Clinical Validation of a Novel T-cell Receptor Sequencing Assay for Identification of Recent or Prior SARS-CoV-2 Infection

Sudeb C. Dalai,[1,2*] Jennifer N. Dines,[1*] Thomas M. Snyder,[3] Rachel M. Gittelman,[3] Tera Eerkes,[4*] Pashmi Vaney,[4] Sally Howard,[4*] Kipp Akers,[5] Lynell Skewis,[5*] Anthony Monteforte,[5] Pamela R. Witte,[5*] Cristina Wolf,[5*] Hans Nesse,[5] Megan Herndon,[5] Jia Qadeer,[1] Sarah Duffy,[1*] Emily Svejnoha,[1] Caroline Taromino,[1] Ian M. Kaplan,[6*] John Alsobrook,[7] Thomas Manley,[1*] Lance Baldo[1*]

[1]Medical Affairs and Clinical Development, Adaptive Biotechnologies, Seattle, Washington, USA

[2]Division of Infectious Diseases and Geographic Medicine, Department of Medicine, Stanford University School of Medicine, Stanford, California, USA

[3]Research, Adaptive Biotechnologies, Seattle, Washington, USA

[4]Regulatory Affairs, Adaptive Biotechnologies, Seattle, Washington, USA

[5]Molecular Product Development, Adaptive Biotechnologies, Seattle, Washington, USA

[6]T-Detect Product Management, Adaptive Biotechnologies, Seattle, Washington, USA

[7]Molecular Lab Management, Adaptive Biotechnologies, Seattle, Washington, USA

*Author was employed by Adaptive Biotechnologies at time of research.

SUPPLEMENTARY MATERIALS

SUPPLEMENTAL METHODS

Clinical Cohorts

Clinical remnant samples used for classifier training and primary validation analyses were collected from cohorts described in Supplemental Table 1 (5 training case cohorts and 4 training control cohorts). Secondary validation was performed on holdout cases from the ImmuneRACE cohort and holdout controls from 4 cohorts listed in Supplemental Table 1. Clinical validation (PPA/NPA) analyses were conducted using samples collected from the Discovery Life Sciences (DLS), ImmuneRACE, and ImmuneSense™ COVID-19 cohorts described below, with additional details provided in Tables 1 and 2.

ImmuneRACE Cohort

The ImmuneRACE study is a prospective, multicohort, exploratory study of participants exposed to, infected with, or recovering from coronavirus disease (COVID-19) (NCT04494893). Residual samples collected under prospective study protocols obtained informed consent from participants under a separate protocol: "ImmuneRACE" (ADAP-006/WIRB# 20200625/NCT04494893). Participants from across the United States were consented and enrolled via a virtual study design, with cohorting based on participant-reported clinical history following the completion of both a screening survey and study questionnaire. Whole blood, serum, and a nasopharyngeal or oropharyngeal swab were collected from participants by trained mobile phlebotomists. Participants with a confirmed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) test were included as residual, retrospective samples in the clinical validation study.

ImmuneSense Cohort

The ImmuneSense COVID-19 Study's prospective study arm enrolled individuals with symptoms suggestive of COVID-19 who were being tested for SARS-CoV-2 at 2 drive-thru testing sites in New Jersey. All samples were collected pursuant to an Institutional Review Board (IRB)-approved clinical study protocol, "ImmuneSense COVID-19 Study" (PRO-00781/ADAP-007/WIRB#20202820/NCT04583982). Whole blood, serum, and a nasopharyngeal swab were collected from participants at study sites. An electronic questionnaire was administered by study staff. Individuals testing positive via Abbott's reverse transcription-polymerase chain reaction (RT-PCR) were included in the secondary percent positive agreement (PPA) analysis. Individuals testing negative for SARS-CoV-2 using Emergency Use Authorized (EUA) RT-PCR tests, BioFire RP V2.1, and EUA antibody tests were included in the negative percent agreement (NPA) analysis.

Sample Allocation

PPA Study Cohorts

The primary PPA study evaluated residual blood samples (n=222) from individuals in the DLS cohort diagnosed with SARS-CoV-2 infection based on the EUA Abbott RealTime SARS-CoV-2 RT-PCR test from a single United States reference lab (New York) (Table 1). Secondary PPA assessments were performed using both retrospectively and prospectively collected samples from multiple cohorts (n=77; ImmuneRACE and ImmuneSense COVID-19 cohorts) and identified as positive based on a variety of EUA testing methods performed by a number of different laboratories. Given the potential for variability in RT-PCR performance given the use of numerous tests by multiple laboratoriess, samples were categorized by days since symptom onset

(Table 1). PPA analyses included samples from individuals of any age/sex/race who were confirmed to be positive for SARS-CoV-2 infection using the Abbott RealTime SARS-CoV-2 assay. Criteria for exclusion included no diagnosis of SARS-CoV-2 infection, insufficient sample volume, previous use of the sample for algorithm training, and sample collection more than 100 days post diagnosis. If multiple samples were collected for a unique individual, the following selection process was applied prior to sample testing: if all samples fell within the same bin for days post diagnosis (0–7, 8–14, or ≥15), the sample closest to diagnosis was selected for analysis; if the samples fell within different bins for days post diagnosis, selection was prioritized based on binning as follows: ≥15, 8-14, < 8 days. Samples with matched sera were prioritized for selection prior to testing, and when more samples were available than needed, samples were selected randomly prior to testing.

NPA Study Cohorts

The primary NPA analysis included 124 retrospective frozen clinical remnant blood samples from the DLS cohort collected prior to December 2019 (Table 2) and thus presumed negative for SARS-CoV-2 infection. These samples were collected over 2 years, during all months (including cold/flu season), and from diverse geographical areas in the United States (Table 2). The secondary NPA study included blood samples from participants enrolled prospectively (ImmuneSense COVID-19) from Oct through Nov 2020 who presented with SARS-CoV-2 symptoms but tested negative for SARS-CoV-2 using EUA RT-PCR tests, BioFire RP V2.1, and EUA antibody tests (Table 2). The NPA analysis included samples from individuals of any age/sex/race collected prior to December 2019. Samples were excluded if the sample volume

was insufficient for analysis. No selection criteria were applied to either cohort beyond the inclusion and exclusion criteria above.

Clinical Specimens

From all sources, whole blood samples were collected in EDTA tubes, frozen, and shipped to Adaptive for immunosequencing. When paired serum samples were collected, they were tested using two different EUA antibody assays: 1) Elecsys® AntiSARS-CoV-2; Roche: qualitative detection of high affinity antibodies to SARS-CoV-2 including all isotypes, but preferentially detects IgG antibodies (https://www.labcorp.com/tests/164068/sars-cov-2-antibodies); and 2) SARS-CoV-2 Antibody, IgG; LabCorp: qualitative detection of IgG antibodies to SARS-CoV-2 (https://www.labcorp.com/tests/164055/sars-cov-2-antibody-igg).

Supplemental Table 1: Summary of Cohorts Used as Training and Holdout Sets for the Development of the T-Detect COVID Classifier

Training Cases				
Study	**Participants (n)**	**Median age**	**Female (%)**	**Study Description**
DLS[a]	337	70	50.7	Whole blood samples collected during routine care in acute and convalescent phases, procured through Discovery Life Sciences (Huntsville, AL)
NIH/NIAID	146	68	30.8	Whole blood samples collected in Brescia and Monza, Italy during active infection, and provided to the NIAID (Bethesda, MD) for DNA extraction
ISB	83	63	55.4	Whole blood samples collected under the INCOVE project at Providence St. Joseph Health (Seattle, WA); participants were enrolled during the active phase and monitored through disease

H12O	156	64	37.2	Whole blood samples were collected at the Hospital Universitario 12 de Octubre (Madrid, Spain) during the active or convalescent phase
BWNW	62	54	48.4	Whole blood samples from convalescent individuals collected at Bloodworks Northwest (Seattle, WA)

Training Controls

Study	Participants (n)	Median age	Female (%)	Study Description
DLS controls batch 11	192	35	41.67	Samples collected from healthy individuals by Discovery Life Sciences
DLS controls batch 12	192	33	45.31	Samples collected from healthy individuals by Discovery Life Sciences

DLS controls batch 1	933	34	61.63	Samples collected from healthy individuals by Discovery Life Sciences
DLS controls batch 2	463	34	50.97	Samples collected from healthy individuals by Discovery Life Sciences
Bay Area, Lyme	252	47	66.67	Whole blood samples from healthy participants collected by primary care physicians in Lyme-endemic regions in the US between 2016 and 2018
OHSU, Pancreatic	55	70	45.45	Pancreatic tumor samples collected at the Oregon Health and Science University (Portland, Oregon)
Hospital Saint Louis, Crohn's	360	NA	NA	Whole blood samples collected from patients receiving surgery for Crohn's disease, collected at Hôpital Saint-Louis (Paris, France) between 2009 and 2019

Holdout Cases				
Study	**Participants (n)**	**Median age**	**Female (%)**	**Study Description**
ImmuneRACE	100	43	72.55	Samples collected by Adaptive Biotechnologies in 2020 as part of the ImmuneRACE study

Holdout Controls				
Study	**Participants (n)**	**Median age**	**Female (%)**	**Study Description**
Johns Hopkins, Lyme	226	52	45.13	Lyme disease samples collected as part of the SLICE study
Johns Hopkins, Lyme	226	52	45.13	Lyme disease samples collected as part of the SLICE study
FHCRC, lung cancer	556	65	46.94	Lung cancer samples collected by researchers in Seattle, WA

Moffitt, pancreatic cancer	492	70	45.53	Pancreatic cancer samples collected from H. Lee Moffitt Cancer Center in Florida
DLS controls batch 8	191	37	19.9	Samples collected from healthy individuals by Discovery Life Sciences
DLS controls batch 10	192	36	57.29	Samples collected from healthy individuals by Discovery Life Sciences

Abbreviations: BWNW, Bloodworks Northwest; COVID, coronavirus disease 2019; DLS, Discovery Life Sciences; FHCRC, Fred Hutchinson Cancer Research Center; H12O, Hospital Universitario 12 de Octubre; ISB, Institute for Systems Biology; NIAID, National Institute of Allergy and Infectious Diseases; NIH, National Institutes of Health; OHSU, Oregon Health and Science University; SLICE, Studies of Lyme disease Immunology and Clinical Events.

[a]Distinct DLS cohort samples were used for classifier development and primary PPA and NPA analyses.

Supplemental Table 2: Summary of COVID Cases and Controls in Cohorts Used for Training and Holdout Sets for T-Detect COVID

	Train (n)	Holdout (n)
Controls	2	1657
Cases	784	100

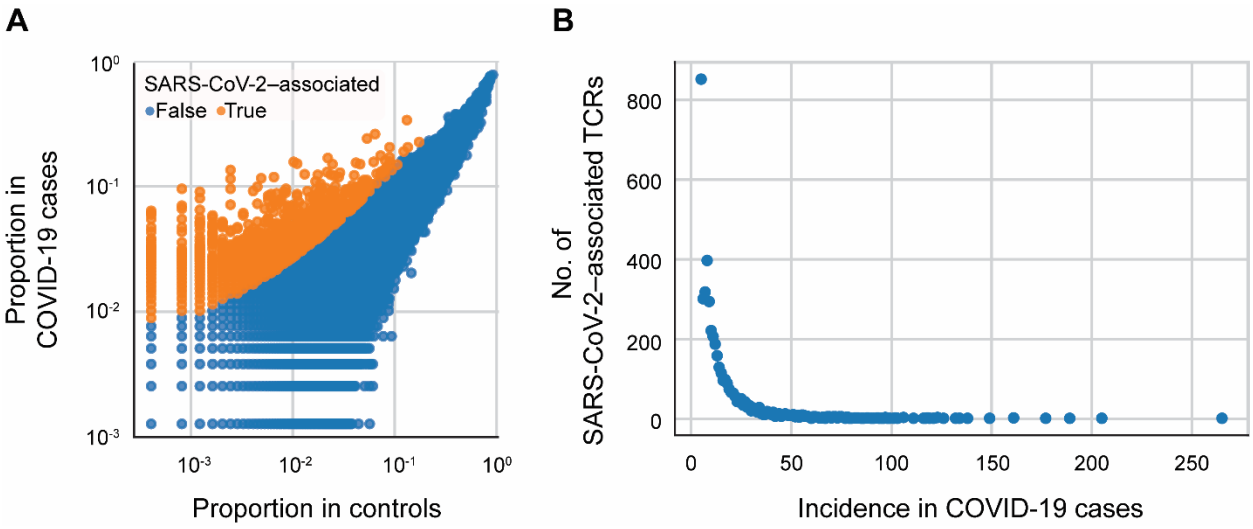
Abbreviations: COVID, coronavirus disease 2019.

Supplemental Table 3: Summary of the Effect of Case Versus Control Status, Age, and Sex on Model Log-Odds Scores

	Train; Cross-validation		Holdout	
Term	Coefficient	*P*-value	Coefficient	*P*-value
Intercept	-13.9003	<0.001	-8.102	<0.001
Status: case	111.4763	<0.001	22.7341	<0.001
Sex	1.2888	0.316	0.6387	0.141
Status: sex	-4.3133	0.08	-1.4556	0.465
Age	0.1263	0.005	0.0678	<0.001
Status: age	-0.8689	<0.001	0.5621	<0.001

The model included interactions between case status and age, and case status and sex.

Ordinary least squares regression was performed.



Supplementary Figure 1. Incidence and publicity of SARS-CoV-2–associated sequences.

(A) All unique TCRs were plotted based on incidence (defined as the proportion of the total cohort in which they appear) in cases vs. controls. Sequences found to be significantly enriched in cases that were included in the classifier are shown in orange, and others are shown in blue.

(B) Numbers of significant SARS-CoV-2-associated sequences by incidence in cases.