

# Highly sensitive scent-detection of COVID-19 patients in vivo by trained dogs

Omar Vesga<sup>1,2\*</sup>, Maria Agudelo<sup>1,2</sup>, Andrés F. Valencia-Jaramillo<sup>2,3</sup>, Alejandro Mira-Montoya<sup>2,3</sup>, Felipe Ossa-Ospina<sup>2,3,4</sup>, Esteban Ocampo<sup>3</sup>, Karl Čiuoderis<sup>5</sup>, Laura Pérez<sup>5</sup>, Andrés Cardona<sup>5</sup>, Yudy Aguilar<sup>2</sup>, Yuli Agudelo<sup>1</sup>, Juan P. Hernández-Ortiz<sup>5,6</sup>, Jorge E. Osorio<sup>5,6</sup>.

<sup>1</sup> Section of Infectious Diseases, Hospital Universitario San Vicente Fundación, Medellín, Colombia.

<sup>2</sup> GRIPE, Universidad de Antioquia, Medellín, Colombia.

<sup>3</sup> Colina K-9, La Ceja, Colombia.

<sup>4</sup> Undergraduate School of Veterinary Medicine, Universidad de Antioquia, Medellín, Colombia.

<sup>5</sup> Colombia/Wisconsin One-Health Consortium, Departamento de Materiales, Facultad de Minas, Universidad Nacional de Colombia, Sede Medellín, Colombia.

<sup>6</sup> Department of Pathobiology, School of Veterinary Medicine, University of Wisconsin, Madison, USA.

\* Corresponding author

Email: [omar.vesga@udea.edu.co](mailto:omar.vesga@udea.edu.co)

## 32      **Abstract**

33

34      Timely and accurate diagnostics are essential to fight the COVID-19 pandemic, but no test  
35      satisfies both conditions. Dogs can scent-identify the unique odors of the volatile organic  
36      compounds generated during infection by interrogating specimens or, ideally, the body of a  
37      patient. After training 6 dogs to detect SARS-CoV-2 in human respiratory secretions (in  
38      vitro scent-detection), we retrained 5 of them to diagnose the infection by scenting the  
39      patient directly (in vivo scent-detection). Then, efficacy trials were designed to compare the  
40      diagnostic performance of the dogs against that of the rRT-PCR in 848 human subjects:  
41      269 hospitalized patients (COVID-19 prevalence 30.1%), 259 hospital staff (prevalence  
42      2.7%), and 320 government employees (prevalence 1.25%). The limit of detection in vitro  
43      was lower than  $10^{-12}$  copies ssRNA/mL. In vivo, all dogs detected 92 COVID-19 patients  
44      present among the 848 study subjects. Detection was immediate, and independent of  
45      prevalence, time post-exposure, or presence of symptoms, with 95.2% accuracy and high  
46      sensitivity (95.9%; 95% C.I. 93.6-97.4), specificity (95.1%; 94.4-95.8), positive predictive  
47      value (69.7%; 65.9-73.2), and negative predictive value (99.5%; 99.2-99.7). To determine  
48      real-life performance, we waited 75 days to carry out an effectiveness assay among the  
49      riders of the Metro System of Medellin, deploying the human-canine teams without  
50      previous training or announcement. Three dogs (one of each breed) scent-interrogated 550  
51      citizens who volunteered for simultaneous canine and rRT-PCR testing. Negative  
52      predictive value remained at 99.0% (95% C.I. 98.3-99.4), but positive predictive value  
53      dropped to 28.2% (95% C.I. 21.1-36.7). Canine scent-detection in vivo is a highly accurate  
54      screening test for COVID-19, and it detects more than 99% of infected individuals

55 independently of the key variables. However, real-life conditions increased substantially the  
56 number of false positives, indicating the necessity of training a threshold for the limit of  
57 detection to discriminate environmental odoriferous contamination from infection.

## 58 **Introduction**

59

60 The only effective measure to ameliorate the impact of the COVID-19 pandemic is early  
61 and accurate identification of people infected with SARS-CoV-2 [1]. A key aspect of any  
62 pandemic is that, in order to prevent contagion, diagnostic tests must detect the pathogen in  
63 asymptomatic, pre-symptomatic and symptomatic patients [2]. The reference standard is the  
64 real-time reverse transcriptase-polymerase chain reaction (rRT-PCR); it is highly specific  
65 (~100%), but lacks sensitivity during the first 5 days post exposure (0% on day 1, 33% on  
66 day 4, 62% on day 5), and its availability is limited [3]. Lateral flow antigen tests are cheap,  
67 instrument-free (easily accessible), provide results faster than rRT-PCR, and have ~100%  
68 specificity as well, but sensitivity is lost 5-7 days after exposure. However, antigen tests are  
69 very sensitive (84%-98%) during the pre-symptomatic period (days 1-5), complementing  
70 very well the insensitive period of the rRT-PCR [4]. Antibody tests are useless to prevent  
71 the dissemination of the virus, as they peak after the infectious period [5]. Several nations  
72 demonstrated eloquently that massive testing to diagnose early everyone infected, followed  
73 by immediate isolation in designated areas away from home, and rigorous contact-tracing,  
74 were the only measures that effectively stopped the pandemic [6]. Quarantines provide time  
75 to respond for health authorities, but benefit is doubtful [7], while cost is catastrophic [8].  
76 Vaccines offer the solution [9], but the immunization of the world's population will take

77 years, time enough for the virus to mutate and adapt [10]. Therefore, finding strategies to  
78 balance prevention with income is an emergency [11].  
79

80 Humans have been using dogs - *Canis lupus familiaris* - for scent-detection since the  
81 beginnings of domestication [12]. The great power of their sense of smell is exceedingly  
82 useful, and the first study of their olfactory capabilities was published more than 130 years  
83 ago by George J. Romanes [13], the research associate of Charles Darwin. Today, highly  
84 trained dogs are invaluable not only for their service [14], but also because their accuracy is  
85 definitely superior to analytical instruments [15]. Medical diagnosis by trained canines  
86 triggers hope and enthusiasm among journalists [16], but it receives no attention from  
87 practicing clinicians, who rely exclusively on semiology and sophisticated instruments to  
88 determine what afflicts their patients [17]. The use of scent-specialized dogs to detect  
89 specific conditions has been published, but most are anecdotic reports instead of formal  
90 protocols designed to validate a diagnostic test for clinical use [18]. However, a few studies  
91 have demonstrated that with appropriate training and strict adherence to the scientific  
92 method, it is possible to obtain consistent results [19]. Recently, a comprehensive method  
93 was published to validate canine diagnosis of the plant pathogens *Candidatus Liberibacter*  
94 *asiaticus* [20] and *Xanthomonas citri* pv. *citri* [21], demonstrating that detection of infected  
95 citrus trees by dogs was superior to quantitative PCR.  
96

97 Dogs detect and differentiate unique odors that result from the emission of volatile organic  
98 compounds (VOCs) that constitute the “smell print” of the target [22]. In the case of SARS-  
99 CoV-2, ethyl butanoate was reported recently as the most abundant VOC in the breath of 10  
100 COVID-19 patients [23]. Dogs are inherently resistant to SARS-CoV-2 [24], and even in

101 experimentally induced infections or after very close contact with an infected owner, the  
102 virus cannot replicate in, cause clinical disease to, or be transmitted from canines [25].  
103 After proper training of six dogs, we compared canine diagnostic performance against the  
104 reference standard to determine the sensitivity (*SEN*), specificity (*SPC*), positive predictive  
105 value (*PPV*), negative predictive value (*NPV*), accuracy (*ACC*), and likelihood ratio (*LR*) of  
106 our dogs to detect by scent COVID-19 *in vivo*, i.e., by direct olfaction of the patient. The  
107 product was a very fast, reliable, and cost-effective screening method for infection by  
108 SARS-CoV-2 in human patients.

109

## 110 Materials and Methods

111 Detailed methodology is available in the Supporting Information file.

### 112 Study Objectives

113 Although any healthy dog can sniff and follow an odor, it cannot be forced to do it. Scent-  
114 detection demands intense concentration, and it exhausts the dog mentally and physically.  
115 For maximal output, working dogs must be rewarded for each positive finding with a prize  
116 that conveys an extremely high value for them, and their performance depends heavily on  
117 the intensity of the expectations that such reward generates in the brain of the dog during  
118 training [26, 27]. Even within optimal training conditions, not all canine individuals will  
119 give their best to gain a reward, and a rigorous selection process is needed before training  
120 canines for delicate missions, like explosive-detectors and military working dogs [28]. For  
121 instance, the Bloodhound is a working breed that excels in scent power, but it is not  
122 suitable for medical detection because most individuals do not enjoy the repetitive task of  
123 searching and signaling one specific odor among many human subjects [29]. Actually,

124 medical screening can be a very boring task for the dog, and most experts agree that the  
125 ideal dog for scent-detection must display superlative amounts of motivation, stamina,  
126 determination, and resilience [14, 15]. Animal behavior scientists avoid this  
127 anthropomorphic jargon, but cognition research is providing solid evidence that dogs  
128 indeed have unusual minds compared with other nonhuman animals, and that such  
129 advantage comes from sharing their lives with us [30].

130

131 The principal objective of this study was to determine the performance of scent-detection  
132 dogs as a screening tool *in vivo* for immediate detection of COVID-19 patients under a  
133 variety of circumstances [31]. The design aimed at answering five research questions: One,  
134 if working dogs with the above-mentioned attributes, but belonging to breeds created for  
135 tasks other than olfaction, would succeed as medical detectors; a positive result would  
136 increase significantly the canine population from which dogs could be selected. Two, the  
137 minimal number of COVID-19 patients required to train the dogs *in vitro*; such number  
138 must be enough for the dogs to make the inference that any human being with the same  
139 smell-print is a positive. Three, the size effects of the diagnostic metrics *in vitro* and *in vivo*  
140 under controlled experimental conditions, i.e., efficacy. Four, the canine limit of detection,  
141 in copies of single stranded viral RNA per milliliter (ssRNA/mL). And five, the real-life  
142 performance of the dogs *in vivo*, i.e., screening effectiveness.

143

#### 144 **Design and sample size**

145 Screening and diagnostic tests differ in their applications, and validation of the former  
146 requires a much smaller sample size [32]. In order to produce dogs capable of screening  
147 humans for SARS-CoV-2 infection, teaching them to identify the virus *in vitro* was a

148 mandatory prerequisite. The first step was to ask written informed consent from 12 patients  
149 hospitalized with COVID-19 to aliquot and ultra-freeze (-70°C) their respiratory secretions.  
150 The specimens were thawed and used as needed. Demographic information of the  
151 participants who provided specimens for phases 1 and 2 can be seen in Table 1. The  
152 training work was planned in three phases, each followed by its corresponding  
153 experimental aspect (Fig 1). Phase 1 (“in vitro recognition) lasted 28 days during which we  
154 trained the dogs to recognize in vitro the scent-print of SARS-CoV-2 under a wide variety  
155 of environmental modifications. One aspect of training that remained constant was the  
156 error-free discrimination learning protocol developed by Terrace [33], which consists in  
157 always presenting the animal a marked contrast between positive and negative stimuli [34].  
158 The “stimulus” is the problem presented to the dog, which was, for in vitro diagnosis,  
159 sterile saline solution (phase 1) or human saliva (phases 2), and for in vivo screening, the  
160 body of a person. A stimulus can be “positive” when it leads to a reward (SARS-CoV-2), or  
161 “negative”, if it does not represent a reward for the dog (controls). To recognize SARS-  
162 CoV-2, we trained the dogs to find their food (the reward) using their olfaction, always  
163 hiding with it a respiratory specimen from Patient 1 (the positive stimuli). The amount of  
164 food was diminished progressively until only the SARS-CoV-2 specimen was left in the  
165 hiding place, while the number of hides with saline increased in number. We marked the  
166 correct behavior (i.e., identification of SARS-CoV-2) pressing a clicker device, and  
167 immediately rewarded the dog. It took one day for all dogs to understand that finding the  
168 SARS-CoV-2 specimen meant a prize for them, and that saline conveyed no reward. The  
169 following 27 days of the first phase the dogs were trained with respiratory secretions from  
170 Patients 1, 2 and 3 under the above-mentioned variations, but keeping constant the negative  
171 stimuli (saline). The other 9 positive specimens (Patients 4-12) were reserved exclusively

172 for experimentation, which only took place after training in each of the first two phases had  
173 taught the dogs the error-free skills necessary to identify SARS-CoV-2 with Patients 1-3.

174 **Table 1.** Human subjects who provided specimens for in vitro training and  
175 experimentation (phases 1 and 2).

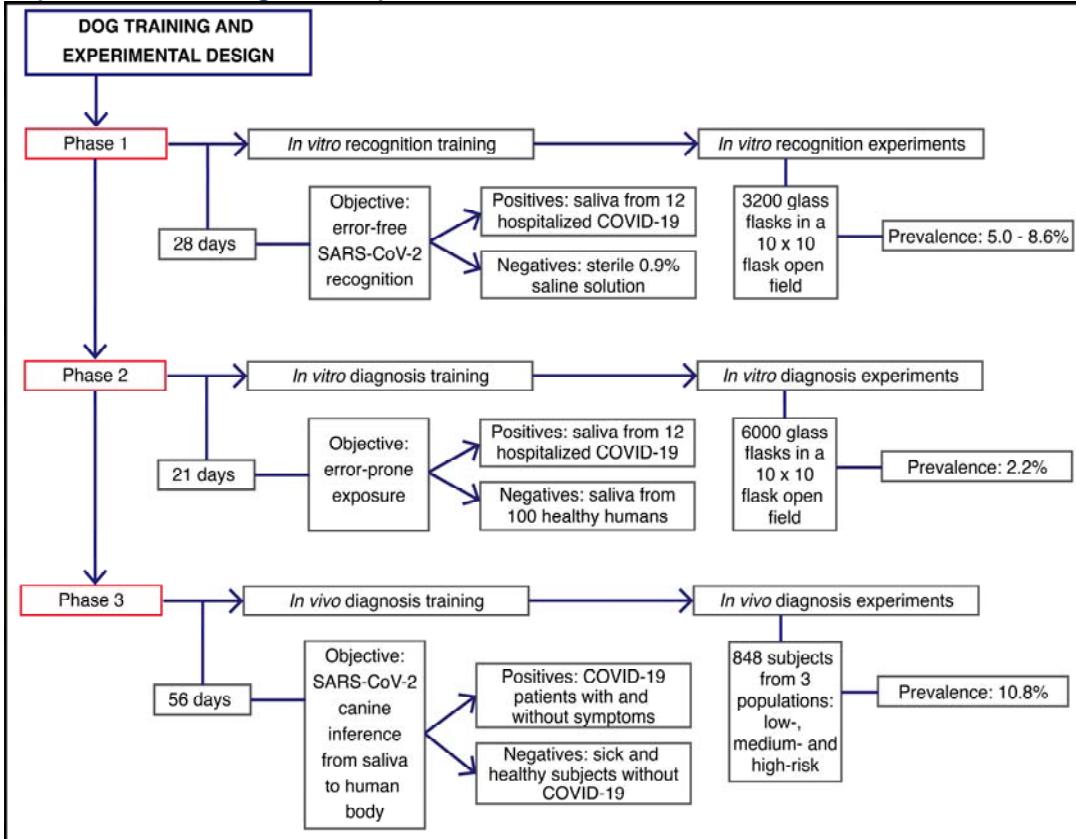
176

Patient #	Sex	Specimen	Days Sick	SARS-CoV-2 rRT-PCR	Viral Load ( $\log_{10}$ copies ssRNA/mL)
1	Female	NPS & saliva	12	Positive	5.42
2	Male	NPS & saliva	10	Positive	ND
3	Female	NPS & saliva	5	Positive	6.90
4	Male	TA	16	Positive	5.15
5	Male	NPA	10	Positive	ND
6	Male	TA	10	Positive	5.09
7	Male	NPS	3	Positive	10.2
8	Female	Sputum	7	Positive	ND
9	Female	Sputum & saliva	5	Positive	ND
10	Female	Sputum & saliva	9	Positive	5.07
11	Male	Saliva	8	Positive	ND
12	Male	TA	13	Positive	ND
13-112	59F, 41M	Saliva	0	Negative	NA

177 NPS: nasopharyngeal swab; TA: tracheal aspirate; NPA: nasopharyngeal aspirate; F: female; M:  
178 male; R: range; ND: not determined; NA: not applicable.

179

180 **Fig 1.** Efficacy studies. Flow chart depicting the order in which training phases and  
181 experimental design were planned.



182  
183 The number of days after each phase indicate the time employed training the dogs, before starting  
184 the corresponding experimentation process. COVID-19 prevalence was set up as desired for in vitro  
185 experiments, introducing a more difficult scenario by minimizing prevalence during phase 2 (in vitro  
186 diagnosis). Prevalence during phase 3 (in vivo diagnosis) was spontaneous, given by the pandemic  
187 epidemiology in our hospital, which is a large reference institution.  
188

189 During phase 2 (“in vitro diagnosis”), we trained the dogs for 21 days keeping constant the  
190 positive stimuli (specimens from Patients 1, 2 and 3), but changing the negative stimuli for  
191 human saliva specimens donated by 100 human volunteers (Patients 13-112, Table 1).  
192 These 100 donors were healthy, ambulatory people belonging to the general population,  
193 whose saliva specimens were negative for SARS-CoV-2 by rRT-PCR the same day that we  
194 aliquoted and froze them at -70°C (collection took place in March 2020, when the  
195 pandemic was just starting in Colombia). To prevent replication of the microbiota within  
196 each saliva sample, working specimens were thawed and kept at 4°C between uses, and

197 multiple aliquots of each specimen frozen at -70°C to replace discarded samples, granting  
198 fresh samples as needed (specimens were heat-sterilized before appropriate disposal).  
199 Training in phase 3 was undertaken only after experimentation demonstrated that the dogs  
200 had acquired the skills provided by phase 2.

201

202 In the third phase of training (“in vivo screening”, 56 days), the dogs learned to identify  
203 COVID-19 patients by scenting the human body; they always preferred the hands, then  
204 search other parts of the anatomy. The rRT-PCR was the reference standard, but we also  
205 included 51 seriously ill COVID-19 patients diagnosed with antigen tests and admitted to  
206 the hospital with life-threatening respiratory distress. The sample sizes required (and  
207 obtained) for sensitivity during experimental phases 1, 2 and 3 were 2140 (3200), 2140  
208 (6000) and 310 (848), respectively; specificity requirements were much lower (Table S1)  
209 [35]. The protocol was approved by the Research Ethics Committee of *Hospital*  
210 *Universitario San Vicente Fundación*, and all human participants gave written informed  
211 consent before enrollment.

212

## 213 **Dog training**

214 Using operant conditioning based on clicker-training and rewarding with food [36], six  
215 canines were trained to detect the odor print of SARS-CoV-2 in saliva and in the human  
216 body (Fig 2): four Belgian Shepherd Malinois (a herding breed), one first-generation cross  
217 Alaskan Malamute by Siberian Husky (a Nordic sled-dog), and one pit bull (a fighting  
218 breed). For every experiment, the position of the samples in the field (1 to 100) and disease  
219 prevalence (1% to 10%) were randomized with a mobile phone app. For each in vitro  
220 experiment, the dogs went through an open field arrangement of 10 x 10 samples (100)

221 distanced 2 m in all directions. An illustration of the experimental field and the scent-  
222 detection work in vitro can be seen in Video S1.

223 **Fig 2. Pictures and identification of the six dogs trained for the scent-detection of**  
224 **SARS-CoV-2.**

225



226 (1) Andromeda, intact female, 6-mo, Belgian Malinois (BM). (2) Nina, intact female, 25-mo, BM, (3)  
227 Niño, castrated male, unknown age, American Pit Bull Terrier. (4) Timo, intact male, 31-mo, BM. (5)  
228 Vika, intact female, 36-mo, BM. (6) Vita, intact female, 36-mo, first generation Alaskan Malamute x  
229 Siberian Husky.

230

231 Generalization (i.e., the dog recognizes the scent-print of SARS-CoV-2 in a specimen from  
232 any infected individual) in phases 1 and 2 was achieved with a small set of three COVID-  
233 positive samples (Patients 1 to 3; Table 1). To demonstrate that the dogs knew that they  
234 were looking for the scent-print of SARS-CoV-2 instead of the specific scents from the

235 three individuals used to train them, we used for experimentation after phases 1 and 2 the  
236 specimens from Patients 4-12, to which they had never been exposed. The same principle  
237 was applied to in vivo screening: we tested the dogs in patients only after obtaining very  
238 high diagnostic metrics in vitro. Dogs could scent any part of the anatomy and were  
239 allowed to touch with their noses the body of the patients, who were instructed to present  
240 their hands opened with palms facing the dog. The relatively long time we spent in training  
241 for phase 3 (56 days) had the only objective of eliminating false positives, because  
242 recognition of true positives was straightforward. The dogs trained with 400 subjects who  
243 did not participate in the experiments: 100 hospitalized patients (40% with COVID-19) and  
244 300 health-care workers (7% with COVID-19). These training was done to improve  
245 specificity and positive predictive value, because sensitivity and negative predictive value  
246 never represented an obstacle. Our focus was in vivo screening, and we used in vitro  
247 training and experimentation as means to arrive to our main goal, the canine detection of  
248 COVID-19 by scenting the human body. To ensure that the dogs were not fixed on the  
249 hospital scent, we obtained saliva samples from each of the 300 health-care workers who  
250 volunteered for in vivo training, and made the dogs scent-interrogate saliva once they had  
251 finished each training session in vivo: *NPV* in vitro was close to 100%, and *PPV* were  
252 above 80%, indicating that they were not looking for a scent other than SARS-CoV-2.  
253

254 **Dog-trainer teams biosafety: evaluation of the SARS-CoV-2  
255 containment devices**

256 To prevent contagion of canine and human individuals participating in this study with  
257 SARS-CoV-2 specimens, we contrived two devices (D1 and D2) made with the fabric of  
258 the Dupont<sup>TM</sup> Tychem 2000. D1 was a used for scent-detection in saliva or respiratory

259 specimens; it was a 130-mL glass flask with a metallic lid in which we perforated a 1 cm  
260 hole in the middle, and the lid allowed a hermetic closure that remained intact after placing  
261 a 10x10 cm piece of Tychem 2000 between the bottle and its lid. D2 was used for the same  
262 purpose but offered greater versatility than D1; it was a waterproof bag made of two 18x8  
263 cm pieces of Tychem 2000, heat-sealed in its four sides after placing inside it a sterile  
264 gauze impregnated with the specimen.

265

266 In order to determine if any of the dog-trainer teams got infected by, or could have been at  
267 risk of exposure to SARS-CoV-2 during the project, were used two approaches: one,  
268 running rRT-PCR tests in saliva of dogs and trainers by at the end of the second and third  
269 phases; and two, evaluating experimentally the efficiency of our containment devices in the  
270 Syrian hamster (*Mesocricetus auratus*) COVID-19 model (Fig S1). After an acclimatization  
271 period of 4 weeks, we exposed during four days 15 animals of both sexes (6 females and 9  
272 males) to SARS-CoV-2 in groups of 3 animals of the same sex, 2 groups for the  
273 experimental arms and 3 groups as control arms, each contained in a HEPA filtered One  
274 System cage. Each of the two experimental groups had inside their respective cage one of  
275 the containment devices (D1 in group 1, D2 in group 2) protected by a metallic welded wire  
276 mesh enclosure that allowed hamsters to smell the contraption without touching it. Each of  
277 the three control groups had free access to an unprotected D1 flask (group A), a sterile  
278 gauze impregnated with a fresh specimen from a different COVID-19 patient (group B), or  
279 and unprotected D2 bag (group C). D1, D2 and the virus-impregnated gauze were replaced  
280 with fresh SARS-CoV-2 specimens every 12 hours in the 5 groups. All hamsters were  
281 sampled for rRT-PCR by saliva swabs before and after SARS-CoV-2 exposure.

282

## 283 **Experimentation after scent-detection training**

284 Three kinds of 2-mL specimens were prepared under a biosafety class III laminar flow  
285 cabinet using 209 sterile, scent-free flasks. One-hundred flasks had 0.9% sterile saline  
286 solution (phase 1, control arm), 100 had rRT-PCR-negative saliva (phase 2, control arm),  
287 and 9 flasks had respiratory secretions from COVID-19 Patients 4-12 (phases 1 and 2,  
288 experimental arm). The positive specimens were diluted (1:1 volume) in 0.9% sterile saline  
289 solution to preserve the virus [37]. Each of the three phases of training was followed by its  
290 respective experiments; phases 1 and 2 involved in vitro scent-detection of COVID-19 in  
291 saliva specimens, and phase 3 concerned in vivo screening. During in vitro experiments,  
292 each dog had to interrogate by scent a field with 100 flasks, the vast majority (90%-99%)  
293 containing negative stimuli; the rest would have the positive stimuli. After finishing a 100-  
294 flask field, the dog was offered abundant water and placed to rest in its individual kennel.  
295 Before the next search, each dog was scheduled to have an unrestricted play session and to  
296 take a long walk with its trainer. Once ready, the field was rearranged for a new experiment  
297 changing at random the position of the specimens and the prevalence of COVID-19.

298

299 Although the majority of the experiments related to phase 1 training were unblinded, all  
300 experiments for phases 2 and 3 were blinded, i.e., the dog handlers did not know the  
301 position and number of positive specimens (phase 2) or patients (phase 3). Such  
302 information was managed exclusively by the training director (AFVJ), whom actioned the  
303 clicker immediately upon any correct indication by the canine, delivering an unequivocal  
304 auditory signal that informed the human-canine team of every correct finding. This design  
305 implied that the training director never interacted with the dogs during the experimental  
306 process.

307

308 Immediately after ending experimentation for phase 2, we started phase 3 training at  
309 *Hospital Universitario San Vicente Fundación* (HUSVF). Phase 3 experiments aimed to  
310 determine the efficacy of the dogs as a COVID-19 screening tool *in vivo*, and included  
311 inpatients from HUSVF (high-risk group), health-care workers serving at HUSVF  
312 (intermediate-risk group), and employees from the office of the Governor of the  
313 Department of Antioquia (low-risk group).

314

315 The effectiveness assay was executed 75 days after the last experiment of the efficacy trial,  
316 and involved general population riding the Metro System of Medellin ( $n = 550$ , 3 dogs of 3  
317 breeds). In contrast with efficacy trials, the human-canine teams were deployed to the field  
318 without previous announcement or environmental training; researchers wear personal  
319 protective equipment, but did not know the clinical status or particular risk factors of the  
320 participants, who were informed about the test on site and recruited without further delay.  
321 Since we had only three trainers, it was impossible to include more than 3 dogs in this trial.

322

323 Data input into 2x2 contingency tables generated the metrics *SEN*, *SPC*, *PPV*, *NPV* (each  
324 with its 95% confidence interval), *ACC*, and *LR*. Since pooling results from experiments  
325 with 100 specimens violates the independence assumption of the Fisher's exact test, we  
326 performed latent class analysis for *in vitro* data. For *in vivo* data, where all assumptions  
327 were rigorously checked, we applied the two-tailed Fisher's Exact Test to challenge the null  
328 hypothesis that the dogs detected COVID-19 by chance.

329

### 330 **Limit of canine scent-detection**

331 Freshly collected saliva specimens from four COVID-19 patients (unknown to the dogs)

332 were serially diluted in sterile physiologic saline solution in 1:10 steps down to  $1 \times 10^{-12}$

333 copies ssRNA/mL. The viral loads had been determined beforehand and ranged from 47 to

334 475 copies ssRNA/mL, therefore 15 dilutions were prepared for each of the four specimens.

335 Then, we randomized the dilutions from each patient by placing two COVID-19 dilutions

336 along with 8 saline controls in D1 contraptions (10 flasks per row), and commanded every

337 dog to search them until they finished the scent-interrogation of all 60 dilutions. The limit

338 of detection (LOD) was the mean of the most diluted specimens that each dog was able to

339 identify without failing a single one of the more concentrated dilutions.

340

### 341 **rRT-PCR assay and RNA quantification, RNA transcript 342 standard generation, assay efficiency, and analytical sensitivity**

343 The SARS-CoV-2 molecular diagnosis was conducted at the Genomic One Health

344 Laboratory, Universidad Nacional de Colombia. Details provided in Supporting

345 Information file.

346

## 347 **Results**

348 Detailed methodology is available in the Supporting Information file.

349

### 350 **Phase 1: in vitro recognition of SARS-CoV-2**

351 Once the dogs were fully proficient identifying the specimens from Patients 1, 2 and 3

352 without errors, we evaluated their performance experimentally with specimens from

353 patients to which they had not been exposed (Patients 4 to 12, Table 1). In phase 1, the  
354 number of experiments varied for each dog because the required sample size (3200) was  
355 reached early, all six recognized COVID-19 specimens with accuracy >95.0%, and we tried  
356 to minimize the dog's exposure to the specimens from each of the 9 patients. The mean  
357 prevalence of SARS-CoV-2 positive samples for these experiments was 7.56% (range,  
358 5.0%-8.6%), and the effect sizes were very high for each of the six dogs (Table S2). To  
359 determine if diagnostic performance would improve by increasing prevalence to 20%, we  
360 set up an experiment with 40 flasks in a 10 x 4 field allocating randomly 8 positive samples  
361 within 32 saline distractors. All six dogs identified correctly every sample without a single  
362 mistake. With these results, dogs were ready for phase 2 training, designed for lower  
363 prevalence (1%-4%) and greater difficulty to discriminate the positive from the negative  
364 stimuli (saliva from 100 non-COVID subjects).

365

## 366 **Phase 2: in vitro diagnosis of SARS-CoV-2**

367 After phase 2 training with specimens from Patients 1-3 (positive stimuli) and 13-112  
368 (negative stimuli) under a prevalence no greater than 4%, diagnostic performance was  
369 determined experimentally with saliva from Patients 4-12 (experimental arm) disguised at  
370 an average prevalence of 2.2% among saliva samples from healthy volunteers 13-112  
371 (control arm). Compared with phase 1, there was a significant improvement in the effect  
372 sizes of all metrics for every dog (Table S3). As a group, the 6 dogs achieved *SEN* 95.5%  
373 (95% C.I. 90.4 - 97.9), *SPC* 99.6% (99.5 - 99.8), *PPV* 85.7% (79.2 - 90.5), *NPV* 99.9%  
374 (99.8 - 100), *ACC* 99.6%, and *LR* 267. The *PPV* improved 12 percentile points, while the  
375 *NPV* was close to perfection, thereby suggesting a very low probability that any of our dogs  
376 would miss a positive case in vitro (Fig S2).

377

378 **Phase 3 (efficacy trial): in vivo diagnosis of SARS-CoV-2 by  
379 direct body-scenting**

380 When the dogs interrogated the first positive patient, all six recognized the scent-print of  
381 SARS-CoV-2 and went down without hesitation, identifying COVID-19. Further training  
382 was needed to teach the dogs that the target odor should not be marked as positive when  
383 found in fomites like beds, chairs, bedside tables, flip flops, cell phones, or other personal  
384 items of the patient, but it proved to be a difficult endeavor. Contrary to false negatives,  
385 always close to zero, false positives represented a serious obstacle during phase 3 training.  
386 This response from the dogs proved that saliva from just 12 COVID-19 patients provided  
387 enough variety for them to infer that the scent-print of the disease could be found in vivo  
388 (as they had learned in vitro), and that marking patients or fomites (contaminated with the  
389 target scent-print) as positive was the behavior we were asking from them. That inference  
390 was our main training goal, but it was evident from day 1 that all training during phase 3  
391 had to focus in reducing false positives, because the scent-print of COVID-19 was spread  
392 throughout the surfaces of HUSVF, and the dogs detected it everywhere. To determine the  
393 reason, we moved training to the Emergency Room (ER), which is not only the port of  
394 entry for all patients, but also the busiest and most crowded service of our institution. In  
395 hospitalization wards we obtained the rRT-PCR results ahead of the training session, but it  
396 was impossible in the ER because patients were admitted the same day of training and  
397 results came afterwards. In three independent training sessions with 50-60 patients each, we  
398 found that 10% of the ER population was made of asymptomatic COVID-19 patients that  
399 had passed undiagnosed after being admitted for other reasons, mostly traumatic injuries.  
400 These patients were being admitted in the different surgical and medical wards,

401 contaminating all surfaces and infecting other patients, explaining the universal presence of  
402 the COVID-19 scent-print. This situation forced to change our training goal to teach the  
403 dogs that surfaces and fomites were not to be marked as positive, but at the cost of great  
404 confusion for the dogs.

405

406 Once the training goals were achieved, we started phase 3 experimentation including 5 dogs  
407 (Vika was excluded due to advanced pregnancy) and 848 human subjects belonging to  
408 three groups: 269 hospitalized patients (high risk group), 259 hospital staff individuals  
409 (intermediate risk group), and 320 government employees (low risk group). Experimental  
410 conditions were closely controlled for the three groups, which demographics are described  
411 in Table 2.

412

413 **Table 2.** Phase 3: in vivo screening, efficacy trial. Demographic and clinical  
414 characteristics of 848 participants in the scent-detection experiments.  
415

Variable		n (%)
Sex	All Participants	848 (100)
	Females	514 (60.6)
	Males	334 (39.4)
Age (years-old)	Median	56
	Mean	53
	Youngest	15
	Oldest	92
COVID-19 Prevalence	All Participants	92 of 848 (10.85)
	Government Employees	4 of 320 (1.25)
	Health-Care Workers, HUSVF	7 of 259 (2.70)
	Hospitalized Patients, HUSVF	81 of 269 (30.1)
COVID-19 Status	SARS-CoV-2 positive	92 (10.85)
	SARS-CoV-2 negative	753 (88.8)
	SARS-CoV-2 indeterminate	3 (0.35)
Result by Reference Standard	rRT-PCR positive	41 (4.83)
	rRT-PCR negative	753 (88.8)
	rRT-PCR indeterminate	3 (0.35)
	Antigen positive	51 (6.01)
	Antigen negative	0 (0)
Clinical Status at K9 Test	COVID-19, asymptomatic	18 (2.12)
	COVID-19, pre-symptomatic	0 (0)
	COVID-19, symptomatic	74 (8.73)
	Not COVID-19, but sick	188 (22.2)
	Not COVID-19, healthy	565 (66.6)
	Indeterminate, asymptomatic	2 (0.24)
	Indeterminate, symptomatic	1 (0.12)

416 HUSVF: Hospital Universitario San Vicente Fundación.  
417

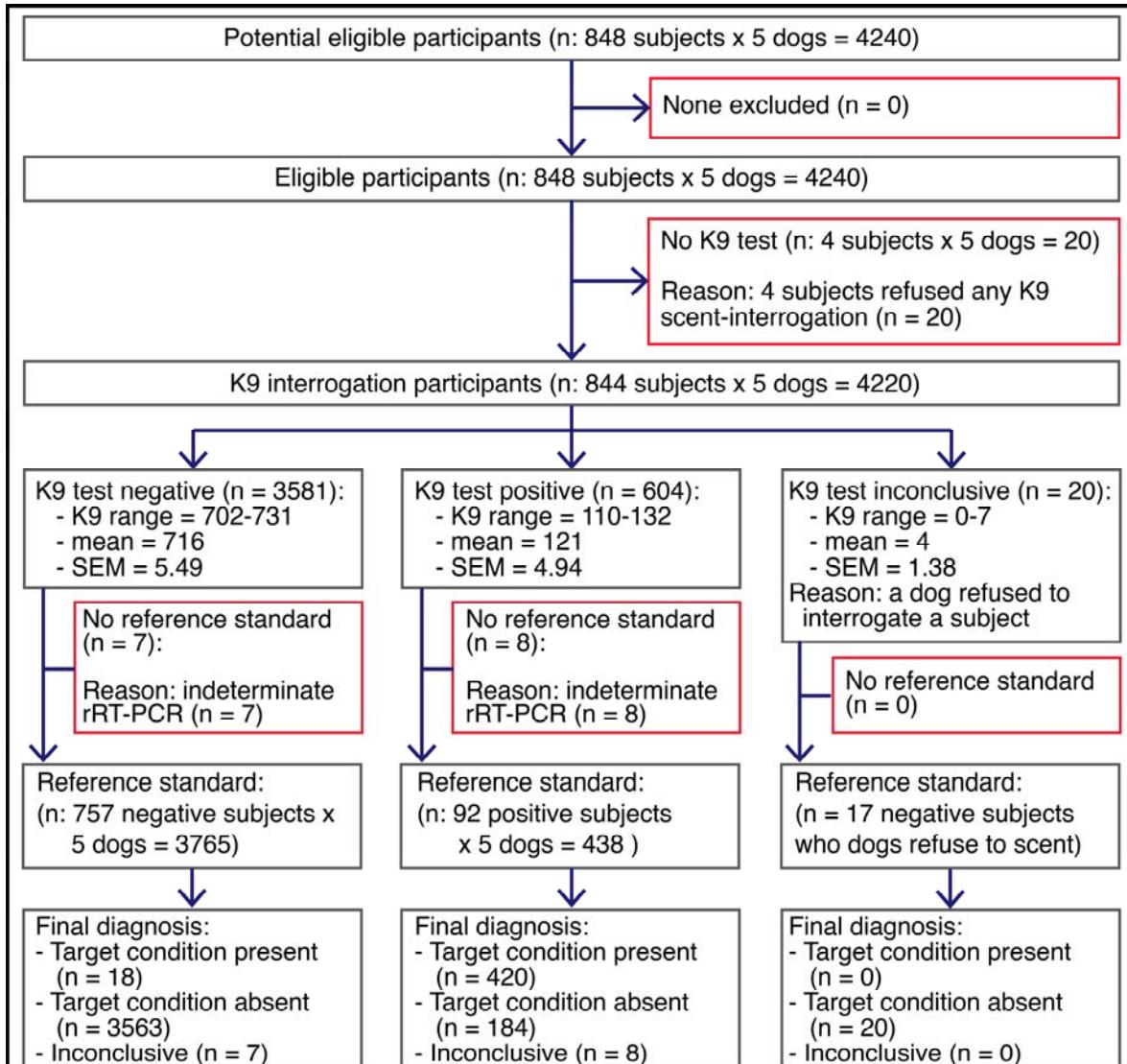
418 Before canine scent-interrogation, we sampled the 848 participants to determine their  
419 COVID-19 status by molecular and antigen testing (Fig 3). COVID-19 was confirmed in 92  
420 patients (10.85%) and discarded in 753 (88.8%). The other 3 (0.35%) were asymptomatic  
421 subjects with “indeterminate” rRT-PCR results after repeated testing, and had to be  
422 excluded from the analysis (Fig 4). The average cycle threshold ( $C_T$ ) of the rRT-PCR  
423 positive patients was 32.0 (range, 20.6-38.8). COVID-19 was diagnosed by antigen test  
424 (Standard Q COVID-19 Ag Test, SD Biosensor) in 51 patients that had bee admitted to the

425 ER with acute respiratory distress, fever, sinus pain, cough, anosmia, or dysgeusia. Out of  
426 753 COVID-19 negative patients, 188 were hospitalized for other diseases that included  
427 respiratory conditions (23%, half had bacterial infections), malignancy (19%),  
428 autoimmunity (8%), coronary or peripheral atherosclerosis (7%), diabetes mellitus (6%), or  
429 chronic osteomyelitis (6%), and the rest had traumatic injuries, peritonitis, HIV, or  
430 cholangitis, among other pathologies. Of note, the dogs did not mark as positive any of the  
431 patients with respiratory diseases other than COVID-19. The prevalence of COVID-19 was  
432 10.85% for the study population (92 of 848 subjects), distributed this way based on pre-test  
433 risk: 30.1% (81 of 269), 2.70% (7 of 259), and 1.25% (4 of 320) for the high, intermediate,  
434 and low-risk groups, respectively. As a group, the five dogs achieved *SEN* 95.9% (95% CI  
435 93.6 - 97.4), *SPC* 95.1% (94.4 - 95.8), *PPV* 69.7% (65.9 - 73.2), *NPV* 99.5% (99.2 – 99.7),  
436 *ACC* 95.2%, and *LR* 19.6 (Table 3). Individual performance mirrored closely the group  
437 metrics (Fig 5). Four of 320 participants in the low-risk group had positive rRT-PCR tests  
438 (prevalence 1.25%), but none did show up for the canine scent test. It produced zero values  
439 in two cells of the 2x2 contingency tables, preventing the computation of size effects (Fig  
440 4).

441

442

443 **Fig 3.** Efficacy trial: diagram illustrating the flow of human participants in the third  
444 phase of the study (in vivo screening).  
445



446

447

448

449 **Fig 4.** Efficacy trial: data analysis by risk group of all participants in experiments  
 450 designed to determine performance metrics of the dogs during in vivo screening.  
 451

848 subjects tested for COVID-19 ↓											
High Risk: 269 HUSVF Inpatients ↓						Intermediate Risk: 259 HUSVF HCWs ↓			Low Risk: 320 employees Governor of Antioquia ↓		
COVID-19 PREVALENCE: 10.8% of 848			30.1%			2.70%			1.25%		
COVID-19 Test Result (rRT-PCR) ↓			COVID-19 Test Result (rRT-PCR) ↓			COVID-19 Test Result (rRT-PCR) ↓					
(+)	(-)	INDET.	(+)	(-)	INDET.	(+)	(-)	INDET.	(+)	(-)	INDET.
Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)	Effect Size (95% C.I.)
81	186	2	7	252	0	4	315	1	0	39	1
POSITIVE →	81	1	1	SEN=0.9547-1.0)	SPC=0.9946 (0.9702-0.9997)	5	6	0	SEN=0.7143 (0.3589-0.9492)	SPC=0.9762 (0.9490-0.9890)	NAN
NEGATIVE →	0	185	1	PPV=0.9878 (0.9341-0.9994)	NPV=1 (0.9797-1.0)	2	246	0	PPV=0.4545 (0.2127-0.7199)	NPV=0.9919 (0.9711-0.9986)	PPV=0 (0.0-0.0897)
Refused to Scent Subject →	0	0	0	ACC=0.9963	LR=186	0	0	0	ACC=0.9691	LR=30.0	NPV=1 (0.9859-1.0)
Subject Refused K9 Test →	0	0	0	P<0.0001		0	0	0	P<0.0001		ACC=0.8734 NAN
ANDROMEDA						0	269	0	0	0	P>0.9999
NINA						0	7	0	0	0	PPV=0 (0.0-0.1380)
NIÑO						0	24	1	0	0	NPV=1 (0.9867-1.0)
TIMO						0	24	0	0	0	ACC=0.9221 NAN
VITA						0	19	0	0	0	P>0.9999
5K9						0	294	1	0	0	PPV=0 (0.0-0.1682)
						0	2	0	0	0	NPV=1 (0.9871-1.0)
						4	0	0	0	0	ACC=0.9393 NAN
						0	15	1	0	0	P>0.9999
						0	300	0	0	0	PPV=0 (0.0-0.2039)
						0	0	0	0	0	NPV=1 (0.9874-1.0)
						4	0	0	0	0	ACC=0.9524 NAN
						0	22	0	0	0	P>0.9999
						0	293	1	0	0	PPV=0 (0.0-0.1487)
						0	0	0	0	0	NPV=1 (0.9871-1.0)
						4	0	0	0	0	ACC=0.9302 NAN
						0	119	3	0	0	P>0.9999
						0	1440	2	0	0	PPV=0 (0.0-0.0313)
						0	16	0	0	0	NPV=1 (0.9973-1.0)
						20	0	0	0	0	ACC=0.9237 NAN
						0	0	0	0	0	P>0.9999
						Formula Bar					

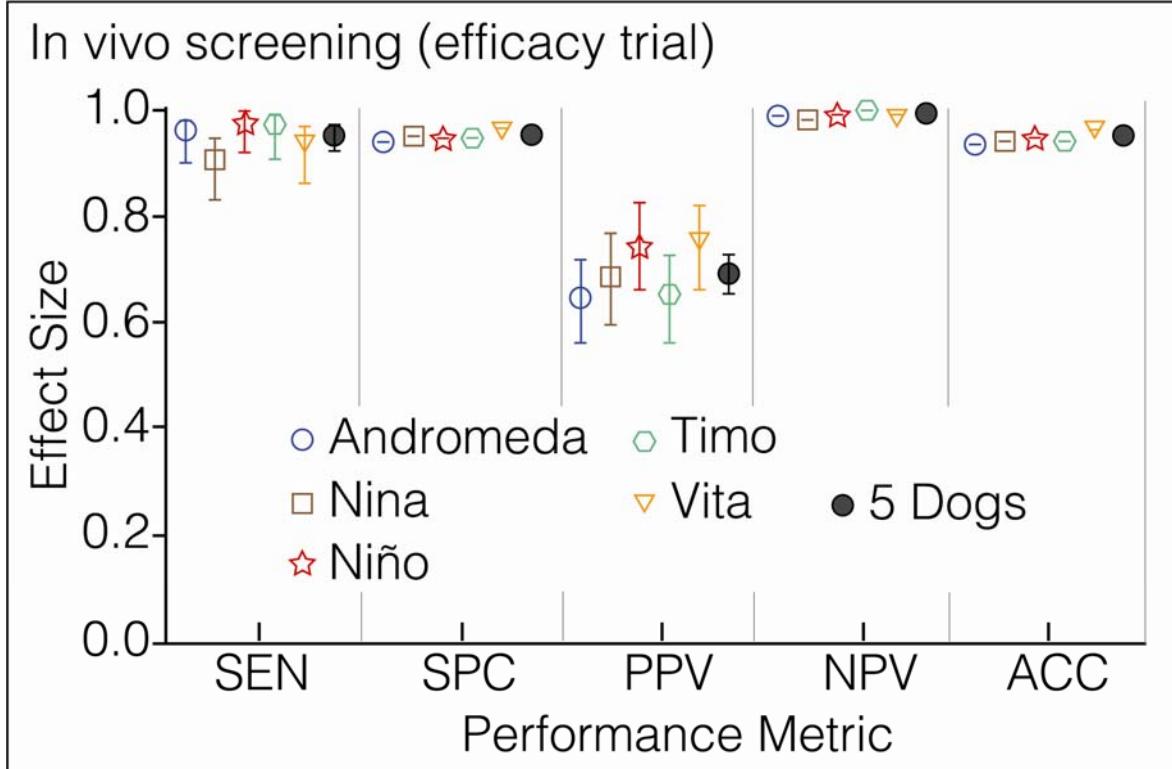
452  
 453 Green, yellow, orange and purple cells contain true positives, false positives, false negatives, and  
 454 true negatives, respectively. Cells not enhanced contain the number of participants with  
 455 “indeterminate” rRT-PCR (3), subjects who declined K9 olfaction (4), and those rare occasions  
 456 where the dogs refused to scent an individual, which happened 7 times with Andromeda and Nina  
 457 and 2 times with Niño. Sensitivity could not be computed in the low risk group (NAN: not a number)  
 458 because all 4 COVID-19 patients declined K9 scent-detection, resulting in 0 in two cells of the 2x2  
 459 contingency table and not significant P values in the two-tailed Fisher's Exact Test (enhanced in  
 460 salmon color).  
 461  
 462

463      **Table 3. In vivo screening (efficacy trial).** Performance metrics of five dogs after 13 weeks of  
464      scent-detection training. Dogs scent-interrogated 848 subjects that included 269 hospitalized patients, 259  
465      health-care workers from the same institution, and 320 government employees; n varies slightly between dogs  
466      because a participant did not show up for the dog test or the dog refused to interrogate an individual. Dog  
467      trainers were blinded regarding the COVID-19 status of all subjects. P values (Fisher's Exact Test) express  
468      the probability that the dogs identified the patients infected by SARS-CoV-2 by chance alone.  
469

Metric	Dog Name (breed) and Effect Sizes [95% Confidence Intervals]					
	Andromeda (BM)	Nina (BM)	Niño (PB)	Timo (BM)	Vita (AMxSH)	All 5 Dogs
Prevalence (%)	10.6	10.6	10.3	10.5	10.5	10.5
n	834	834	836	839	841	4184
TP	86	81	84	86	83	420
TN	700	708	724	705	726	3563
FP	46	38	26	46	27	183
FN	2	7	2	2	5	18
SEN (%)	97.7 [92.1-99.6]	92.1 [92.1-99.6]	97.7 [91.9-99.6]	97.3 [92.1-99.6]	94.3 [87.4-97.6]	95.9 [93.6-97.4]
SPC (%)	93.8 [91.9-95.4]	94.9 [91.9-95.4]	96.5 [95.0-97.6]	93.9 [91.9-95.4]	96.4 [94.8-97.5]	95.1 [94.4-95.8]
PPV (%)	65.2 [56.7-72.8]	68.1 [56.7-72.8]	76.4 [67.6-83.3]	65.2 [56.7-72.8]	75.5 [66.6-82.6]	69.7 [65.9-73.2]
NPV (%)	99.7 [99.0-100]	99 [99.0-100]	99.7 [99.0-100]	99.7 [99.0-100]	99.3 [98.4-99.7]	99.5 [99.2-99.7]
ACC (%)	94.2	94.6	96.7	94.3	96.2	95.2
LR	15.9	18.1	28.2	16.0	26.3	19.6
P	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

470      TP: true positives; TN: true negatives; FP: false positives; FN: false negatives; BM: Belgian  
471      malinois; PB: Pit bull; AMxSH: Alaskan malamute by Siberian husky first generation cross.  
472  
473

474 **Fig 5. In vivo screening (efficacy trial).** Prevalence, 10.5%. Each symbol has a different color to  
475 ease visualization of the dogs. The vertical lines above and below the symbols represent the 95% confidence  
476 interval for each metric, which is contained within the symbol for *SPC*, *NPV* and *ACC*.  
477



478  
479  
480 **Effectiveness assay: in vivo screening of citizens riding the**  
481 **Metro System of Medellin**  
482 The mass transit service of Medellin transports 1.5 million passengers every day. Without  
483 prior notification to the San Antonio station users or the dog-trainer teams, three canines  
484 screened over two days 550 individuals who also volunteered to provide saliva specimens  
485 for rRT-PCR testing. Despite the environmental impact on the concentration of the dogs,  
486 they detected 17 COVID-19 cases with high *SPC* and *NPV*. During the first hours of the  
487 assay, effect sizes for *SEN* and *PPV* dropped significantly in comparison with the efficacy  
488 trial (Table 4, Fig 6), but the dogs adjusted within hours to the new environment and  
489 improved their performance until reaching a plateau (Fig 7).

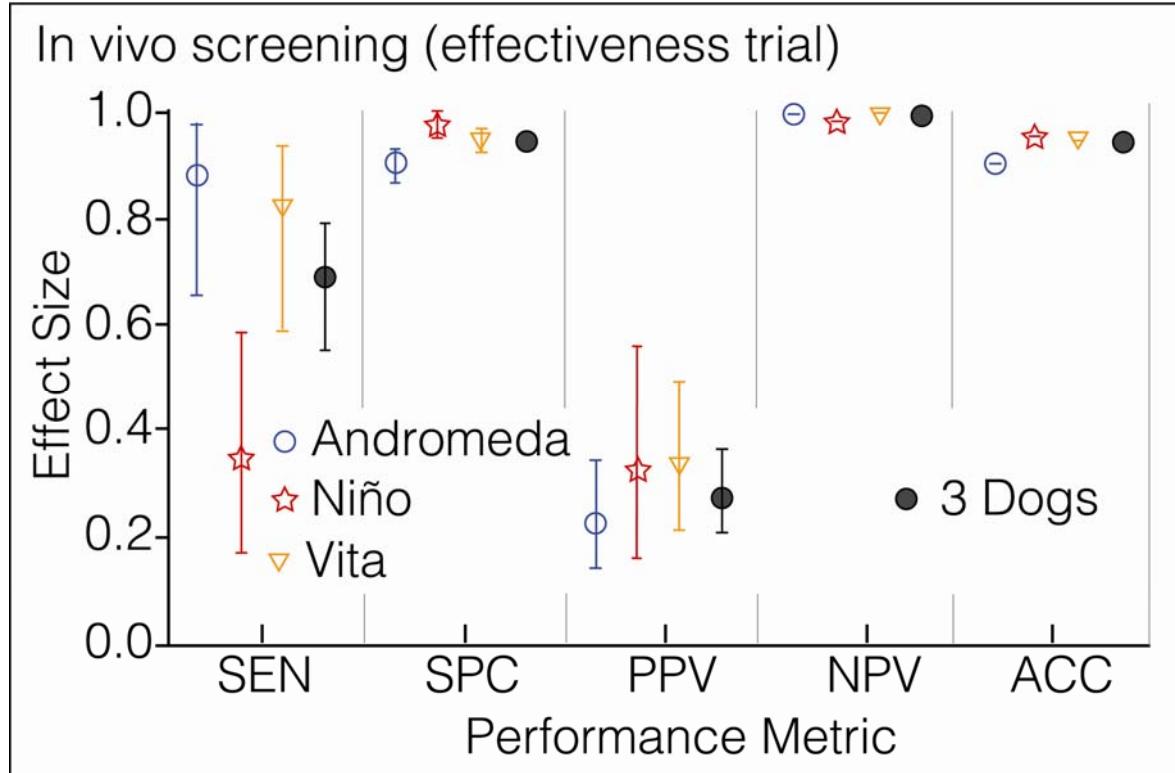
490

491 **Table 4. Effectiveness assay to determine dog performance during in vivo**  
492 **screening under real-life conditions.** Performance metrics of three dogs in the Metro System of  
493 Medellin. Without new training or environmental habituation, dogs scent-interrogated 550 volunteers  
494 recruited on site. Simultaneous rRT-PCR in saliva led to detection of 17 COVID-19 patients, all  
495 asymptomatic or mildly symptomatic. Obviously, the research team was blinded regarding the diagnosis of all  
496 subjects. *P* values (Fisher's Exact Test) express the probability that the dogs identified the patients infected by  
497 SARS-CoV-2 by chance alone.  
498

Metric	Dog Name (breed) and Effect Sizes [95% Confidence Intervals]							
	Andromeda (BM)		Niño (PB)		Vita (AMxSH)		All 3 Dogs	
	Effect Size	95% C.I.	Effect Size	95% C.I.	Effect Size	95% C.I.	Effect Size	95% C.I.
Prevalence (%)	3.1		3.1		3.1		3.1	
n	550		550		550		1650	
TP	15		6		14		35	
TN	483		521		506		1510	
FP	50		12		27		89	
FN	2		11		3		16	
SEN (%)	88.24	65.7 to 97.9	35.3	17.3 to 58.7	82.4	59.0 to 93.8	68.6	55.0 to 79.7
SPC (%)	90.62	87.9 to 92.8	97.7	96.1 to 98.7	94.9	92.7 to 96.5	94.4	93.2 to 95.5
PPV (%)	23.08	14.5 to 34.6	33.3	16.3 to 56.3	34.1	21.6 to 49.5	28.2	21.1 to 36.7
NPV (%)	99.59	98.5 to 99.9	97.9	96.3 to 98.8	99.4	98.3 to 99.8	99.0	98.3 to 99.4
ACC (%)	90.6		95.8		94.5		93.6	
LR	9.41		15.7		16.3		12.3	
P	<0.0001		<0.0001		<0.0001		<0.0001	

499 TP: true positives; TN: true negatives; FP: false positives; FN: false negatives; BM: Belgian  
500 malinois; PB: Pit bull; AMxSH: Alaskan malamute by Siberian husky first generation cross.  
501  
502

503 **Fig 6. In vivo screening (effectiveness assay). Performance metrics of three dogs**  
504 **in the Metro System of Medellin; prevalence, 3.1%.** Each symbol has a different color to ease  
505 visualization of the dogs. The vertical lines above and below the symbols represent the 95% confidence  
506 interval for each metric, which is contained within the symbol for *NPV* and *ACC*.  
507

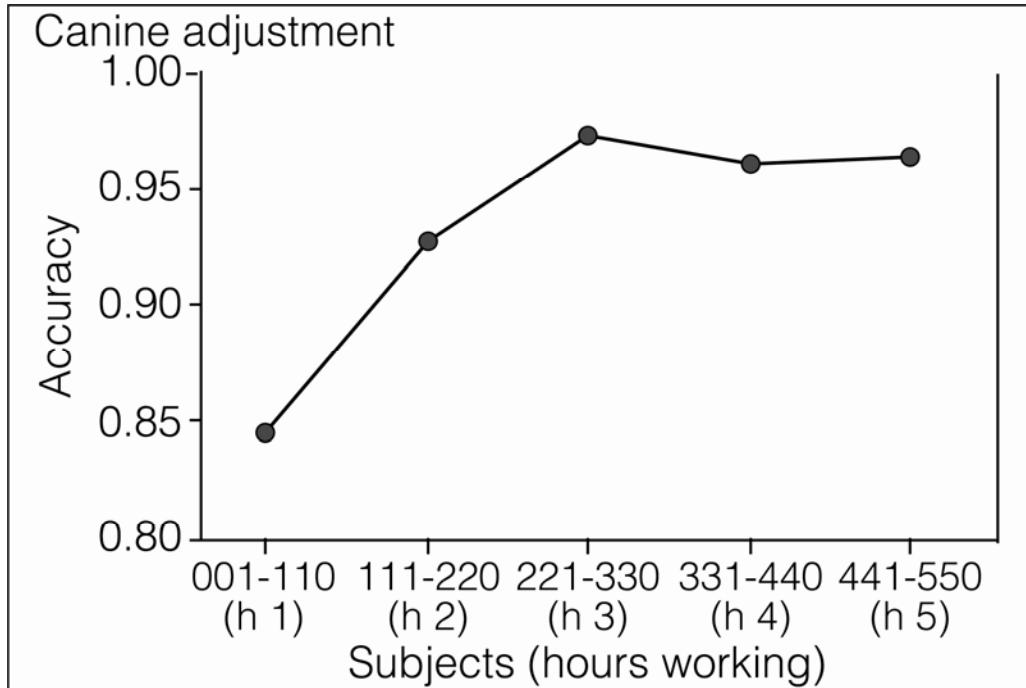


508  
509

510

511 **Fig 7. In vivo screening (effectiveness assay). Canine adjustment to a real-life**  
512 **situation.** Accuracy started much lower under real-life conditions, but improved with time  
513 as the dogs adjusted to the new environment. Numbers labeling the abscissa represent the  
514 order in which subjects were screened by the dogs, divided in groups of 110 individuals.  
515 Scent-interrogation of each group took approximately one hour of work for the dogs.

516



517

518

### 519 **Limit of canine scent-detection**

520 The LOD was determined in vitro using freshly collected saliva specimens from four  
521 COVID-19 patients new to the dogs. The moment of this assay coincided with the estrus  
522 cycle of several females, which caused the exclusion of the males from this experiment  
523 because both refused to work. The LOD for Andromeda, Nina, Vika, and Vita was lower  
524 than  $2.61 \times 10^{-12}$  copies ssRNA/mL (Table S4), the equivalent of detecting a drop (0.05  
525 mL) of any odorous substance dissolved in a volume of water greater than the capacity of  
526 10.5 Olympic swimming pools ( $2.6 \times 10^{10}$  mL).

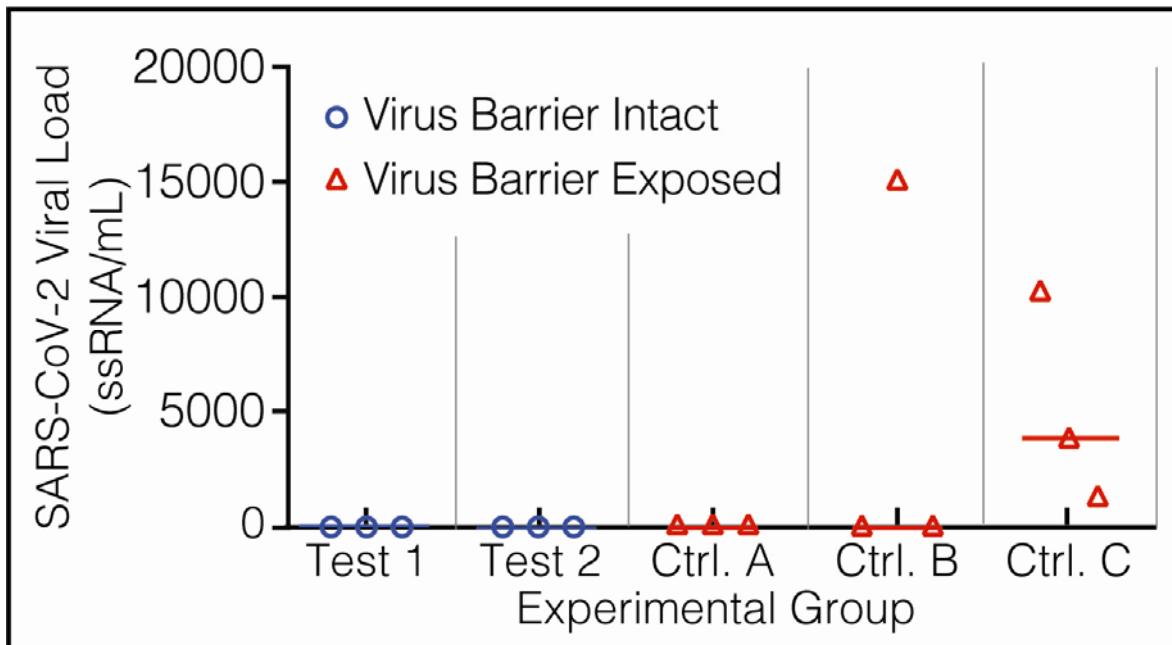
527

528 **Biosafety of the canine and human team handling the virus**  
529 None of the dogs, their trainers, or the physician-scientists in charge of sampling and taking  
530 care of the patients contracted COVID-19 during this study. The rRT-PCR tests for SARS-  
531 CoV-2 from canines and humans resulted negative twice, once after ending the in vitro  
532 phase, and again after finishing the in vivo phase of the study (Table S5). Experimental  
533 testing of the contraptions to contain SARS-CoV-2 showed that both devices worked as  
534 intended, allowing the scent to evaporate while holding the virus secured inside (Fig 8).  
535 Test hamsters climbed and smelled the mesh-protected contraptions D1 (group 1) and D2  
536 (group 2), but no animal contracted SARS-CoV-2. Control hamsters group A did climb on  
537 D1but could not damage the Tychem 2000 fabric covering the flask, and none got infected;  
538 group C did bite the Tychem of D2, and 1 animal was infected; and group B played, bit,  
539 nested, and slept in the gauze impregnated with SARS-CoV-2, and the three hamsters in the  
540 cage acquired SARS-CoV-2 (Table S6).

541

542

543 **Fig 8. Biosafety data. Experimental evaluation of the contraptions devised to**  
544 **contain SARS-CoV-2.** After testing negative for SARS-CoV-2 in saliva, 5 groups of 3 golden Syrian  
545 hamsters each were exposed during 4 days to SARS-CoV-2 directly (Group B, virus control) or enclosed in  
546 devices 1 (D1) and 2 (D2). Animals in test groups 1 (D1) and 2 (D2) were allowed to smell their devices but  
547 could not touch them, while the hamsters allocated to control groups A (D1) and C (D2) had direct access to  
548 the containment fabric. The ordinate represents the viral load in saliva of each hamster after exposure to  
549 SARS-CoV-2 in 5 experimental groups.  
550



551  
552

## 553 Discussion

554 This study shows that canine scent-detection of COVID-19 is immediate, accurate,  
555 applicable anytime, and deployable anywhere as a diagnostic test in saliva or respiratory  
556 secretions, or as a screening tool in the patient directly. In any of those two roles, the dogs  
557 missed very few infected individuals, as demonstrated by  $NPV > 99\%$  in vitro and in vivo,  
558 and independently of the experimental design (in vivo efficacy and effectiveness trials).  
559 COVID-19 severity, ranging from asymptomatic to pre-symptomatic, sick and very sick  
560 patients, had no impact on performance. Prevalence from three populations of diverse  
561 levels of risk showed, as expected, that  $PPV$  went down when the presence of the disease in

562 the population is very low, but *NPV* remained close to 100% across low and high  
563 prevalence. The error-free training system prepared the dogs for in vitro diagnosis in  
564 improvised, open fields, making sophisticated and expensive equipment superfluous.  
565  
566 In vivo screening generated very encouraging results in both, efficacy and effectiveness  
567 trials: the dogs detected more than 99% of the infected individuals spending <5 seconds per  
568 subject. Such output is ideal, because it allows immediate identification and isolation of all  
569 contagious subjects. However, the effectiveness assay also demonstrated that, under real-  
570 life conditions, odor contamination causes a substantial increment in the false positive rate  
571 driving down the *PPV*: only 28 of 100 dog-positive subjects would have simultaneously a  
572 positive rRT-PCR result; the other 72 would be rRT-PCR negative. It means that for each  
573 true positive, the dogs produce 2.5 false positives, an acceptable error rate for any screening  
574 test offering a very high *NPV* [38, 39]. The ultra-sensitive limit of detection ( $<10^{-12}$  copies  
575 ssRNA/mL) suggests that at least a fraction of the “false positives” are actually pre-  
576 symptomatic COVID-19 patients: during training, the dogs marked as positive three nurses  
577 which rRT-PCR was negative, but the three of them had symptomatic COVID-19  
578 demonstrated 4-7 days later. We also observed several times during training that the dogs  
579 spontaneously marked as positive the scientists that had touched any COVID-19 patient, or  
580 the cell phones of nurses and physicians in care of COVID-19 patients. It means that  
581 trained canines detect the scent-print of SARS-CoV-2 in contaminated individuals or in  
582 their belongings and, since contamination could lead to infection [40], the dogs actually  
583 identify potential COVID-19 cases before infection takes place. Although such capacity  
584 would result quite useful to follow-up contacts, it caused the high rate of false positives in  
585 the effectiveness assay. Inasmuch as the index test (rRT-PCR) is not very sensitive in the

586 pre-symptomatic stage [3, 41], it is possible that some of the dog's false positives actually  
587 represent true positives, a scenario where the canine test exceeds the gold standard.  
588 However, this trial was not designed to test such hypothesis.

589

590 The data provided the answer for the other three research questions. First, the limit of  
591 detection in vitro was lower than  $2.61 \times 10^{-12}$  copies ssRNA/mL, close to previous  
592 concentration thresholds determined with pure chemicals [42]. Second, all six dogs  
593 succeeded as medical detectors despite that their breeds are not intended for that purpose. It  
594 supports recent data showing that canines, independently of breed, could serve as a medical  
595 detection dogs if they exhibit motivation, stamina, determination, and resilience [43, 44].  
596 And third, only three COVID-19 patients sufficed for our dogs to recognize the scent-print  
597 of this particular disease in fresh saliva specimens, as demonstrated before with wildlife  
598 detection dogs [45]. With that knowledge, the dogs made the generalization necessary to  
599 diagnose COVID-19 in saliva of other nine patients disguised among saliva from 100  
600 negative controls. And that generalization was all they needed to make the inference that  
601 the same odor, emanating from a human being, implied to lay down, the behavior trained to  
602 indicate that such individual was positive for COVID-19.

603

604 Some experts believe that training with positive stimulus from a few COVID-19 patients  
605 (just three in this case) could lead the dog to learn by memory the scent-print of the  
606 individuals instead of that of the virus, and that "at least a hundred" positive donors are  
607 needed [46]. We found that such belief is unfounded: on the one hand, our dogs identified  
608 the positive specimens of 9 unknown patients in phase 2, and almost all 105 patients  
609 infected by SARS-CoV-2 in phase 3. On the other hand, research on the concept of

610 working memory demonstrated that expert detection dogs remembered a new scent-print  
611 98% of the time as long as it was located first in a line-up with five distractors, but  
612 performance went down when the positive stimulus was located farther, dropping to 11.5%  
613 at the sixth location [29]. Therefore, it seems impossible for a dog to remember the odor of  
614 up to 12 different individuals randomly allocated among 100 distractors. These data suggest  
615 that dogs are not relying in memory when they detect SARS-CoV-2 and discriminate it  
616 among so many other odoriferous cues. In fact, dogs are capable not only of generalization  
617 [47], they do make inferences [48]. Furthermore, dogs have the capacity, without previous  
618 training, to categorize new objects based on their functionality instead of on their  
619 perceptual similarity [49]. Not long ago, cognitive skills like making inferences and  
620 categorization, or the understanding of the meaning of words [50], were considered  
621 exclusive of *Homo sapiens*, but studies with dogs exposed to new toys a single time and  
622 only during one minute clearly showed that we have been underestimating the cognition  
623 capacity of canines as study subjects [51]. These reductionist preconceptions also explain  
624 the lack of interest displayed by the medical and scientific communities, as well as the  
625 funding agencies, in this approach to the diagnosis of infectious diseases [52].

626

627 Although nasopharyngeal swabs were first established as the method of choice to take the  
628 specimens for rRT-PCR for SARS-CoV-2, we used saliva because there is solid evidence  
629 demonstrating that it is more sensitive than nasopharyngeal swabs [53], sampling is faster  
630 and much easier, the need for sterile swabs is eliminated as well as the uncomfortable  
631 introduction of a foreign object through the nose of the patient and, contrary to the  
632 nasopharynx, the viral loads in saliva predict severity [54]. Also, we refrained from using  
633 saliva from non-COVID hospitalized patients as negative stimuli to adhere to error-free

634 discrimination learning [33, 34]. The microbiological environment of hospitals is  
635 dominated by human pathogens, while healthy people are inhabited by myriads of functional  
636 commensals that play an irreplaceable role in physiology [55]. Soon after admission,  
637 nosocomial pathogens start replacing the microbiota of the inpatients [56], and COVID-19  
638 victims are not an exception [57]. Therefore, negative stimuli (saliva) from hospitalized  
639 patients would provide much less contrast with the positive stimuli than saliva from healthy  
640 people, increasing the probability of error during training.

641

642 This study has some limitation that deserve attention. First, there are no human coronavirus  
643 strains involved, therefore it is impossible to predict if the dogs can discriminate them from  
644 SARS-CoV-2 or, even less, from other non-human coronavirus. Since human coronavirus  
645 usually cause mild upper respiratory infections, these patients rarely need hospitalization  
646 and were not part of our sample. We did find that none of the 43 hospitalized patients with  
647 respiratory conditions other than COVID-19 was positive for the dogs despite the fact that  
648 half of them had pneumonia caused by bacterial or viral pathogens like influenza virus.

649 Second, the dogs did not have the opportunity to scent-interrogate any of the four COVID-  
650 19 subjects from the low-risk group because they relinquished that part of the study. This  
651 prevented statistical calculations necessary to determine the size effects of the different  
652 performance metrics under very low prevalence (1.25%). The third limitation arose during  
653 dog training, and it caused the sharp decline in *PPV* from 69.7% to 28.2% between efficacy  
654 and effectiveness trials. To improve *PPV*, it is necessary to teach the dogs that there is a  
655 cut-off value in odor intensity below which they should disregard the COVID-19 scent-  
656 print. However, more research is needed to identify such threshold value and to know if it

657 is worth to train for that, because it might be an asset to have dogs that can detect the virus  
658 immediately after a patient gets contaminated with it.

659

660 After our preprint [58], at least four studies on canine scent-detection of SARS-CoV-2 in  
661 vitro have been formally published [59-63], and despite substantial methodological  
662 differences with our work, results are reproducible. The main difference with those studies  
663 is that we chose to scent-interrogate the human body instead of a specimen, and did it  
664 because of the many obvious advantages that such approach brings: results are immediate,  
665 can be obtained anywhere, do not require equipment, and allow in situ separation of  
666 contagious individuals. The use of trained dogs as medical detectors was safe for the human  
667 participants during training and experimentation regardless of the breed, a point of major  
668 importance considering that deployment would require the participation of many canines  
669 [64, 65], and the possibility of training dogs for real-time diagnosis of many other  
670 infectious diseases may help humanity be better prepared to confront the next pandemic  
671 [66]. These data suggest that well-trained dogs may aid the safe re-opening of economies  
672 and educational systems, while offering an efficient way to control the pandemic.

673

## 674 **Acknowledgments**

675 We thank sincerely the patients and staff of HUSVF and the officers of the Governor of  
676 Antioquia for participating in the study.

677

## 678 **References**

- 679 1. Botti-Lodovico Y, Rosenberg E, Sabeti PC. Testing in a Pandemic - Improving Access,  
680 Coordination, and Prioritization. *N Engl J Med.* 2021; 384:197-199.
- 681 2. Gandhi M, Yokoe DS, Havlir DV. Asymptomatic transmission, the Achilles' heel of  
682 current strategies to control Covid-19. *N Engl J Med.* 2020; 382:2158-2160.
- 683 3. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in false-negative  
684 rate of reverse transcriptase polymerase chain reaction-based SARS-CoV-2 tests by  
685 time since exposure. *Ann Intern Med.* 2020; 173:262-267.
- 686 4. Centers for Disease Control and Prevention. Interim guidelines for rapid antigen testing  
687 for SARS-CoV-2. [https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antigen-](https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antigen-tests-guidelines.html)  
688 [tests-guidelines.html](https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antigen-tests-guidelines.html).
- 689 5. Centers for Disease Control and Prevention. Interim guidelines for COVID-19 antibody  
690 testing. [https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-](https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html)  
691 [guidelines.html](https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html).
- 692 6. EndCoronavirus. Countries beating COVID-19. Available at:  
693 <https://www.endcoronavirus.org/countries>
- 694 7. Allen DW. Covid Lockdown Cost/Benefits: A Critical Assessment of the Literature.  
695 April 2021. Available at: <http://www.sfu.ca/~allen/LockdownReport.pdf>.
- 696 8. Yu KDS, Aviso KB, Santos JR, Tan RR. The Economic Impact of Lockdowns: A  
697 Persistent Inoperability Input-Output Approach. *Economies.* 2020; 8:109. doi:  
698 10.3390/economies8040109.
- 699 9. Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, et al. BNT162b2 mRNA  
700 Covid-19 Vaccine in a Nationwide Mass Vaccination Setting. *N Engl J Med.* 2021;  
701 384:1412-1423.

- 702 10. Wang P, Nair MS, Liu L, Iketani S, Luo Y, Guo Y, et al. Antibody resistance of SARS-  
703 CoV-2 variants B.1.351 and B.1.1.7. *Nature*. 2021 Mar 8. doi: 10.1038/s41586-021-  
704 03398-2.
- 705 11. Barnes M, Sax PE. Challenges of “return to work” in an ongoing pandemic. *N Engl J  
706 Med.* 2020; 383:779-786.
- 707 12. Shipman, P. *The invaders: how humans and their dogs drove Neanderthals to  
708 extinction*. Cambridge: The Belknap Press; 2015.
- 709 13. Romanes GJ. Experiments on the sense of smell in dogs. *Nature*. 1887; 36:273–274.
- 710 14. Jezierski T, Ensminger J, Papet LE, editors. *Canine olfaction science and law: advances  
711 in forensic science, medicine, conservation, and environmental remediation*. New York:  
712 CRC Press; 2016.
- 713 15. Helton WS. *Canine ergonomics: the science of working dogs*. New York: CRC Press;  
714 2009.
- 715 16. Goodavage M. *Doctor dogs: how our best friends are becoming our best medicine*. New  
716 York: Penguin Random House LLC; 2019.
- 717 17. Bijland LR, Bomers MK, Smulders YM. Smelling the diagnosis. A review on the use of  
718 scent in diagnosing disease. *Neth J Med.* 2013; 71:300-307.
- 719 18. Cambau E, Poljak M. Sniffing animals as a diagnostic tool in infectious diseases. *Clin  
720 Microbiol Infect.* 2020; 26:431-435.
- 721 19. Bomers MK, van Agtmael MA, Luik H, Vandebroucke-Grauls CMJE, Smulders YM.  
722 A detection dog to identify patients with *Clostridium difficile* infection during a hospital  
723 outbreak. *J Infect.* 2014; 69:456-461.

- 724 20. Gottwald T, Poole G, McCollum T, Hall D, Hartung J, Bai J, et al. Canine olfactory  
725 detection of a vectored phytobacterial pathogen, *Liberibacter asiaticus*, and integration  
726 with disease control. *Proc Natl Acad Sci.* 2020; 117:3492-3501.
- 727 21. Gottwald T, Poole G, Taylor E, Luo W, Posny D, Adkins S, et al. Canine olfactory  
728 detection of a non-systemic phytobacterial citrus pathogen of international quarantine  
729 significance. *Entropy (Basel)*. 2020; 22:1269. doi: 10.3390/e22111269.
- 730 22. Angle C, Angle C, Waggoner LP, Ferrando A, Haney P, Passler T. Canine detection of  
731 the volatileome: a review of implications for pathogen and disease detection. *Front Vet  
732 Sci.* 2016; 3:47.
- 733 23. Chen H, Qi X, Ma J, Zhang C, Feng H, Yao M. Breath-borne VOC Biomarkers for  
734 COVID-19. *MedRxiv [Preprint]*. 2020 medRxiv 2020.06.21.20136523 [posted 2020  
735 June 24; cited 2021 Apr 17]: [11 p]. Available from:  
736 <https://doi.org/10.1101/2020.06.21.20136523>.
- 737 24. Damas J, Hughes GM, Keough KC, Painter CA, Persky NS, Corbo M, et al. Broad host  
738 range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in  
739 vertebrates. *Proc Natl Acad Sci.* 2020; 117:22311-22322.
- 740 25. Bosco-Lauth AM, Hartwig AE, Porter SM, Gordy PW, Nehring M, Byas AD, et al.  
741 Experimental infection of domestic dogs and cats with SARS-CoV-2: Pathogenesis,  
742 transmission, and response to reexposure in cats. *Proc Natl Acad Sci.* 2020; 117:26382-  
743 26388.
- 744 26. Thorndike EL. *Animal Intelligence: Experimental Studies*. New York: The McMillian  
745 Company; 1911.

- 746 27. Porritt F, Shapiro M, Waggoner P, Mitchell E, Thomson T, Nicklin S, et al.  
747        Performance decline by search dogs in repetitive tasks, and mitigation strategies. Appl  
748        Anim Behav Sci. 2015; 166:112-122.
- 749 28. Cobb M, Branson N, McGreevy P, Lill A, Bennett P. The advent of canine performance  
750        science: offering a sustainable future for working dogs. Behav Process. 2015; 110:96-  
751        104.
- 752 29. Gadbois S, Reeve C. Canine olfaction: scent, sign, and situation. In: Horowitz A, editor.  
753        Domestic dog cognition and behavior: the scientific study of *Canis familiaris*. Berlin:  
754        Springer-Verlag; 2014. pp. 2-30.
- 755 30. Hare B, Ferrans M. Is cognition the secret to working dog success? Anim Cogn. 2021  
756        Mar 9. doi: 10.1007/s10071-021-01491-7.
- 757 31. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD  
758        2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ  
759        (Clinical research ed.). 2015; 351:h5527. doi:10.1136/bmj.h5527.
- 760 32. Bujang MA, Adnan TH. Requirements for minimum sample size for sensitivity and  
761        specificity analysis. J Clin Diagn Res. 2016; 10:YE01-YE06. doi:  
762        10.7860/JCDR/2016/18129.8744
- 763 33. Terrace HS. Discrimination learning with and without “errors”. J Exp Anal Behav.  
764        1963; 6:1–27.
- 765 34. Terrace HS. Wavelength generalization after discrimination learning with and without  
766        errors. Science. 1964; 144:78-80.
- 767 35. Cohen J. Statistical power analysis for the behavioral sciences. Mahwah, NJ: Lawrence  
768        Erlbaum Associates; 1988.
- 769 36. Spector M. Clicker Training for Obedience. Waltham, MA: Sunshine Books; 2005.

- 770 37. Radbel J, Jagpal S, Roy J, Brooks A, Tischfield J, Sheldon M, et al. Detection of Severe  
771 Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is comparable in clinical  
772 samples preserved in saline or viral transport medium. *J Mol Diagn.* 2020; 22:871-875.  
773 doi: 10.1016/j.jmoldx.2020.04.209.
- 774 38. National Academies of Sciences, Engineering, and Medicine 2020. Rapid Expert  
775 Consultation on Critical Issues in Diagnostic Testing for the COVID-19 Pandemic  
776 (November 9, 2020). Washington, DC: The National Academies Press.  
777 <https://doi.org/10.17226/25984>.
- 778 39. Mina MJ, Parker R, Larremore DB. Rethinking Covid-19 test sensitivity - A strategy  
779 for containment. *N Engl J Med.* 2020 Nov 26;383(22):e120. doi:  
780 10.1056/NEJMp2025631.
- 781 40. Behzadinasab S, Chin AWH, Hossein M, Poon LLM, Ducker WA. SARS-CoV-2 virus  
782 transfers to skin through contact with contaminated solids. *MedRxiv [Preprint].* 2021  
783 medRxiv 2021.04.24.21256044 [posted 2021 April 28; cited 2021 May 2]: [18 p].  
784 Available from: <https://doi.org/10.1101/2021.04.24.21256044>.
- 785 41. Woloshin S, Patel N, Kesselheim AS. False negative tests for SARS-CoV-2 infection -  
786 challenges and implications. *N Engl J Med.* 2020. doi:10.1056/NEJMp2015897.
- 787 42. Walker DB, Walker JC, Cavnar PJ, Taylor JL, Pickel DH, Hall SB, et al. Naturalistic  
788 quantification of canine olfactory sensitivity. *Appl Anim Behav Sci.* 2006; 97:241–254.
- 789 43. Turcsán B, Kubinyi E, Miklósi A. Trainability and boldness traits differ between dog  
790 breed clusters based on conventional breed categories and genetic relatedness. *Appl*  
791 *Anim Behav Sci.* 2011; 132:61–70. doi: 10.1016/j.applanim.2011.03.006.

- 792 44. Leighton EA, Hare E, Thomas S, Waggoner LP, Otto CM. A solution for the shortage  
793 of detection dogs: a detector dog center of excellence and a cooperative breeding  
794 program. *Front Vet Sci.* 2018; 5:284. doi: 10.3389/fvets.2018.00284.
- 795 45. Oldenburg Jr C, Schoon A, Heitkönig IMA. Wildlife detection dog training: A case  
796 study on achieving generalization between target odor variations while retaining  
797 specificity. *J Vet Behav.* 2016; 13: 34e38. doi: 10.1016/j.jveb.2016.03.008
- 798 46. WHO R&D Blueprint [Internet]: Geneva, Switzerland: COVID-19 Consultation on the  
799 use of trained dogs for screening COVID-19 cases. 8th March 2021 [cited 2021 Apr  
800 14]. Available from: [https://cdn.who.int/media/docs/default-source/blue-print/who-consultation-screening-dogs--8th-march-2021.pdf?sfvrsn=a0d8cbda\\_1&download=true](https://cdn.who.int/media/docs/default-source/blue-print/who-consultation-screening-dogs--8th-march-2021.pdf?sfvrsn=a0d8cbda_1&download=true)
- 801 47. Moser AY, Bizo L, Brown WY. Olfactory generalization in detector dogs. *Animals.*  
802 2019; 9:702. doi:10.3390/ani9090702.
- 803 48. Pilley JW, Reid AK. Border collie comprehends object names as verbal referents.  
804 Behav Process. 2011; 86:184-195.
- 805 49. Fugazza C, Miklosi A. Depths and limits of spontaneous categorization in a family  
806 dog. *Sci Rep.* 2020; 10:3082. doi: 10.1038/s41598-020-59965-6.
- 807 50. Pilley JW. Border collie comprehends sentences containing a prepositional object, verb,  
808 and direct object. *Learn Motiv.* 2013; 44:229-240.
- 809 51. Hare B, Woods V. The genius of dogs – How dogs are smarter than you think. New  
810 York: Dutton; 2013.
- 811 52. Wyllie AL, Fournier J, Casanovas-Massana A, Campbell M, Tokuyama M,  
812 Vijayakumar P, et al. Saliva or nasopharyngeal swab specimens for detection of SARS-  
813 CoV-2. *N Engl J Med.* 2020; 383:1283-1286.

- 815 53. Silva J, Lucas C, Sundaram M, Israelow B, Wong P, Klein J, et al. Saliva viral load is a  
816 dynamic unifying correlate of COVID-19 severity and mortality. medRxiv [Preprint].  
817 2021 Jan 10:2021.01.04.21249236. doi: 10.1101/2021.01.04.21249236.
- 818 54. Klopfer PH. Animal cognition and the new anthropomorphism. Int J Comp Psychol.  
819 2005; 18:202-206.
- 820 55. Lynch SV, Pedersen O. The human intestinal microbiome in health and disease. N Engl  
821 J Med. 2016; 375:2369-2379.
- 822 56. Ojima M, Motooka D, Shimizu K, Gotoh K, Shintani A, Yoshiya K, et al. Metagenomic  
823 analysis reveals dynamic changes of whole gut microbiota in the acute phase of  
824 intensive care unit patients. Dig Dis Sci. 2016; 61:1628-34.
- 825 57. Zuo T, Zhang F, Lui GCY, Yeoh YK, Li AYL, Zhan H, et al. Alterations in gut  
826 microbiota of patients with COVID-19 during time of hospitalization.  
827 Gastroenterology. 2020; 159:944-955.e8. doi: 10.1053/j.gastro.2020.05.048.
- 828 58. Vesga O, Valencia AF, Mira A, Ossa F, Ocampo E, Agudelo M, et al. Dog savior:  
829 immediate scent-detection of SARS-CoV-2 by trained dogs. BioRxiv [Preprint]. 2020  
830 bioRxiv 2020.06.17.158105 [posted 2020 June 19; revised 2021 Apr 16; cited 2021 Apr  
831 17]: [17 p]. Available  
832 from: <https://www.biorxiv.org/content/10.1101/2020.06.17.158105v1.full>.
- 833 59. Jendrny P, Schulz C, Twele F, Meller S, von Köckritz-Blickwede M, Osterhaus  
834 ADME, et al. Scent dog identification of samples from COVID-19 patients - a pilot  
835 study. BMC Infect Dis. 2020; 20:536. doi: 10.1186/s12879-020-05281-3.
- 836 60. Grandjean D, Sarkis R, Lecoq-Julien C, Benard A, Roger V, Levesque E, et al. (2020)  
837 Can the detection dog alert on COVID-19 positive persons by sniffing axillary sweat

- 838 samples? A proof-of-concept study. PLoS ONE. 2020; 15:e0243122. doi:  
839 10.1371/journal.pone.0243122.
- 840 61. Essler JL, Kane SA, Nolan P, Akaho EH, Berna AZ, DeAngelo A, et al. Discrimination  
841 of SARS-CoV-2 infected patient samples by detection dogs: A proof of concept study.  
842 PLoS One. 2021; 16:e0250158. doi: 10.1371/journal.pone.0250158.
- 843 62. Jones RT, Guest C, Lindsay SW, Kleinschmidt I, Bradley J, Dewhirst S, et al. Could  
844 bio-detection dogs be used to limit the spread of COVID-19 by travellers? Journal of  
845 Travel Medicine 2020; 27:taaa131. doi: 10.1093/jtm/taaa131
- 846 63. Dickey T, Junqueira H. Toward the use of medical scent detection dogs for COVID-19  
847 screening. J Osteopath Med. 2021; 121: 141–148.
- 848 64. Duffy DL, Hsub Y, Serpell JA. Breed differences in canine aggression. Appl Anim  
849 Behav Sci. 2008; 114:441-460.
- 850 65. Mikkola S, Salonen M, Puurunen J, Hakanen E, Sulkama S, Araujo C, et al. Aggressive  
851 behaviour is affected by demographic, environmental and behavioural factors in  
852 purebred dogs. Sci Rep. 2021; 11:9433. doi: 10.1038/s41598-021-88793-5.
- 853 66. Horton R. Offline: The case for No-COVID. Lancet. 2021; 397:359. doi:  
854 10.1016/S0140-6736(21)00186-0.

855  
856

857 **Conflicts of Interest**

858 The authors declare no conflict of interest.  
859

## 860 Authors Contributions

861 OV came up with the idea, obtained the funding, designed and supervised the training and  
862 the experimental process, took medical care of the patients with COVID-19 who  
863 participated in the study, and wrote the manuscript.  
864 MA coordinated the experimental process; obtained, processed, and made available SARS-  
865 CoV-2 for experimentation; and took medical care of the patients with COVID-19 who  
866 participated in the study.  
867 AFV directed the training process and supervised the dog experiments in the field.  
868 AM, FO and EO trained the dogs and participated in all the experiments with them.  
869 AFV, AM and FO obtained and processed the saliva specimens from 100 healthy  
870 volunteers.  
871 KC, LP, AC and YA performed the molecular biology work.  
872 YA directed the logistics and organization of the experimental process with COVID-19 and  
873 control patients within *Hospital Universitario San Vicente Fundación*.  
874 JPH-O and JO supervised the virology and molecular biology work, wrote the manuscript,  
875 and made invaluable comments to improve its quality.  
876

## 877 Ethical Statement: animal usage and human samples

878 The study protocol was approved by the Ethical Committee for Human Research of  
879 HUSVF and the Animal Research Ethics Committee of Colina-K9. All human subjects read  
880 and signed their informed consent. We did not subject our dogs to any kind of pressure for  
881 training. We did not starve the dogs and did not need to make them fanatics for food, toys,  
882 games, or anything else. Since it is impossible to force a dog to do scent-work, our methods  
883 are exclusively positive, rewarding every correct response, and being indifferent to any  
884 mistake.  
885  
886