Microscopy-based assay for semi-quantitative detection of SARS-CoV-2 specific antibodies in human sera

Authors: Constantin Pape[1,2,#], Roman Remme[1,#], Adrian Wolny[1,2], Sylvia Olberg[3], Steffen Wolf[1], Lorenzo Cerrone[1], Mirko Cortese[4], Severina Klaus[5], Bojana Lucic[6], Stephanie Ullrich[3], Maria Anders-Össwein[3], Stefanie Wolf[3], Berati Cerikan[4], Christopher J. Neufeldt[4], Markus Ganter[5], Paul Schnitzler[3], Uta Merle[7], Marina Lusic[6,9], Steeve Boulant[3,8], Megan Stanifer[4,8], Ralf Bartenschlager[4,9], Fred A. Hamprecht[1], Anna Kreshuk[2], Christian Tischer[2], Hans-Georg Kräusslich[3,9], Barbara Müller[3,*] and Vibor Laketa[3,9,*]

Affiliations:

[1]HCI/IWR, Heidelberg University, Heidelberg, Germany.

[2]European Molecular Biology Laboratory, Heidelberg, Germany.

[3]Department of Infectious Diseases, Virology, University Hospital Heidelberg, Heidelberg, Germany.

[4]Department of Infectious Diseases, Molecular Virology, University Hospital Heidelberg, Heidelberg, Germany.

[5]Department of Infectious Diseases, Parasitology, University Hospital Heidelberg, Heidelberg, Germany.

[6]Department of Infectious Diseases, Integrative Virology, University Hospital Heidelberg, Heidelberg, Germany.

[7]Department of Gastroenterology and Hepatology, University Hospital of Heidelberg, Heidelberg, Germany.

[8]Research Group "Cellular polarity and viral infection", German Cancer Research Center (DKFZ), Heidelberg, Germany

[9]German Center for Infection Research, Heidelberg, Germany.

*Corresponding authors: Vibor Laketa, Department of Infectious Diseases, Virology, University Hospital Heidelberg, Im Neuenheimer Feld 344, 69120 Heidelberg, Germany, Email: vibor.laketa@med.uni-heidelberg.de and Barbara Müller, Department of Infectious Diseases, Virology, University Hospital Heidelberg, Im Neuenheimer Feld 344, 69120 Heidelberg, Germany, Email: barbara.mueller@med.uni-heidelberg.de

32   #equal contribution

33

34   **Subtitle**: A semi-quantitative, high throughput, microscopy-based assay expands existing
35   approaches to measure SARS-CoV-2 specific antibody levels in human sera

Abstract

37   Emergence of the novel pathogenic coronavirus SARS-CoV-2 and its rapid pandemic
38   spread presents numerous questions and challenges that demand immediate attention. Among
39   these is the urgent need for a better understanding of humoral immune response against the
40   virus as a basis for developing public health strategies to control viral spread. For this, sensitive,
41   specific and quantitative serological assays are required. Here we describe the development of
42   a semi-quantitative high-content microscopy-based assay for detection of three major classes
43   (IgG, IgA and IgM) of SARS-CoV-2 specific antibodies in human samples. The possibility to
44   detect antibodies against the entire viral proteome together with a robust semi-automated image
45   analysis workflow resulted in specific, sensitive and unbiased assay which complements the
46   portfolio of SARS-CoV-2 serological assays. The procedure described here has been used for
47   clinical studies and provides a general framework for the application of quantitative high-
48   throughput microscopy to rapidly develop serological assays for emerging virus infections.
49

Keywords:

51   SARS-CoV-2, antibody, serological test, quantitative microscopy, immunofluorescence,
52   machine learning image analysis

53

1. Introduction

55   The recent emergence of the novel pathogenic coronavirus SARS-CoV-2 [1–3] and the
56   rapid pandemic spread of the virus has dramatic consequences in all affected countries. In the
57   absence of a protective vaccine or a causative antiviral therapy for COVID-19 patients, testing
58   for SARS-CoV-2 infection and tracking of transmission and outbreak events are of paramount
59   importance to control viral spread and avoid the overload of healthcare systems. The sequence
60   of the viral genome became publicly available only weeks after the initial reports on COVID-19

61   via the community online resource *virological.org* and allowed rapid development of reliable and

62   standardized quantitative RT-PCR (qPCR) based tests for direct virus detection in

63   nasopharyngeal swab specimens [4–6]. These tests are the key to identify acutely infected

64   individuals and monitor virus load as a basis for the implementation of quarantine measures and

65   treatment decisions.

66   In response to the initial wave of COVID-19 infection many countries implemented more

67   or less severe lockdown strategies, resulting in a gradual decrease in the rate of new infections

68   and deaths [7]. With gradual release of these lockdown strategies, monitoring and tracking of

69   SARS-CoV-2 specific antibody levels becomes highly important. Many critical aspects of the

70   humoral immune response against SARS-CoV-2 are currently not well understood [8]. In addition,

71   levels of infection in the general population in different areas remain largely unknown due to

72   proportion of undocumented cases arising from asymptomatic individuals [9,10] which had not

73   been subjected to RNA testing, or to limitations in testing capacity especially in areas of relatively

74   high prevalence. Public health control strategies aiming at regulating human mobility and social

75   behaviour in order to suppress the infection rate will have to take into account the proportion of

76   seropositive individuals in the general population, or in specific population groups [11]. Information

77   on the level of antiviral antibodies, as well as on the serological response against different viral

78   proteins, is also a key element of understanding the nature, development and durability of the

79   antiviral immune response. Therefore, specific, sensitive and reliable methods for the

80   quantitative detection of virus specific antibodies in human specimens are urgently needed from

81   the beginning of an emerging pandemic.

82   Compared to approaches for direct virus diagnostics by PCR, development of test

83   systems for detection of SARS-CoV-2 specific antibodies proved to be more challenging. In

84   particular, cross reactivity of antibodies against circulating common cold coronaviruses (strains

85   OC43, NL63, 229E and HKU1) are of concern in this respect as it was observed in case of

86   serological tests developed for closely related SARS-CoV and MERS-CoV [12]. Developments in

87   the past months yielded well validated, commercially available ELISA or

88   (electro)chemoluminescence-based kits for SARS-CoV-2 serological diagnostics. However,

89   initially marketed test kits underwent a very rapid development and approval process due to the

90   emergency of the situation, with low numbers of samples used for validation; consequently,

91   sensitivity and specificity of the test systems often failed to meet the practical requirements [13].

92   Furthermore, the disruption of supply chains and high demand for tests during pandemic

93   situations can lead to shortage of commercially available test kits and/or required reagents, as

94   witnessed in the early phases of the ongoing SARS-CoV-2 pandemic. Thus, complementary
95   strategies to test for antiviral antibodies that can be rapidly deployed in situations where
96   commercially available kits are either not yet developed or not available are an important addition
97   to the diagnostic toolkit.

98   Immunofluorescence (IF) using virus infected cells as a specimen is a classical
99   serological approach in virus diagnostics and has been applied to coronavirus infections,
100   including the closely related virus SARS-CoV [14–16]. The advantages of IF are (i) that it does not
101   depend on specific diagnostic reagent kits or instruments, (ii) that the specimen contains all viral
102   antigens expressed in the cellular context and (iii) that the method has the potential to provide
103   high information content (differentiation of staining patterns and intensities due to reactivity
104   against various viral proteins). A mayor disadvantage of the IF approach as it is typically used in
105   serological testing is its limited throughput capacity due to the involvement of manual microscopy
106   handling steps and sample evaluation based on visual inspection of micrographs. Furthermore,
107   visual classification is subjective and thus not well standardized and yields only binary results.
108   Here, we address those limitations, making use of advanced automated microscopy and image
109   analysis strategies developed for basic research. We present the establishment and validation
110   of a semi-quantitative, semi-automated workflow for SARS-CoV-2 specific antibody detection.
111   With its 96-well format, semi-automated microscopy and automated image analysis workflow it
112   combines advantages of IF with a reliable and objective semi-quantitative readout and high
113   throughput compatibility. The protocol described here was developed in response to the
114   emergence of SARS-CoV-2, but it represents a general approach that can be adapted for the
115   study of other viral infections and is suitable for rapid deployment to support diagnostics of
116   emerging viral infections in the future.

117 2. Results

118 2.1 Setup of the IF assay for SARS-CoV-2 antibody detection

119   We decided to use cells infected with SARS-CoV-2 as samples for our IF analyses,
120   since this setup provides the best chance for detection of antibodies targeted at the different
121   viral proteins expressed in the host cell context. African green monkey kidney epithelial cells
122   (VeroE6 cell line) have been used for infection with SARS-CoV-2, virus production and IF[3,17].
123   In preparation for our analyses we compared different cell lines for use in infection and IF

124   experiments, but all tested cell lines were found to be inferior to VeroE6 cells for our purposes
125   (see Materials and Methods and Fig. S1). All following experiments were thus carried out using
126   the VeroE6 cell line.

127   In order to allow for clear identification of positive reactivity in spite of a variable and
128   sometimes high nonspecific background from human sera, our strategy involves a direct
129   comparison of the IF signal from infected and non-infected cells in the same sample. Preferential
130   antibody binding to infected compared to non-infected cells indicates the presence of specific
131   SARS-CoV-2 antibodies in the examined serum. Under our conditions, infection rates of ~40-
132   80% of the cell population were achieved, allowing for a comparison of infected and non-infected
133   cells in the same well of the test plate. An antibody that detects dsRNA produced during viral
134   replication was used to distinguish infected from non-infected cells within the same field of view
135   (Fig. 1A).

136   In order to define the conditions for immunostaining using human serum, we selected a
137   small panel of negative and positive control sera. Four sera from healthy donors collected before
138   November 2019 were chosen as negative controls, and eight sera from PCR confirmed COVID-
139   19 inpatients collected at day 14 or later post symptom onset were employed as positive controls.
140   Sera from this test cohort were used for primary staining, and bound antibodies were detected
141   using fluorophore-coupled secondary antibodies against human IgG, IgA or IgM.

142   No difference between infected and non-infected cells in serum IgG antibody binding was
143   observed when sera collected before the onset of the SARS-CoV-2 pandemic were examined
144   (Fig. 1B, Fig. S2). In contrast, COVID-19 patient sera were clearly characterized by higher serum
145   IgG antibody binding to infected compared to non-infected cells (Fig. 1B). All eight COVID-19
146   patient serum samples yielded higher IgG binding to infected compared to non-infected cells as
147   assessed by visual inspection (Fig. S2). Similar results were obtained when an IgA or IgM
148   specific secondary antibody was used for detection (Fig. S3). In order to allow for the parallel
149   assessment of IgG and IgA or IgM antibodies, we established conditions for the parallel detection
150   of anti-IgG coupled to AlexaFluor488 and anti-IgA or anti-IgM coupled to DyLight650 or
151   AlexaFluor647 secondary antibodies, respectively, without signal bleedthrough. Using this
152   approach, it was possible to implement detection of SARS-CoV-2 specific IgG and IgA or IgM
153   antibodies in a single experimental setup (Fig. S4).

154   Titration experiments were performed with positive control sera to determine the optimal

155   range of serum concentration in the IF experiments. All eight positive control samples showed

156   visually detectable specific labelling of infected cells over the range of $1:10^2$ and $1:10^5$,

157   demonstrating robustness of the assay (Fig. S5). Serum concentrations of less than $1:10^5$ did

158   not yield detectable signals in all cases. We decided to employ a dilution of $1:10^2$ in the further

159   experiments

160   2.2 Image analysis

161   Our next aim was to establish a semi-automated analysis workflow for image acquisition and

162   analysis for a medium to high throughput setting. VeroE6 cells were seeded into 96-well plates

163   infected and immunostained using anti-dsRNA antibody and patient serum, followed by indirect

164   detection using a mixture of anti-IgG and anti-IgA/IgM secondary antibodies. Images were

165   acquired using an automated widefield microscope (see Materials and Methods section for more

166   detail).

167   To obtain a measure for specific antibody binding we performed automated

168   segmentation of cells and classified them into infected and non-infected cells based on the

169   dsRNA staining. We then measured fluorescence intensities in the serum channel per cell as a

170   proxy for the amount of bound antibodies for both infected and non-infected cells and calculated

171   the ratio between these values for infected and non-infected cells in a given specimen. To enable

172   training of a machine learning approach for cell segmentation and to directly evaluate infected

173   cell classification, we manually labelled cells and annotated them as infected/non-infected in 10

174   images chosen from 5 positive and 5 control specimens. Fig. 2 presents a graphical overview of

175   all analysis steps; the full description of every step can be found in Materials and Methods.

176   Briefly, our approach works as follows:

177   First, we manually discarded all images that contained obvious artefacts such as large

178   dust particles or dirt and out-of-focus images. Then, images were processed to correct for the

179   uneven illumination profile in each channel. Next, we segmented individual cells with a seeded

180   watershed algorithm [18], using nuclei segmented via StarDist [19] as seeds and boundary

181   predictions from a U-Net [20,21] as a heightmap. We evaluated this approach using leave-one-

182   image-out cross-validation on the manual annotations and measured an average precision[22] of

183   0.77 +- 0.08 (i.e., on average 77% of segmented cells are matched correctly to the

184   corresponding cell in the annotations). Combined with extensive automatic quality control which

185 discards outliers in the results, the segmentation was found to be of sufficient quality for our
186 analysis, especially since robust intensity measurements were used to reduce the effect of
187 remaining errors.

188   We then classified the segmented cells into infected and non-infected, by measuring
189 the 95th percentile intensities in the dsRNA channel and classifying cells as infected if this value
190 exceeded 4.8 times the noise level, determined by the mean absolute deviation. This factor and
191 the percentile were determined empirically using grid search on the manually annotated images
192 (see above). Using leave-one-out cross validation on the image level, we found that this
193 approach yields an average F1-score of 84.3%.

194   In order to make our final measurement more reliable, we then discarded whole wells,
195 images or individual segmented cells based on quality control criteria that were determined by
196 inspection of initial results. Those criteria include a minimal number of non-infected cells per
197 well; minimal and maximal number of cells per image; minimal cell intensities for images; and
198 minimal and maximal sizes of individual cells (see Materials and Methods for full details).

199 To score each sample, we computed the intensity ratio r :

$$r = \frac{m_I}{m_N}$$
Eq. 1

200 Here, m_I is the median serum intensity of infected cells and m_N the median serum intensity of
201 non-infected cells. For each cell, we compute its intensity by computing the mean pixel intensity
202 in the serum channel (excluding the nucleus area where we typically did not observe serum
203 binding) and then subtracting the background intensity, which is measured on two control wells
204 that did not contain any serum.

205   We used efficient implementations for all processing steps and deployed the analysis
206 software on a computer cluster in order to enhance the speed of imaging data processing. For
207 visual inspection, we have further developed an open-source software tool (PlateViewer) for
208 interactive visualization of high-throughput microscopy data [23]. PlateViewer was used in a final
209 quality control step to visually inspect positive hits. For example, PlateViewer inspection allowed
210 identifying a characteristic spotted pattern co-localizing with the dsRNA staining (Fig. S6) that
211 was sometimes observed in the IgA channel upon staining with negative control serum. In
212 contrast, sera from COVID-19 patients typically displayed cytosol, ER-like and plasma
213 membrane staining patterns in this channel (Fig. 1B, Fig. S3). The dsRNA co-localizing pattern

214 observed for sera from the negative control cohort is by definition non-specific for SARS-CoV-2,

215 but would be classified as a positive hit based on staining intensity alone. Using PlateViewer,

216 we performed a quality control on all IgA positive hits and removed those displaying the spotted

217 pattern colocalising with the dsRNA signal from further analysis.

2.3 Assay characterization and validation

219 With the immunofluorescence protocol and automated image analysis in place we

220 proceeded to test a larger number of control samples in a high throughput compatible manner

221 for assay validation. All samples were processed for IF as described above, and in parallel

222 analysed by a commercially available semi-quantitative SARS-CoV-2 ELISA approved for

223 diagnostic use (Euroimmun, Lübeck, Germany) for the presence of SARS-CoV-2 specific IgG

224 and IgA antibodies.

225 As outlined above, a main concern regarding serological assays for SARS-CoV-2

226 antibody detection is the occurrence of false positive results. A particular concern in this case is

227 cross-reactivity of antibodies that originated from infection with any of the four types of common

228 cold Corona viruses (ccCoV) circulating in the population. The highly immunogenic major

229 structural proteins of SARS-CoV-2 nucleocapsid (N) and spike (S) protein, have an overall

230 homology of ~30% [3] to their counterparts in ccCoV and subdomains of these proteins display a

231 higher degree homology; cross-reactivity with ccCoV has been discussed as the major reason

232 for false positive detection in serological tests for closely related SARS-CoV and MERS-CoV [12].

233 Also, acute infection with Epstein-Barr virus (EBV) or cytomegalovirus (CMV) may result in

234 unspecific reactivity of human sera [24,25]. We therefore selected a negative control panel

235 consisting of 218 sera collected before the fall of 2019, comprising samples from healthy donors

236 (n=105, cohort B), patients that tested positive for ccCoV several months before the blood

237 sample was taken (n=34, all four types of ccCoV represented; cohort A), as well as patients with

238 diagnosed *Mycoplasma pneumoniae* (n=22; cohort Z), EBV or CMV infection (n=57, cohort E).

239 We further selected a panel of 57 sera from 29 RT-PCR confirmed COVID-19 patients collected

240 at different days post symptom onset as a positive sample set (cohort C, see below).

241 Sera were employed as primary antisera for IF staining using IgM, IgA or IgG specific

242 secondary antibodies, and samples were imaged and analysed as described above. This

243 procedure yielded a ratiometric intensity score for each serum sample. Based on the scores

244 obtained for the negative control cohort and the patient sera, we defined the threshold separating

245   negative from positive scores for each of the antibody channels. For this, we performed ROC

246   curve analysis [26–28] on a subset of the data (cohorts A, B, C, Z). Using this approach, it is

247   possible to take the relative importance of sensitivity versus specificity as well as seroprevalence

248   in the population (if known) into account for optimal threshold definition. By giving more weight

249   to false positive or false negative results, one can adjust the threshold dependent on the context

250   of the study. Whereas high sensitivity is of importance for e.g. monitoring seroconversion of a

251   patient known to be infected, high specificity is crucial for population based screening

252   approaches, where large study cohorts characterized by low seroprevalence are tested. Since

253   we envision the use of the assay for screening approaches, we decided to assign more weight

254   to specificity at the cost of sensitivity for our analyses (see Materials and Methods for an in-depth

255   description of the analysis). Optimal separation in this case was given using threshold values of

256   1.39, 1.31 and 1.27 for IgA, IgG and IgM channels respectively (Fig. S7). We validated the

257   classification performance on negative control cohort E (n=57) which was not seen during

258   threshold selection, and detected no positive scores. Results from the analysis of the negative

259   control sera are presented in Fig. 4 and Table 1.

260   While the majority of these samples tested negative in ELISA measurements as well

261   as in the IF analyses, some positive readings were obtained in each of the assays, in particular

262   in the IgA specific analyses (Fig. 4 and Table 1). Since samples from these cohorts were

263   collected between 2015 and 2019, and donors were therefore not exposed to SARS-CoV-2

264   before sampling, these readings represent false positives. Of note, negative control cohort E

265   displayed a particularly high rate of false positives in ELISA measurements, but not in IF (Table

266   1). We conclude that the threshold values determined achieve our goal of yielding highly specific

267   IF results (at the cost of sub-maximal sensitivity).

268   Roughly 10.6% (IgA) or 3% (IgG) of the samples were classified as positive or

269   potentially positive by ELISA (Table 1). The notably lower specificity of the IgA determination in

270   a seronegative cohort observed here is in accordance with findings in other studies [29,30] and

271   information provided by the manufacturer of the test (90,5% for IgA vs. 99,3% for IgG;

272   Euroimmun SARS-CoV-2 data sheet, April 24, 2020; in response to these findings, an improved

273   version of the test has been recently developed). The respective proportion of false-positives

274   obtained based on IF, 0% for IgA and 0,9% for IgG, were lower, indicating higher specificity of

275   the IF readout compared to the ELISA measurements. Importantly, however, false positive

276   readings did not correlate between ELISA and IF (Fig. 4). Thus, classifying only samples that

277   test positive in both assays as true positives resulted in the elimination of false positive results

278   (0 of 218 positives detected). We conclude that applying both methods in parallel and using the

279   'double positive' definition for classification notably improves specificity of SARS-CoV-2 antibody

280   detection.

281   In order to determine the sensitivity of our IF assay, we employed 57 sera from 29

282   symptomatic COVID-19 patients that had been RT-PCR confirmed for SARS-CoV-2 infection.

283   Archived sera from these patients had been collected in the range between day 5 and 27 post

284   symptom onset. Again, samples were measured both in IF and ELISA, and the correlation

285   between the semi-quantitative values was assessed as shown in Fig. 5. While there were

286   deviations in the height of the values, positive correlation was evident in both cases, with values

287   for the IgG readout being more congruent than those for the less specific IgA determination

288   (Pearson r: 0,847 for IgG; 0,655 for IgA).

289   For an assessment of sensitivity, we stratified the samples according to the day post

290   symptom onset, as shown in Fig. 6. and Table 2. For both methods, and for all antibody classes,

291   mean values and the proportion of positive samples increased over time. In all cases, only

292   positive values were obtained for samples collected later than day 14 post symptom onset, in

293   accordance with other reports [30–32]. Consistent with other reports [32], SARS-CoV-2 specific IgM

294   was not detected notably earlier than the two other antibody classes in our measurements. At

295   the earlier time points (up to day 14), a similar or higher proportion of positive samples was

296   detected by IF compared to ELISA for IgG. Although the sample size used here is too small to

297   allow a firm conclusion, these results suggest that the sensitivity of IgG detection by the semi-

298   quantitative IF approach is higher than that of an approved semi-quantitative ELISA assay

299   routinely used in diagnostic labs. In the case of IgA detection at earlier time points (< day 11)

300   ELISA performed slightly better (11/17 samples scored positive) compared to IF (9/17 scored

301   positive) however that came with the price of a very low specificity of ELISA IgA assay (10.6%

302   false negative detection) compared to IF (0.5%).

3. Discussion

Here, we describe the development of a semi-quantitative IF based assay for detection of SARS-CoV-2 specific antibodies in human samples that complements available ELISA-based testing systems [33,34]. Alternatives to ELISA-based commercial test kits are important in situations where those kits are not available either because they are not yet developed in early days of the pandemic or due to high global demands for tests and required reagents. The microscopy-based assay described here has been developed during the early phase of the COVID-19 pandemic to support the serological testing needs of the University Hospital Heidelberg, Germany and is employed as a confirmatory assay in clinical studies [35] and ongoing studies]. The assay displayed comparable or slightly better sensitivity and specificity than a commercially available semi-quantitative SARS-CoV-2 ELISA approved for diagnostic use at the time. More importantly, combining two technically different serological assays, IF and ELISA, and classifying as "positive hits" only those that scored positive in both assays was instrumental to minimize false positive results while maintaining high sensitivity, and thus serves as a principle for serological studies or diagnostics where specificity of detection is of critical importance. Specificity of detection is essential in settings of relatively low SARS-CoV-2 antibody prevalence [36–38] in conjunction with high prevalence of potentially cross-reactive anti-ccCoV antibodies in a global population [39].

One advantage of the IF based assay presented here is that the specimens used for detection present the entire viral proteome, while ELISA or chemiluminescent approaches use a single recombinantly expressed antigen. Both the N and S protein of coronaviruses are highly immunogenic, and antibodies binding to the receptor binding domain on the S1 subunit are considered most relevant for neutralization. However, the relative importance of antibodies directed against the N protein for potential protective immunity against SARS-CoV-2 and the possible relevance of the overall breadth of the antibody response is currently unclear. Other SARS-CoV-2 structural and non-structural proteins might play a role in immune response as it was shown for proteins 3a and 9b of the closely related SARS-CoV [40]. In addition, expression of the viral proteome in permissive cells ensures correct protein folding and post-translational modification patterns. Alterations in post-translational modifications are likely to influence the ability of serum antibodies to bind to different viral epitopes as it was shown for other viruses such as HIV-1 [41]. It has to be noted that the detection of viral RNA requires fixation and permeabilization of cells, which has the potential to affect epitope preservation. However, based

336 on the high sensitivity of antibody detection and the good correlation to ELISA measurements
337 observed we conclude that this was no major concern in this case.

338 Two major disadvantages of typical IF-based serological assays as applied in the past
339 are manual microscopy acquisition steps and evaluation of samples based on a visual
340 inspection. This procedure is incompatible with high throughput approaches and results are
341 subjective, not quantitative and difficult to standardize. We have addressed these disadvantages
342 by implementing automated microscopy acquisition and developing a robust software platform
343 that is able to identify individual cells, classify infected and non-infected cells and take into
344 account specific and non-specific background in order to generate semi-quantitative results.
345 Depending on the context of a study and the questions to be addressed, sensitivity or specificity
346 may be of higher importance. The automated image analysis protocol developed here allows the
347 user to adapt the classification according to the study needs, putting more weight on either one
348 of the parameters.

349 Automated image acquisition and image analysis presented here are compatible with a
350 high throughput approach. Plates with fixed samples of infected cells can be prepared in
351 advance and stored at 4°C for several weeks. In the manual workflow used here, four 96-well
352 plates (384 samples) could easily be analysed within a typical work day (1.5 h for
353 immunofluorescence, 1.5 h for image acquisition, 2 h of image analysis). This is already the
354 throughput in the range of some ELISA-based automated systems used in diagnostics and is
355 sufficient for urgent applications in an early phase of disease response. The major disadvantage
356 of the procedure described here for a virus like SARS-CoV-2 is the requirement of a BSL3
357 containment area to generate virus stocks and produce infected cell specimens. Recombinant
358 cell lines expressing key viral antigens can address this drawback and also allows to easily
359 implement already established automated cell seeding and immunostaining pipelines for a true
360 high-throughput application [42,43]. Combining such cell lines with spectral unmixing microscopy
361 [44] would not only enable simultaneous determination of levels of all three major classes of
362 antibodies (IgM, IgG and IgA), but also identification of the viral antigens recognized, in a single
363 multiplexed approach. The high information content of the IF data (differential staining patterns)
364 together with a machine learning-based approach [45] and the implementation of stable cell lines
365 expressing selected viral antigens in the IF assay will provide additional parameters for
366 classification of patient sera and further improve sensitivity and specificity of the presented IF
367 assay.

368 The described analysis pipeline can be readily applied for serological analysis of other
369 virus infections, provided that an infectable cell line and a staining procedure that allows

370 differentiating between infected and non-infected cells are available. The assay described here
371 thus offers potential as an immediate response to any future virus pandemic, as it can be rapidly
372 deployed from the moment the first isolate of the pathogen has been obtained without requiring
373 information on the expression of immunogenicity of viral proteins.
374

375 4. Materials and Methods

376 *4.1 Human material*

377 Negative control serum samples (n=218) were collected for various serological testing in the
378 routine laboratory of the Center of Infectious Diseases, University Hospital Heidelberg between
379 2015 and 2019, before the start of the SARS-CoV-2 outbreak. Samples used corresponded to
380 pseudonymized remaining material from the archive of the Center of Infectious Diseases
381 Heidelberg. SARS-CoV-2 positive sera were collected from 29 PCR confirmed symptomatic
382 COVID-19 inpatients (n=17) or outpatients (n=12) treated at the University Hospital Heidelberg
383 under general informed consent (ethics votum no S-148/2020, University Hospital Heidelberg).
384 Days post symptom onset were defined based on the anamnesis carried out upon admission.
385 Serum samples were stored at -20°C until use.

386 *4.2 Virus stock production*

387 VeroE6 cells were cultured in Dulbecco's modified Eagle medium (DMEM, Life Technologies)
388 containing 10% fetal bovine serum, 100 U/mL penicillin, 100 µg/mL streptomycin and 1% non-
389 essential amino acids (complete medium).

390 SARS-CoV-2 virus stocks were produced by amplification of the BavPat1/2020 strain (European
391 Virus Archive) in VeroE6 cells. To generate the seed virus (passage 3), VeroE6 cells were
392 infected with the original virus isolate, received as passage 2, at an MOI of 0.01. At 48 h post
393 infection (p.i.), the supernatant was harvested and cell debris was removed by centrifugation at
394 800xg for 10 min. For production of virus stocks (passage 4), 500µl of the seed virus was used
395 to infect $9x10^6$ VeroE6 cells. The resulting supernatant was harvested 48h later as described
396 above. Virus titers were determined by plaque assay. Briefly, $2.5x10^6$ VeroE6 cells were plated
397 into 24 well plates. 24 h later, cells were infected with serial dilutions of SARS-CoV-2 for 1 h.
398 Inoculum was then removed and the cells were overlaid with serum free DMEM containing 0.8%

399 carboxymethylcellulose. At 72 h. p.i., cells were fixed with 5% formaldehyde for 1 h followed by

400 staining with 1% Crystal violet solution. Plaque forming units per ml (PFU/ml) were estimated by

401 manual counting of the viral plaques. Stock solutions were stored in aliquots at -80°C until use

402 for infection experiments.

403 *4.3 Infection of cells and immunofluorescence staining*

404 In order to find a suitable cell line for our application, we performed pre-experiments comparing

405 different cell lines with respect to their susceptibility to SARS-CoV-2 infection. Cells were seeded

406 on glass coverslips and infected on the following day with SARS-CoV-2 strain BavPat1/2020 for

407 16h at MOI 0.01. Cells were fixed with 6%PFA in PBS, followed by permeabilisation with 0.5%

408 Triton X100 in PBS and then subjected to a standard immunofluorescence staining protocol as

409 described in materials and methods. Only very few infected calls were detected in the case of

410 hepatocyte-derived carcinoma cells (HUH-7), human embryonic kidney (HEK293T) and human

411 alveolar basal epithelial (A549) cells (Fig. S1). Calu-3 cells grew in small clumps, often on top of

412 each other which impacted our microscopy-based readout. In contrast, VeroE6 cells grew as a

413 monolayer and were viable for at least 24 h p.i. Based on these results, VeroE6 cells were

414 chosen for all experiments in this manuscript.

415 For serum screening by IF microscopy, VeroE6 cells were seeded at a density of 7,000 cells per

416 well into a black-wall glass-bottom 96 well plates (Corning, Product Number 353219) or on glass

417 coverslips placed in a 24-well plate. 24 h after seeding, cells were infected with SARS-CoV-2 at

418 an MOI of 0.01 for 16 h. Cells were then fixed with 6% Formaldehyde for 1 h followed by washing

419 3x with phosphate buffered saline (PBS) under biosafety level 3. Afterwards, samples were

420 handled under biosafety level 2. Cells were washed once in PBS containing 0,02% Tween 20

421 (Sigma) and permeabilised using 0,5% Triton X100 (Sigma) for 10 minutes. Samples were

422 washed again and blocked using 2% powdered milk (Roth) in PBS for 20 min followed by two

423 additional washing steps. All washing steps in a 96-well format were performed using the

424 HydroFlex microplate washer (Tecan). Next, cells were incubated with patient serum (prediluted

425 1:1 in 0,4% Triton-X100 in PBS; further dilution 1:50 in PBS if not stated otherwise) and anti-ds-

426 RNA mouse monoclonal J2 antibody (Scicons, 1:4000) in PBS for 30 min at room temperature.

427 After 3 washing steps, cognate secondary antibodies were applied for 20 min at room

428 temperature. Goat anti-human IgG-AlexaFluor 488 (Invitrogen, Thermofisher Scientific), goat

429 anti-human IgA DyLight 650 (Abcam), goat anti-human IgM u chain (Invitrogen, Thermofisher

430 Scientific), for detecting immunoglobulins in human serum together with goat anti-mouse IgG-

431   AlexaFluor 568 (Invitrogen, Thermofisher Scientific) for dsRNA detection, all at 1:2000 dilution

432   in PBS, have been used. After incubation with secondary antibodies cells were washed twice,

433   stained with Hoechst (0,002µg/ml in PBS) for 3 minutes, washed again twice and stored at +4°C

434   until imaging.

435   *4.4 Microscopy*

436   Samples were imaged on motorized Nikon Ti2 widefield microscope using a Plan Apo lambda

437   20x/0.75 air objective and a back-illuminated EM-CCD camera (Andor iXon DU-888). To

438   automatically acquire images in 96-well format, the JOBS module was used. The system was

439   configured to acquire 9 images per well (in a regular 3 x 3 pattern centered in the middle of each

440   well). The Perfect Focus System was used for autofocusing followed by a software-based fine

441   focusing using the Hoechst signal in an axial range of 40um. Images were acquired in 4 channels

442   using the following excitation/emission settings: Ex 377/50, Em 447/60 (Hoechst); Ex 482/35,

443   Em 536/40 (AlexaFluor 488); Ex 562/40, Em 624/40 (AlexaFluor 568) and Ex 628/40, Em 692/40

444   (AlexaFluor 647 and DyLight 650). Exposure times were in the range between 50 and 100ms

445   with EM gain between 50 and 150.

446

447   *4.5 Enzyme linked immuno- sorbent assay (ELISA)*

448   ELISA measurements for determination of reactivity against the S1 domain of the viral spike

449   protein were carried out using the Euroimmun Anti-SARS-CoV-2-ELISA (IgA) and Anti-SARS-

450   CoV-2-ELISA (IgG) test kits (Euroimmun, Lübeck, Germany; EI 2606-9601 A and EI 2606-9601

451   G) run on an Euroimmun Analyzer I instrument according to the manufacturer's instructions.

452   Optical densities measured for the samples were normalized using the value obtained for a

453   calibrator sample provided in the test kit. The interpretation of the semi-quantitative ratiometric

454   values obtained followed the manufacturer's protocol: values <0.8 were classified as negative,

455   0.8-1.1 as borderline, and values of 1.1 or higher as positive.

456   *4.6 Image Analysis*

457   *Manual Annotations*

458   Two of our processing steps require manually annotated data: in order to train the convolutional

459   neural network used for boundary and foreground prediction, we needed label masks for the

460   individual cells. To determine suitable parameters for the infected cell classification, we needed

461   a set of cells classified as being infected or non-infected. We have produced these annotations

462   for 10 images with the following steps. First, we created an initial segmentation following the

463   approach outlined in the Segmentation subsection, using boundary and foreground predictions

464   from the ilastik [46] pixel classification workflow, which can be obtained from a few sparse

465   annotations. We then corrected this segmentation using the annotation tool BigCat

466   (https://github.com/saalfeldlab/bigcat). After correction, we manually annotated these cells as

467   infected or non-infected. Note that this mode of annotations can introduce two types of bias: the

468   segmentation labels are derived from an initial segmentation. Small systematic errors in the

469   initial segmentation that were not found during correction, could influence the boundary

470   prediction network. More importantly, when annotating the infected / non-infected cells, both the

471   serum channel and the virus marker channel have to be available to the annotators, in order to

472   visually delineate the cells. This may result in subconscious bias, with the observed intensity in

473   the serum channel influencing the decision on the infection status of a cell.

474

475   *Preprocessing*

476   On all acquired images, we performed minimal preprocessing (i.e., flat-field correction) in order

477   compensate for uneven illumination of the microscope system [47]. First, we subtract a constant

478   CCD camera offset ($\mathrm{ccd\_offset}$). Secondly, we correct uneven illumination by dividing each

479   channel by a corresponding corrector image ($\mathrm{flatfield}(x, y)$), which was obtained as a normalized

480   average of all images of that channel, smoothed by a normalized convolution with a Gaussian

481   filter with a bandwidth of 30 pixels.

$$\mathrm{processed}(x, y) = \frac{\mathrm{raw}(x, y) - \mathrm{ccd\_offset}}{\mathrm{flatfield}(x, y) - \mathrm{ccd\_offset}} \qquad \text{Eq. 2}$$

482   This corrector image was obtained for all images of a given microscope set-up. Full background

483   subtraction is performed later in the pipeline using either the background measured on wells that

484   (deliberately) do not contain any serum or, if not available, using a fixed value that was

485   determined manually.

486

487   *Segmentation*

488   Cell segmentation forms the basis of our analysis method. In order to obtain an accurate

489   segmentation, we make use of both the DAPI and the serum channel. First, we segment the

490   nuclei on the DAPI channel using the StarDist method [19] trained on data from Caicedo et al.

491   2019 [48]. Note that this method yields an *instance segmentation*: each nucleus in the image is

492   assigned a unique ID. In addition, we predict per pixel probabilities for the boundaries between

493   cells and for the foreground (i.e. whether a given pixel is part of a cell) using a 2D U-Net [20]

494   based on the implementation of Wolny et al. 2020 [21]. This method was trained using the 9

495   annotated images, see above. The cells are then segmented by the seeded watershed algorithm

496   [18]. We use the nucleus segmentation, dilated by 3 pixels, as seeds and the boundary predictions

497   as the height map. In addition, we threshold the foreground predictions, erode the resulting

498   binary image by 20 pixels and intersect it with the binarised seeds. The result is used as a

499   foreground mask for the watershed. The dilation / erosion is performed to alleviate issues with

500   very small nucleus segments / imprecise foreground predictions. In order to evaluate this

501   segmentation method, we train 9 different networks using leave-one-out cross-validation,

502   training each network on 8 of the manually annotated images and evaluating it on the remaining

503   one. We measure the segmentation quality using average precision [22] at an intersection over

504   union (IoU) threshold of 0.5 as described in https://www.kaggle.com/c/data-science-bowl-

505   2018/overview/evaluation. We measure a value of 0.77 +- 0.08 with the optimum value being

506   1.0.

507

508   *Quantitation and Scoring*

509   *Infection classification*

510   To distinguish infected cells from control cells we use the dsRNA virus marker channel: infected

511   cells show a signal in this channel while the non-infected control cells should ideally be invisible

512   (see Fig. 3). We classified each cell in the cell segmentation (see above) individually, using the

513   following procedure. First, we denoised the marker channel using a white tophat filter with a

514   radius of 20 pixels. To account for inaccuracies in the cell segmentation (the exact position of

515   cell borders is not always clear), we then eroded all cell masks with a radius of 5 pixels and

516   thereby discard pixels close to segment boundaries. This step does not lead to information loss,

517   since the virus marker is mostly concentrated around the nuclei. On the remaining pixels of each

518   cell, we compute the 0.95 quantile (q) of the intensity in the marker channel. For the pixels that

519   the neural network predicts to belong to the background (b), we compute the median intensity

520   of the virus marker channel across all images in the current plate. Finally, we classify the cell as

521   infected if the 0.95 quantile of its intensity exceeds the median background by more than a given

522   threshold:

$$q - \mathrm{median}(b) > t$$ Eq. 3

523 For additional robustness against intensity variations we adapt the threshold based on the

524 variation in the background in the plate. Hence, we define it as a multiple of the mean absolute

525 deviation of all background pixels of that plate with N=4.8:

$$t = N \cdot \mathrm{mad}(b)$$ Eq. 4

526 To determine the optimal values of the parameters used in our procedure, we used the cells

527 manually annotated as infected / non-infected (see above). We performed grid search over the

528 following parameter ranges:

529   ● Quantile: 0.9, 0.93, 0.95, 0.96, 0.97, 0.98, 0.99, 0.995

530   ● N: 0 to 10 in intervals of 0.1

531 To estimate the validation accuracy, we performed leave-one-out cross-validation on the image

532 level. This yields an average validation F1 score of 84.3%, precision of 84.3% and recall of

533 84.8%. These values are the arithmetic means of the individual results per split.

534

535 *Immunoglobulin intensity measurements*

536

537 In order to obtain a relative measure of antibody binding, we determined the mean intensity and

538 the integrated intensity in each segmented cell from images recorded in the IgG, IgA or IgM

539 channel. A comparative analysis revealed that the mean intensity was more robust against the

540 variability of cell sizes, whereas using the integrated intensity as a proxy yielded a higher

541 variance in non-infected cells. Thus, mean intensity per cell was chosen as a proxy for the

542 amount of antibody bound. Non-specific auto-fluorescence signals required a background

543 correction of the measured average serum channel intensities. For background normalization,

544 we used cells (one well per plate) which were not immunostained with primary antiserum. From

545 this we computed the background to be the median serum intensity of all pixels of images taken

546 from this well. This value was subtracted from all images recorded from the respective plate. In

547 case this control well was not available, background was subtracted manually by selecting the

548 area outside of cells in randomly selected wells and measuring the median intensity.

549

550 *Scoring*

551 The core interest of the assay is to measure the difference of antibody binding to cells infected

552 with the coronavirus in comparison to non-infected control cells. To this end, utilizing the results

553 of the image analysis, we compute the following summary statistics of the background corrected

554 antibody binding of infected cells, I, and of non-infected cells, N:

$$m_I = \text{median}(I) \qquad\qquad \text{Eq. 5}$$

$$m_N = \text{median}(N) \qquad\qquad \text{Eq. 6}$$

$$\sigma_N = \text{mad}(N) \qquad\qquad \text{Eq. 7}$$

Using these, the ratio r, difference d and robust z score z are computed: $\quad r = \dfrac{m_I}{m_N} \qquad$ Eq. 8

$$d = m_I - m_N \qquad\qquad \text{Eq. 9}$$

$$z = \frac{m_I - m_N}{\sigma_N} \qquad\qquad \text{Eq. 10}$$

555 We compute above scores for each well and each image, taking into account only the cells that
556 passed all quality control criteria (see below). While the final readout of the assay is well-based,
557 image scores are useful for quality control.

558

559 *Decision threshold selection*

560 In order to determine the presence of SARS-CoV-2 specific antibodies in patient sera, it was
561 necessary to define a decision threshold r*. If a measured intensity ratio r is above a decision
562 threshold r* than the serum would be characterized as positive for SARS-CoV-2 antibodies. For
563 this an ROC analysis was performed [28]. Each possible choice of r* for a test corresponds to a
564 particular sensitivity/specificity pair. By continuously varying the decision threshold, we
565 measured all possible sensitivity/specificity pairs, known as ROC curves (Fig. S7). To determine
566 the appropriate r* we considered two factors [26]:

567

568 ● The undesirability of errors or relative cost of false-positive and false-negative
569   classifications
570 ● The prevalence, or prior probability of disease

571

572 These factors can be combined to calculate a slope in the ROC plot[26–28]

$$m = \frac{(falsepositivecost)}{(falsenegativecost)} \frac{(1 - P)}{P} \qquad\qquad \text{Eq. 11}$$

573 where P is the prevalence or prior probability of disease.

574

575 The optimal decision threshold r*, given the false-positive/false-negative cost ratio and
576 prevalence, is the point on the ROC curve where a line with slope m touches the curve. As

577   discussed in the main text, a major concern regarding serological assays for SARS-CoV-2

578   antibody detection is the occurrence of false-positive results. Therefore, we choose m to be

579   larger than one in our analysis. In particular, we determine r* for the choice of m=10 (see Fig.

580   S7).

581   *Quality control*

582   We performed quality control of the images and analysis results at the level of wells, images and

583   cells. The entities that did not pass quality control are not taken into account when computing

584   the score during final analysis. We exclude wells that contain less than 100 non-infected cells,

585   that have a median serum intensity of infected cells smaller than 3 times the noise level

586   (measured by the median absolute deviation), or that have negative intensity ratios, which can

587   happen due to the background subtraction. Out of 1.736 wells, 94 did not pass the quality control,

588   corresponding to 5.4 % of wells. At the image level, we visually inspect all images and mark

589   those that contain imaging artifacts using a viewer based on napari [49]. We distinguish the

590   following types of artifacts during the visual inspection: empty, unstained or over-saturated

591   images, as well as images covered by a large bright object. In addition, we automatically exclude

592   images that contain less than 10 or more than 1000 cells. These thresholds are motivated by

593   the observation that too few or too many cells often result from a problem in the assay. Thus,

594   296 of the total 15.624 images were excluded from further analysis, corresponding to 1.9 % of

595   images. Out of these, 295 were manually marked as outliers and only a single one did not pass

596   the subsequent automatic quality control. Finally, we automatically exclude segmented cells with

597   a size smaller than 250 pixels or larger than 12.500 pixels that most likely correspond to

598   segmentation errors. These limits were derived by the histogram of cell sizes investigated for

599   several plates. Two percent of the approximate 5.5 million segmented cells did not pass this

600   quality control. In addition, we have also inspected all samples scored as positives. For the IgA

601   channel, we have found a dotty staining pattern in ten cases that produced positive hits based

602   on intensity ratio in negative control cohorts, but does not appear to indicate a specific antibody

603   response. We have also excluded these samples from further analysis.

604   *Implementation*

605   In order to scale the analysis workflow to the large number of images produced by the assay,

606   we implemented an open-source python library to run the individual analysis steps. This library

607   allows rerunning experiments for a given plate for newly added data on demand and caches

608   intermediate results in order to rerun the analysis from checkpoints in case of errors in one of

609   the processing steps. To this end, we use a file layout based on hdf5 [50] to store multi-resolution

610   image data and tabular data. The processing steps are parallelized over the images of a plate if

611   possible. We use efficient implementations for the U-Net [21], StarDist [19] and the watershed

612   algorithm (http://ukoethe.github.io/vigra/) as well as other image processing algorithms [51]. We

613   use pytorch (https://pytorch.org/) to implement GPU-accelerated cell feature extraction. The

614   total processing time for a plate (containing around 800 images) is about two hours and thirty

615   minutes using a single GPU and 8 CPU cores. In addition, the results of the analysis as well as

616   meta-data associated with individual plates are automatically saved in a centralized MongoDB

617   database (https://www.mongodb.com) at the end of the workflow execution. Apart from keeping

618   track of the analysis outcome and meta-data, a user can save additional information about a

619   given plate/well/image in the database conveniently using the PlateViewer (see below). All

620   source code is available open source under the permissive MIT license at https://github.com/hci-

621   unihd/batchlib.

622

623   *Data visualization*

624   In order to explore the numerical results of our analysis together with the underlying image data

625   we further developed a Fiji [52] based open-source software tool for interactive visualization of

626   high-throughput microscopy data [23]. The PlateViewer links interactive results tables and

627   configurable scatter plots (image and well based) with a plate view of all raw, processed and

628   segmentation images. The PlateViewer is connected to the centralised database such that also

629   image and well based metadata can be accessed. The viewer thus enables efficient visual

630   inspection and scientific exploration of all relevant data of the presented assay.

631

632   *Data availability*

633   The data from the IF assay are available in the BioImage Archive

634   (http://www.ebi.ac.uk/bioimage-archive) under accession number S-BIAD24. This includes raw

635   microscopy images, intermediate segmentation and infected cell classification results as well as

636   quality control and final score results.

637   Acknowledgements

656 Conflict of Interest

657   The authors declare they have no conflicts of interest.

658

659 Author Contributions

660   Model system development: M. Cortese, B. Lucic, B. Cerikan, C.J. Neufeldt, M. Lusic, S.

661   Boulant, M. Stanifer, R. Bartenschlager

662   Microscopy development: S. Olberg, V. Laketa

663   Image analysis development: C. Pape, R. Remme, A. Wolny, S. Wolf, L. Cerrone, S. Klaus, M.

664   Ganter, F.A.Hamprecht, A. Kreshuk, C. Tischer, V. Laketa

665   ELISA assay: S. Wolf, P. Schnitzler

666   Sera selection and processing: S. Ullrich, M. Anders-Össwein, B. Müller, P. Schnitzler, U. Merle

667   Data interpretation: C. Pape, R. Remme, A. Wolny, S. Wolf, F. A. Hamprecht, A. Kreshuk, C.

668   Tischer, H-G. Kräusslich, B. Müller, V. Laketa

669   Study design: F. A. Hamprecht, H-G. Kräusslich, B. Müller, V. Laketa

670   Manuscript writing, figures, tables: C. Pape, R. Remme, A. Wolny, S. Wolf, L. Cerrone, F. A.

671   Hamprecht, A. Kreshuk, C. Tischer, B. Müller, V. Laketa

672   All authors have read and approved the final version of the manuscript.

References

[1] F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E. C. Holmes, Y.-Z. Zhang, *Nature* **2020**, *579*, 265.

[2] A. E. Gorbalenya, S. C. Baker, R. S. Baric, R. J. de Groot, C. Drosten, A. A. Gulyaeva, B. L. Haagmans, C. Lauber, A. M. Leontovich, B. W. Neuman, D. Penzar, S. Perlman, L. L. M. Poon, D. V. Samborskiy, I. A. Sidorov, I. Sola, J. Ziebuhr, *Nat. Microbiol.* **2020**, *5*, 536.

[3] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, *Nature* **2020**, *579*, 270.

[4] V. M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D. K. Chu, T. Bleicker, S. Brünink, J. Schneider, M. L. Schmidt, D. G. Mulders, B. L. Haagmans, B. van der Veer, S. van den Brink, L. Wijsman, G. Goderski, J.-L. Romette, J. Ellis, M. Zambon, M. Peiris, H. Goossens, C. Reusken, M. P. Koopmans, C. Drosten, *Eurosurveillance* **2020**, *25*, 1.

[5] V. L. Dao Thi, K. Herbst, K. Boerner, M. Meurer, L. P. Kremer, D. Kirrmaier, A. Freistaedter, D. Papagiannidis, C. Galmozzi, M. L. Stanifer, S. Boulant, S. Klein, P. Chlanda, D. Khalid, I. Barreto Miranda, P. Schnitzler, H.-G. Kräusslich, M. Knop, S. Anders, *Sci. Transl. Med.* **2020**, *12*, eabc7075.

[6] M. Venter, K. Richter, *J. Clin. Pathol.* **2020**, *73*, 370.

[7] S. Hsiang, D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L. Y. Huang, A. Hultgren, E. Krasovich, P. Lau, J. Lee, E. Rolf, J. Tseng, T. Wu, *Nature* **2020**, 1.

[8] E. Callaway, H. Ledford, S. Mallapaty, *Nature* **2020**, *583*, 178.

[9] K. Mizumoto, K. Kagaya, A. Zarebski, G. Chowell, *Eurosurveillance* **2020**, *25*, 2000180.

[10] H. Qiu, J. Wu, L. Hong, Y. Luo, Q. Song, D. Chen, *Lancet Infect. Dis.* **2020**, *20*, 689.

[11] J. S. Weitz, S. J. Beckett, A. R. Coenen, D. Demory, M. Dominguez-Mirazo, J. Dushoff, C. Y. Leung, G. Li, A. Măgălie, S. W. Park, R. Rodriguez-Gonzalez, S. Shivam, C. Y. Zhao, *Nat. Med.* **2020**, 1.

[12] B. Meyer, C. Drosten, M. A. Müller, *Virus Res.* **2014**, *194*, 175.

[13] J. D. Whitman, J. Hiatt, C. T. Mowery, B. R. Shy, R. Yu, T. N. Yamamoto, U. Rathore, G. M. Goldgof, C. Whitty, J. M. Woo, A. E. Gallman, T. E. Miller, A. G. Levine, D. N. Nguyen, S. P. Bapat, J. Balcerek, S. A. Bylsma, A. M. Lyons, S. Li, A. W. Wong, E. M. Gillis-Buck, Z. B. Steinhart, Y. Lee, R. Apathy, M. J. Lipke, J. A. Smith, T. Zheng, I. C. Boothby, E. Isaza, J. Chan, D. D. Acenas, J. Lee, T. A. Macrae, T. S. Kyaw, D. Wu, D. L. Ng, W. Gu, V. A. York, H. A. Eskandarian, P. C. Callaway, L. Warrier, M. E. Moreno, J. Levan, L. Torres, L. A. Farrington, R. P. Loudermilk, K. Koshal, K. C. Zorn, W. F.

Garcia-Beltran, D. Yang, M. G. Astudillo, B. E. Bernstein, J. A. Gelfand, E. T. Ryan, R. C. Charles, A. J. Iafrate, J. K. Lennerz, S. Miller, C. Y. Chiu, S. L. Stramer, M. R. Wilson, A. Manglik, C. J. Ye, N. J. Krogan, M. S. Anderson, J. G. Cyster, J. D. Ernst, A. H. B. Wu, K. L. Lynch, C. Bern, P. D. Hsu, A. Marson, *Nat. Biotechnol.* **2020**, DOI 10.1038/s41587-020-0659-0.

[14] P. K. S. Chan, K. C. Ng, R. C. W. Chan, R. K. Y. Lam, V. C. Y. Chow, M. Hui, A. Wu, N. Lee, F. H. Y. Yap, F. W. T. Cheng, J. J. Y. Sung, J. S. Tam, *Emerg. Infect. Dis.* **2004**, *10*, 530.

[15] I. J. Liu, P. J. Chen, S. H. Yeh, Y. P. Chiang, L. M. Huang, M. F. Chang, S. Y. Chen, P. C. Yang, S. C. Chang, W. K. Wang, *J. Clin. Microbiol.* **2005**, *43*, 2444.

[16] I. Manopo, L. Lu, Q. He, L. L. Chee, S. W. Chan, J. Kwang, *J. Immunol. Methods* **2005**, *296*, 37.

[17] N. S. Ogando, T. J. Dalebout, J. C. Zevenhoven-Dobbe, R. W. A. L. Limpens, Y. van der Meer, L. Caly, J. Druce, J. J. C. de Vries, M. Kikkert, M. Bárcena, I. Sidorov, E. J. Snijder, *J. Gen. Virol.* **2020**, DOI 10.1099/jgv.0.001453.

[18] J. B. T. M. Roerdink, A. Meijster, *Fundam. Informaticae* **2000**, *41*, 187.

[19] U. Schmidt, M. Weigert, C. Broaddus, G. Myers, in *Med. Image Comput. Comput. Assist. Interv. – MICCAI 2018. MICCAI 2018. Lect. Notes Comput. Sci.*, Springer Verlag, **2018**, pp. 265–273.

[20] O. Ronneberger, P. Fischer, T. Brox, in *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, Springer Verlag, **2015**, pp. 234–241.

[21] A. Wolny, L. Cerrone, A. Vijayan, R. Tofanelli, A. V. Barro, M. Louveaux, C. Wenzl, S. Strauss, D. Wilson-Sánchez, R. Lymbouridou, S. S. Steigleder, C. Pape, A. Bailoni, S. Duran-Nebreda, G. W. Bassel, J. U. Lohmann, M. Tsiantis, F. A. Hamprecht, K. Schneitz, A. Maizel, A. Kreshuk, *Elife* **2020**, *9*, 1.

[22] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, A. Zisserman, *Int. J. Comput. Vis.* **2010**, *88*, 303.

[23] C. Tischer, R. Pepperkok, **2019**, DOI 10.5281/zenodo.3522688.

[24] H. A. T. Goossens, M. K. E. Nohlmans, A. E. J. M. Van den Bogaard, *Infection* **1999**, *27*, 231.

[25] B. Ladizinski, C. Sankey, *Am. J. Med.* **2014**, *127*, e9.

[26] M. H. Zweig, G. Campbell, *Clin. Chem.* **1993**, *39*, 561.

[27] K. Linnet, *Clin. Chem.* **1988**, *34*, 1379.

[28] B. J. McNeil, E. Keeler, S. J. Adelstein, *N. Engl. J. Med.* **1975**, *293*, 211.

[29] R. Lassaunière, A. Frische, Z. B. Harboe, A. C. Nielsen, A. Fomsgaard, K. A. Krogfelt, C. S. Jørgensen, *medRxiv* **2020**, 2020.04.09.20056325.

747 [30] N. M. A. Okba, M. A. Müller, W. Li, C. Wang, C. H. GeurtsvanKessel, V. M. Corman, M.
748   M. Lamers, R. S. Sikkema, E. de Bruin, F. D. Chandler, Y. Yazdanpanah, Q. Le Hingrat,
749   D. Descamps, N. Houhou-Fidouh, C. B. E. M. Reusken, B.-J. Bosch, C. Drosten, M. P.
750   G. Koopmans, B. L. Haagmans, *Emerg. Infect. Dis.* **2020**, *26*, DOI
751   10.3201/eid2607.200841.

752 [31] R. Wölfel, V. M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M. A. Müller, D.
753   Niemeyer, T. C. Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brünink, J.
754   Schneider, R. Ehmann, K. Zwirglmaier, C. Drosten, C. Wendtner, *Nature* **2020**, *581*,
755   465.

756 [32] Q. X. Long, B. Z. Liu, H. J. Deng, G. C. Wu, K. Deng, Y. K. Chen, P. Liao, J. F. Qiu, Y.
757   Lin, X. F. Cai, D. Q. Wang, Y. Hu, J. H. Ren, N. Tang, Y. Y. Xu, L. H. Yu, Z. Mo, F.
758   Gong, X. L. Zhang, W. G. Tian, L. Hu, X. X. Zhang, J. L. Xiang, H. X. Du, H. W. Liu, C.
759   H. Lang, X. H. Luo, S. B. Wu, X. P. Cui, Z. Zhou, M. M. Zhu, J. Wang, C. J. Xue, X. F.
760   Li, L. Wang, Z. J. Li, K. Wang, C. C. Niu, Q. J. Yang, X. J. Tang, Y. Zhang, X. M. Liu, J.
761   J. Li, D. C. Zhang, F. Zhang, P. Liu, J. Yuan, Q. Li, J. L. Hu, J. Chen, A. L. Huang, *Nat.*
762   *Med.* **2020**, 1.

763 [33] F. Amanat, D. Stadlbauer, S. Strohmeier, T. H. O. Nguyen, V. Chromikova, M.
764   McMahon, K. Jiang, G. A. Arunkumar, D. Jurczyszak, J. Polanco, M. Bermudez-
765   Gonzalez, G. Kleiner, T. Aydillo, L. Miorin, D. S. Fierer, L. A. Lugo, E. M. Kojic, J.
766   Stoever, S. T. H. Liu, C. Cunningham-Rundles, P. L. Felgner, T. Moran, A. García-
767   Sastre, D. Caplivski, A. C. Cheng, K. Kedzierska, O. Vapalahti, J. M. Hepojoki, V.
768   Simon, F. Krammer, *Nat. Med.* **2020**, DOI 10.1038/s41591-020-0913-5.

769 [34] F. Krammer, V. Simon, *Science* **2020**, *368*, 1060.

770 [35] "Prevalence of COVID-19 in children in Baden-Württemberg Preliminary study report,"
771   can be found under https://www.klinikum.uni-
772   heidelberg.de/fileadmin/pressestelle/Kinderstudie/Prevalence_of_COVID-
773   19_in_BaWu__.pdf, **2020**.

774 [36] E. Bendavid, B. Mulaney, N. Sood, S. Shah, E. Ling, R. Bromley-Dulfano, C. Lai, Z.
775   Weissberg, R. Saavedra, J. Tedrow, D. Tversky, A. Bogan, T. Kupiec, D. Eichner, R.
776   Gupta, J. Ioannidis, J. Bhattacharya, *medRxiv* **2020**, 2020.04.14.20062463.

777 [37] E. Lavezzo, E. Franchin, C. Ciavarella, G. Cuomo-Dannenburg, L. Barzon, C. Del
778   Vecchio, L. Rossi, R. Manganelli, A. Loregian, N. Navarin, D. Abate, M. Sciro, S.
779   Merigliano, E. De Canale, M. C. Vanuzzo, V. Besutti, F. Saluzzo, F. Onelia, M. Pacenti,
780   S. G. Parisi, G. Carretta, D. Donato, L. Flor, S. Cocchio, G. Masi, A. Sperduti, L.
781   Cattarino, R. Salvador, M. Nicoletti, F. Caldart, G. Castelli, E. Nieddu, B. Labella, L.
782   Fava, M. Drigo, K. A. M. Gaythorpe, A. R. Brazzale, S. Toppo, M. Trevisan, V. Baldo, C.
783   A. Donnelly, N. M. Ferguson, I. Dorigatti, A. Crisanti, *Nature* **2020**, *584*, 425.

784 [38] S. Stringhini, A. Wisniak, G. Piumatti, A. S. Azman, S. A. Lauer, H. Baysson, D. De
785   Ridder, D. Petrovic, S. Schrempft, K. Marcus, I. Arm-Vernez, S. Yerly, O. Keiser, S.
786   Hurst, K. Posfay-Barbe, D. Trono, D. Pittet, L. Getaz, F. Chappuis, I. Eckerle, N.
787   Vuilleumier, B. Meyer, A. Flahault, L. Kaiser, I. Guessous, *medRxiv* **2020**,
788   2020.05.02.20088898.

789 [39] E. G. Severance, I. Bossis, F. B. Dickerson, C. R. Stallings, A. E. Origoni, A. Sullens, R.

790   H. Yolken, R. P. Viscidi, *Clin. Vaccine Immunol.* **2008**, *15*, 1805.

791 [40] M. Qiu, Y. Shi, Z. Guo, Z. Chen, R. He, R. Chen, D. Zhou, E. Dai, X. Wang, B. Si, Y.
792   Song, J. Li, L. Yang, J. Wang, H. Wang, X. Pang, J. Zhai, Z. Du, Y. Liu, Y. Zhang, L. Li,
793   J. Wang, B. Sun, R. Yang, *Microbes Infect.* **2005**, *7*, 882.

794 [41] R. Pejchal, K. J. Doores, L. M. Walker, R. Khayat, P.-S. Huang, S.-K. Wang, R. L.
795   Stanfield, J.-P. Julien, A. Ramos, M. Crispin, R. Depetris, U. Katpally, A. Marozsan, A.
796   Cupo, S. Maloveste, Y. Liu, R. McBride, Y. Ito, R. W. Sanders, C. Ogohara, J. C.
797   Paulson, T. Feizi, C. N. Scanlan, C.-H. Wong, J. P. Moore, W. C. Olson, A. B. Ward, P.
798   Poignard, W. R. Schief, D. R. Burton, I. A. Wilson, *Science* **2011**, *334*, 1097.

799 [42] R. Pepperkok, J. Ellenberg, *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 690.

800 [43] M. Bickle, *Anal. Bioanal. Chem.* **2010**, *398*, 219.

801 [44] T. Zimmermann, J. Rietdorf, R. Pepperkok, *FEBS Lett.* **2003**, *546*, 87.

802 [45] M. Doan, A. E. Carpenter, *Nat. Mater.* **2019**, *18*, 414.

803 [46] S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, M. Schiegg, J.
804   Ales, T. Beier, M. Rudy, K. Eren, J. I. Cervantes, B. Xu, F. Beuttenmueller, A. Wolny, C.
805   Zhang, U. Koethe, F. A. Hamprecht, A. Kreshuk, *Nat. Methods* **2019**, *16*, 1226.

806 [47] D. Tomazevic, B. Likar, F. Pernus, *J. Microsc.* **2002**, *208*, 212.

807 [48] J. C. Caicedo, A. Goodman, K. W. Karhohs, B. A. Cimini, J. Ackerman, M. Haghighi, C.
808   Heng, T. Becker, M. Doan, C. McQuin, M. Rohban, S. Singh, A. E. Carpenter, *Nat.*
809   *Methods* **2019**, *16*, 1247.

810 [49] N. Sofroniew, K. Evans, T. Lambert, J. Nunez-Iglesias, A. C. Solak, kevinyamauchi, G.
811   Buckley, T. Tung, J. Freeman, H. Har-Gil, P. Boone, L. Royer, S. Axelrod, jakirkham, R.
812   Dunham, P. Vemuri, M. Huang, Bryant, A. Rokem, S. Li, R. Anderson, M. Bussonnier, J.
813   Kiggins, Hector, H. Patterson, G. Gay, E. Perlman, D. Bennett, C. Gohlke, A. de
814   Siqueira, **2020**, DOI 10.5281/ZENODO.3689231.

815 [50] S. Koranne, in *Handb. Open Source Tools*, Springer US, Boston, MA, **2011**, pp. 191–
816   200.

817 [51] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N.
818   Yager, E. Gouillart, T. Yu, *PeerJ* **2014**, *2*, e453.

819 [52] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S.
820   Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein,
821   K. Eliceiri, P. Tomancak, A. Cardona, *Nat. Methods* **2012**, *9*, 676.

822

Figures and tables

| Negative cohort	IF IgM	IF IgA	IF IgG	ELISA IgA	ELISA IgG
B (n=105)	1	0	1	7	5
A (n=34)	0	0	1	3	1
Z (n=22)	0	0	0	2	0
E (n=57)	0	0	0	11	1
Total (n=218)	1 (0,5%)	0 (0,0%)	2 (0,9%)	23[a] (10,6%)	7[a] (3,2%)

Table 1: Summary of positive results for the negative control samples obtained by ELISA and IF. The classification of positive or borderline results in ELISA followed the definition of the test manufacturer. The classification in IF is described in materials and methods. Positive IgA and IgG ELISA readings were derived from the same sample. Cohort B = healthy donors, cohort A = patients that tested positive for ccCoV (all four types of ccCoV represented), cohort Z = patients with diagnosed *Mycoplasma pneumoniae*, cohort E = patient with diagnosed EBV or CMV infection. [a] – borderline values were considered positive.

days post symptom onset	IF IgM	IF IgA	IF IgG	ELISA IgA	ELISA IgG
< 11 (n=17)	7 (41%)	9 (53%)	7 (41%)	11 (65%)	3 (18%)
11-14 (n=24)	18 (75%)	19 (79%)	19 (79%)	19 (79%)	16 (67%)
>14 (n=16)	16 (100%)	16 (100%)	16 (100%)	16 (100%)	16 (100%)
Total (n=57)	42 (73%)	44 (77%)	42 (73%)	46 (80%)	34 (60%)

833 **Table 2: Positive results obtained for sera from COVID-19 patients collected at the**
834 **indicated days post symptom onset.**

835   Figure 1.

836
837
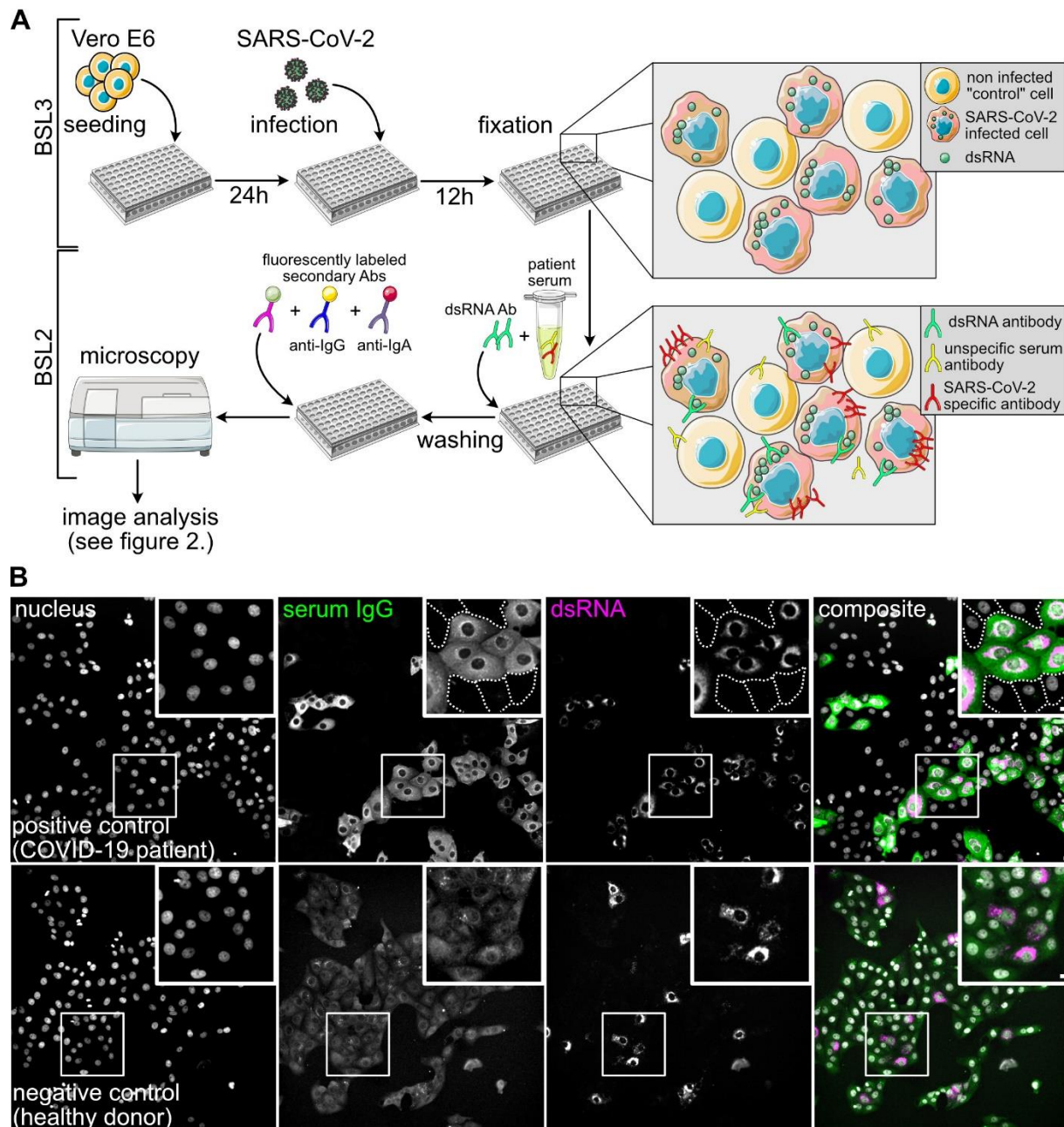
Figure 1: Principle of the immunofluorescence assay for SARS-CoV-2 antibody detection.

839   (**A**) Scheme of the IF workflow and the concept for SARS-CoV-2 antibody detection. (**B**)

840   Representative images showing immunofluorescence results using a COVID-19 patient serum

841   (positive control, upper panels) and a negative control serum (lower panels), followed by staining

842   with an AlexaFluor488-coupled anti-IgG secondary antibody. Nuclei (grey), IgG (green), dsRNA

843   (magenta) channels and a composite image are shown. White boxes mark the zoomed areas.

844   Dashed lines mark borders of non-infected cells which are not visible at the chosen contrast

845   setting. Note that the upper and lower panels are not recorded and displayed with the same

846   brightness and contrast settings. In the lower panels the brightness and contrast scales have

847   been expanded in order to visualize cells in the IgG serum channel where only background

848   staining was detected. Scale bar is 20 µm in overview and 10 µm in the insets.

849   Figure 2.
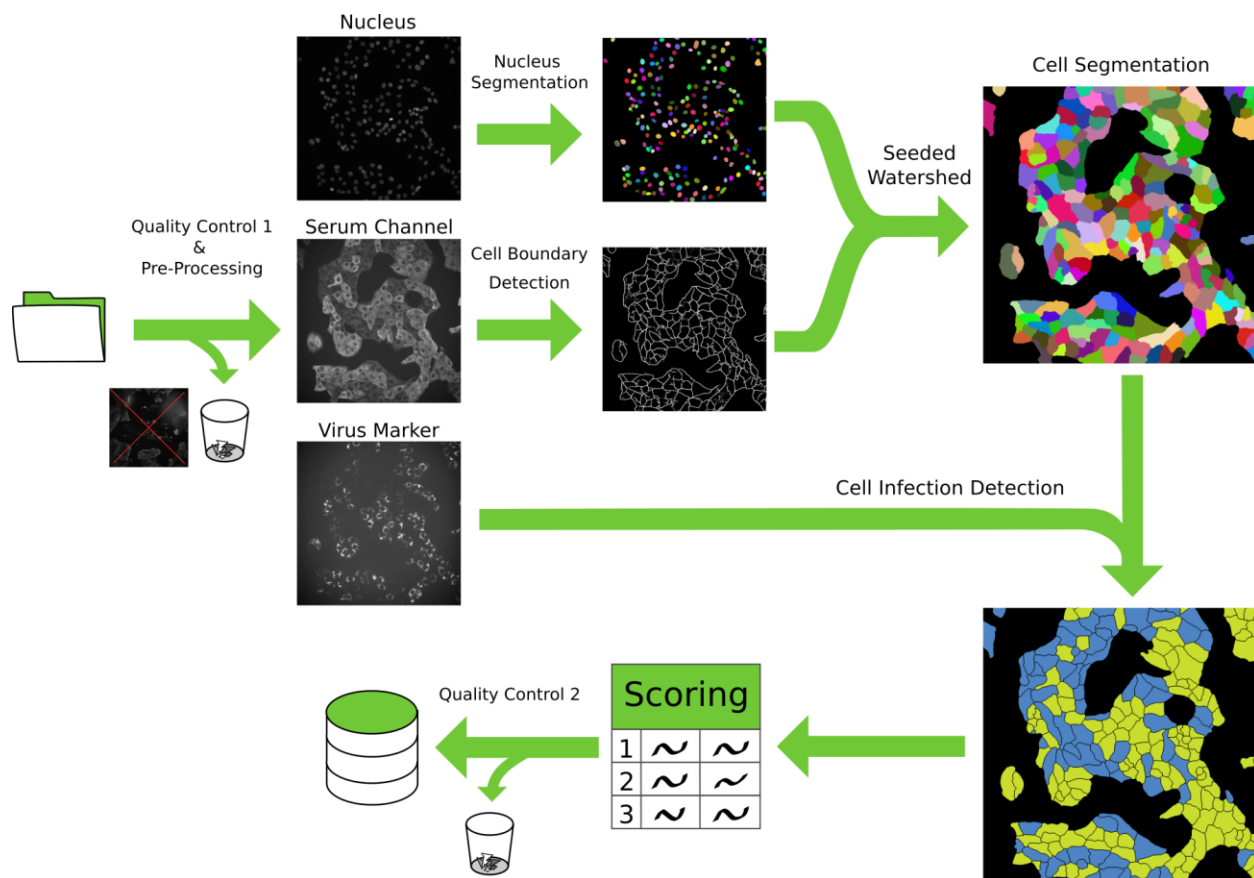
850

851

852   **Figure 2: Schematic overview of the image processing pipeline.** Initially, images are
853   subjected to the first manual quality control, where images with acquisition defects are
854   discarded. A pre-processing step is then applied to correct for barrel artifacts. Subsequently,
855   segmentation is obtained via seeded watershed, this algorithm requires seeds obtained from
856   StarDist segmentation of the nuclei and boundary evidence computed using a neural network.
857   Lastly, using the virus marker channel we classify each cell as infected or not infected and we
858   computed the scoring. A final automated quality control identifies and automatically discards
859   non-conform results. All intermediate results are saved in a database for ensuring fully
860   reproducibility of the results.

861   Figure 3.

862
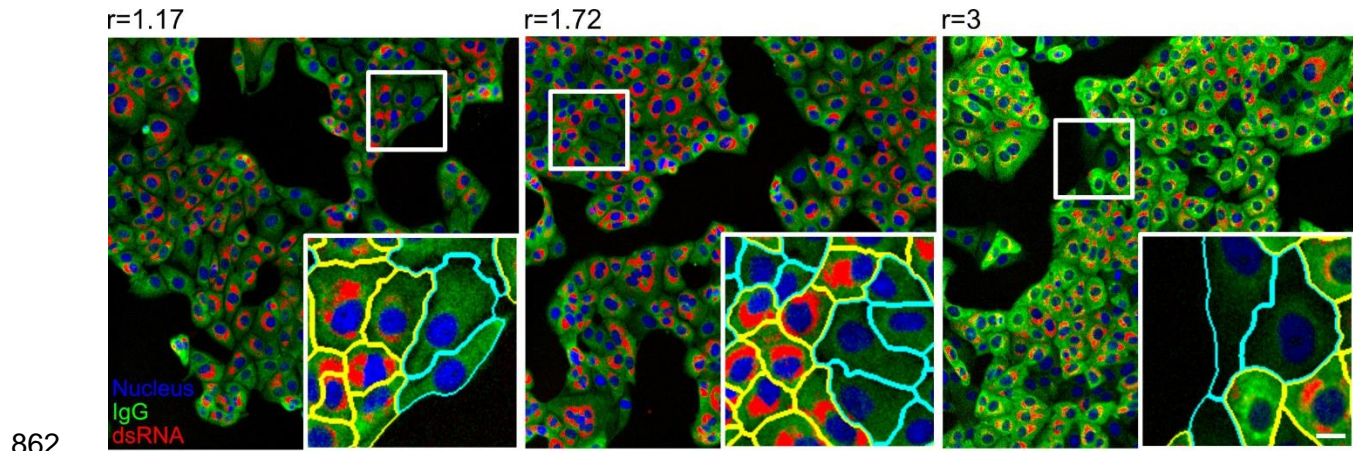
863

864   **Figure 3: Examples of results from the automated image analysis pipeline**. Panels display

865   images that correspond to three different ratio scores (ratio score is indicated above the image)

866   determined from samples stained with three different human sera, followed by staining with an

867   anit-IgG secondary antibody coupled to AlexaFluore488. Images represent overlays of three

868   channels - nuclei (blue), IgG (green) and dsRNA (red). White boxes mark the zoomed area.

869   Cells in the insets are highlighted with yellow or cyan boundaries, indicating infected and non-

870   infected cells, respectively. Scale bar = 10 μm.

871   Figure 4.

872   healthy donor   ccCoV pos   EBV pos   CMV pos   mycoplasma pos
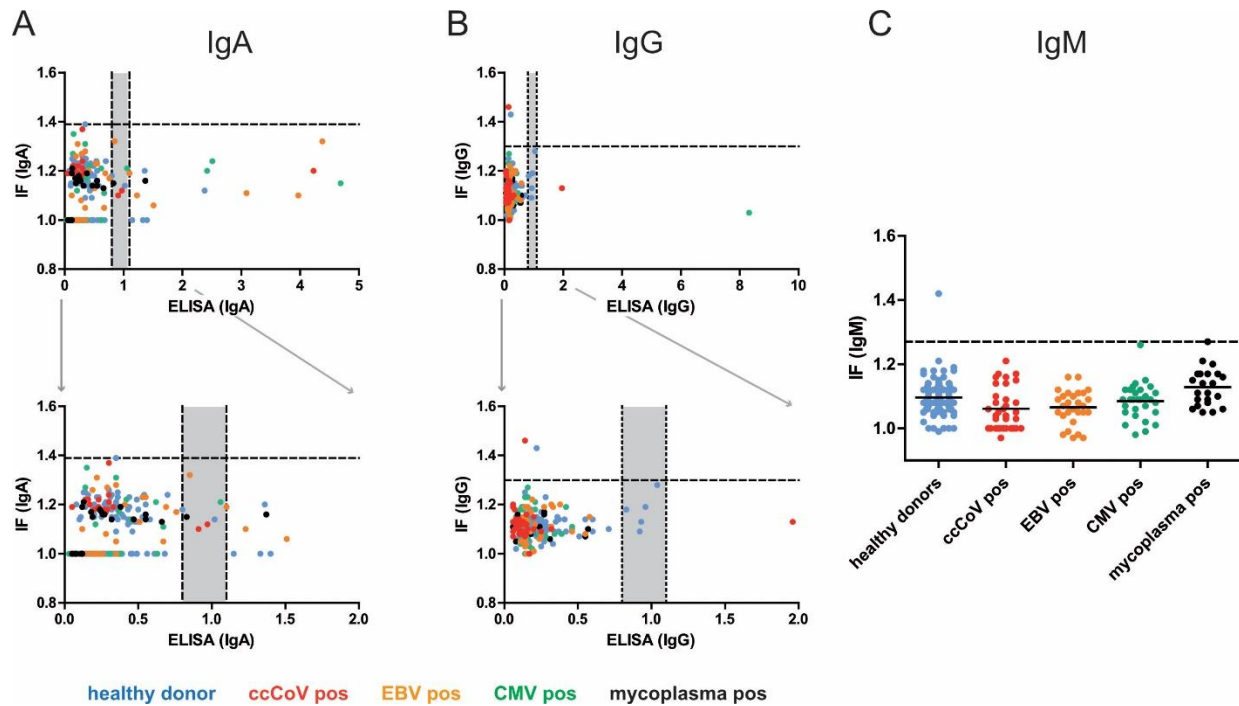
873

874   **Figure 4: Correlation between SARS-CoV-2 specific IF and ELISA results for the negative**
875   **control panel obtained in IgA (A) or IgG (B) measurements**. Each dot represents one serum
876   sample. Blue, healthy donors; red, ccCoV positive; green, CMV positive; orange, EBV positive;
877   black, mycoplasma positive. Bottom panels represent zoomed-in versions of the respective top
878   panel to illustrate the borderline region. **(C)** IgM values for the indicated negative control cohorts
879   determined by IF. Since a corresponding IgM specific ELISA kit from Euroimmun was not
880   available, correlation was not analysed in this case. In some cases, antibody binding above
881   background was undetectable by IF in non-infected as well as in infected cells, indicating low
882   unspecific cross-reactivity and lack of specific reactivity of the respective serum. In order to allow
883   for inclusion of these data points in the graph, the IF ratio was set to 1,0. Dotted lines indicate
884   the optimal separation cut-off values defined for sample classification, grey areas indicate
885   borderline results in ELISA.

886

887   Figure 5.
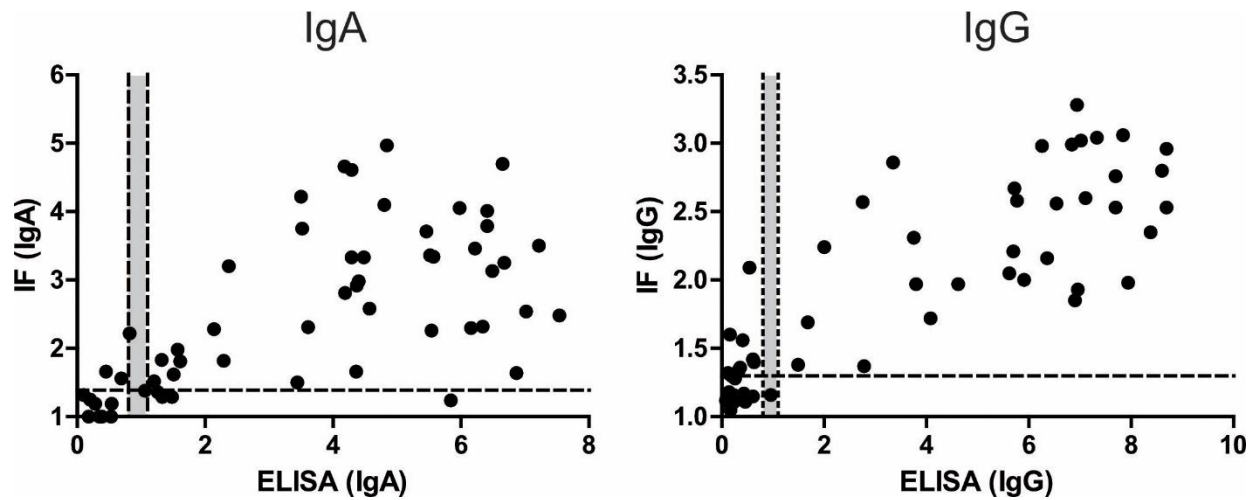
888

889

890   **Figure 5: Correlation between IgA or IgG values obtained by ELISA and IF for sera from**

891   **29 COVID-19 patients collected at different days post infection.** In some cases, antibody

892   binding above background was undetectable by IF in non-infected as well as in infected cells,

893   indicating low unspecific cross-reactivity and lack of specific reactivity of the respective serum.

894   In order to allow for inclusion of these data points in the graph, the IF ratio was set to 1,0. Dotted

895   lines indicate the cut-off values defined for classification of readouts, grey areas indicate

896   borderline values.

897

898   Figure 6.

899

900
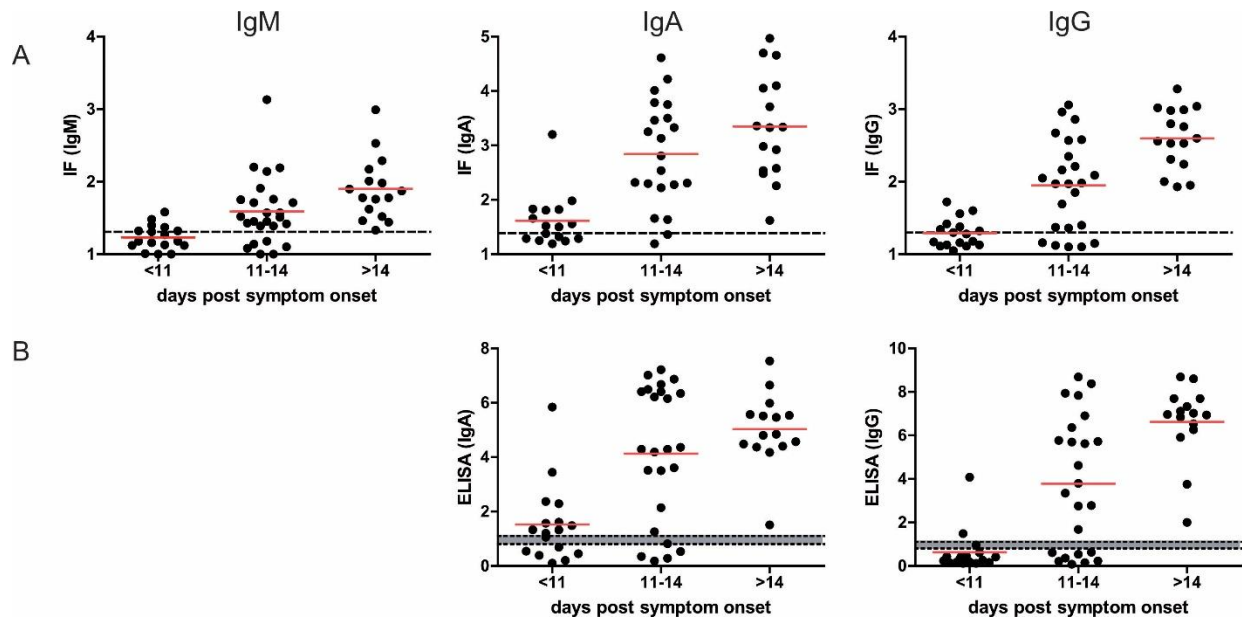
Figure 6: Detection of SARS-CoV-2 specific antibodies in sera from COVID-19 patients.
(**A**) Fifty-seven serum samples from 29 PCR confirmed patients collected at the indicated times post symptom onset were analysed by the IF workflow for the presence of SARS-CoV-2 specific IgM, IgA and IgG antibodies. Each dot represents one serum sample. Red line: mean value; dotted line: cut-off between negative and positive values. (**B**) The same samples as in A were analysed by ELISA for the presence of SARS-CoV-2 specific IgA and IgG antibodies. Each dot represents one serum sample. Red line: mean value; dotted lines: cut-off; grey zone: borderline.

908