Title

Development and validation of blood-based prognostic biomarkers for severity of COVID disease outcome using EpiSwitch 3D genomic regulatory immuno-genetic profiling.

Authors

Ewan Hunter[1], Christina Koutsothanasi[1], Adam Wilson[1], Francisco C. Santos[1], Matthew Salter[1], Jurjen W. Westra[1], Ryan Powell[1], Ann Dring[1], Paulina Brajer[1], Benedict Egan[1], Matthew Parnall[1], Catriona Williams[1], Aemilia Katzinski[1], Thomas Lavin[1], Aroul Ramadass[1], William Messer[2], Amanda Brunton[2], Zoe Lyski[2], Rama Vancheeswaran[3], Andrew Barlow[3], Dmitri Pchejetski[4], Peter A. Robbins[5, 7], Jane Mellor[6, 7], Alexandre Akoulitchev[1]

1 Oxford BioDynamics Plc, Oxford UK
2 Oregon Health & Science University, Portland, OR
3 West Hertfordshire NHS Trust, Watford, UK
4 Norwich Medical School, University of East Anglia
5 Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, UK
6 Department of Biochemistry, University of Oxford, Oxford, UK
7 The Queen's College, University of Oxford, Oxford, UK

Address correspondence to: alexandre.akoulitchev@oxfordbiodynamics.com

Author e-mails:
Ewan Hunter: ewan.hunter@oxfordbiodynamics.com
Christina Koutsothanasi: christina.koutsothanasi@oxfordbiodynamics.com
Adam Wilson: adam.wilson@oxfordbiodynamics.com
Francisco Coroado Santos: francisco.santos@oxfordbiodynamics.com
Matthew Salter: matthew.salter@oxfordbiodynamics.com
Jurjen Westra: willem.westra@oxfordbiodynamics.com
Ryan Powell: ryan.powell@oxfordbiodynamics.com
Ann Dring: ann.dring@oxfordbiodynamics.com
Paulina Brajer: paulina.brajer@oxfordbiodynamics.com
Benedict Egan: benedict.egan@oxfordbiodynamics.com
Matthew Parnall: matthew.parnall@oxfordbiodynamics.com
Catriona Williams: catriona.williams@oxfordbiodynamics.com
Aemilia Katzinski: aemilia.katzinski@oxfordbiodynamics.com
Thomas Lavin: tomas.lavin@oxfordbiodynamics.com
William Messner: messner@ohsu.edu
Amanda Brunton: brunton@ohsu.edu
Zoe Lyski: lyski@ohsu.edu
Peter Robbins: peter.robbins@dpag.ox.ac.uk

Jane Mellor: jane.mellor@bioch.ox.ac.uk
Rama Vancheeswaran: rama.vancheeswaran @nhs.net
Andrew Barlow: a.barlow1@nhs.net
Dmitri Pchejetski: d.pshezhetskiy@uea.ac.uk
Alexandre Akoulitchev: alexandre.akoulitchev@oxfordbiodynamics.com

Abstract

The COVID-19 pandemic has raised several global public health challenges to which the international medical community have responded. Diagnostic testing and the development of vaccines against the SARS-CoV-2 virus have made remarkable progress to date. As the population is now faced with the complex lifestyle and medical decisions that come with living in a pandemic, a forward-looking understanding of how a COVID-19 diagnosis may affect the health of an individual represents a pressing need. Previously we used whole genome microarray to identify 200 3D genomic marker leads that could predict mild or severe COVID-19 disease outcomes from blood samples in a multinational cohort of COVID-19 patients. Here, we focus on the development and validation of a qPCR assay to accurately predict severe COVID-19 disease requiring intensive care unit (ICU) support and/or mechanical ventilation. From 200 original biomarker leads we established a classification model containing six markers. The markers were qualified and validated on 38 COVID-19 patients from an independent cohort. Overall, the six-marker model obtained a positive predictive value of 93% and balanced accuracy of 88% across 116 patients for the prognosis of COVID-19 severity requiring ICU care/ventilation support. The six-marker signature identifies individuals at the highest risk of developing severe complications in COVID-19 with high predictive accuracy and can assist in patient prognosis and clinical management decisions.

Background

The COVID-19 outbreak, which the World Health Organization (WHO) declared a pandemic in March 2020, represents one of the greatest global health crises the world has faced in recent history [1]. In addition to the estimated 130+ million people that have been infected with the SARS-CoV-2 virus to date and the more than 3 million deaths attributed to COVID-19 related causes; the pandemic has placed tremendous strain on healthcare systems, caused devastating mental health crises, and tested global economic resiliency [2, 3]. In response to the COVID-19 crisis, the global research and medical communities responded quickly with the rapid development and deployment of diagnostic tests to identify individuals who are currently infected or have previously been infected with the SARS-CoV-2 virus. As of March 2021, over 250 different rapid antigen or molecular tests to detect SARS-CoV-2 have received Emergency Use Authorization by the US FDA alone [4, 5]. For several reasons, including the early immediate unmet need to identify SARS-CoV-2 positive individuals and the lack of availability of real-world data on how different individuals respond to the virus, research on features that are predictive for different manifestations of COVID-19 have lagged behind the development of diagnostic testing. The few predictive measures that have been assessed suffer from low certainty, high bias, and insufficient predictive accuracy [6]. As clinicians have gained experience in treating COVID-19 patients, one of the more striking observations has been the wide range of clinical manifestations of the disease in different individuals; many of which could not be predicted by

98 basic clinical parameters such as age, gender, or pre-existing conditions/co-
99 morbidities alone [7–9]. It is now clear that the variability in host response,
100 rather than viral genetics or load, is the primary determinant of the wide range
101 of observed disease manifestations.
102
103 While diagnostic testing is now routine, there remains a pressing need to
104 develop molecular readouts that can reveal individual susceptibility to
105 developing severe COVID-19 symptoms in response to the SARS-CoV-2 virus
106 with high predictive accuracy to support clinical decision making and patient
107 management [10]. Specifically, there is a need to develop accurate, non-
108 invasive tests to identify prognostic profiles for clinical cases where patients
109 develop hyperinflammation, do not respond to the standard of care, and
110 progress to needing ICU-level support and/or mechanical ventilation in
111 accordance with the WHO severity definitions [10]. Individualized knowledge
112 of disease severity risk can help people take appropriate precautions in their
113 daily lives, understand how a COVID-19 diagnosis may impact a pre-existing
114 medical condition, anticipate and plan for medical treatment if they become
115 infected, and decrease their personal risk of contracting severe symptoms by
116 managing risk factors. All of these are critical tools for managing the global
117 health crisis, even in the face of the recent development of COVID-19 vaccines.
118 The first vaccines against the SARS-CoV-2 virus became available in the late
119 2020 and represented a major step forward in containing the spread of the virus
120 and its myriad social tolls [11–13]. However, several factors may limit the
121 effectiveness of universal vaccination: 1) vaccine shortages and limited national
122 vaccination rates 2) contraindications in certain patient populations (e.g.

123 children under 18, individuals who have certain health conditions, such as
124 recent infections and anaphylaxis) 3) individual decisions to forego vaccination
125 and 4) reductions in vaccine efficacy in newly emerging SARS-CoV-2 variants
126 as demonstrated recently in a Chilean population with persistently high infection
127 rates despite the availability of the CoronaVac vaccine and >40% national
128 vaccination rates [14]. As such, the ability for individuals and their physicians
129 to understand the risk of developing severe symptoms if infected with the
130 SARS-CoV-2 virus remains a critical unmet healthcare need.
131
132 Recently, we reported on the use of *EpiSwitch*®, a chromosome conformation
133 capture (3C) methodology that has been reduced to practice for the discovery
134 of blood-based 3D genomic biomarkers and patient stratification in a variety of
135 immune-related diseases, to identify prognostic profiles that could serve as a
136 potential blood-based molecular classifier for the prediction of COVID-19
137 disease severity [15–22]. Specifically, we used *EpiSwitch* Explorer 3D whole
138 genome arrays to identify changes at immune-related loci in blood samples
139 from a multi-national cohort of COVID-19 patients. We identified 200 array-
140 based 3D genomic marker leads that could differentially predict mild or severe
141 COVID-19 disease severity, with a high degree of sensitivity and specificity.
142 The proteins encoded at the locations of the 3D genomic marker leads were
143 involved in biological pathways with direct relevance to immune system function
144 including T-cell signalling, macrophage-stimulating protein (MSP)-RON
145 signalling, and calcium signalling [15]. The aim of this study was an extension
146 of our previous work, using the established *EpiSwitch* platform to translate the
147 array-derived 3D genomic biomarkers into a qPCR-based testing format, refine

and reduce the biomarker panel, build a classification model and perform validation on independent COVID-19 patient cohorts. Here we report on the development of a 3D genomic biomarker panel with clinical utility in predicting disease severity in COVID-19.

Materials and Methods

Patient characteristics

Clinical whole blood samples from consented patients were supplied from Boca Biolistics LLC (FL, USA) and Reprocell USA Inc. (MD, USA). A total of 116 patients from three sample cohorts were used in this study, comprising a multinational set of COVID-19 cases from the United States, Peru, and the Dominican Republic. Patient annotations are listed in (**Supplemental Tables 1,2)**. In line with WHO guidelines, the patient annotations provided were used to classify the severe outcome group on the basis of a confirmed admission to the Intensive Care Unit (ICU) and/or advanced clinical interventions such as mechanical ventilation [10]. Patients that required a lower level of clinical care, such as administration of supplemental oxygen only, were classified as the mild outcome group. All samples were collected within 72 hours of a patient being admitted to a hospital for treatment of a PCR-confirmed COVID infection. The age of the patients ranged from 28 to 92, with median of 64.5 years (**Supplemental Tables 1, 2**).

Preparation of 3D genomic templates

EpiSwitch® 3C libraries, with chromosome conformation analytes converted to sequence based tags, were prepared from frozen whole blood samples using

EpiSwitch® protocols following the manufacturer's instructions (Oxford BioDynamics Plc) [15]. All samples were processed under biological containment level CL2+. Initial sample processing was performed manually in a Category 3 microbial safety cabinet with the remainder performed on the Freedom EVO 200 robotic platform (Tecan Group Ltd). Briefly, 50 μL of whole blood sample was diluted and fixed with a formaldehyde containing *EpiSwitch* buffer. Density cushion centrifugation was used to purify intact nuclei. Following a short detergent-based step to permeabilise the nuclei, restriction enzyme digestion and proximity ligation were used to generate the 3C libraries. Samples were centrifuged to pellet the intact nuclei before purification with an adapted protocol from the QIAmp DNA FFPE Tissue kit (Qiagen) Eluting in 1x TE buffer pH7.5. 3C libraries were quantified using the Quant-iT™ Picogreen dsDNA Assay kit (Invitrogen) and normalised to 5 ng/μl prior to interrogation by PCR.

Translation of array-based 3D genomic markers to PCR readouts

The top array-derived markers identified in our previous study were interrogated using OBD's proprietary primer design software package to identify genomic positions suitable for a hydrolysis probe based real time PCR (RT-PCR) assay [15]. Briefly, the top array-derived markers associated with prognostic potential to differentiate between mild and severe COVID disease outcomes were filtered on fold change and adjusted p value. PCR primer probes were ordered from Eurofins genomics as salt-free primers. The probes were designed with a 5' FAM fluorophore, 3' IABkFQ quencher and an additional internal ZEN quencher and ordered from iDT (integrated DNA

198 Technologies) [23]. Each assay was optimised using a temperature gradient
199 PCR with an annealing temperature range from 58-68° C. Individual PCR
200 assays were tested across the temperature gradient alongside negative
201 controls including soluble and unstructured commercial TaqMan human
202 genomic DNA control (Life Technologies) and used a TE buffer only negative
203 control. Assay performance was assessed based on Cq values and reliability
204 of detection and efficiency based on the slope of the individual amplification
205 curves. Assays that passed the quality criteria and presented with reliable
206 detection differences between the pooled samples associated with Severe and
207 Mild COVID disease outcomes were used to screen individual patient samples.
208
209 *EpiSwitch PCR*
210 Each patient sample was interrogated using RT-PCR in triplicate. Each reaction
211 consisted of 50 ng of EpiSwitch library template, 250 mM of each of the primers,
212 200 mM of the hydrolysis probe and a final 1X Kapa Probe Force Universal
213 (Roche) concentration in a final 25 μl volume. The PCR cycling and data
214 collection was performed using a CFX96 Touch Real-Time PCR detection
215 system (Bio-Rad). The annealing temperature of each assay was changed to
216 the optimum temperature identified in the temperature gradients performed
217 during translation for each assay. Otherwise, the same cycling conditions were
218 used: 98°C for 3 minutes followed by 45 cycles of 95°C for 10 seconds and 20
219 seconds at the identified optimum annealing temperature. The individual well
220 Cq values were exported from the CFX manager software after baseline and
221 threshold value checks. All Cq values obtained for individual samples and
222 markers are available online

[https://github.com/oxfordBiodynamics/medrxiv/tree/main/CST%20publication]. A total of 21 3D genomic markers that passed the translation phase were screened on 78 individual samples from the Training cohort. A marker reduction step based on statistical criteria were used to identify the top six discriminating markers which were used to screen the remaining set of 38 samples in the Test cohort.

Genomic mapping

The 21 3D genomic markers from the statistically filtered list with the greatest and lowest abundance scores were selected for genome mapping. Mapping was carried out using Bedtools *closest* function for the 3 closest protein coding loci (Gencode v33). All markers were visualized using the *EpiSwitch* Data Portal.

Statistical analysis

The 21 markers screened on 78 individual patient samples were subject to permutated logistic modelling with bootstrapping for 500 data splits and non-parametric Rank Product analysis (EpiSwitch® RankProd R library). Two machine learning procedures (eXtreme Gradient Boosting: XGBoost and CatBoost) were used to further reduce the feature pool and identify the most predictive/prognostic, 3D genomic markers. The resulting markers were then used to build the final classifying models using CatBoost and XGBoost. All analysis was performed using R statistical language with Caret, XGBoost, SHAPforxgboost and CatBoost libraries.

Biological network/pathway analysis

Pathway enrichment analysis was performed using the Reactome Pathway Browser [24]. Protein interaction networks were generated using the Search Tool for the Retrieval of Interacting proteins (STRING) database [24, 25].

Results

Identification of the top prognostic 3D genomic markers for severe COVID-19 disease outcomes

In this study we employed a sequential stepwise strategy to identify a minimal set of biomarkers that were predictive of COVID-19 disease severity (**Figure 1**). Starting with an initial set of 200 array-derived 3D genomic marker leads identified in our previously published work, here we used independent sets of 116 whole blood samples collected at the time of confirmed COVID-19 infection from patients in the United States, Peru, and the Dominican Republic to refine and test the prognostic 3D genomic biomarker panel. Whole blood samples were procured for a Training cohort used to build and refine the classifier model, and Test cohort to assess the predictive performance of the model. Clinical characteristics of the patients are shown in **Table 1 and Supplemental Tables 1, 2**. To translate the *EpiSwitch* Explorer Array markers to a PCR-detectable assay, primers to detect individual 3D genomic markers were generated and validated (Materials & Methods). Using a 78 sample Training set of COVID-19 patient whole blood samples, feature reduction using machine learning methods on the initial pool of 200 3D genomic biomarkers identified 21 markers with predictive power to differentiate between COVID-19 patients that required a high degree of medical disease management (e.g. admission to the intensive

care unit (ICU), mechanical ventilation) and those that were hospitalized but required less interventional care and support (Supplemental **Table 1, for Cq data see online link –** *https://github.com/oxfordBiodynamics/medrxiv/tree/main/CST%20publication*). We found the top 21 markers to be non-randomly distributed throughout the human genome, with notable enrichment on larger chromosomes and a hotspots on chromosome 11 (**Figure 2A**). Four out of the 21 markers associated with ICU outcomes occurred within an approximately 265 kb region on the p-arm of chromosome 11 containing the switching B cell complex subunit SWAP70 (SWAP70) locus (**Figure 2B**). While some of the 3D genomic markers were localized within protein coding regions of single genes (**Figure 2C**), other markers spanned the protein coding regions of multiple genes (**Figure 2B**). Pathway enrichment for genes contained within the 21 3D genomic markers revealed the top two pathways to be related to downstream signalling mediated by B-cell receptor activation (**Table 3**). Importantly, genomic loci encoding proteins involved in hemostasis/clotting were also enriched (**Figure 3, Table 3)**. The 21 3D genomic markers were further refined to a set of 6 markers (**Table 4**) with predictive ability for COVID severity and applied to an independent Test cohort (Supplemental **Table 2**).

Testing of the prognostic 3D genomic biomarker panel for severe COVID-19 disease outcomes on independent patient cohorts.

To assess the predictive power of the model, the 6-marker 3D genomic panel was validated on an independent (samples that were not used to build and refine the model) Test cohort (**Figure 4, Supplemental Table 2**). Samples were

298 collected upon admission to COVID hospital wards in Peru, the USA, and the
299 Dominican Republic and shipped to OBD's processing facility in Oxford, UK.
300 The *EpiSwitch* platform read outs for the six-marker classifier model were
301 uploaded to the *EpiSwitch* Analytical Portal for analysis. Classifier calls for high-
302 risk COVID-19 disease outcomes are shown in **Table 5**. Clinical outcomes for
303 the Test cohort included 10 mild cases or 28 severe cases requiring ventilation
304 and/or ICU support. *EpiSwitch* prognostic calls based on the 6-marker model
305 demonstrated performance of 90.9% positive predictive value for high-risk
306 disease outcomes in the Test cohort (**Figure 4A**). Interestingly, two of the mild
307 case patients (COVID 0696 and 0213) (**Table 5**), identified as high risk by the
308 *EpiSwitch* test subsequently died in the hospital within 28 days of admission.
309 This suggests an early, pre-symptomatic detection of a hyperinflammatory state
310 leading to fatal outcomes and is being investigated further. Across all 116
311 patients used in this study, the test demonstrated positive predictive value for
312 high-risk disease outcomes of 92.9 with, 88% sensitivity, 87% specificity, and
313 a balanced accuracy of 87.9% (**Table 4B and Supplementary Table 3**).
314
315
316
317
318
319
320
321
322

Figures & Tables

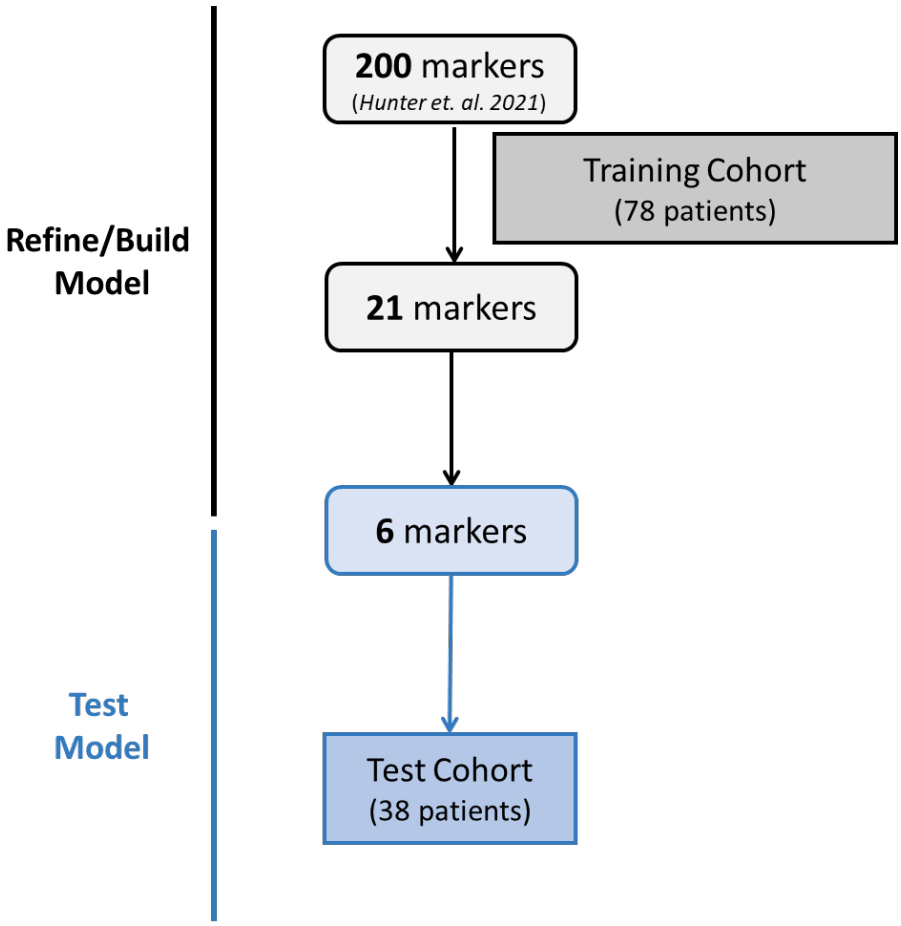


Figure 1. Simplified workflow used to develop and test the prognostic 3D genomic classifier model for prediction of COVID-19 disease severity. Starting from a list of 200 3D genomic markers identified by Hunter et al. 2021, we used a sequential, stepwise approach employing a 78-patient Training cohort to refine the marker set and build a predictive classifier model containing six 3D genomic markers. The 6-marker model/assay was tested on an independent Test cohort of 38 COVID-19 patient blood samples.

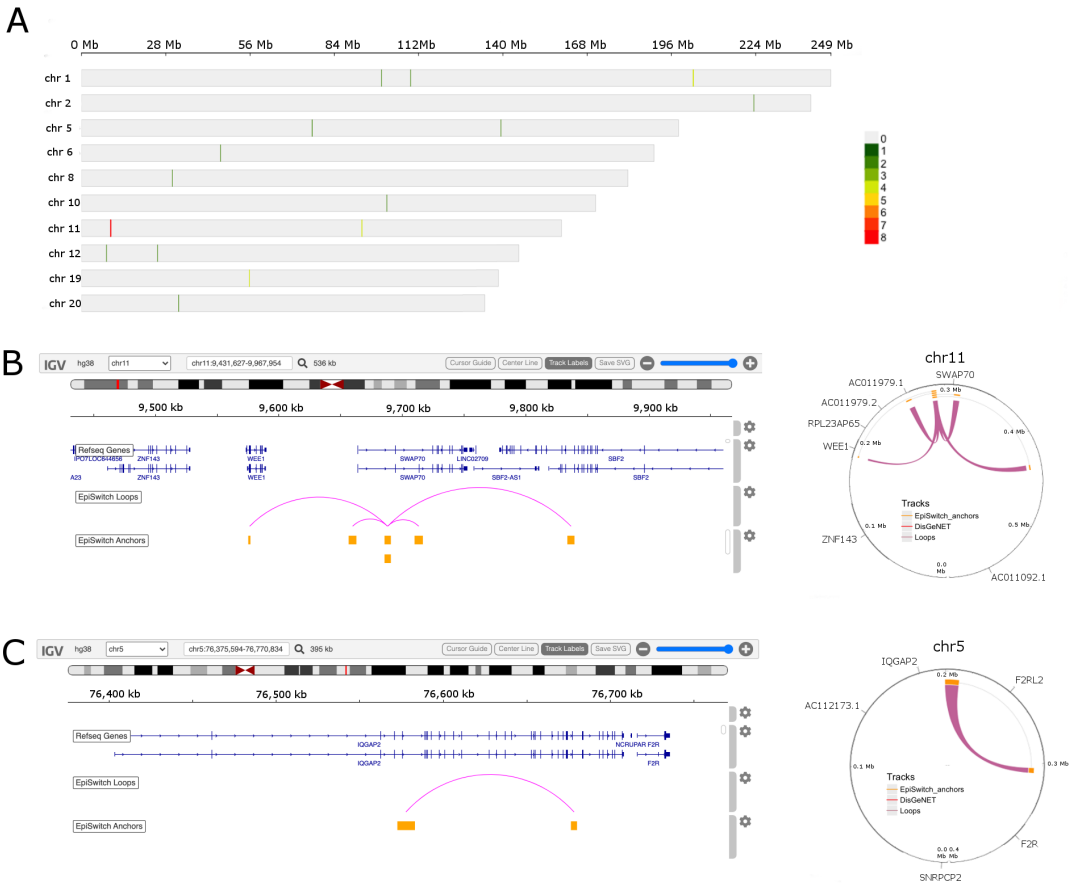


Figure 2. Genomic detail view of top prognostic 3D genomic biomarkers. **A.** Genomic locations and distribution of the top 21 3D genomic markers for Hospitalized and ICU clinical outcomes. Individual human chromosomes where the top 21 markers were found are shown on the y-axis. The heatmap shows the number of markers within a 0.3Mb genomic window with green representing a low density of markers and red indicating a high density of markers. A region on chromosome 11 shows a high density of markers associated with ICU outcomes. **B.** Linear and circos plot views of a ~500 kb region of chromosome 11 containing the SWAP70 locus showing the genomic location for four markers. **C**. Linear and circos views of a ~400 kb region of chromosome 5 containing the IQGAP2 and F2RL2 loci showing the genomic location for one of the markers in the final 6-marker set.

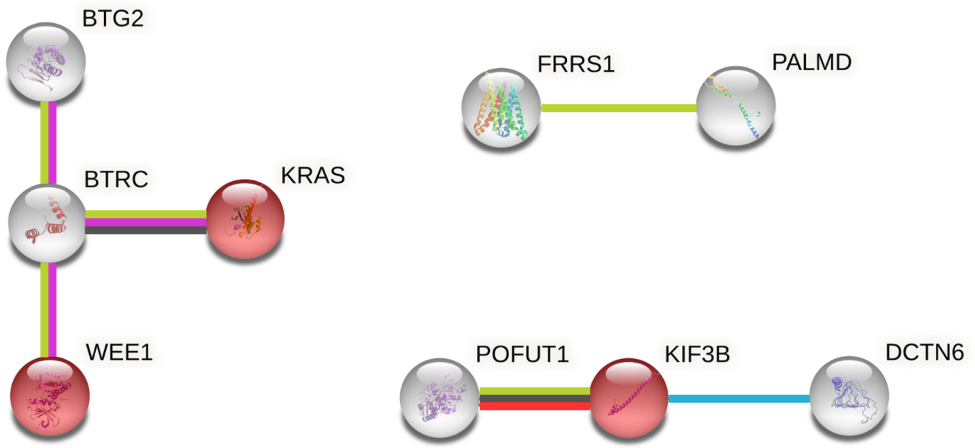


Figure 3. STRING network for top 3D genomic markers. Protein-protein interaction network for proteins encoded by genes spanning the top 21 3D genomic marker set. Edges are colored by protein-protein association type (blue = known interactions from curated databases, magenta = known interactions from experiments, light green = interactions derived from literature text mining, orange = gene fusions, black = association by co-expression). Nodes highlighted in red are associated with the Reactome 'Hemostasis' pathway (HSA-109582).

A

Test	Severe Outcome: Present	n	Absent	n	Total
Yes	True Positive	20	False Positive	2	22
No	False Negative	8	True Negative	8	16
Total		28		10	

Results — Statistic	Value	95% CI
Sensitivity	71.43%	51.33% to 86.78%
Specificity	80.00%	44.39% to 97.48%
Positive Likelihood Ratio	3.57	1.01 to 12.61
Negative Likelihood Ratio	0.36	0.18 to 0.69
Disease prevalence*	73.68%	56.90% to 86.60%
Positive Predictive Value*	90.91%	73.90% to 97.25%
Negative Predictive Value*	50.00%	34.02% to 65.98%
Accuracy*	73.68%	56.90% to 86.60%

B

Test	Severe Outcome: Present	n	Absent	n	Total
Yes	True Positive	66	False Positive	5	71
No	False Negative	9	True Negative	36	45
Total		75		41	

Results — Statistic	Value	95% CI
Sensitivity	88.00%	78.44% to 94.36%
Specificity	87.80%	73.80% to 95.92%
Positive Likelihood Ratio	7.22	3.16 to 16.48
Negative Likelihood Ratio	0.14	0.07 to 0.25
Disease prevalence*	64.66%	55.24% to 73.31%
Positive Predictive Value*	92.96%	85.25% to 96.79%
Negative Predictive Value*	80.00%	68.20% to 88.18%
Accuracy*	87.93%	80.58% to 93.24%

Figure 4. Performance of the prognostic biomarker classifier for calling high risk of severe disease outcome on COVID-19 patient cohorts. (A)

Confusion matrix and test performance statistics for the 6-marker classifier on the 38 patients of the Test cohort (B) and the 116 patients in the combined Training and Test cohorts. Note: (*) These values are dependent on disease prevalence

Table 1. Summary of clinical characteristics by cohort for COVID-19 patient samples used in this study. Clinical features of the 116 COVID-19 patient samples used in this study by cohort. SD: Standard deviation; PEC: pre-existing condition; Hosp.: Hospitalized; ICU: Intensive Care Unit; SO: Received supplemental oxygen; Vent.: Received mechanical ventilation.

Cohort	*N*	*% male*	*% female*	*Age (mean)*	*Age (SD[a])*	*% w/PEC[b]*	*% Hosp.[c]*	*% ICU[d]*	*% SO[e]*	*% Vent.[f]*
Training	78	64	36	64.8	14.5	76	46	54	37	60
Test	38	79	21	65.1	16.2	90	53	47	29	74

[a] Standard deviation
[b] pre-existing condition
[c] Hospitalized
[d] Intensive Care Unit
[e] Received supplemental oxygen
[f] Received mechanical ventilation

Table 2. List of the top 21 3D genomic markers for prediction of COVID-19 disease severity. List of the top 21 markers with predictive ability for COVID-19 severity. Markers are listed by the OBD internal ID. The six markers in the final 3D genomic panel are outlined in red. The closest protein-coding genes near the 3D genomic markers are also listed (Closest Genes).

Marker[a]	Closest Genes[b]
hg38_10_101411215_101490136_RF	BTRC,DPCD,POLL
hg38_11_9577172_9685884_FR	AC011979.1,AC011979.2,RPL23AP65,SWAP70,WEE1
hg38_20_32238035_32290178_FF	KIF3B,PLAGL2,POFUT1
hg38_11_9685855_9716901_RF	AC011979.1,AC011979.2,SWAP70
hg38_1_109341941_109359750_RR	MYBPHL,PSMA5,SORT1
hg38_11_9663012_9685884_FR	AC011979.1,AC011979.2,SWAP70
hg38_5_139331499_139356679_FF	MATR3,PAIP2,SLC23A1

hg38_2_223395100_223450604_FF	AP1S3,HIGD1AP4,KCNE4,SCG2
hg38_1_99670351_99714401_FF	AGL,FRRS1,HMGB3P10,PALMD
hg38_8_30132538_30177089_RR	DCTN6,LEPROTL1,MBOAT4
hg38_12_8219312_8342000_RR	AC092745.2,AC092745.3,ALG1L10P,CLEC4A,ENPP7P5, FAM86FP,FAM90A1
hg38_11_9685855_9839717_RF	AC011979.1,SBF2,SWAP70
hg38_19_55694909_55778461_RF	AC008749.1,AC010525.2,EPN1,NLRP9,RFPL4A,RFPL4A L1,RFPL4AP1
hg38_6_46139224_46175482_FF	ACTG1P9,ENPP4,ENPP5
hg38_12_25206967_25256704_FR	CASC1,ETFRF1,KRAS
hg38_11_93198707_93237221_RR	SLC36A4;MTNR1B;DEUP1
hg38_5_76572659_76680168_RF	F2R,F2RL2,IQGAP2
hg38_1_203182882_203350382_FR	BTG2,CHI3L1,CHIT1,FMOD,NPM1P40
hg38_19_55711884_55778461_RF	NLRP9,RFPL4A,RFPL4AL1,RFPL4AP1
hg38_1_203182882_203368482_FF	AL359837.1,BTG2,CHI3L1,CHIT1,FMOD,NPM1P40
hg38_11_93057516_93237221_RR	SLC36A4;MTNR1B;DEUP1

384

Pathway identifier[a]	Pathway name[b]	#Entities found[c]	#Entities total[d]	Entities ratio[e]	Entities pValue[f]	Entities FDR[g]	#Reactions found[h]	#Reactions total[i]	Reactions ratio[j]	Species name[j]
R-HSA-1168372	Downstream signaling events of B Cell Receptor (BCR)	3	91	0.006266354	8.80E-04	0.008475369	3	15	0.001117235	Homo sapiens
R-HSA-983705	Signaling by the B Cell Receptor (BCR)	3	189	0.013014736	0.006883631	0.019502853	3	44	0.003277223	Homo sapiens
R-HSA-5621481	C-type lectin receptors (CLRs)	3	203	0.013978791	0.008365905	0.019502853	7	68	0.0050648	Homo sapiens
R-HSA-69275	G2/M Transition	3	212	0.01459854	0.009410889	0.019502853	8	78	0.005809623	Homo sapiens
R-HSA-453274	Mitotic G2-G2/M phases	3	214	0.014736262	0.009653045	0.019502853	8	80	0.005958588	Homo sapiens
R-HSA-2454202	Fc epsilon receptor (FCERI) signaling	3	235	0.016182344	0.012417469	0.024834939	2	63	0.004692388	Homo sapiens
R-HSA-416476	G alpha (q) signalling events	3	283	0.019487674	0.020298035	0.040596071	4	35	0.002606882	Homo sapiens
R-HSA-109582	Hemostasis	5	801	0.055157692	0.022922661	0.045845322	13	334	0.024877104	Homo sapiens
R-HSA-8856688	Golgi-to-ER retrograde transport	2	148	0.010191434	0.037413554	0.069833207	4	18	0.001340682	Homo sapiens
R-HSA-2132295	MHC class II antigen presentation	2	148	0.010191434	0.037413554	0.069833207	1	26	0.001936541	Homo sapiens
R-HSA-199991	Membrane Trafficking	4	665	0.045792591	0.046627545	0.069833207	12	219	0.016311634	Homo sapiens
R-HSA-168249	Innate Immune System	6	1333	0.091791764	0.052086077	0.069833207	14	708	0.052733502	Homo sapiens
R-HSA-1280218	Adaptive Immune System	5	1004	0.069136483	0.052885643	0.069833207	9	264	0.01966334	Homo sapiens
R-HSA-9013149	RAC1 GTPase cycle	2	191	0.013152458	0.059025208	0.069833207	2	6	4.47E-04	Homo sapiens
R-HSA-983231	Factors involved in megakaryocyte development and platelet production	2	194	0.013359041	0.060664833	0.069833207	5	43	0.003202741	Homo sapiens
R-HSA-6811442	Intra-Golgi and retrograde Golgi-to-ER traffic	2	218	0.015011706	0.074331183	0.074331183	4	48	0.003575153	Homo sapiens
R-HSA-6798695	Neutrophil degranulation	3	480	0.033053298	0.075629054	0.075629054	4	10	7.45E-04	Homo sapiens
R-HSA-168256	Immune System	9	2681	0.184616444	0.086870431	0.086870431	29	1621	0.120735886	Homo sapiens
R-HSA-5653656	Vesicle-mediated transport	4	824	0.056741496	0.087750436	0.087750436	12	252	0.018769552	Homo sapiens
R-HSA-71240	Tryptophan catabolism	1	48	0.00330533	0.094550828	0.094550828	1	16	0.001191718	Homo sapiens

385 [a] The Reactome's Pathway Identifier.

386 [b] The Reactome's Pathway Name.

387 [c] The number of curated molecules of the type selected with Results Type that

388 are common between the submitted data set and the pathway named in column

389 1.

390 [d] The total number of curated molecules of the type selected with Results Type

391 within the pathway named in column 1.

[e] The proportion of Reactome pathway molecules represented by this pathway. Calculated as the ratio of entities from this pathway that are molecules of the type selected with Results Type Vs.

[f] The result of the statistical test for over-representation, for molecules of the results type selected.

[g] False discovery rate. Corrected over-representation probability.

[h] The number of reactions in the pathway that are represented by at least one molecule in the submitted data set, for the molecule type selected with Results Type.

[i] The number of reactions in the pathway that contain molecules of the type selected with Results Type.

[j] The proportion of Reactome reactions represented by this pathway. Calculated as the ratio of reactions from this pathway that contain molecules of the type selected with Results Type Vs. all Reactome reactions that contain molecules of the type selected with Results Type.

[k] The Species name

Table 4. List of the top 6 3D genomic markers for prediction of COVID-19 disease severity. List of the 6 markers that comprise the predictive COVID-19 severity biomarker panel. Markers are listed by the OBD internal ID. The closest protein-coding genes near the 3D genomic markers are also listed in the 'Closest Genes' column.

Marker[a]	Closest Genes[b]
hg38_1_109341941_109359750_RR	MYBPHL,PSMA5,SORT1
hg38_5_76572659_76680168_RF	F2R,F2RL2,IQGAP2
hg38_2_223395100_223450604_FF	AP1S3,HIGD1AP4,KCNE4,SCG2
hg38_6_46139224_46175482_FF	ACTG1P9,ENPP4,ENPP5
hg38_8_30132538_30177089_RR	DCTN6,LEPROTL1,MBOAT4

hg38_19_55694909_55778461_RF	AC008749.1,AC010525.2,EPN1,NLRP9,RFPL4A,RFPL4AL1,RFPL4AP1

[a] **The marker's ID**

[b] **The genes around the genomic loci of the marker**

Table 5. Prognostic calls for high-risk of severe outcome with the 6-marker *EpiSwitch* classifier model. Prognostic calls of average and high risk for the 38 patients against disclosed attributes of severity: ventilation and /or ICU support

Cohort	SampleID	COVID Severity		EpiSwitch Prognostic Call for High-Risk		
		Ventilation	ICU[a]	No	Yes	Final Call
Test	COVID0732	No	No	0.624535561	0.3754644	No
Test	COVID0129	No	No	0.989352465	0.0106475	No
Test	COVID0636	No	No	0.810631394	0.1893686	No
Test	COVID0189	No	No	0.96364671	0.0363533	No
Test	COVID0708	No	No	0.918016613	0.0819834	No
Test	COVID0117	No	No	0.760194659	0.2398053	No
Test	COVID0207	No	No	0.740656555	0.2593434	No
Test	COVID0380	No	No	0.990677834	0.0093222	No
Test	COVID0696	No	No	0.020404769	0.9795952	**Yes***
Test	COVID0213	No	No	0.04568797	0.954312	**Yes***
Test	COVID0606	**Yes**	No	0.809160769	0.1908392	No
Test	COVID0648	**Yes**	No	0.987436414	0.0125636	No
Test	COVID0642	**Yes**	No	0.665544152	0.3344558	No
Test	COVID0516	**Yes**	**Yes**	0.601811945	0.3981881	No
Test	COVID0564	**Yes**	**Yes**	0.942398548	0.0576015	No
Test	COVID0450	**Yes**	**Yes**	0.885789573	0.1142104	No
Test	COVID0714	**Yes**	No	0.888814926	0.1111851	No
Test	COVID0408	**Yes**	**Yes**	0.700852036	0.299148	No
Test	COVID0558	**Yes**	**Yes**	0.056852765	0.9431472	**Yes**
Test	COVID0540	**Yes**	**Yes**	0.26985541	0.7301446	**Yes**
Test	COVID0444	**Yes**	**Yes**	0.012335699	0.9876643	**Yes**
Test	COVID0456	**Yes**	**Yes**	0.34420839	0.6557916	**Yes**
Test	COVID0468	**Yes**	**Yes**	0.26985541	0.7301446	**Yes**
Test	COVID0498	**Yes**	**Yes**	0.045760725	0.9542393	**Yes**
Test	COVID0576	**Yes**	**Yes**	0.057154838	0.9428452	**Yes**
Test	COVID0504	**Yes**	**Yes**	0.006351133	0.9936489	**Yes**
Test	COVID0600	**Yes**	No	0.106978044	0.893022	**Yes**
Test	COVID0672	**Yes**	No	0.08792568	0.9120743	**Yes**
Test	COVID0588	**Yes**	No	0.028880829	0.9711192	**Yes**
Test	COVID0654	**Yes**	No	0.029438535	0.9705615	**Yes**
Test	COVID0666	**Yes**	No	0.124919437	0.8750806	**Yes**
Test	COVID0726	**Yes**	No	0.198130682	0.8018693	**Yes**
Test	COVID0474	**Yes**	**Yes**	0.077650517	0.9223495	**Yes**
Test	COVID0432	**Yes**	**Yes**	0.145361423	0.8546386	**Yes**
Test	COVID0462	**Yes**	**Yes**	0.06204395	0.9379561	**Yes**
Test	COVID0510	**Yes**	**Yes**	0.248548523	0.7514515	**Yes**
Test	COVID0768	**Yes**	**Yes**	0.377659917	0.6223401	**Yes**

Test	COVID0427	**Yes**	**Yes**	0.349406302	0.6505937	**Yes**

[a] Intensive Care Unit

Note: (*) these patients were called as high-risk an initially annotated clinicaly as mild-COVID cases. They subsequently died in the hospital within 28 days of admission

Supplemental Table 1 & 2. Annotations and clinical characteristics for COVID-19 patient samples used in this study. Individual characteristics and clinical features for the 116 COVID-19 patient samples used in this study (78 Training Set & 38 Test Set). CV: cardiovascular; DR: Dominican Republic; PEC: pre-existing condition; T2DM: Type 2 diabetes mellitus.

Supplemental Table 3. Prognostic calls for high-risk with 6-marker *EpiSwitch* classifier model. Prognostic calls of average and high risk for the 78 patients against disclosed attributes of severity: ventilation and /or ICU support.

Discussion

The COVID-19 pandemic will represent a major public health crisis for months to come. As a corollary, there remains a pressing need for prognostic testing that can give individuals a readout of likely disease severity if infected with SARS-CoV2, despite the rigorous social measures put in place to slow the spread of the disease and the recent availability of vaccination options. Despite the development and approval of vaccines by Pfizer-BioNTech, Moderna, Janssen, and AstraZeneca for use in the US and EU, there are still many individuals that will not be vaccinated due to 1) lack of access 2) ineligibility or

3) a personal choice. The worldwide vaccine roll-out itself has faced difficulties. Despite initial optimism, vaccination rates in the EU are low [26, 27]. In the US, the vaccine rollout has also faced adoption challenges. Around one third of military personnel and a similar percentage of healthcare workers have opted against vaccination, citing a variety of factors in their decision [28, 29]. Even as the vaccine is available in many parts of the U.S., several states including Minnesota, Michigan, Maryland, and Oregon have seen spikes in average new daily COVID-19 cases in the month of April 2021 [30]. Even in countries where vaccination campaigns have gone well from the standpoint of adoption rates such as Israel and the United Kingdom, severe complications due to COVID-19 are still seen [31, 32]. The emergence of novel coronavirus variants that render some vaccines less effective also threatens to prolong the pandemic. Currently there are six main variants circulating in the U.S.: the B.1.1.7 originating in the U.K.; the B.1.351 from South Africa; the P.1. first detected in Brazil; and the B.1.427 and B.1.429, both originating in California and the highly infectious B.1.617.2 (Delta) variant originating in India [33–36]. Based on decades of research on viral evolution, the emergence of other strains in the future is likely [37, 38]. As SARS-CoV-2 variants can spread faster than the original strain and some variants may be resistant to the neutralizing antibodies produced by individuals after vaccination, the risk of a widespread disease resurgence and the continued societal burdens of an extended pandemic are real [39–41].

While many molecular tests have been developed to detect the presence of the SARS-CoV-2 virus, a reduced to practice molecular test with high prognostic

accuracy for severe outcomes following a SARS-CoV-2 infection represents an unmet healthcare need. Here we extended on a previously published study and describe the validation of a blood-based biomarkers for the prediction of severe symptoms of COVID-19 disease, which can place individuals in the intensive care support [15]. We identified a set of six biomarkers associated with systemic immune-genetic dysregulation, that when processed and measured by qPCR on human blood samples provided an accurate predictive readout on likely COVID-19 disease severity. While to the best of our knowledge, a comparable assay has not yet been reported on, the performance of the *EpiSwitch* assay meets the urgent need for many individuals to have *a priori* knowledge of likely disease severity in the event they are infected with the SARS-CoV-2 virus.

Interestingly, we found a significant enrichment of predictive markers at the SWAP70 gene locus. The SWAP70 locus encodes a B-cell specific ATP and calcium binding protein that is recruited to the plasma membrane following B-cell activation with subsequent translocation to the nucleus. SWAP70 has been shown to mediate immunoglobulin isotype switching in B-cells following activation of the B-cell antigen receptor complex [42, 43]. This finding is consistent with the recent reports of ongoing isotype switching in patients who are critically ill with COVID-19 and the association of differential immunoglobulin M (IgM)/IgG/IgA epitope diversity in mild or severe COVID-19, especially in patients who succumbed to SARS-CoV-2 infection [44]. We also observed that genes involved in haemostasis and blood clotting were found in the regions covered by the predictive marker panel. This is consistent with clinical reports of severe COVID-19 patients presenting clinically with a

'microvascular injury syndrome' with an associated procoagulant state as well as clinical reports of hypercoagulation in patients with severe COVID-19 [45, 46]. Interestingly, it remains to be investigated if the identified microvascular markers may have potential functional relevance to prothrombotic vaccination responses in individual patients. The involvement of B-cell activation and haemostasis are both in line with a prevailing view that systemic inflammation and the cardiovascular clinical manifestations it triggers, all lie at the root of the clinical symptomology seen in severe COVID-19 cases [47–49].

Extension of the current study to a wider distribution and larger number of individuals could help define the regional, racial, and epigenetic prevalence of high-risk biomarkers in these populations. A longitudinal observational study with collections before and after resolution of the acute and chronic phases of COVID disease will provide further invaluable insights into the mechanisms and the long-term stability of the identified systemic biomarker signature. Early evidence indicates that blood samples collected from patients before the onset of the COVID pandemic reveal high-risk profiles in some individuals. This would suggest that the biomarker profiles identified in this study are not emerging in response to COVID infection, but rather represent a pre-existing default state on the spectrum of outcome susceptibility.

There are several immediate implications of the results reported here. The availability of a simple blood-based assay that provides a readout of likely disease course if infected with SARS-CoV-2 is especially helpful for the triage of individuals who either 1) do not have access to COVID-19 vaccines (due to

underlying medical conditions, location, or age for example) or 2) choose to forgo vaccination for other reasons. It has been well appreciated that the heterogeneity seen in COVID-19 disease outcomes are largely defined by the host response, rather than the virus or its variants [15]. The emergence of viral variants that can render some current vaccination strategies less effective and the unknown longer-term path of the virus may make information about individual risk of developing severe disease valuable for those who may be at risk of re-infection. For individuals at the early stage of a confirmed SARS-CoV-2 infection, the assay described here could assist with patient triage in hospital settings and assist in clinical decision making with a direct impact on clinical outcomes. For the currently non-infected population, the assay could help inform individuals on the potential for high-risk severe outcomes in the event of an infection. This could enable a review of preventative lifestyle decisions on vaccinations, guide appropriate social distancing measures, and inform return to work planning. Advanced knowledge of likely disease severity can also aid patients and their physicians in making preparatory medical planning decisions where appropriate. Last, when applied in larger scale, population-wide settings; knowledge of pooled individual risk profiles can help health systems make informed decisions about staffing and infrastructure needs in the event of a pandemic resurgence.

Conclusions

Here we report on the development and validation of a predictive blood-based assay that can identify, with high accuracy, individuals who are at the highest risk of developing severe complications in COVID-19 disease. The 3D genomic

546 biomarker panel described here can be used to evaluate personalized risk of
547 complications prior to exposure to the SARS-CoV-2 virus. The assay described
548 here can assist individuals and their physicians in making informed decisions
549 on lifestyle, social, and return to work activities as well as guiding medical
550 intervention decisions, where necessary.

551

552 **Declarations**

553 C.K., A.W., F.S., M.S., J.W. E.H. and A.A. are full-time employees at Oxford
554 BioDynamics plc and have no other competing financial interests.
555
556

557 **Author Contributions**

558 EH, AA conceived the study, MS, EH and AR assisted with *EpiSwitch*® array
559 design. WM, AB and ZL assisted with design of experiments. MS and AA
560 planned and reviewed experiments. CK, AW, FCS and EH analysed the data.
561 RP, AD, PB, BE, TL performed experiments. RV, AB, DP, PAR, JM helped with
562 interpretation of the data and writing of the manuscript. WW, CK, EH, and AA.
563 wrote and reviewed the manuscript.

564

565 **Acknowledgements**

566 The authors would like to thank members of OBD Reference Facility Morgan
567 Thacker, Louis James, Thomas Lavin, Catriona Williams, Matt Parnell, Aemilia
568 Katzinski, Pieter Koorts, Kalsoom Rana and Jayne Green for direct operational
569 support and taking part in processing of clinical samples, Olly Hunter for
570 assistance in data analysis. In addition, we acknowledge Boca Biolistics LLC.
571 and Reprocell USA Inc. for the timely provision of high-quality clinical blood

samples and clinical annotations, and Agilent Technologies, Inc. for supply of *EpiSwitch*® designed CGH microarray slides.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Written informed consent for publication was obtained from all authors.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

All patients signed informed consent forms prior to providing blood samples. All ethical guidelines were followed.

Funding

This work was funded by Oxford BioDynamics.

References

1. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Biomed. 2020;91.

2. Wang Z, Tang K. Combating COVID-19: health equity matters. Nat Med. 2020;26:458.

3. Pfefferbaum B, North C. Mental Health and the Covid-19 Pandemic. N Engl J Med. 2020;383:510–2.

4. FDA EUAs for Molecular COVID-19 Diagnostic Tests. https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/in-vitro-diagnostics-euas-molecular-diagnostic-tests-sars-cov-2.

5. FDA EUAs for Anitgen COVID-19 Diagnostic Tests. https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/in-vitro-diagnostics-euas-antigen-diagnostic-tests-sars-cov-2.

6. Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nat Mach Intell. 2021;3:199–217.

7. G G, Zangrillo A, A Z, Antonelli M. Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. JAMA - J Am Med Assoc. 2020;323:1574–81.

8. Zangrillo A, Beretta L, Scandroglio A, G M. Characteristics, treatment, outcomes and cause of death of invasively ventilated patients with COVID-19 ARDS in Milan, Italy. Crit Care Resusc. 2020;22:200–11.

628 9. Yang X, Yu Y, Xu J, Shu H, Xia J. Clinical course and outcomes of critically
629 ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered,
630 retrospective, observational study. Lancet Respir Med. 2020;8:475–81.
631 10. WHO. COVID-19 Clinical Management, Living Guidance. 2021.
632 https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2021-1.
633 11. Polack F, Thomas S, Kitchin N, Absalon J. Safety and Efficacy of the
634 BNT162b2 mRNA Covid-19 Vaccine. N Engl J Med. 2020;383:2603–15.
635 12. Baden L, Sahly H, Essink B, Kotloff K, Frey S. Efficacy and Safety of the
636 mRNA-1273 SARS-CoV-2 Vaccine. N Engl J Med. 2021;384:403–16.
637 13. Voysey M, Clemens S, Madhi S, Weckx L. Safety and efficacy of the
638 ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim
639 analysis of four randomised controlled trials in Brazil, South Africa, and the
640 UK. Lancet. 2021;397:99–111.
641 14. Tapia M. University of Chile-COVID-19.
642 https://www.uchile.cl/noticias/174186/resultados-primer-estudio-de-
643 efectividad-de-las-vacunas-en-chile. Accessed 6 Nov 2021.
644 15. Hunter E, Koutsothanasi C, Wilson A, Santos F, Salter M, Powell R, et al.
645 3D genomic capture of regulatory immuno-genetic profiles in COVID-19
646 patients for prognosis of severe COVID disease outcome. bioRxiv. 2021.
647 16. Yan H, Hunter E, Akoulitchev A, Park P, Winchester DJ, Moo-Young TA,
648 et al. Epigenetic chromatin conformation changes in peripheral blood can
649 detect thyroid cancer. Surg (United States). 2019.
650 17. Hunter E, McCord R, Ramadass AS, Greene J, Westra JW, Akoulitchev
651 A, et al. Comparative molecular cell-of-origin classification of diffuse large B-
652 cell lymphoma based on liquid and tissue biopsies. Transl Med Commun.

653 2020;5.

654 18. Alshaker H, Mills R, Hunter E, Salter M, Ramadass A, Skinner BM, et al.
655 Chromatin conformation changes in peripheral blood can detect prostate
656 cancer and stratify disease risk groups. J Transl Med. 2021.

657 19. Carini C, Hunter E, Ramadass AS, Green J, Akoulitchev A, McInnes IB, et
658 al. Chromosome conformation signatures define predictive markers of
659 inadequate response to methotrexate in early rheumatoid arthritis. J Transl
660 Med. 2018.

661 20. Salter M, Corfield E, Ramadass A, Grand F, Green J, Westra J, et al.
662 Initial Identification of a Blood-Based Chromosome Conformation Signature
663 for Aiding in the Diagnosis of Amyotrophic Lateral Sclerosis. EBioMedicine.
664 2018.

665 21. Jakub JW, Grotz TE, Jordan P, Hunter E, Pittelkow M, Ramadass A, et al.
666 A pilot study of chromosomal aberrations and epigenetic changes in
667 peripheral blood samples to identify patients with melanoma. Melanoma Res.
668 2015;:1. doi:10.1097/CMR.0000000000000182.

669 22. Shah P, Hunter E, Potluri S, Zhang S, Dezfouli M, Back J, et al.
670 Development and validation of baseline predictive biomarkers for response to
671 avelumab in second-line (2L) non-small cell lung cancer (NSCLC) using
672 EpiSwitchTM epigenetic profiling. J Immunother Cancer. 2019;7.

673 23. Tsourkas A, Behlke M, Xu Y, Bao G. Spectroscopic features of dual
674 fluorescence/luminescence resonance energy-transfer molecular beacons.
675 Anal Chem. 2003;75:3697–703.

676 24. Fabregat A, Sidiropolous K, Viteri G, Forner O. Reactome pathway
677 analysis: a high-performance in-memory approach. BMC Bioinformatics.

678 2017;18:142.

679 25. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al.

680 The STRING database in 2017: Quality-controlled protein-protein association

681 networks, made broadly accessible. Nucleic Acids Res. 2017.

682 26. COVID-19 Vaccination in the EU. https://graphics.reuters.com/HEALTH-

683 CORONAVIRUS/EU-VACCINES/qmypmrelyvr/.

684 27. European Centre for Disease Prevention and Control: EU Vaccination

685 Rates. https://www.ecdc.europa.eu/en/publications-data/data-covid-19-

686 vaccination-eu-eea.

687 28. Full Committee Hearing: “Update on the Department of Defense’s

688 Evolving Roles and Mission in Response to the COVID-19 Pandemic.”

689 https://armedservices.house.gov/2021/2/full-committee-hearing-update-on-

690 the-department-of-defense-s-evolving-roles-and-mission-in-response-to-the-

691 covid-19-pandemic.

692 29. Shaw J, Stewart T, Anderson K, Hanley S. Assessment of US Healthcare

693 Personnel Attitudes Towards Coronavirus Disease 2019 (COVID-19)

694 Vaccination in a Large University Healthcare System. Clin Infect Dis.

695 2021;ciab054.

696 30. United States Centers for Disease Control and Prevention: COVID Data

697 Tracker. https://covid.cdc.gov/covid-data-tracker/#compare-trends_newcases.

698 31. Rinott E, Youngster I, Lewis Y. Reduction in COVID-19 Patients Requiring

699 Mechanical Ventilation Following Implementation of a National COVID-19

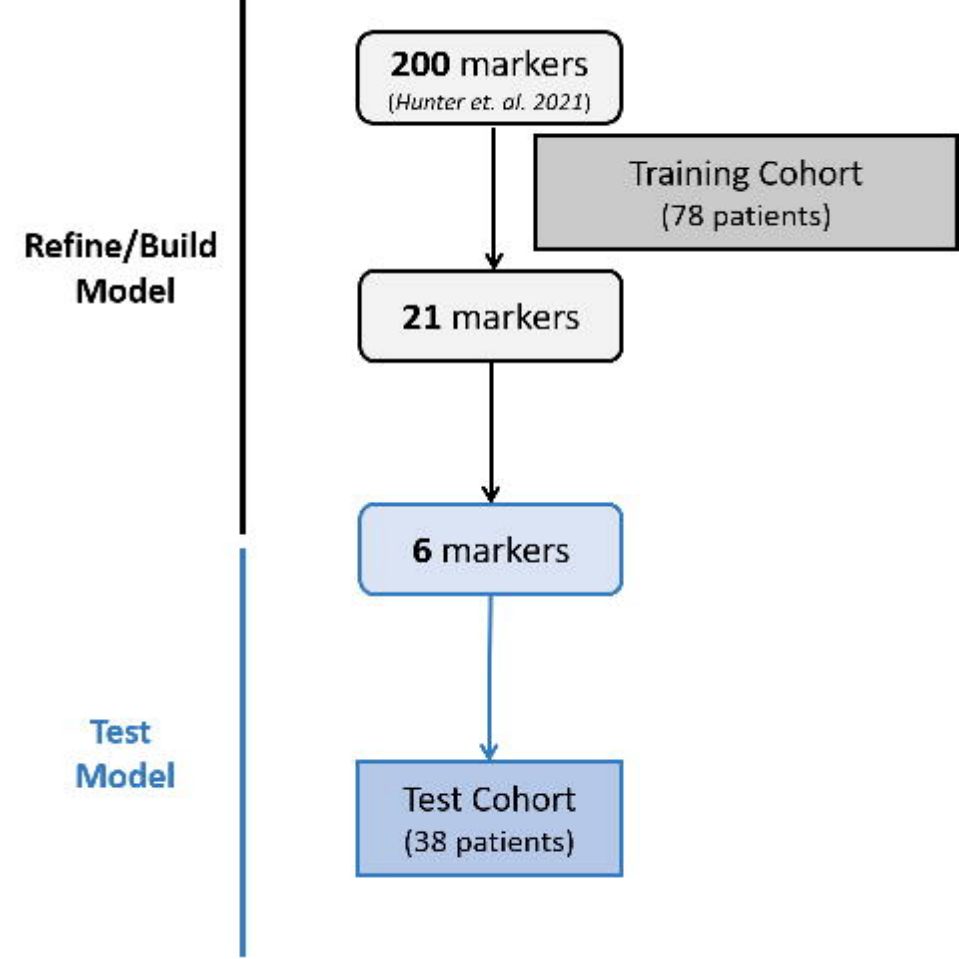
700 Vaccination Program — Israel, December 2020–February 2021. MMWR Morb

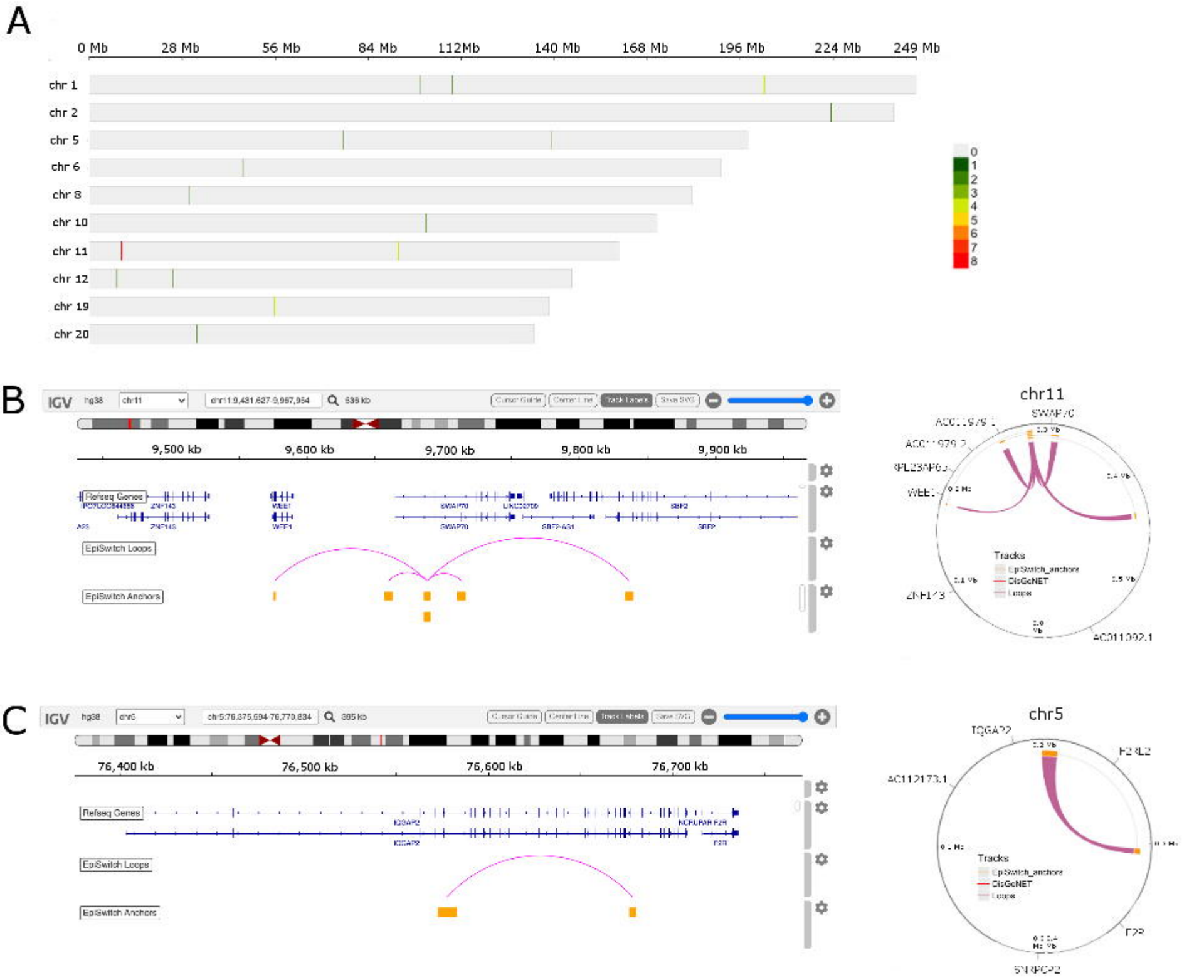
701 Mortal Wkly Rep. 2021;70:326–8.

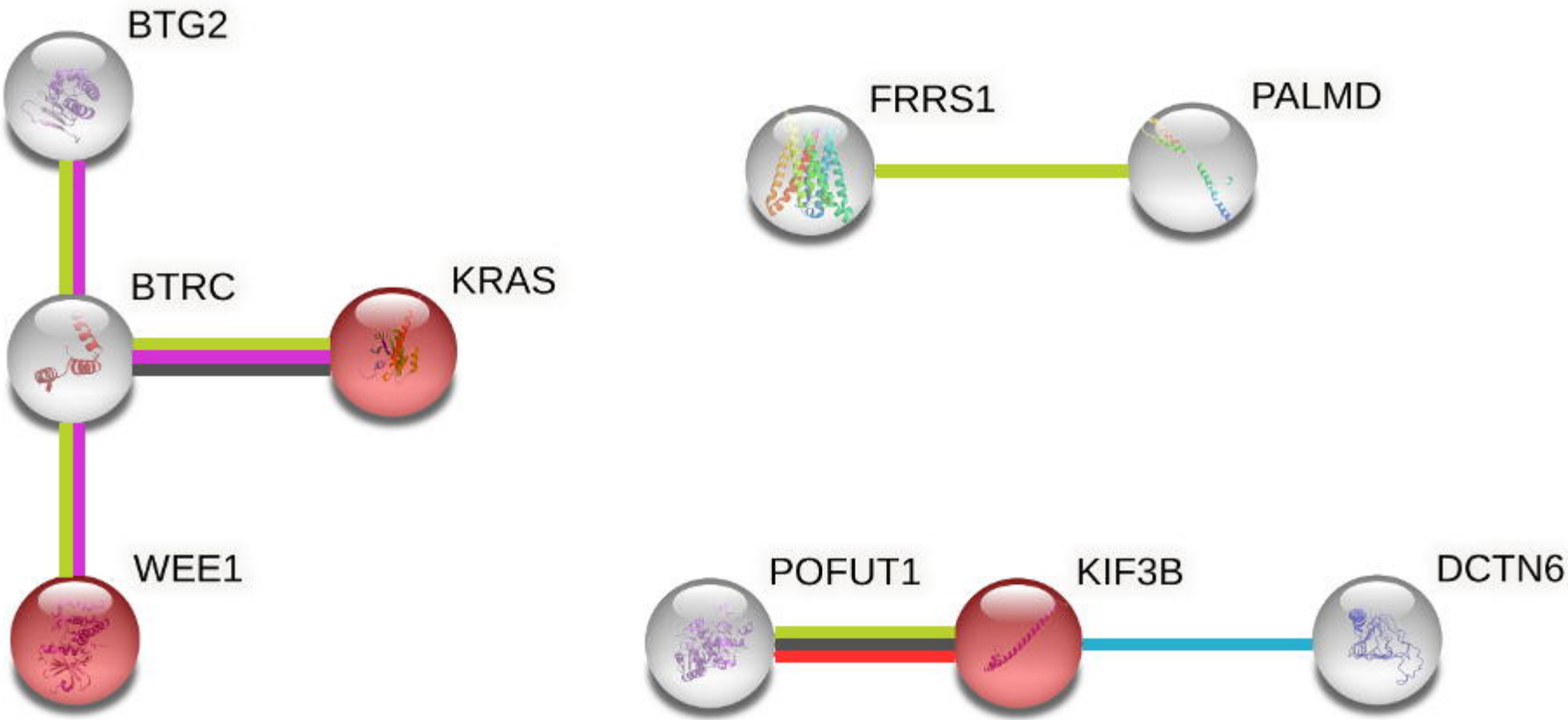
702 32. Frampton D, Rampling T, Cross A, Bailley H, Heaney J, Byott M. Genomic

703 characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7
704 lineage in London, UK: a whole-genome sequencing and hospital-based
705 cohort study. Lancet. 2021.
706 33. Lauring A, Hodcroft E. Genetic Variants of SARS-CoV-2—What Do They
707 Mean? JAMA - J Am Med Assoc. 2021;325:529–31.
708 34. M M, Bassi J, De Marco A, Chen A, Walls A, Di Iulio J. SARS-CoV-2
709 immune evasion by variant B.1.427/B.1.429. bioRxiv. 2021.
710 35. Davies N, Abbott S, Barnard R, Jarvis C. Estimated transmissibility and
711 impact of SARS-CoV-2 lineage B.1.1.7 in England. Science (80-). 2021;372.
712 36. Torjesen I. Covid-19: Delta variant is now UK's most dominant strain and
713 spreading through schools. BMJ. 2021;373.
714 37. Sanjuan R, Nebot M, Chirico N, Mansky L, Belshaw R. Viral Mutation
715 Rates. Jounal Virol. 2010;84:9733–48.
716 38. Koonin E, Dolja V, Krupovic M. Origins and evolution of viruses of
717 eukaryotes: The ultimate modularity. Virology. 2015;479:2–25.
718 39. Ikegame S, Siddiquey M, Hung C, Haas G, Brambilla L. Qualitatively
719 distinct modes of Sputnik V vaccine-neutralization escape by SARS-CoV-2
720 Spike variants. medRxiv. 2021;:03.31.21254660.
721 40. Huang S, Wang S. SARS-CoV-2 Entry Related Viral and Host Genetic
722 Variations: Implications on COVID-19 Severity, Immune Escape, and
723 Infectivity. Int J Mol Sci. 2021;22:3060.
724 41. Groves D, Rowland-Jones S, Angyal A. The D614G mutations in the
725 SARS-CoV-2 spike protein: Implications for viral infectivity, disease severity
726 and vaccine design. Biochem Biophys Res Commun. 2021;538:104–7.
727 42. Masat L, Caldwell J, Armstrong R, Khoshevenian H. Association of

728 SWAP-70 with the B cell antigen receptor complex. Proc Natl Acad Sci.
729 2000;97:2180–4.
730 43. Borggrefe T, Wabl M, Akhmendov A, Jessberger R. A B-cell-specific DNA
731 recombination complex. J Biol Chem. 1998;273:17025–35.
732 44. Ravichandran S, Lee Y, Grubbs G, Coyle E. Longitudinal antibody
733 repertoire in “mild” versus “severe” COVID-19 patients reveals immune
734 markers associated with disease severity and resolution. Sci Adv. 2021;7.
735 45. Helms J, Tacquard C, Severac F, Leonard-Lorant I, Ohana M. High risk of
736 thrombosis in patients with severe SARS-CoV-2 infection: a multicenter
737 prospective cohort study. Intensive Care Med. 2020;46:1089–98.
738 46. Magro C, Mulvey J, Berlin D, Nuovo G. Complement associated
739 microvascular injury and thrombosis in the pathogenesis of severe COVID-19
740 infection: A report of five cases. Transl Res. 2020;220:1–13.
741 47. Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F,
742 et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in
743 Covid-19. N Engl J Med. 2020.
744 48. Babapoor-Farrokhran S, Gill D, Walker J. Myocardial injury and COVID-
745 19: Possible mechanisms. Life Sci. 2020;July:117723.
746 49. Magadum A, Kishore R. Cardiovascular Manifestations of COVID-19
747 Infection. Cells. 2020;9:2508.
748


200 markers
(Hunter et. al. 2021)
Training Cohort
(78 patients)
Refine/Build Model
21 markers
6 markers
Test Model
Test Cohort
(38 patients)


A
0 Mb
28 Mb
56 Mb
84 Mb
112Mb
140 Mb
168 Mb
196 Mb
224 Mb
249 Mb
chr 1
chr 2
chr 5
chr 6
chr 8
chr 10
chr 11
chr 12
chr 19
chr 20
0
1
2
3
4
5
6
7
8
B
IGV
9,500 kb
9,600 kb
9,700 kb
9,800 kb
9,900 kb
Refseq Genes
EpiSwitch Loops
EpiSwitch Anchors
chr11
Tracks
EpiSwitch_anchors
DisGeNET
Loops
C
IGV
76,400 kb
76,500 kb
76,600 kb
76,700 kb
Refseq Genes
EpiSwitch Loops
EpiSwitch Anchors
chr5
Tracks
EpiSwitch_anchors
DisGeNET
Loops


BTG2
BTRC
KRAS
WEE1
FRRS1
PALMD
POFUT1
KIF3B
DCTN6

A

Test	Severe Outcome: Present	n	Absent	n	Total
Yes	True Positive	20	False Positive	2	22
No	False Negative	8	True Negative	8	16
Total		28		10	

Results

Statistic	Value	95% CI
Sensitivity	71.43%	51.33% to 86.78%
Specificity	80.00%	44.39% to 97.48%
Positive Likelihood Ratio	3.57	1.01 to 12.61
Negative Likelihood Ratio	0.36	0.18 to 0.69
Disease prevalence*	73.68%	56.90% to 86.60%
Positive Predictive Value*	90.91%	73.90% to 97.25%
Negative Predictive Value*	50.00%	34.02% to 65.98%
Accuracy*	73.68%	56.90% to 86.60%

B

Test	Severe Outcome: Present	n	Absent	n	Total
Yes	True Positive	66	False Positive	5	71
No	False Negative	9	True Negative	36	45
Total		75		41	

Results

Statistic	Value	95% CI
Sensitivity	88.00%	78.44% to 94.36%
Specificity	87.80%	73.80% to 95.92%
Positive Likelihood Ratio	7.22	3.16 to 16.48
Negative Likelihood Ratio	0.14	0.07 to 0.25
Disease prevalence*	64.66%	55.24% to 73.31%
Positive Predictive Value*	92.96%	85.25% to 96.79%
Negative Predictive Value*	80.00%	68.20% to 88.18%
Accuracy*	87.93%	80.58% to 93.24%