. 2020;15(12):e0244126.
764 43. Reche PA. Potential Cross-Reactive Immunity to SARS-CoV-2 From Common Human Pathogens
765 and Vaccines. Frontiers in immunology. 2020;11:586984.
766 44. Doshi P. Covid-19: Do many people have pre-existing immunity? 2020;370:m3563.
767 45. Steiner S, Sotzny F, Bauer S, Na I-K, Schmueck-Henneresse M, Corman VM, et al. HCoV- and
768 SARS-CoV-2 Cross-Reactive T Cells in CVID Patients. 2020;11(3347).
769 46. Mateus J, Grifoni A, Tarke A, Sidney J, Ramirez SI, Dan JM, et al. Selective and cross-reactive
770 SARS-CoV-2 T cell epitopes in unexposed humans. 2020;370(6512):89-94.
771 47. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, et al. Targets of T Cell
772 Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed
773 Individuals. Cell. 2020;181(7):1489-501.e15.
774 48. Braun J, Loyal L, Frentsch M, Wendisch D, Georg P, Kurth F, et al. SARS-CoV-2-reactive T cells
775 in healthy donors and patients with COVID-19. Nature. 2020;587(7833):270-4.
776 49. Kim H, Seiler P, Jones JC, Ridout G, Camp KP, Fabrizio TP, et al. Antibody Responses to
777 SARS-CoV-2 Antigens in Humans and Animals. Vaccines. 2020;8(4).
778 50. Anderson EM, Goodwin EC, Verma A, Arevalo CP, Bolton MJ, Weirick ME, et al. Seasonal
779 human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with
780 protection. Cell. 2021;184(7):1858-64.e10.
781 51. Song G, He W-t, Callaghan S, Anzanello F, Huang D, Ricketts J, et al. Cross-reactive serum and
782 memory B-cell responses to spike protein in SARS-CoV-2 and endemic coronavirus infection. Nature
783 Communications. 2021;12(1):2938.
784 52. Fraley E, LeMaster C, Banerjee D, Khanal S, Selvarangan R, Bradley T. Cross-reactive antibody
785 immunity against SARS-CoV-2 in children and adults. Cellular & molecular immunology. 2021:1-3.
786 53. Nguyen-Contant P, Embong AK, Kanagaiah P, Chaves FA, Yang H, Branche AR, et al. S
787 Protein-Reactive IgG and Memory B Cell Production after Human SARS-CoV-2 Infection Includes
788 Broad Reactivity to the S2 Subunit. mBio. 2020;11(5).

54. Zhao Q, Elson CO. Adaptive immune education by gut microbiota antigens. Immunology. 2018;154(1):28-37.
55. Bradt V, Malafa S, von Braun A, Jarmer J, Tsouchnikas G, Medits I, et al. Pre-existing yellow fever immunity impairs and modulates the antibody response to tick-borne encephalitis vaccination. NPJ vaccines. 2019;4:38.
56. Lee IC, Huo TI, Huang YH. Gastrointestinal and liver manifestations in patients with COVID-19. Journal of the Chinese Medical Association : JCMA. 2020;83(6):521-3.
57. Jin X, Lian JS, Hu JH, Gao J, Zheng L, Zhang YM, et al. Epidemiological, clinical and virological characteristics of 74 cases of coronavirus-infected disease 2019 (COVID-19) with gastrointestinal symptoms. Gut. 2020;69(6):1002-9.
58. Mao R, Qiu Y, He JS, Tan JY, Li XH, Liang J, et al. Manifestations and prognosis of gastrointestinal and liver involvement in patients with COVID-19: a systematic review and meta-analysis. The lancet Gastroenterology & hepatology. 2020;5(7):667-78.

833 **Figure legends**

834

835 **Figure 1. Pre-existing cross-reactive antibodies against SARS-CoV-2 S**
836 **protein observed in both healthy human and naïve SPF mice**
837 **predominantly targeted S2**

838 The pre-existing cross-reactive antibodies against S1 and S2 were measured
839 using an in-house ELISA (Sample dilution factor: 100). (**A**) Plasma samples of
840 healthy individuals collected in 2020 (n=95). (**B**) Plasma samples of healthy
841 individuals collected in 2016 (n=78). (**C**) Sera of naïve C57BL/6J mice (n=12).
842 (**D**) Sera of naïve BALB/c mice (n=101). The dot lines showed the threshold of
843 background (3 folds of the background OD). Statistical analyses were
844 performed using the method of paired t-test.

845

846 **Figure 2 The pre-existing S protein binding antibodies in naïve SPF mice**
847 **recognized S2 exclusively and a dominant linear antibody epitope was**
848 **identified on the connector domain**

849 The pre-existing S1 and S2 cross-reactive antibody levels of 6 mice selected
850 for WB assays were shown in (**A**). (**B**) Western-blotting assays of pre-existing
851 cross-reactive antibodies for 6 selected mouse serum samples. The purities of
852 S1 and S2 proteins were shown by coomassie blue staining. (**C**) The minimal
853 epitope of P144 was defined using a method of competitive ELISA (Data
854 shown as mean±SD, n=5). Purified S2 protein was used as the coating antigen
855 and truncated peptides derived from P144 were used as competitors. The
856 decreases of competitive inhibition reflected the necessity of each amino acid
857 for the epitope recognition. Statistical differences among groups was analyzed
858 using One-way ANOVA. ****, P<0.0001. M: molecular weight markers; SP,
859 signal peptide; RBD, receptor-binding domain; FP, fusion peptide; HR, heptad
860 repeat; CH, central helix; CD, connector domain; TM, transmembrane domain;
861 CT, cytoplasmic tail.

862

863 **Figure 3 Cross-reactive antibodies recognizing epitope P144 widely**
864 **existed in both healthy human and naïve SPF mice**

865 P144 specific binding antibodies were detected using a method of competitive
866 ELISA. (**A**) For the detections of P144 specific binding antibodies in naïve SPF
867 mice, purified S2 protein was used as the coating antigen and P144 peptide

868 was used as the competitor. (**B**) For the detections of P144 specific binding
869 antibodies in healthy individuals, purified BSA-P144 conjugate was used as
870 the coating antigen and P144 peptide was used as the competitor. In both
871 experiments, the reduction of OD value reflected the presence of P144 binding
872 antibodies.

873

874 **Figure 4 The generation of the pre-existing S2 cross-reactive antibodies**
875 **might be associated with commensal gut microbiota**
876 (**A**) Phylogenetic analyses of the spike protein sequences of SARS-CoV-2
877 variants, SARS-CoV, MERS-CoV and seasonal human coronaviruses. The
878 sequence accession numbers are as follows, SARS-CoV-2 Wuhan,
879 YP_009724390; sequences of B.1.1.7, D614G, B.1.351, B.1.525, B.1.617 and
880 P.1(B.1.1.281) were obtained from Global Initiative on Sharing Avian Influenza
881 Data (GISAID); SARS-CoV, NP_828851.1; MERS-CoV, QFQ59587.1;
882 HCoV-OC43, QDH43719.1; HCoV-NL63, AKT07952.1; HCoV-229E,
883 AOG74783.1; HCoV-HKU1, YP_173238.1. Boxed fragments represent IEDB
884 predicted linear antibody epitope (http://tools.iedb.org/bcell/). (**B**) Comparison
885 of the levels of S2 specific pre-existing antibodies between naïve mice
886 maintained under SPF condition and mice in a sterile isolation pack. (**C**)
887 Comparison of commensal gut bacteria compositions between mice housed in
888 a SPF facility and mice in a sterile isolation pack by metagenomic sequencing.
889 Differences of bacterial abundance was analyzed by linear discriminant
890 analysis (LDA) analysis and shown as the histogram of LDA scores. It was
891 considered a significant effect size with LDA score > 4.0. (p), Phylum. (c),
892 Class. (o), Order. (f), Family. (g), Genus. (s), Species.

893

894 **Figure 5 P144 specific monoclonal antibodies isolated from naïve SPF**
895 **mice cross-reacted with commensal gut bacteria from both human and**
896 **mouse**

897 Cross-reactivities between P144 specific mAbs and gut microbial antigens
898 were detected using WB assays. A purified mouse IgG was used as the control.
899 (**A**) WB assays of mouse fecal bacteria samples. L: mixed fecal bacteria
900 samples collected from 3 mice with low levels of pre-existing S2 binding
901 antibodies ($OD_{450nm-630nm} \leqslant 0.140$ at serum dilution of 1:100); H: mixed fecal

902 bacteria samples collected from 3 mice with high levels of pre-existing S2
903 binding antibodies ($OD_{450nm-630nm} \geqslant 0.615$ at serum dilution of 1:100). (**B**) WB
904 assays of fecal bacteria samples collected from 7 healthy individuals (Lanes
905 1-7). All the mAbs and the control mouse IgG were diluted at the final
906 concentration of 1μg/ml. Black arrows indicated the locations of protein bands
907 chosen for the Mass spectrometry analysis.

908

909 **Figure 6 Impact of pre-existing antibodies on the humoral immune**
910 **responses elicited by a DNA vaccine encoding SARS-CoV-2 S protein**
911 (**A**) Schematic illustration of the vaccination regimen. 50μg of the candidate
912 DNA vaccine was injected intra muscularly into each mice at week 0, week 2
913 and week 4. Two weeks after the final vaccination, the mice was sacrificed for
914 the measurements of specific immune responses. (**B**) Periphral blood was
915 collected before immunization and levels of pre-existing S2 specific IgG were
916 compared among three groups of mice. (**C**) Comparisons of endpoint IgG titers
917 against S2 in the serum of mice measured at 2 weeks post the last
918 immunization. (**D**) Comparisons of P144 specific IgG titers in the serum of
919 mice as measured using BSA-P144 conjugate as the coating antigen. (**E**)
920 Comparisons of P144 specific binding antibody levels as determined using a
921 method of competitive ELISA. (**F**) Endpoint IgG titers against S1 measured at
922 2 weeks post the final vaccination. (**G**) Neutralizing antibody titers against
923 SARS-CoV-2 D614G pseudo-virus in serum of mice at 2 weeks post the final
924 immunization. The vaccination experiment was repeated twice with 6 mice for
925 each group. Data were shown as mean±SD. Statistical analyses were
926 performed by the method of one-way ANONVA.

927

928 **Figure 7 The impact of pre-existing antibody on the levels of specific IgG**
929 **and IgA in BALF after vaccination**
930 BALF was collected from each mouse after euthanasia. Specific IgG (A and B)
931 and IgA (C and D) against S1 or S2 were detected using in-house ELISA
932 methods. All the BALF samples were adjusted to the initial total protein
933 concentration of 51.9μg/ml. Data are shown as mean± SD. Statistical analyses
934 were performed by the method of one-way ANONVA.
935

936 **Figure 8 The impact of pre-existing antibody on the minimum epitope**
937 **recognition of P144 after vaccination.**
938 The minimal epitope recognized by mouse sera after vaccination were
939 analyzed using a method of competitive ELISA. Purified BSA-P144 conjugate
940 was used as the coating antigen and truncated peptides derived from P144
941 were used as the competitors. The decreases of competitive inhibition
942 reflected the necessity of each amino acid for the epitope recognition.
943 Statistical analyses were performed by the method of two-tailed t-test (*, $P<$
944 0.05, **, $P< 0.01$, ***, $P< 0.001$).
945
946 **Figure 9 The impact of pre-existing antibody on the cellular immune**
947 **responses after vaccination**
948 (**A**) The diagram of the method for cellular immune responses assays. (**B**) S1
949 and S2 specific IFN-γ responses were compared among groups of mice with
950 different levels of pre-existing S2 cross-reactive antibodies. Additionally, S1 (**C**)
951 and S2 (**D**) specific releases of IL-2, IL-6 and TNF-α as measured using the
952 method of multiplex cytokine bead assay were also compared among different
953 groups. Data were shown as mean$\pm$ SD. SFU, spot forming units.
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970

971 **Supplementary figures**

972 **Supplementary Figure 1 A dominant linear epitope recognized by the**
973 **pre-existing antibodies was located on the connector domain of S2**

974 (**A**) The location of P144 on the full length of SARS-CoV-2 spike protein.
975 Synthesized peptides (18-mer, overlapped by 11amino acids) spanning the full
976 length of S2 were divided into 9 peptide pools, each contained 8 peptides. To
977 identify the potential antibody binding epitopes, we first performed competitive
978 ELISA assays using the peptide pools as competitors (Data not shown). The
979 pool showing significant inhibition effect was further delineated by testing the
980 inhibiting efficiencies of each individual peptides (**B**).
981 The data was shown as mean±SD, n=5. Inhibition efficiencies among groups
982 were analyzed using the method of One-way ANOVA. ****, $p<0.0001$. SP,
983 signal peptide; RBD, receptor-binding domain; FP, fusion peptide; HR, heptad
984 repeat; CH, central helix; CD, connector domain; TM, transmembrane domain;
985 CT, cytoplasmic tail.
986

987 **Supplementary Figure 2 The pre-existing cross-reactive antibodies might**
988 **be induced by exposure to certain gut microbial antigens**

989 (**A**) Principle component analysis (PCA) of gut microbial communities between
990 4 mice in SPF condition and 4 mice in a sterile isolation pack. (**B**) The
991 splenocytes and lymphocytes of MLN were harvested from 4 naive C57BL/6J
992 mice with relatively high levels of pre-existing S2 cross-reactive antibodies. S2
993 specific B cells were defined as $CD45^{+}CD19^{+}S2^{+}$ (Left). Statistical analysis
994 was performed by the method of paired t-test (Right).
995

996 **Supplementary Figure 3 Characterizations of the minimal epitope**
997 **recognitions for the 6 monoclonal antibodies isolated from naïve mice**

998 Six P144 specific monoclonal antibodies were isolated from 2 naïve SPF mice
999 using the hybridoma technology. (**A**) The minimum epitope recognitions of
1000 clones H9, E10, G13, F5 and G18 were similar detected by a method of
1001 competitive ELISA. Purified S2 protein was used as the coating antigen and
1002 the truncated peptides derived from P144 were used as the competitors. (**B**)
1003 The epitope recognition of clone M3 was analyzed using purified S2 protein as
1004 the coating antigen and each listed peptide was used as the competitor. SP,
1005 signal peptide; RBD, receptor-binding domain; FP, fusion peptide; HR, heptad

1006 repeat; CH, central helix; CD, connector domain; TM, transmembrane domain;
1007 CT, cytoplasmic tail.
1008
1009 **Supplementary Figure 4 Neutralizing activities of the purified monoclonal**
1010 **antibodies as measured using a pseudo virus neutralization assay**
1011 Neutralization activities of 6 isolated mAbs against 5 major SARS-CoV-2
1012 variants (D614G, B.1.617, B.1.1.7, B.1.351 and P.1) were examined using a
1013 pseudo virus neutralization assay. Purified mouse IgG was used as negative
1014 control and serum of an RBD protein vaccinated goat was use as positive
1015 control. Purified mAbs and control mouse IgG were tested at a concentration
1016 of 1μg/ml; positive goat serum was diluted at 1:90.

Table 1 Demographics of SARS-CoV-2 unexposed healthy individuals

Sample Collection Time	**2016**	**2020**
Number of individuals	78	95
Gender (males, females)	17, 61	78, 17
Age, years (mean ± SD)	35.88 ± 8.40	30.40 ± 7.55

Table 2 Potential cross-reactive antigens identified in mouse fecal bacteria

NCBI Accession #	Protein name	Bacterium	Score	Proteins	Unique Peptides	Peptides	PSMs	Area	MW [kDa]
Q8A470	DNA-directed RNA polymerase subunit beta' OS	Bacteroides thetaiotaomicron	168.01	550	2	6	6	3.536E7	158.3
Q5L897	DNA-directed RNA polymerase subunit beta OS	Bacteroides fragilis	92.15	8	2	3	3	8.315E6	142.4
Q8A1G1	TonB-dependent receptor SusC OS	Bacteroides thetaiotaomicron	145.74	1	1	2	3	1.385E8	111.1
Q46509	Aldehyde oxidoreductase OS	Desulfovibrio gigas	75.93	1	1	1	1	8.506E6	97.0
P22983	Pyruvate, phosphate dikinase OS	Clostridium symbiosum	109.83	7	2	2	2	3.721E7	96.6
P0A9Q8	Aldehyde-alcohol dehydrogenase OS	Escherichia coli	34.29	1	1	1	1	1.085E7	96.1
Q826F6	Chaperone protein dnaK2 OS	Streptomyces avermitilis	20.37	1	1	1	1	1.059E7	67.5
Q892R0	Chaperone protein DnaK OS	Clostridium tetani	51.52	5	1	1	1	1.780E7	66.4
B1I9W8	L-fucose isomerase OS	Streptococcus pneumoniae	57.33	12	2	2	2	2.792E7	65.7
P95334	Chaperone protein DnaK OS	Myxococcus xanthus	39.23	4	1	1	1	1.792E7	65.3
Q1MPZ9	Formate--tetrahydrofolate ligase OS	Lawsonia intracellularis	62.49	1	2	2	2	2.068E7	64.2
P26929	Urease subunit alpha OS	Lactobacillus fermentum	134.06	4	4	4	5	1.526E7	61.8
Q9I165	Periplasmic trehalase OS	Pseudomonas aeruginosa	77.05	1	1	1	1	4.240E8	61.1
Q9EZ02	Pyrophosphate--fructose 6-phosphate 1-phosphotransferase OS	Spirochaeta thermophila	68.21	2	1	1	1	1.192E7	61.0
P14407	Fumarate hydratase class I, anaerobic OS	Escherichia coli	30.79	1	1	1	1	6.580E6	60.1
Q189R2	Formate--tetrahydrofolate ligase OS	Clostridioides difficile	123.52	22	3	4	4	2.621E7	59.9
P22252	Flagellin B OS	Campylobacter jejuni subsp. jejuni serotype O:6	75.42	1	1	1	1	1.478E7	59.7
Q1WTW0	Formate--tetrahydrofolate ligase OS	Lactobacillus salivarius	83.43	23	1	1	1	5.229E7	59.4
A7HLZ4	Formate--tetrahydrofolate ligase OS	Fervidobacterium nodosum	93.49	29	1	2	2	8.036E6	59.2
C4ZBL1	Phosphoenolpyruvate carboxykinase (ATP) OS	Agathobacter rectalis	211.34	4	1	4	5	3.965E7	59.0
A6LFQ4	Phosphoenolpyruvate carboxykinase (ATP) OS	Parabacteroides distasonis	32.29	1	1	1	1	8.523E6	58.9
A3MZI4	Formate--tetrahydrofolate ligase OS	Actinobacillus pleuropneumoniae serotype 5b	78.45	78	1	2	2	1.320E8	58.9
Q2LPJ8	60 kDa chaperonin 1 OS	Syntrophus aciditrophicus	101.22	15	1	2	3	2.217E7	58.6
Q3ALZ3	60 kDa chaperonin 1 OS	Synechococcus sp.	113.34	23	1	2	3	9.973E7	58.5
B8J123	60 kDa chaperonin OS	Desulfovibrio desulfuricans	232.59	27	2	5	6	9.659E7	58.4
Q72AL6	60 kDa chaperonin OS	Desulfovibrio vulgaris	126.02	9	1	3	3	2.402E7	58.4

A0Q2T1	60 kDa chaperonin OS	Clostridium novyi	78.78	16	1	2	2	2.292E7	58.1
B0SCC0	60 kDa chaperonin OS	Leptospira biflexa serovar Patoc	49.58	1	1	1	1	1.882E8	58.1
B2TIX0	60 kDa chaperonin OS	Clostridium botulinum	126.34	35	1	3	3	4.903E7	57.9
A7GZ43	60 kDa chaperonin OS	Campylobacter curvus	105.05	41	2	3	3	5.327E7	57.9
Q67KB8	60 kDa chaperonin OS	Symbiobacterium thermophilum	107.47	126	1	3	3	9.411E7	57.9
B5YDR9	60 kDa chaperonin OS	Dictyoglomus thermophilum	59.61	2	1	1	1	3.345E7	57.9
Q3ADX3	60 kDa chaperonin OS	Carboxydothermus hydrogenoformans	104.07	25	1	2	3	7.715E7	57.6
O50305	60 kDa chaperonin OS	Bacillus halodurans	250.57	55	3	4	6	3.021E8	57.4
P26821	60 kDa chaperonin OS	Clostridium perfringens	131.27	36	1	3	3	4.609E7	57.3
P37282	60 kDa chaperonin OS	Lactococcus lactis subsp. lactis	102.15	18	1	2	2	4.298E7	57.2
C4ZD46	60 kDa chaperonin OS	Agathobacter rectalis	372.10	52	3	5	8	3.352E8	57.1
A6L8C4	Glucose-6-phosphate isomerase OS	Parabacteroides distasonis	88.85	55	1	2	3	5.392E8	48.7
P43793	NADP-specific glutamate dehydrogenase OS	Haemophilus influenzae	96.32	2	1	2	2	7.357E7	48.6
P00370	NADP-specific glutamate dehydrogenase OS	Escherichia coli	161.88	1	1	3	4	3.280E8	48.6
P15111	NADP-specific glutamate dehydrogenase OS	Salmonella typhimurium	185.28	4	1	4	5	1.755E8	48.5
P94316	NAD-specific glutamate dehydrogenase OS	Bacteroides fragilis	192.93	3	3	5	6	1.730E8	48.4
Q1WSY0	Enolase OS	Lactobacillus salivarius	79.89	8	3	3	3	1.340E7	48.0
B7MD95	Trigger factor OS	Escherichia coli O45:K1	78.80	10	2	2	2	4.344E7	47.8
B2GAM0	Enolase OS	Lactobacillus fermentum	79.94	8	2	2	2	1.858E7	47.8
Q6MEY2	Enolase OS	Protochlamydia amoebophila	112.67	1	1	1	1	7.493E7	47.5
Q0SNH5	Enolase OS	Borrelia afzelii	71.86	2	1	1	1	4.654E7	47.4
G3KIM4	Lactoyl-CoA dehydratase subunit alpha (Fragment) OS	Anaerotignum propionicum	63.03	1	1	1	1	1.136E7	47.4
Q9I3S1	Biofilm dispersion protein BdlA OS	Pseudomonas aeruginosa	59.07	22	1	1	1	3.282E7	46.9
B8J4A8	Sulfate adenylyltransferase OS	Desulfovibrio desulfuricans	210.65	1	5	5	5	2.656E7	46.9
A6L3M9	Enolase OS	Bacteroides vulgatus	48.62	2	1	2	2	3.004E7	46.7
Q043Z5	Enolase 1 OS	Lactobacillus gasseri	289.72	4	5	5	5	4.189E7	46.6
Q1ISS7	Enolase OS	Koribacter versatilis	71.57	4	1	1	1	7.884E6	46.5
B8DTI9	Enolase OS	Bifidobacterium animalis subsp. lactis	115.71	4	2	2	2	1.684E7	46.4
A7GUR7	Enolase OS	Bacillus cytotoxicus	122.68	17	2	2	3	2.771E7	46.4
Q2LR33	Enolase OS	Syntrophus aciditrophicus	31.73	1	1	1	1	4.823E7	46.2
Q89Z05	Enolase OS	Bacteroides thetaiotaomicron	140.16	2	2	4	5	3.005E7	46.1

B2RLL7	Enolase OS	Porphyromonas gingivalis	123.91	2	2	3	3	2.636E7	45.8
Q6MPQ2	Enolase OS	Bdellovibrio bacteriovorus	66.50	1	1	1	1	8.918E6	45.7
B0K742	Serine hydroxymethyltransferase OS	Thermoanaerobacter pseudethanolicus	85.67	92	2	2	2	2.140E7	45.6
Q5LFT7	Peptidase T OS	Bacteroides fragilis	89.56	5	1	1	1	7.323E7	45.5
A9NEA9	Serine hydroxymethyltransferase OS	Acholeplasma laidlawii	44.82	3	1	1	1	5.875E6	45.3
Q7MU77	Phosphoglycerate kinase OS	Porphyromonas gingivalis	54.70	1	1	1	1	2.106E7	45.0
Q8A753	Phosphoglycerate kinase OS	Bacteroides thetaiotaomicron	25.30	1	1	1	1	1.320E7	45.0
A5VHR0	Phosphopentomutase OS	Lactobacillus reuteri	159.08	1	4	4	4	4.364E7	44.0
A8EWM4	Phosphoglycerate kinase OS	Arcobacter butzleri	45.52	1	1	1	1	2.577E7	43.8
Q042T5	Elongation factor Tu OS	Lactobacillus gasseri	400.92	36	3	10	13	1.184E8	43.7
Q74JU6	Elongation factor Tu OS	Lactobacillus johnsonii	348.06	36	2	9	12	1.184E8	43.6
Q5L890	Elongation factor Tu OS	Bacteroides fragilis	185.88	297	3	5	7	1.454E8	43.6
Q8R603	Elongation factor Tu OS	Fusobacterium nucleatum subsp. nucleatum	150.02	302	1	3	3	6.170E7	43.4
B8J1A0	Elongation factor Tu OS	Desulfovibrio desulfuricans	139.50	64	2	3	3	4.287E7	43.4
A5VJ92	Elongation factor Tu OS	Lactobacillus reuteri	413.79	60	8	11	13	3.595E8	43.4

Note: *Score:* The Mascot score. Proteins: The total number of proteins contained in the protein group. *Unique Peptides:* The total number of peptides unique to the protein group. *Peptides:* The total number of peptides identified from all included searches for the master protein of the protein group. *PSMs:* The total number of peptide-spectrum matches identified from all included searches for the master protein of the protein group. *Area:* The chromatographic peak area was used to characterize the quantitative abundance of protein.*MW(kDa):* The theoretical molecular weight of the protein.

Table 3 Potential cross-reactive antigens identified in human fecal bacteria

NCBI Accession #	Protein Name	Bacterium	Score	Proteins	Unique Peptides	Peptides	PSMs	Area	MW [kDa]
A6LFK9	Polyribonucleotide nucleotidyltransferase OS	Parabacteroides distasonis	62.68	3	1	2	2	2.332E7	82.0
P75764	Uncharacterized protein YbhJ OS	Escherichia coli	27.35	1	1	1	1	3.198E7	81.5
Q8A4N6	Polyribonucleotide nucleotidyltransferase OS	Bacteroides thetaiotaomicron	226.97	23	6	7	7	6.947E7	78.3
B3DT30	Elongation factor G OS	Bifidobacterium longum	236.98	64	8	10	11	4.407E7	78.1
E1WNR6	Chaperone protein htpG OS	Bacteroides fragilis	131.15	4	1	2	2	4.992E7	77.9
O31673	ATP-dependent Clp protease ATP-binding subunit ClpE OS	Bacillus subtilis	120.19	73	1	3	4	8.448E7	77.9
A7ZSL5	Elongation factor G OS	Escherichia coli O139:H28	133.81	92	2	3	4	2.888E7	77.5
Q5L8A7	Elongation factor G OS	Bacteroides fragilis	431.35	187	3	10	14	1.391E8	77.5
A6KYJ7	Elongation factor G OS	Bacteroides vulgatus	361.61	186	2	9	12	1.342E8	77.4
A6LFP0	Methionine--tRNA ligase OS	Parabacteroides distasonis	65.86	8	1	1	1	4.772E7	77.4
P39396	Pyruvate/proton symporter BtsT OS	Escherichia coli	92.20	1	2	2	2	1.446E7	77.3
A9KNK6	Polyribonucleotide nucleotidyltransferase OS	Lachnoclostridium phytofermentans	68.87	1	1	1	1	1.075E8	76.9
Q5L6S5	Elongation factor G OS	Chlamydia abortus	76.05	58	1	2	3	1.622E7	76.8
Q67JU0	Elongation factor G OS	Symbiobacterium thermophilum	147.39	86	1	4	5	2.631E8	76.8
Q5U8S9	Elongation factor G OS	Staphylococcus intermedius	98.88	65	1	3	4	5.177E7	76.7
Q8A294	Putative K(+)-stimulated pyrophosphate-energized sodium pump OS	Bacteroides thetaiotaomicron	150.65	20	1	3	3	1.276E8	76.5
B9DYA6	Elongation factor G OS	Clostridium kluyveri	136.00	67	1	3	4	8.826E7	76.4
Q5WLR5	Elongation factor G OS	Bacillus clausii	134.85	81	1	4	5	1.119E8	76.4
Q97I51	Translation initiation factor IF-2 OS	Clostridium acetobutylicum	70.53	4	1	1	1	2.330E8	76.3
A0PXU3	Elongation factor G OS	Clostridium novyi	114.21	67	1	3	4	8.905E7	76.1
Q18CF4	Elongation factor G OS	Clostridioides difficile	174.49	59	1	4	5	1.288E8	75.8
Q8AB53	Putative glucosamine-6-phosphate deaminase-like protein BT_0258 OS	Bacteroides thetaiotaomicron	54.10	1	1	1	1	4.629E6	75.2
Q8XJ01	Penicillin-binding protein 1A OS	Clostridium perfringens	39.66	3	1	1	1	1.599E7	75.1
A6L7J7	Threonine--tRNA ligase OS	Bacteroides vulgatus	91.33	9	2	2	2	2.440E7	74.2
P30539	1,4-alpha-glucan branching enzyme GlgB OS	Butyrivibrio fibrisolvens	30.12	1	1	1	1	2.287E7	73.8
B2TIT5	Threonine--tRNA ligase OS	Clostridium botulinum	53.61	2	1	1	1	1.877E7	73.8

P56116	Chaperone protein HtpG OS	Helicobacter pylori	38.02	12	1	1	1	1.269E7	71.2
P0A9P7	ATP-dependent RNA helicase DeaD OS	Escherichia coli O6:H1	85.88	5	3	3	4	1.336E8	70.5
P19410	3-oxocholoyl-CoA 4-desaturase OS	Clostridium scindens	63.87	1	1	1	1	5.844E7	70.2
Q8RHJ2	Putative K(+)-stimulated pyrophosphate-energized sodium pump OS	Fusobacterium nucleatum subsp. nucleatum	140.41	23	1	3	3	1.320E8	68.9
A5CX56	Chaperone protein DnaK OS	Vesicomyosocius okutanii subsp. Calyptogena okutanii	15.38	1	1	1	1	2.961E7	68.7
Q5LG30	Chaperone protein DnaK OS	Bacteroides fragilis	418.39	33	1	9	11	6.399E7	68.6
Q89YW6	Chaperone protein DnaK OS	Bacteroides thetaiotaomicron	372.24	35	1	7	9	6.399E7	68.3
A6L2X7	Chaperone protein DnaK OS	Bacteroides vulgatus	371.71	33	1	8	10	6.399E7	68.3
Q93GF1	Chaperone protein DnaK OS	Prevotella loescheii	185.46	1	4	4	4	3.512E7	68.0
A6LGR5	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin) OS	Parabacteroides distasonis	38.34	5	1	1	1	2.396E7	67.9
B8H444	ATP-dependent zinc metalloprotease FtsH OS	Caulobacter vibrioides	26.45	6	1	1	1	1.241E7	67.7
A9KIA6	Aspartate--tRNA(Asp/Asn) ligase OS	Lachnoclostridium phytofermentans	31.12	1	1	1	1	8.400E6	67.5
P0AG91	Protein translocase subunit SecD OS	Escherichia coli O157:H7	233.70	2	6	6	6	2.164E7	66.6
Q49Y22	Chaperone protein DnaK OS	Staphylococcus saprophyticus subsp. saprophyticus	56.64	206	1	2	2	1.171E7	66.5
A6LBU6	Aspartate--tRNA ligase OS	Parabacteroides distasonis	88.06	49	1	2	2	7.901E6	66.4
P21332	Oligo-1,6-glucosidase OS	Bacillus cereus	53.99	2	1	1	1	1.391E7	66.0
Q8A5W4	Lysine--tRNA ligase OS	Bacteroides thetaiotaomicron	154.84	2	2	2	2	1.304E7	65.9
Q8GBW6	Methylmalonyl-CoA carboxyltransferase 12S subunit OS	Propionibacterium freudenreichii subsp. shermanii	98.01	1	1	1	2	1.386E8	65.9
Q67S54	Chaperone protein DnaK OS	Symbiobacterium thermophilum	96.29	225	1	3	3	4.528E7	65.7
Q9RQ13	L-fucose isomerase OS	Bacteroides thetaiotaomicron	181.20	7	1	5	7	1.038E8	65.7
A6L048	L-fucose isomerase OS	Bacteroides vulgatus	162.35	5	2	6	7	3.672E7	65.6
Q56403	V-type ATP synthase alpha chain OS	Thermus thermophilus	132.50	30	1	1	2	1.117E7	63.6
Q8G7I6	Glucose-6-phosphate isomerase OS	Bifidobacterium longum	531.07	4	14	14	15	3.155E8	63.0
Q4JX51	Glucose-6-phosphate isomerase OS	Corynebacterium jeikeium	30.85	1	1	1	1	3.306E7	62.1
Q8FZC4	2-isopropylmalate synthase OS	Brucella suis biovar 1	51.65	8	1	1	1	1.014E7	61.6
A6TGT4	Glucose-6-phosphate isomerase OS	Klebsiella pneumoniae subsp. pneumoniae	53.60	21	1	1	1	3.337E7	61.3

P0AG69	30S ribosomal protein S1 OS	Escherichia coli O157:H7	88.52	1	2	2	2	8.393E6	61.1
Q9EZ02	Pyrophosphate--fructose 6-phosphate 1-phosphotransferase OS	Spirochaeta thermophila	110.07	2	2	2	2	1.245E8	61.0
O31716	Uncharacterized ABC transporter ATP-binding protein YkpA OS	Bacillus subtilis	48.45	1	1	1	1	2.259E6	61.0
P59173	Probable 2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS	Leptospira interrogans serogroup Icterohaemorrhagiae serovar Lai	61.50	2	1	1	2	6.104E6	61.0
P23843	Periplasmic oligopeptide-binding protein OS	Escherichia coli	368.94	2	11	11	12	1.586E8	60.9
Q0SQ82	Formate--tetrahydrofolate ligase OS	Clostridium perfringens	167.47	26	1	3	4	6.063E7	60.4
P14407	Fumarate hydratase class I, anaerobic OS	Escherichia coli	66.46	4	2	2	2	4.922E7	60.1
Q3A9K2	Formate--tetrahydrofolate ligase OS	Carboxydothermus hydrogenoformans	110.68	55	1	3	3	3.356E8	60.1
Q251P8	Formate--tetrahydrofolate ligase 1 OS	Desulfitobacterium hafniense	72.85	55	1	2	2	5.742E8	60.0
C0QX38	Formate--tetrahydrofolate ligase OS	Brachyspira hyodysenteriae	145.04	9	1	2	3	2.562E7	60.0
Q189R2	Formate--tetrahydrofolate ligase OS	Clostridioides difficile	157.46	1	3	3	3	1.347E9	59.9
C4ZBG8	Formate--tetrahydrofolate ligase OS	Agathobacter rectalis	235.86	56	4	6	6	3.610E8	59.7
A8AQV7	Phosphoenolpyruvate carboxykinase (ATP) OS	Citrobacter koseri	104.30	44	4	5	5	1.300E8	59.6
Q24ZZ8	Formate--tetrahydrofolate ligase 2 OS	Desulfitobacterium hafniense	79.67	55	1	2	2	5.193E8	59.4
B2RHV8	Phosphoenolpyruvate carboxykinase (ATP) OS	Porphyromonas gingivalis	188.39	21	2	5	5	2.446E8	59.4
Q47VD0	Phosphoenolpyruvate carboxykinase (ATP) OS	Colwellia psychrerythraea	111.53	16	1	2	2	1.952E8	59.3
A1R7X2	Arginine--tRNA ligase OS	Paenarthrobacter aurescens	43.93	15	1	1	1	2.994E7	59.2
Q8A414	Phosphoenolpyruvate carboxykinase (ATP) OS	Bacteroides thetaiotaomicron	293.03	21	3	8	8	3.939E8	59.1
C4ZBL1	Phosphoenolpyruvate carboxykinase (ATP) OS	Agathobacter rectalis	501.52	17	5	12	15	4.167E8	59.0
Q5L7N5	Phosphoenolpyruvate carboxykinase (ATP) OS	Bacteroides fragilis	420.91	17	1	6	12	2.103E8	59.0
A3CL27	Formate--tetrahydrofolate ligase 1 OS	Streptococcus sanguinis	145.32	20	1	2	3	3.296E7	59.0
A6LFQ4	Phosphoenolpyruvate carboxykinase (ATP) OS	Parabacteroides distasonis	438.19	17	5	11	14	2.336E8	58.9
C4ZAW6	Dihydroxy-acid dehydratase OS	Agathobacter rectalis	193.87	6	5	6	6	2.765E7	58.9
B9E299	Dihydroxy-acid dehydratase OS	Clostridium kluyveri	85.82	2	1	2	3	7.244E7	58.8
O09460	Phosphoenolpyruvate carboxykinase (ATP) OS	Anaerobiospirillum succiniciproducens	214.42	21	2	5	5	3.702E8	58.6
B2TIR2	Dihydroxy-acid dehydratase OS	Clostridium botulinum	46.21	2	1	1	1	3.131E6	58.5
B3DTV2	ATP synthase subunit alpha OS	Bifidobacterium longum	383.31	393	9	11	12	5.746E7	58.4
B3DRY6	Bifunctional purine biosynthesis protein PurH OS	Bifidobacterium longum	204.27	2	3	3	3	1.243E8	58.4
A6LIG0	60 kDa chaperonin OS	Parabacteroides distasonis	325.96	2	1	7	10	4.117E8	58.3

A5N857	Ribonuclease Y OS	Clostridium kluyveri	85.68	95	1	3	3	4.874E7	58.3
Q8G3N6	Inosine-5'-monophosphate dehydrogenase OS	Bifidobacterium longum	685.57	20	15	16	19	2.005E8	58.2
Q8A6P8	60 kDa chaperonin OS	Bacteroides thetaiotaomicron	591.81	1	2	11	15	2.794E8	58.2
Q5LAF6	60 kDa chaperonin OS	Bacteroides fragilis	690.75	1	4	15	20	5.301E8	58.2
A0Q2T1	60 kDa chaperonin OS	Clostridium novyi	58.42	6	2	2	2	2.212E8	58.1
A6KXA0	60 kDa chaperonin OS	Bacteroides vulgatus	948.24	1	11	22	26	6.146E8	58.1

Note: *Score:* The Mascot score. Proteins: The total number of proteins contained in the protein group. *Unique Peptides:* The total number of peptides unique to the protein group. *Peptides:* The total number of peptides identified from all included searches for the master protein of the protein group. *PSMs:* The total number of peptide-spectrum matches identified from all included searches for the master protein of the protein group. *Area:* The chromatographic peak area was used to characterize the quantitative abundance of protein. *MW(kDa):* The theoretical molecular weight of the protein.

Figure 1 Pre-existing cross-reactive antibodies against SARS-CoV-2 S protein observed in both healthy human and naive SPF mice predominantly targeted S2 subunit

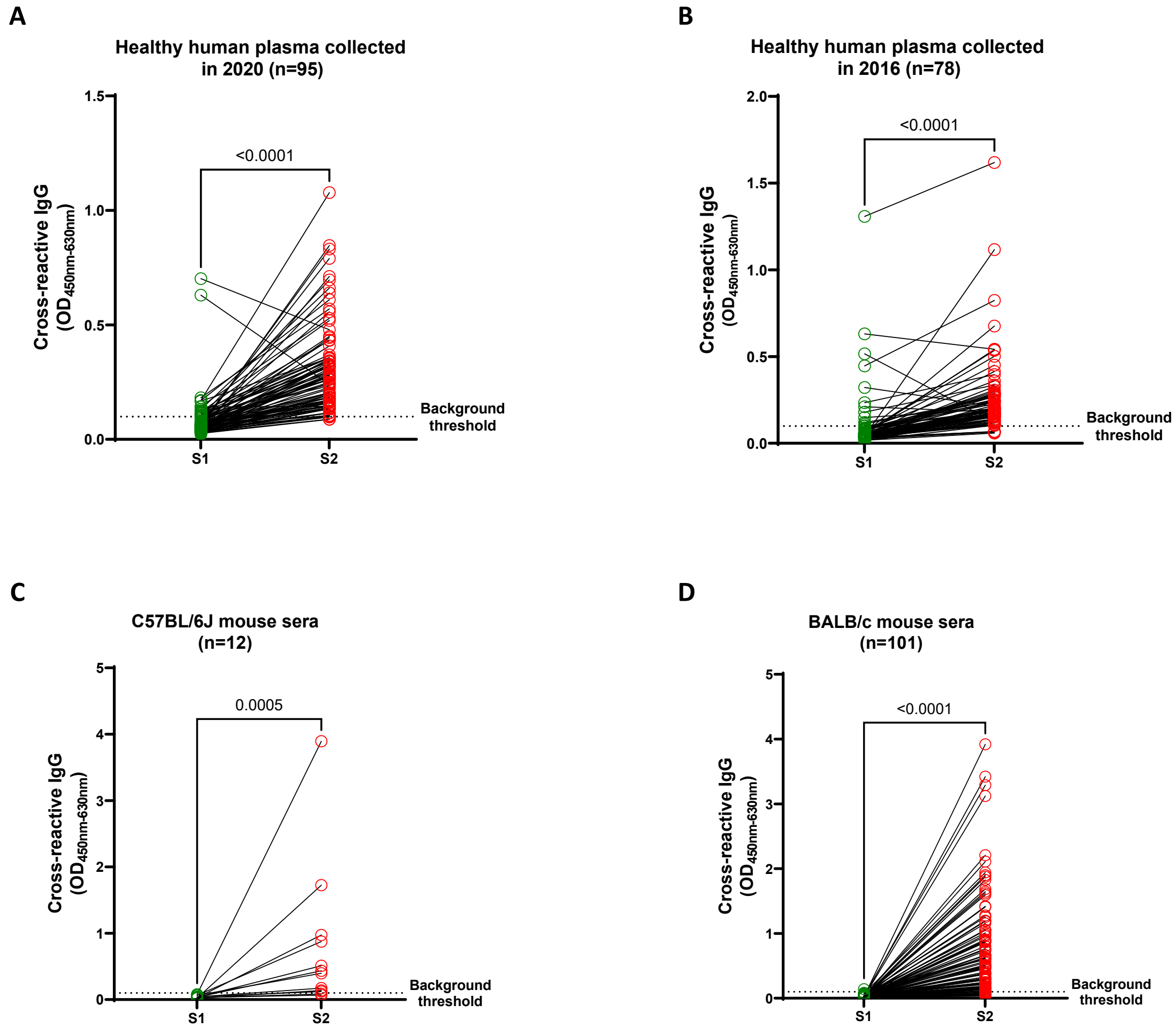


Figure 2 The pre-existing S protein cross-reactive antibodies in naïve SPF mice recognized S2 exclusively and a dominant linear antibody epitope was identified on the connector domain

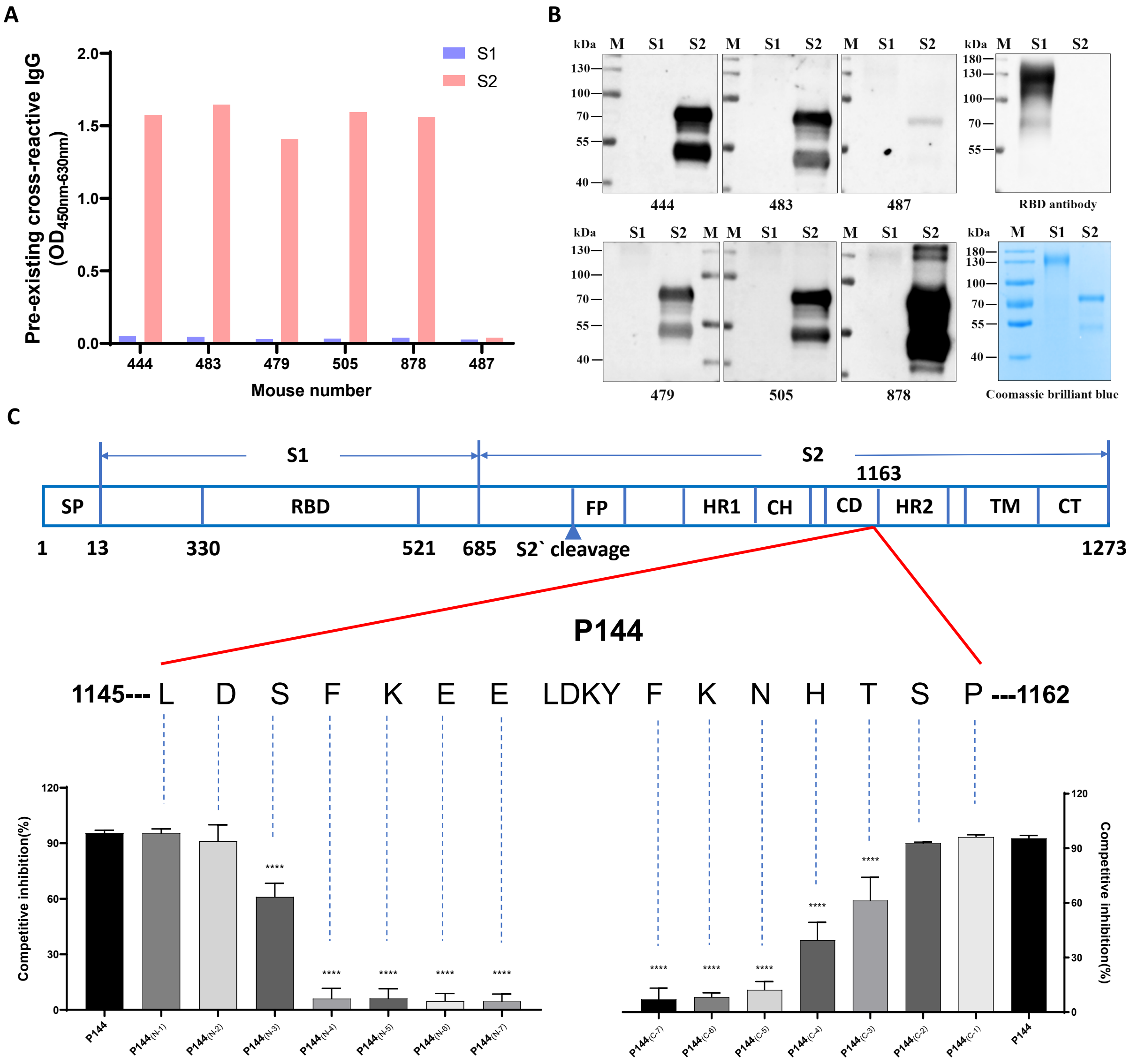


Figure 3 Cross-reactive antibodies recognizing epitope P144 widely existed in both healthy human and naïve SPF mice

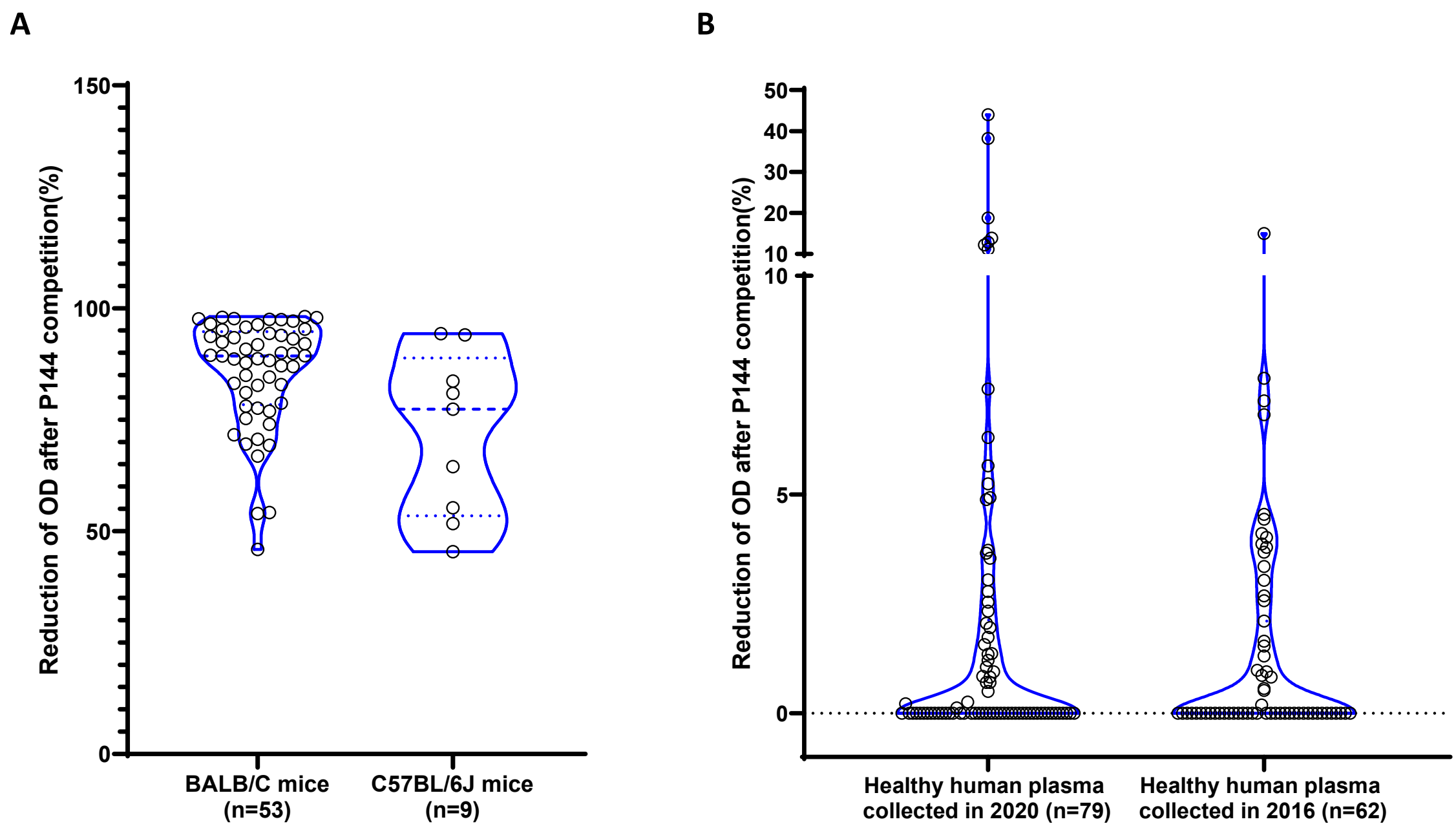
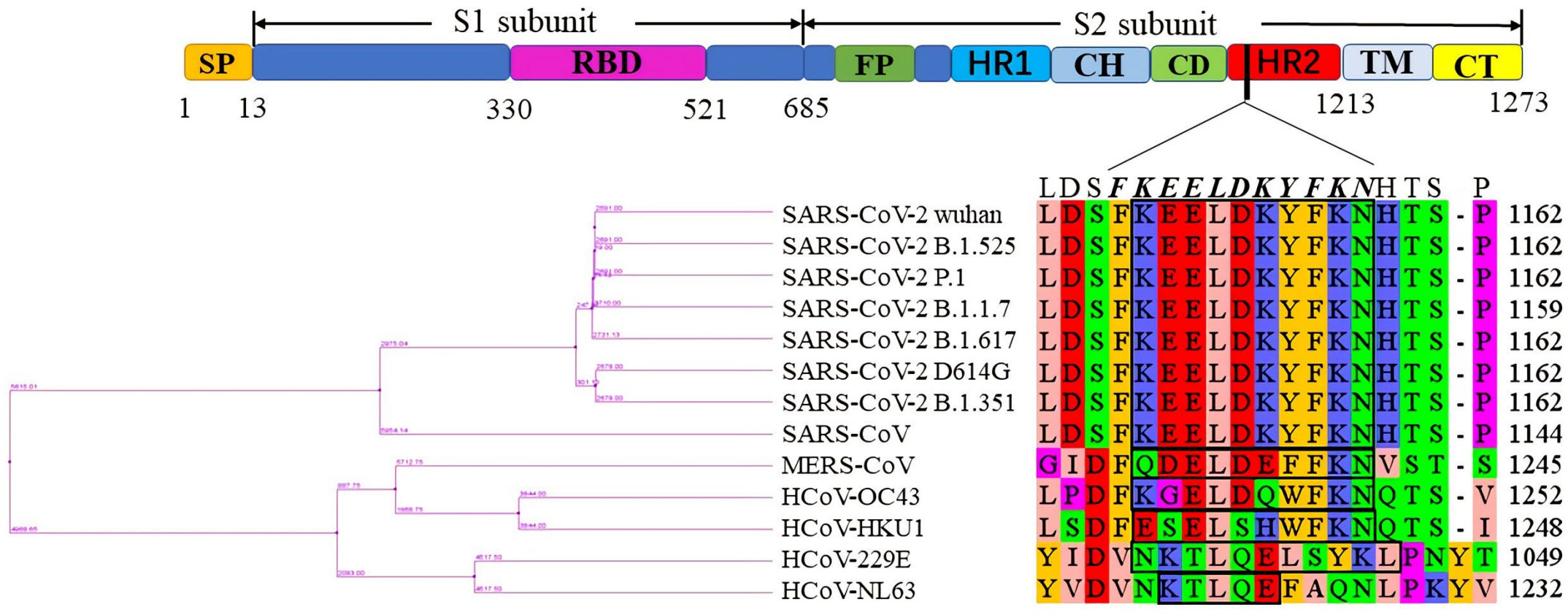
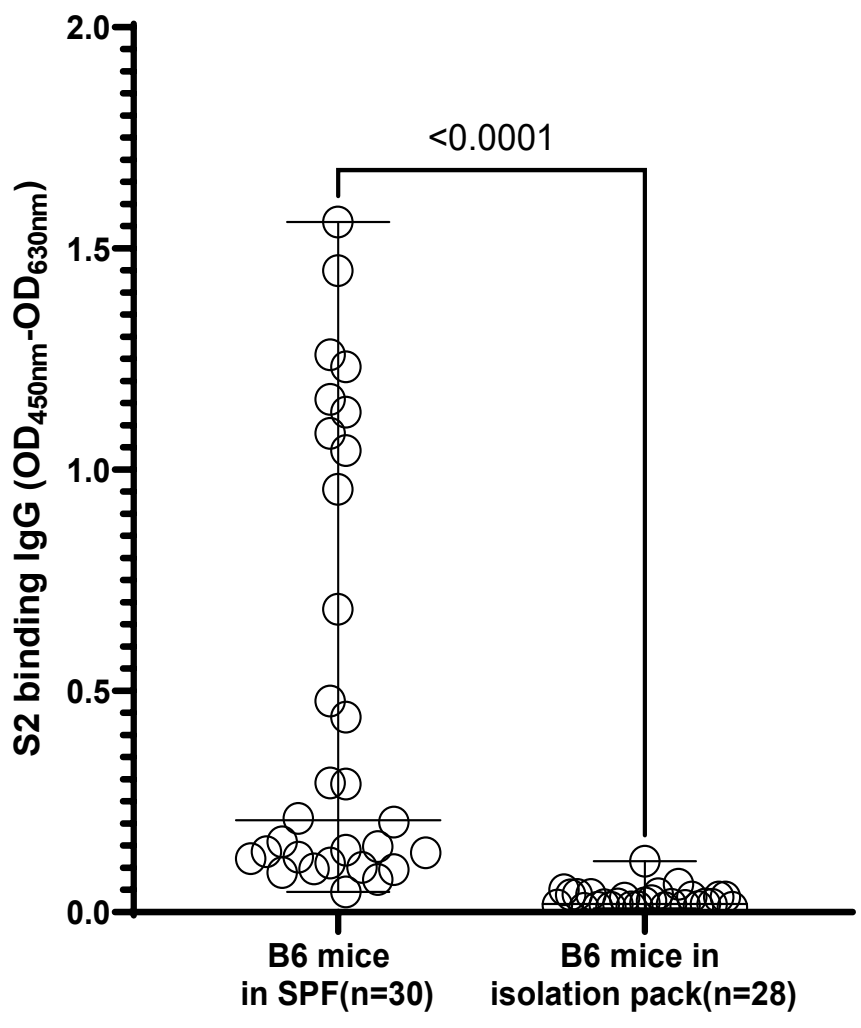


Figure 4 The generation of the pre-existing S2 cross-reactive antibodies might be associated with commensal gut microbiota

A



B



C

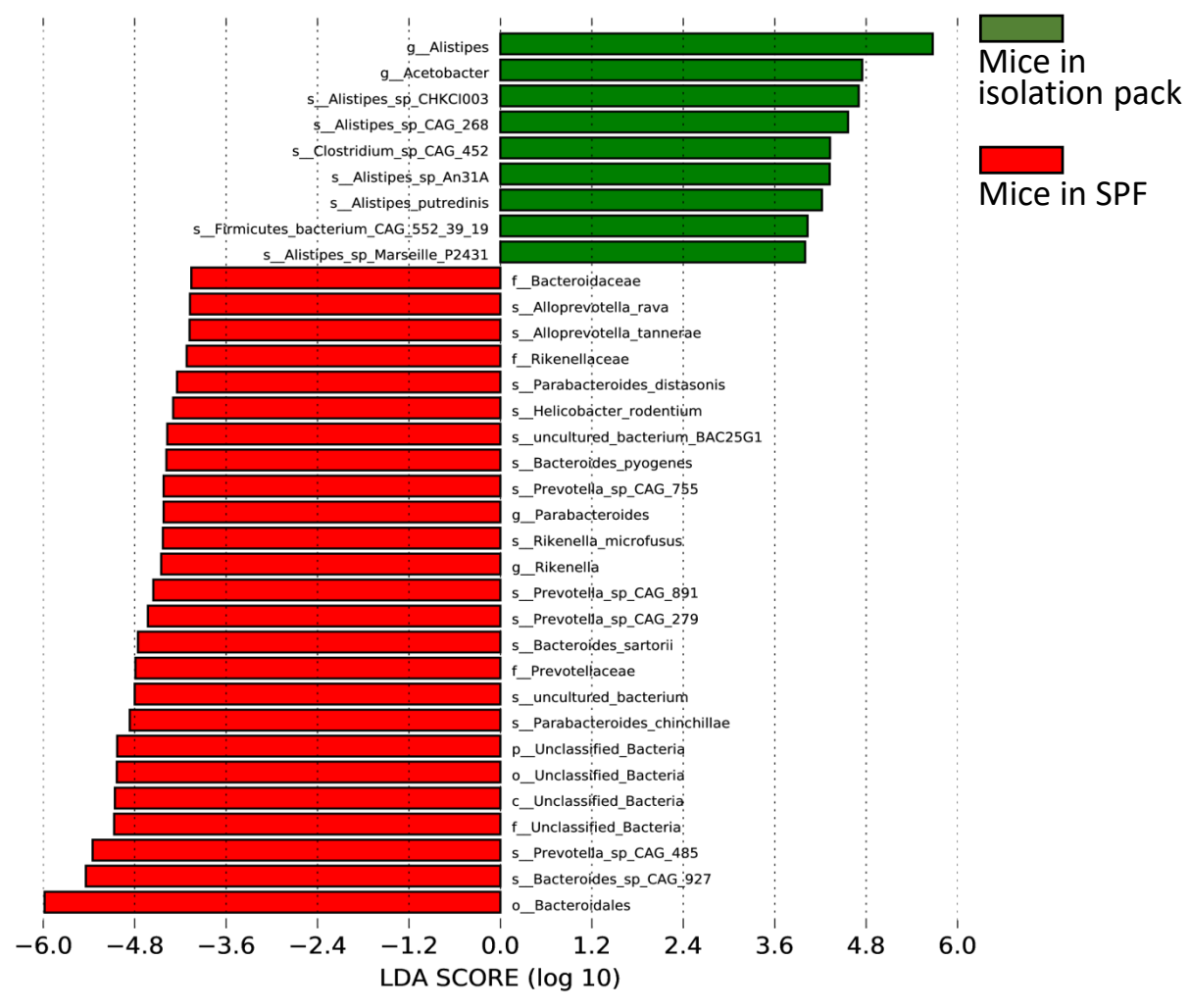


Figure 5 P144 specific monoclonal antibodies isolated from naïve SPF mice cross-reacted with commensal gut bacteria from both human and mouse

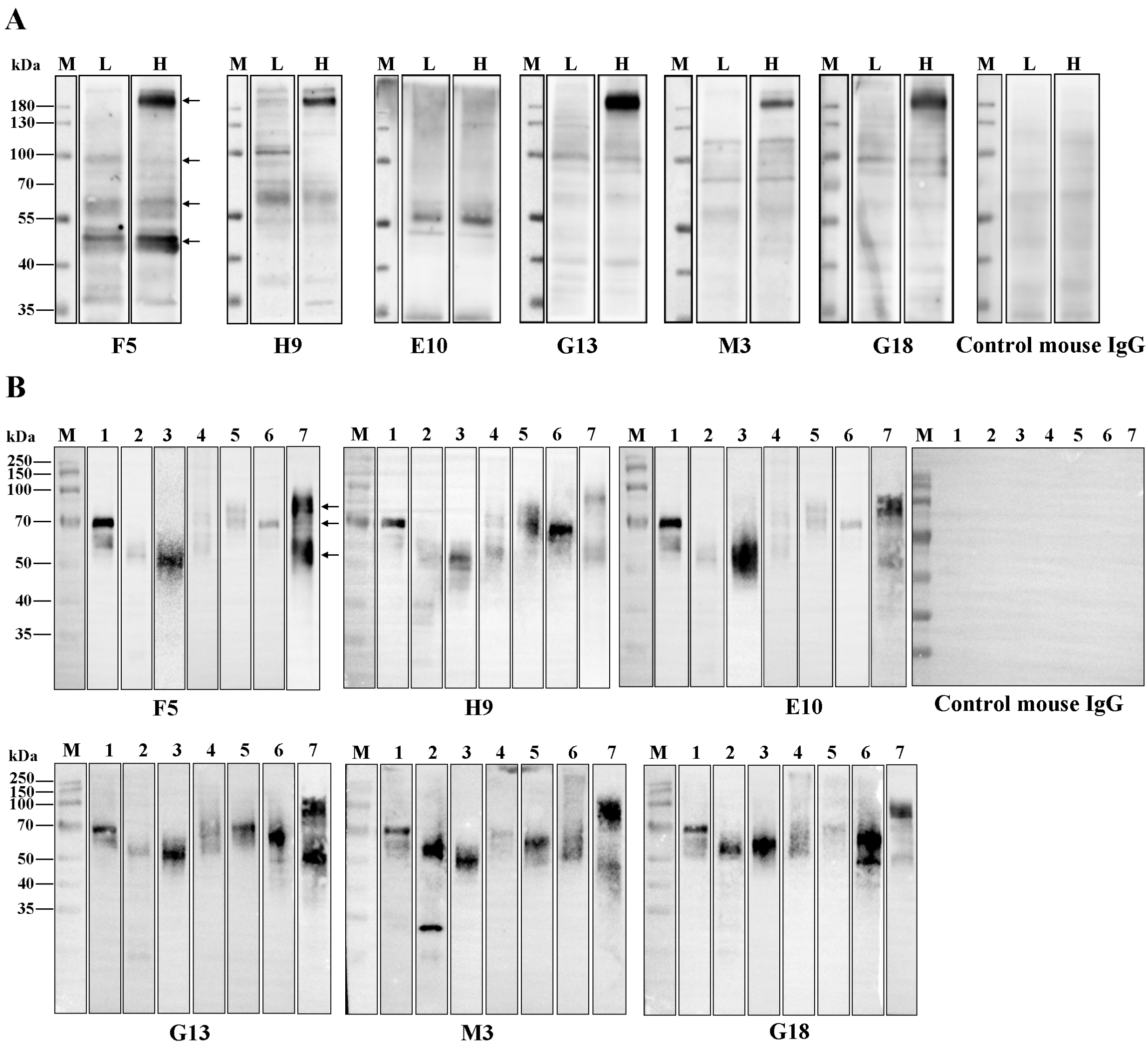


Figure 6 The impact of pre-existing antibody on the humoral immune responses elicited by a DNA vaccine encoding SARS-CoV-2 S protein

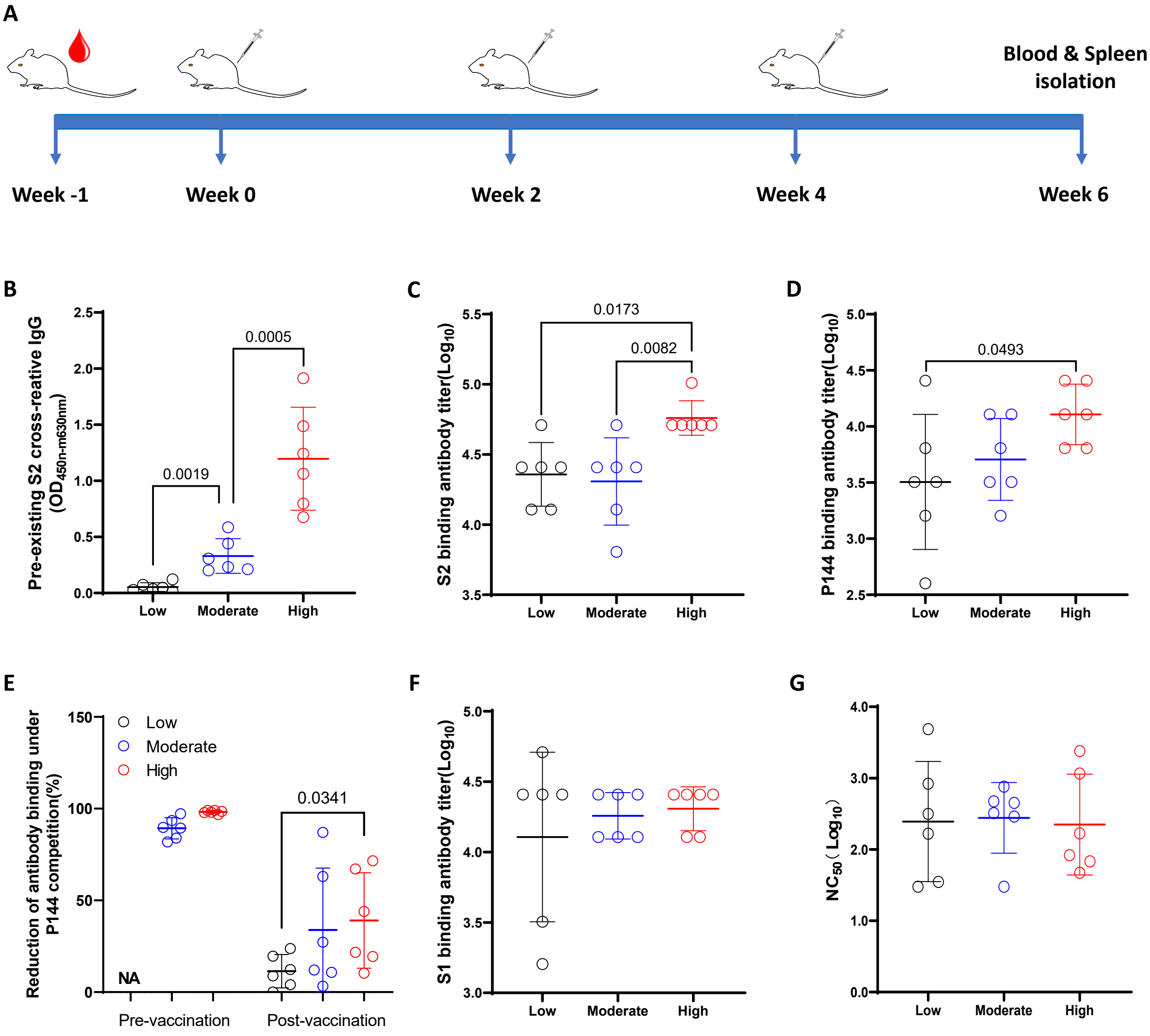


Figure 7 The impact of pre-existing antibody on the levels of specific IgG and IgA in BALF after vaccination

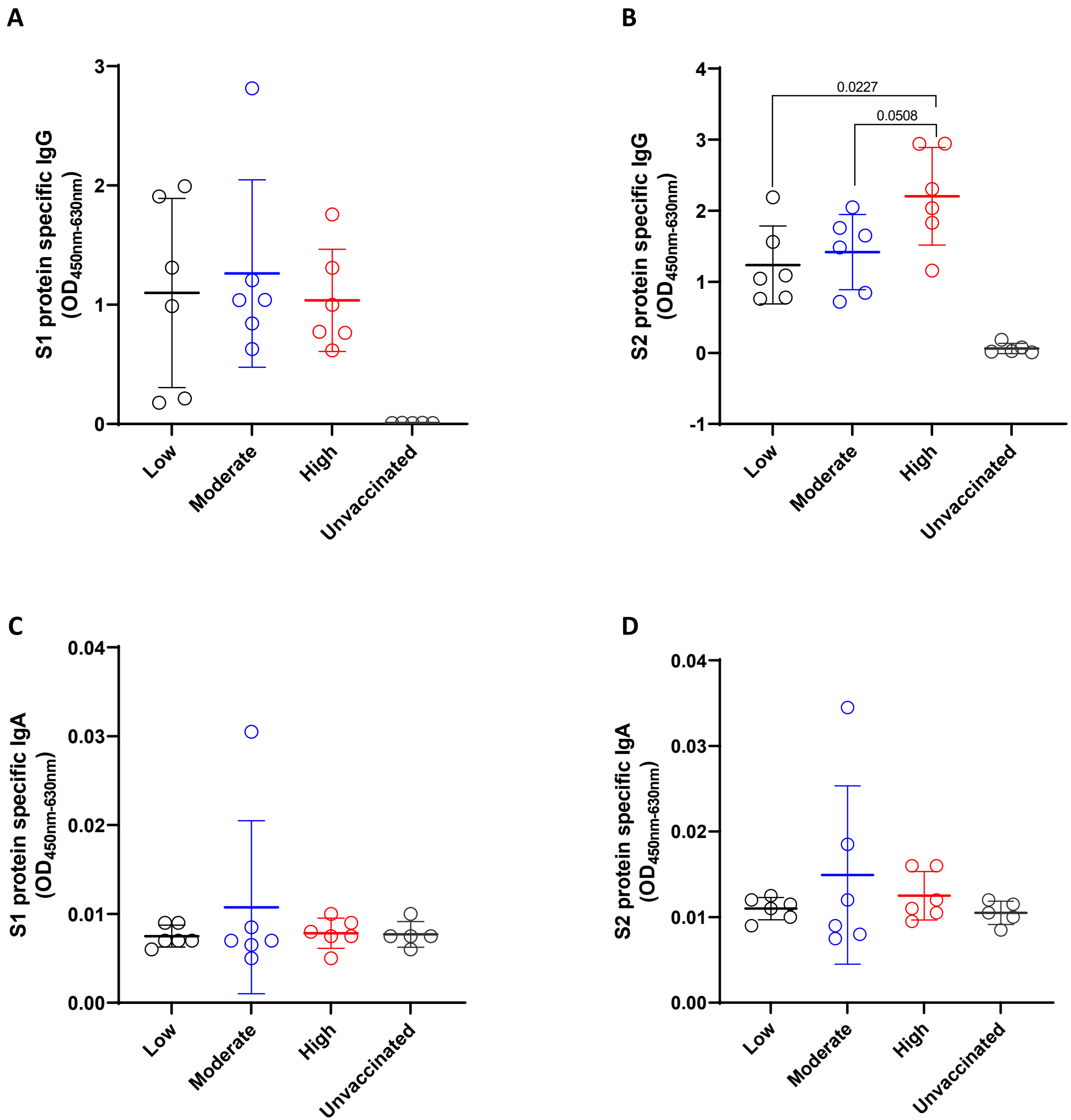


Figure 8 The impact of pre-existing antibodies on the minimum recognition of epitope P144 after vaccination

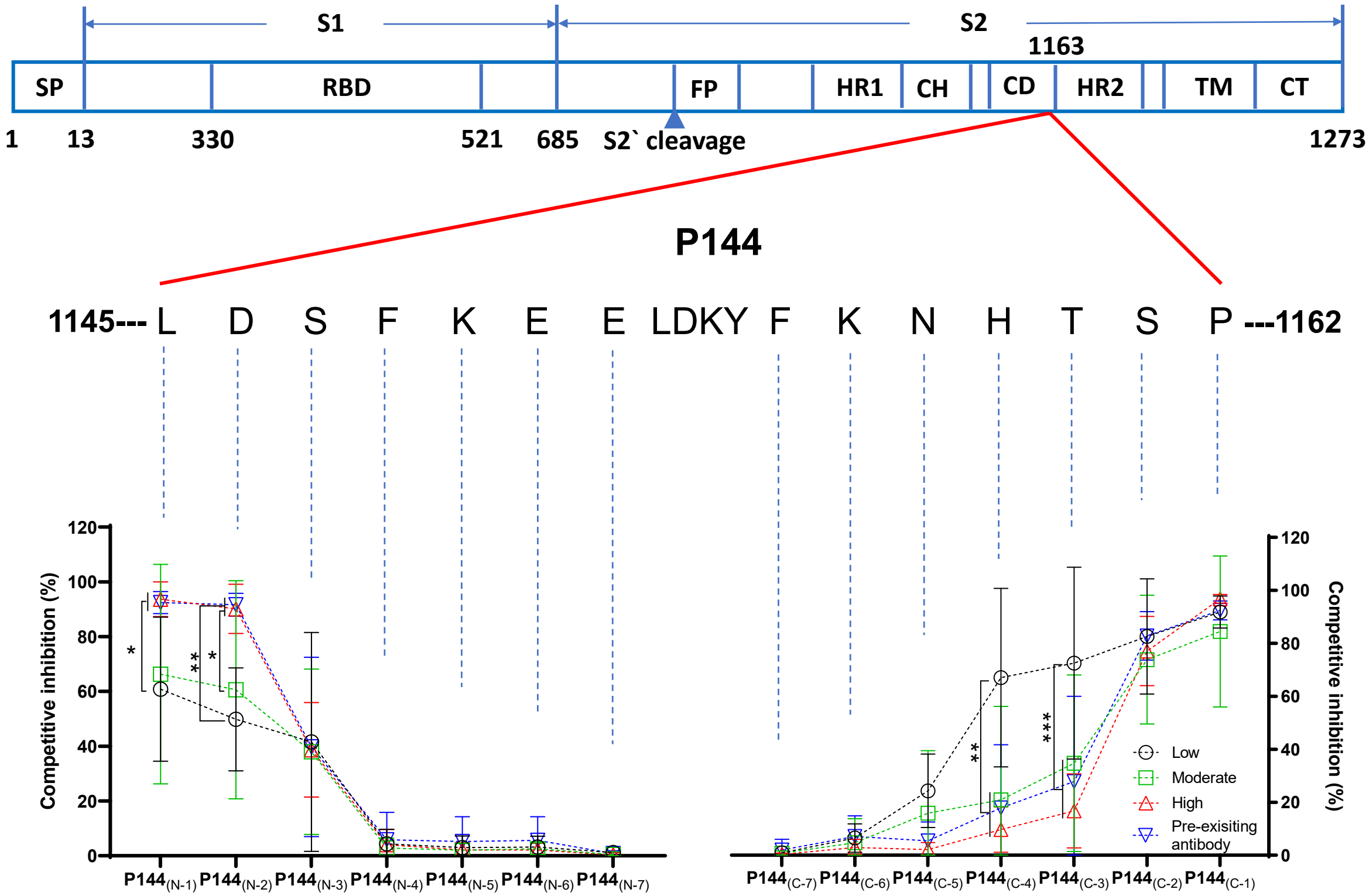
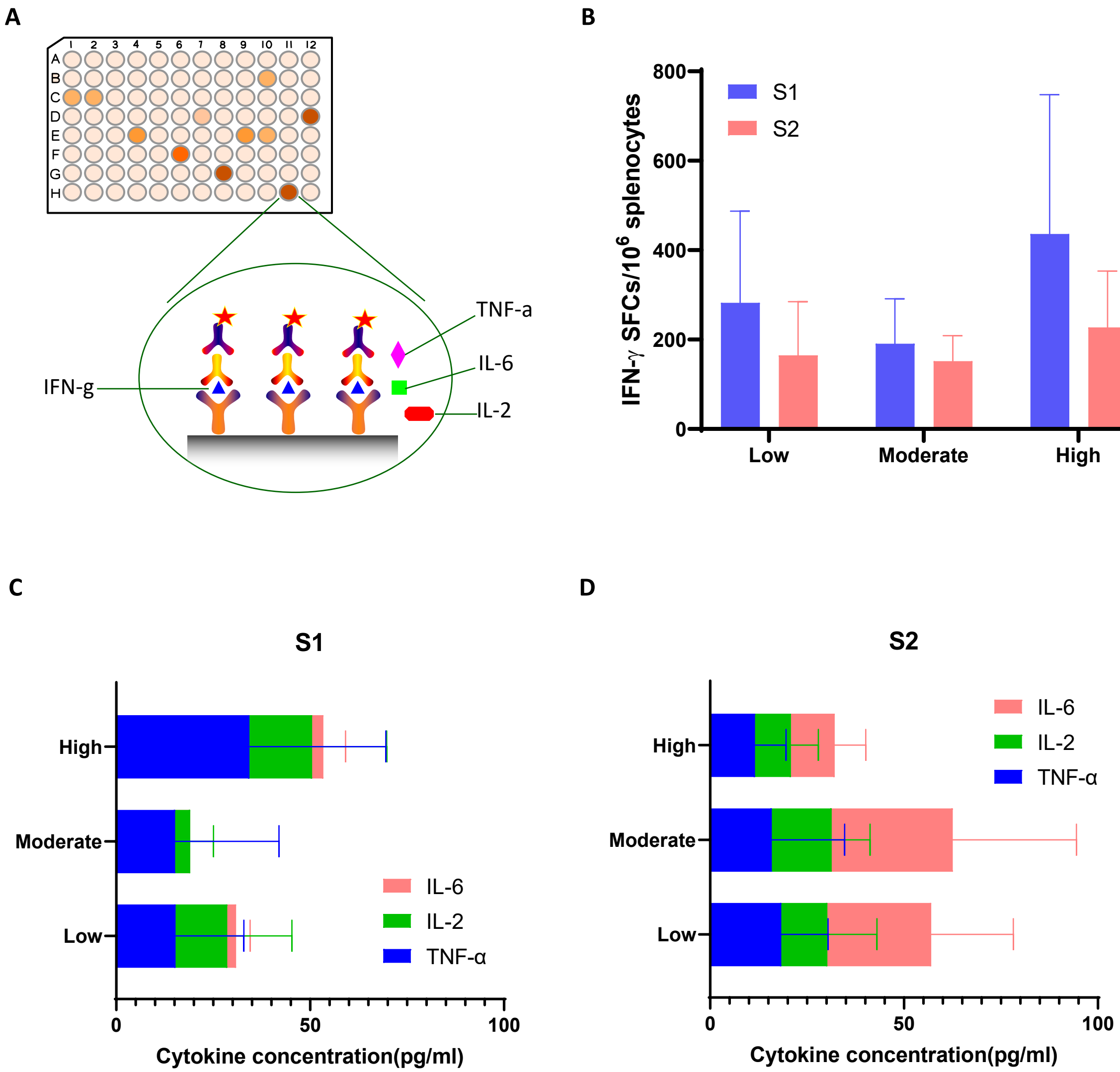



Figure 9 The impact of pre-existing antibody on the cellular immune responses induced by a DNA vaccine encoding SARS-CoV-2 S protein