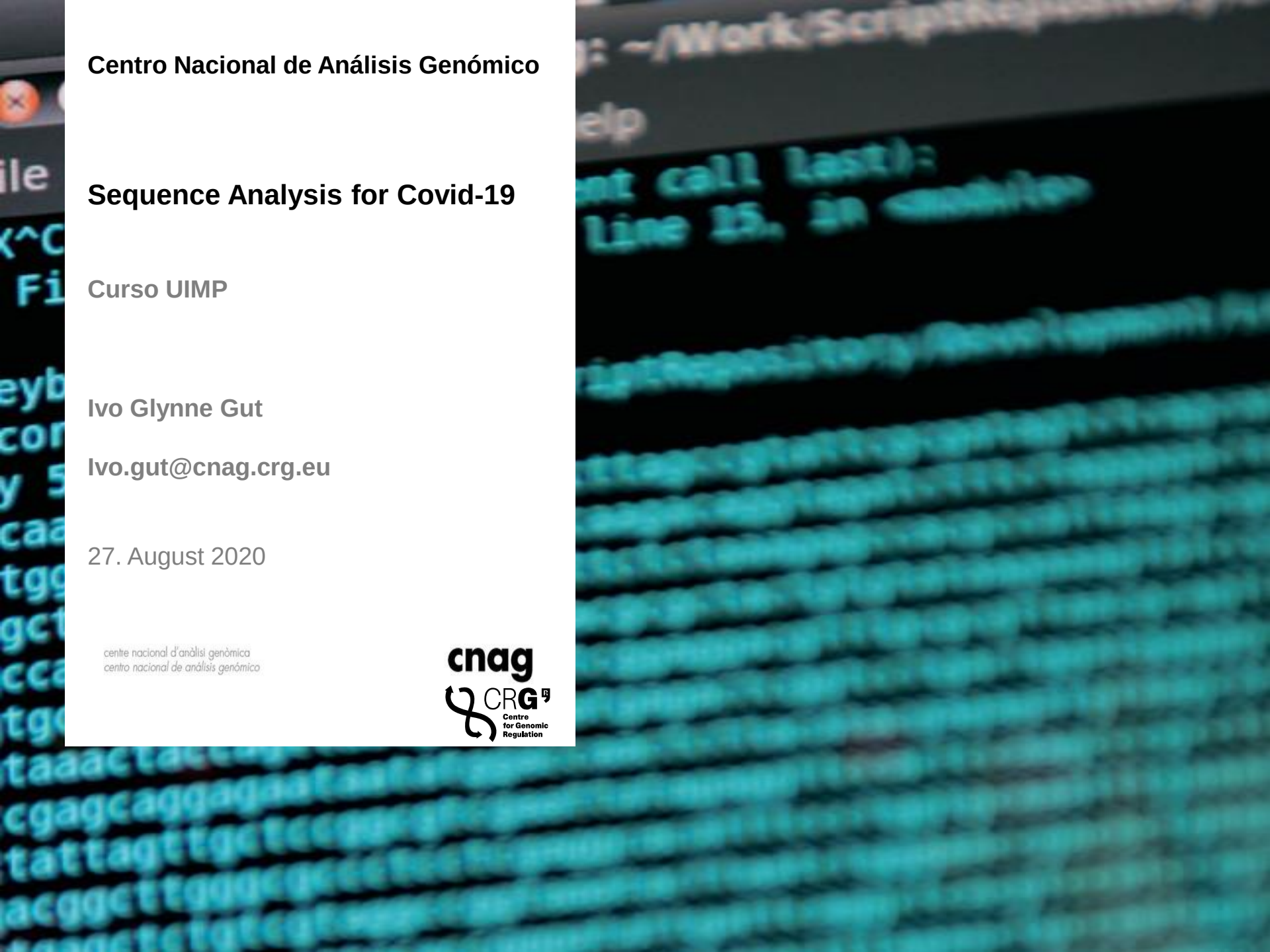


Centro Nacional de Análisis Genómico

Sequence Analysis for Covid-19

Curso UIMP

Ivo Glynne Gut

Ivo.gut@cnag.crg.eu

27. August 2020

centre nacional d'anàlisi genòmica
centro nacional de análisis genómico

cnag

CRG
Centre for Genomic Regulation

Centro Nacional de Análisis Genómico (CNAG-CRG)

- ✓ Created in 2010
- ✓ Funded by MCI and Generalitat de Catalunya
- ✓ Competitive grants & contractual research provide additional funds
- ✓ Since 2015 it is integrated with the CRG
- ✓ > 80 people, directed by Ivo Gut

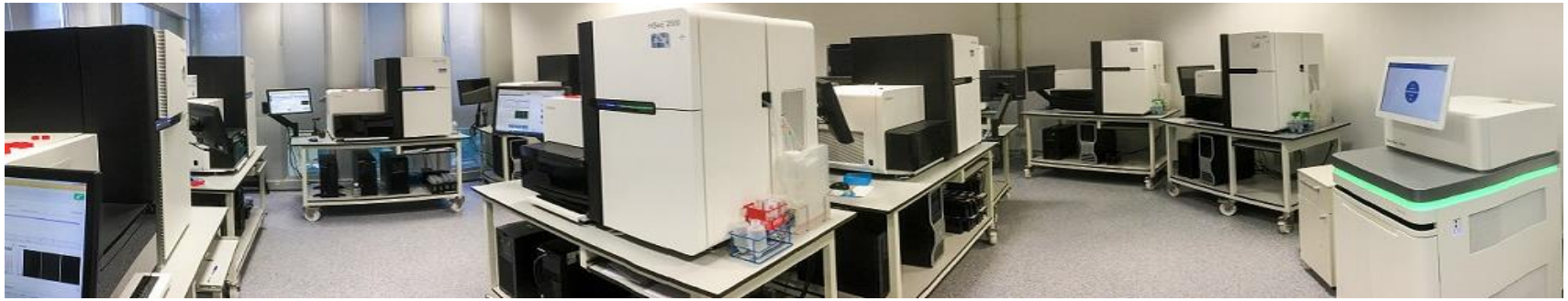
Mission

- ✓ To carry out ~~large-scale~~ projects in genome analysis that will lead to significant improvements in people's health and quality of life, in collaboration with the Spanish, European and International research and clinical community.

Vision

- ✓ To be a high quality sequence analysis center and to be a world reference center for genomic analysis.

The CNAG-CRG's Genomehenge 2019



Sequencing capacity

- >6000 Gbases/day = 70 human genomes/day at 30x coverage

Sequencing

- 2 Illumina NovaSeq6000
- 3 Illumina HiSeq2500
- 2 Illumina HiSeq4000
- 2 Illumina MiSeq
- 3 Oxford Nanopores MinIons
- Oxford Nanopore GridIon
- 10x Genomics Chromium

Computing

- 3552 cores
- 3.7 PB disk + 3 PB tape
- 35,5 TB RAM
- Barcelona Super Computing Center - 10 x 10 Gb/s

cnag
centre nacional d'anàlisi genòmica
centro nacional de análisis genómico

CRG
Centre for Genomic Regulation

The CNAG-CRG Quality Certifications

- ✓ SGS Certification ISO 9001: 2015
- ✓ ENAC ISO 17025 : 2005 Accreditation
- ✓ Illumina Certified Service Provider CSPro
- ✓ Agilent Certified Service Provider CSP
- ✓ Certified Service Provider status under the Roche Sequencing Solutions® Technical Certification Program for CNAG-CRG's expertise in running the Roche SeqCap® EZ target enrichment system
- ✓ BBMRI-ERIC Expert Centre

CNAG Workflow

BIOREPOSITORY

LABORATORY

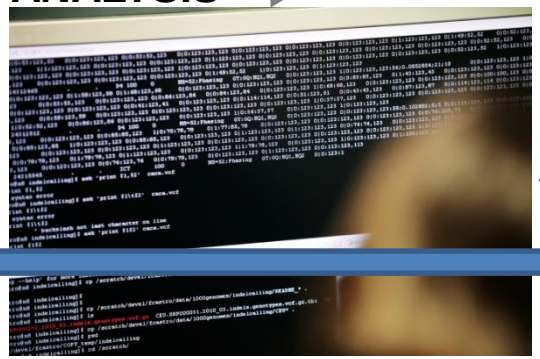
SEQUENCING

LIMS

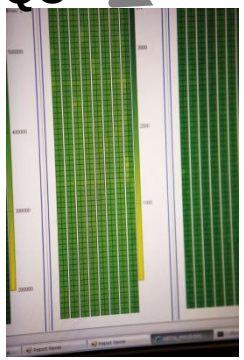
Subproject	Status	Deadline (Countdown)	Details by Sample	Details by Library	Analysis	Sample Type	Application	Expected Capture Protocol	Numbers of Samples	Ready for Sample prep / Sequence	Samples	Sample Reception Dates	Aliquots	Libraries pass / total	Libraries Barcodes	Samples with a pass Library	Flowcells & Dates	Samples above Coverage	Samples above pfmr	Data Transfed by Sample
_11	Open (Experimental)	2012-10-01 (26 days)	Details by Sample / Sample Aggregate View	Details by Library / Library Aggregate View	BAG	gDNA	ExomeCapture-Seq	Agilent Human All Exon 51Mb v4	6 8 8 6 4	yes	E061 E058 E055 E059 E056 E060 E057 E054	2012-07-02 2012-07-02	861C 948C 862C 863C 864C	3 / 4	861C 862C 863C 864C	3				0 sent, of 4 to be seq
_01	Open (Experimental)	2012-10-02 (27 days)	Details by Sample / Sample Aggregate View	Details by Library / Library Aggregate View	BAG	gDNA	WG-Seq		2 4 4 4 2	yes	H367 H368 H368 H367	2012-07-03 2012-07-03	580C 581C	2 / 2	580C 581C	4	C0YHRACXX - 2012-07-27 C11CBACXX - 2012-08-09 C11D3ACXX - 2012-08-09			0 sent, of 2 to be seq
_03	Open (Experimental)	2012-10-03 (28 days)	Details by Sample / Sample Aggregate View	Details by Library / Library Aggregate View	No	Exome library	ExomeCapture-Seq	Nimblegen SeqCap 64 Mb v3	16 80 80 0 80	yes	H349 H365 H346 H362 H343 H359 H340 H356 H337	2012-06-18 2012-07-04		80 / 80	I231 I229 I230 I227 I228 H342 H344 H346 H347	80	C0TRHACXX - 2012-06-19 C0TVGACXX - 2012-06-26 D113HACXX - 2012-06-26 D14PPACXX - 2012-07-25 C0YHRACXX - 2012-07-27	80		80 sent, of 80 to be seq

TRANSFER

ANALYSIS

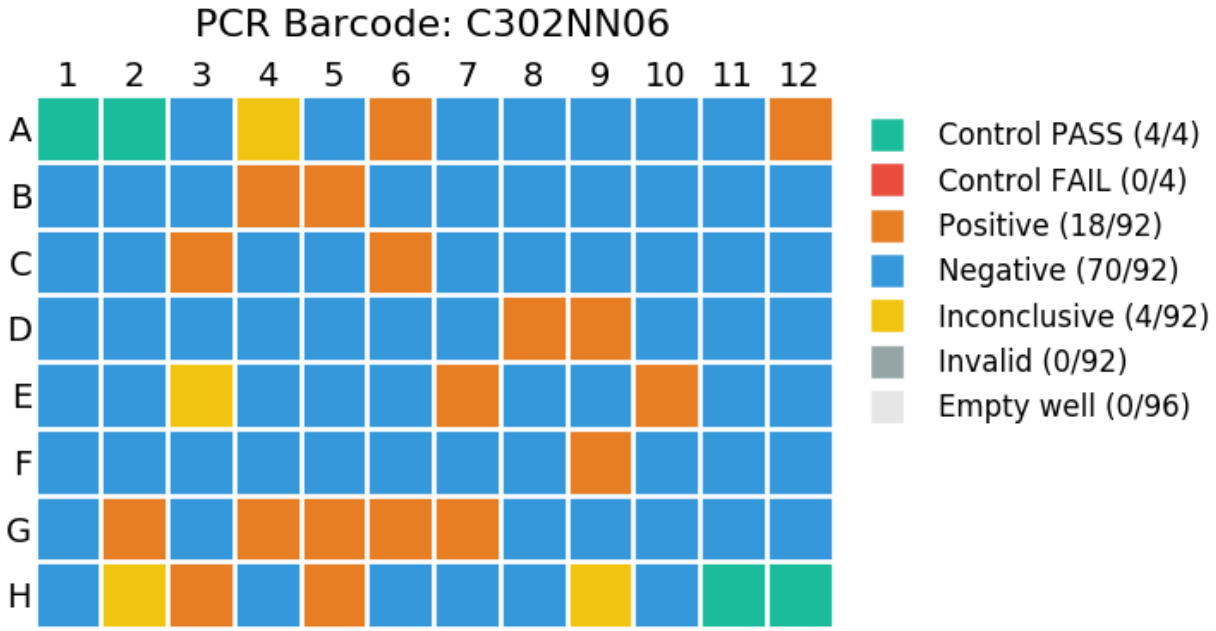


QC



Involvement in ORFEU

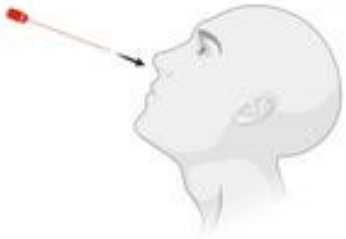
- ✓ SARS-CoV-2 diagnostic testing platforms at the CRG and at the Parc Cientific de Barcelona by CNAG, IRB and IBEC
- ✓ Dedicated BSL2 facility
- ✓ Pipeline
- ✓ LIMS



COVID-19 Molecular Diagnostic Test through RT-PCR

1 Nasopharyngeal (NP) or Oropharyngeal (OP) swab

Cotton swab is inserted into nostril to absorb secretions. <15 min

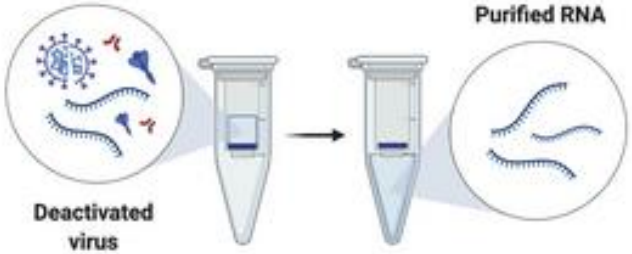


2 Collected specimen 0-72 h

Specimen is stored at 2-8°C for up to 72 hours or proceed to RNA extraction.

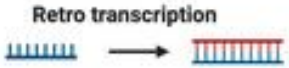
3 RNA extraction ~45 min

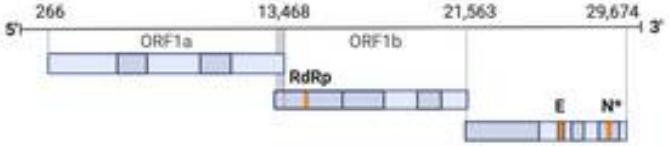
Purified RNA is extracted from deactivated virus.



4 RT-qPCR ~1 h per primer set

Purified RNA is reverse transcribed to cDNA and amplified by qPCR.





Example primers and probes for screening

E gene First-line screening tool

E_Forward: ACAGGTACGTTAATAGTTAATAGCGT
E_Probe1: FAM-ACACTAGCCATCCTTACTGCGCTTCG-BBQ
E_Reverse: ATATTGCAGCAGTACGCACACA

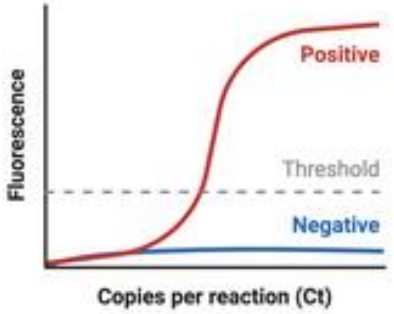
RdRp gene Confirmatory testing

RdRp_Forward: GTGARATGGTCATGTGTGGCGG
RdRp_Probe1: FAM-CCAGGTGGWACRTCATCMGGTGATGC-BBQ
RdRp_Probe2: FAM-CAGGTGGAACCTCATCAGGAGATGC-BBQ
RdRp_Reverse: CARATGTTAAASACACTATTAGCATA

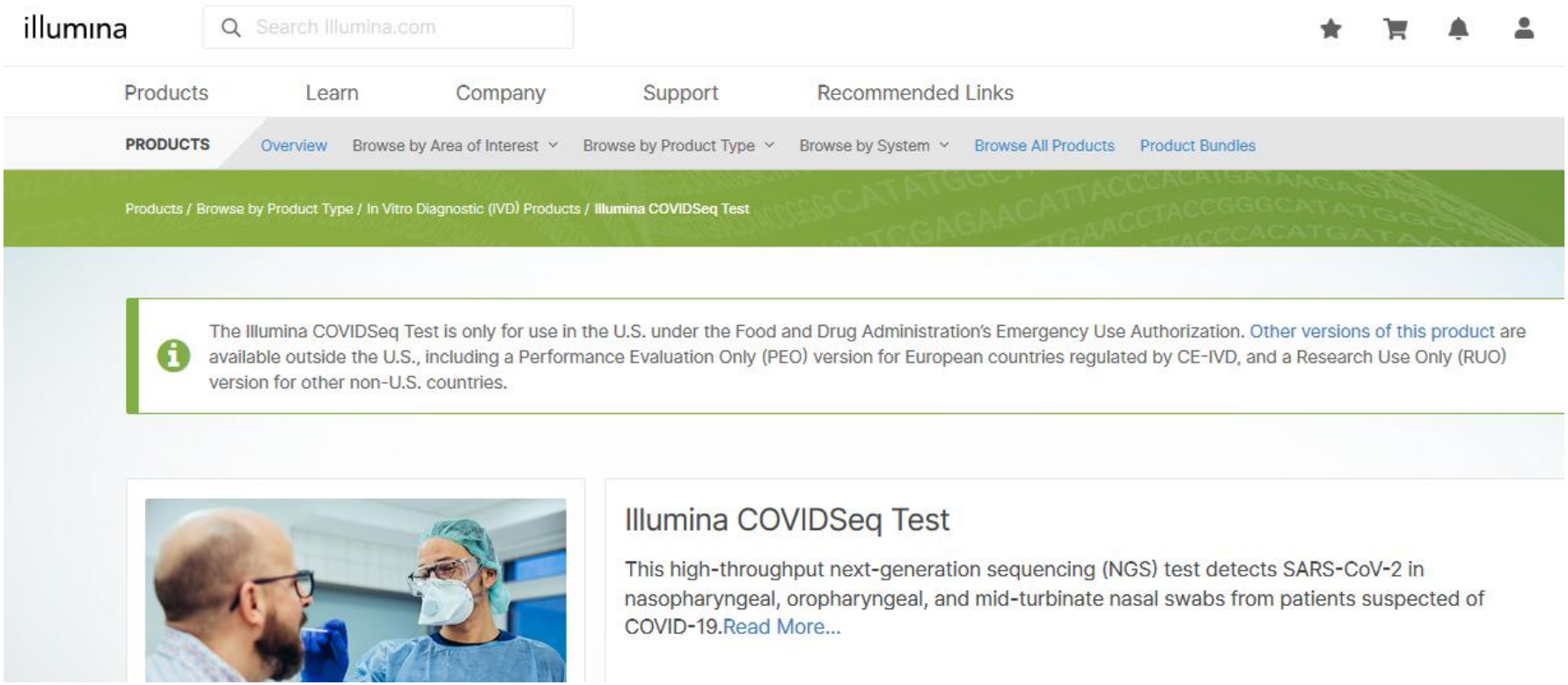
Primer sequences are for *illustrative* purposes only.

5 Test results real-time

Positive SARS-CoV2 patients cross the threshold line within 40.00 cycles (< 40.00 Ct).



Short-read sequencing technology for SARS-CoV-2



Product Highlights

The Illumina COVIDSeq Test is the first NGS test approved for use under the U.S. Food and Drug Administration's Emergency Use Authorization (EUA). This amplicon-based NGS test includes 2019-nCoV primer and probe sets designed to detect RNA from the SARS-CoV-2 virus in nasopharyngeal, oropharyngeal, and mid-turbinate nasal swabs from patients with signs and symptoms of infection who are suspected of COVID-19.

Rapid, Scalable SARS-CoV-2 Detection

The Illumina COVIDSeq Test can be scaled up or down to accommodate different numbers of samples. Up to 3072 results can be processed in 12 hours on the NovaSeq 6000 System using two NovaSeq 6000 S4 Reagent Kits with the Xp workflow.

Design and Quality Control

The Illumina COVIDSeq Test leverages a modified version of the validated, publicly available ARTIC multiplex PCR protocol, with 98 amplicons designed to amplify SARS-CoV-2 virus-specific sequences, combined with proven Illumina sequencing technology. As a quality feature, an internal control consisting of 11 human mRNA targets is included in every sample to monitor for errors.

Long-read sequencing technology for SARS-CoV-2

NANOPORE Products Services Applications Resources

COVID-19 Overview | Community Timeline | Get started / Workflows | LamPORE Assay

Get started | Talk to us | Subscribe

LamPORE

LamPORE is a new assay, in advanced development, for the detection of the SARS-CoV-2 virus.

Status at 7 August 2020: *Oxford Nanopore has spent recent weeks and months collaborating with multiple clinical laboratories in the UK to complete evaluations of analytical performance of LamPORE. Information about this dataset of several hundred samples will be released soon and regulatory submissions are underway. We will keep you informed about the rapid progress of its development here.*

LamPORE is also in development for the detection of multiple respiratory pathogens in a single assay, in preparation for a need to detect SARS-CoV-2 in samples that may also contain influenza or other respiratory viruses.

- **Scalable:** users can run between 1 - 768 samples on a single flow cell (including positive / negative controls)
- Designed to be deployed on MinION Mk1C (one flow cell) or GridION (up to 5 flow cells)
- Simple desktop devices can therefore enable as many as 15,000 tests a day
- **On-demand processing** rather than a requirement to batch
- **Rapid results:** under two hours for 1-96 samples on a single flow cell
- **Near-sample analysis.** Simple hardware requirement enables use in centralised lab environment, or 'pop-up', near-community labs; flexible location is also designed to support rapid turnaround times
- **Control mechanism** inbuilt to identify errors in sample collection, a common source of 'false negatives'
- Assay sequences three genes of the SARS-CoV-2 virus for **precise analysis**

LamPORE is a rapid, low-cost and highly scalable assay for the detection of SARS-CoV-2, the virus that causes COVID-19. The assay combines nanopore sequencing with loop-mediated isothermal amplification to detect the virus, and uses materials outside of current supply chain limitations.

LamPORE is the first assay that Oxford Nanopore has developed with the intention for future diagnostic use.

Interested in finding out more and staying informed?

Register your interest in LamPORE

The LamPORE assay

MinION Mk1C

The fully connected, all-in-one, portable device.

- Users can analyse 1-768 samples at once.
- Due to portable nature, could be deployed in a pop-up lab environment, requiring little space.

GridION Mk1

The desktop sequencer that can run 5 times as many samples as MinION.

- GridION could be used in a high-throughput, centralised lab to test high numbers of samples.

The LamPORE assay

LamPORE is deployable in both high-throughput, centralised settings as well as smaller, local environments for quick turnaround of a large number of samples, opening up opportunities for future routine screening.

1-96 samples can be processed in just over an hour on a single MinION Flow Cell, or using more barcodes can enable the sequencing of 768 samples on a MinION Flow Cell, requiring an additional ~3 hours sequencing time. The GridION can process up to five times this many samples.

How does it work?

LamPORE leverages LAMP (loop-mediated isothermal amplification) upstream of nanopore sequencing to detect the presence of SARS-CoV-2 in a sample.

Isothermal amplification has been used successfully alongside nanopore sequencing previously for the analysis of malaria parasite Plasmodium, leishmaniasis and dengue virus, providing a simple and fast way to amplify a specific target.

The LamPORE assay can be performed on extracted RNA from swabs, and is also in development to enable working directly from saliva.

In a SARS-CoV-2-positive sample, once the viral target has been amplified via LAMP the sample is then prepared for sequencing using Oxford Nanopore's rapid sequencing chemistry. By barcoding the samples at both the amplification and library preparation stages, high multiplexing capacity can be achieved for large sample volumes.

As well as targeting three specific genes of the SARS-CoV-2 virus, a control target is included in the assay (actin). This acts as confirmation of successful sample collection, and is designed to show the user where a negative result is because of sample collection errors rather than the lack of presence of SARS-CoV-2.

The prepared sequencing library is subsequently loaded onto a nanopore sequencing device, such as the GridION Mk1 or MinION Mk1C, and real-time analysis begins. When sequencing reads aligning to the SARS-CoV-2 genome and a control target reach a threshold number per sample, the sample can be classed as positive.

Results can be delivered in under two hours for between 1-96 samples, meaning rapid turnaround of results.

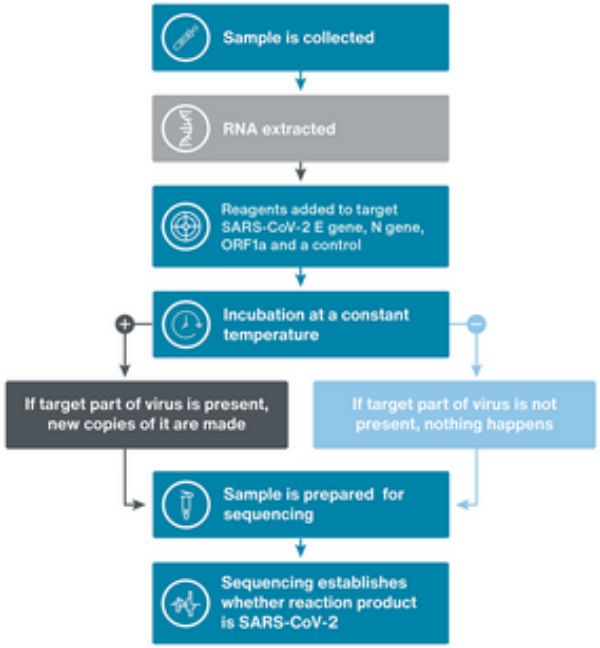
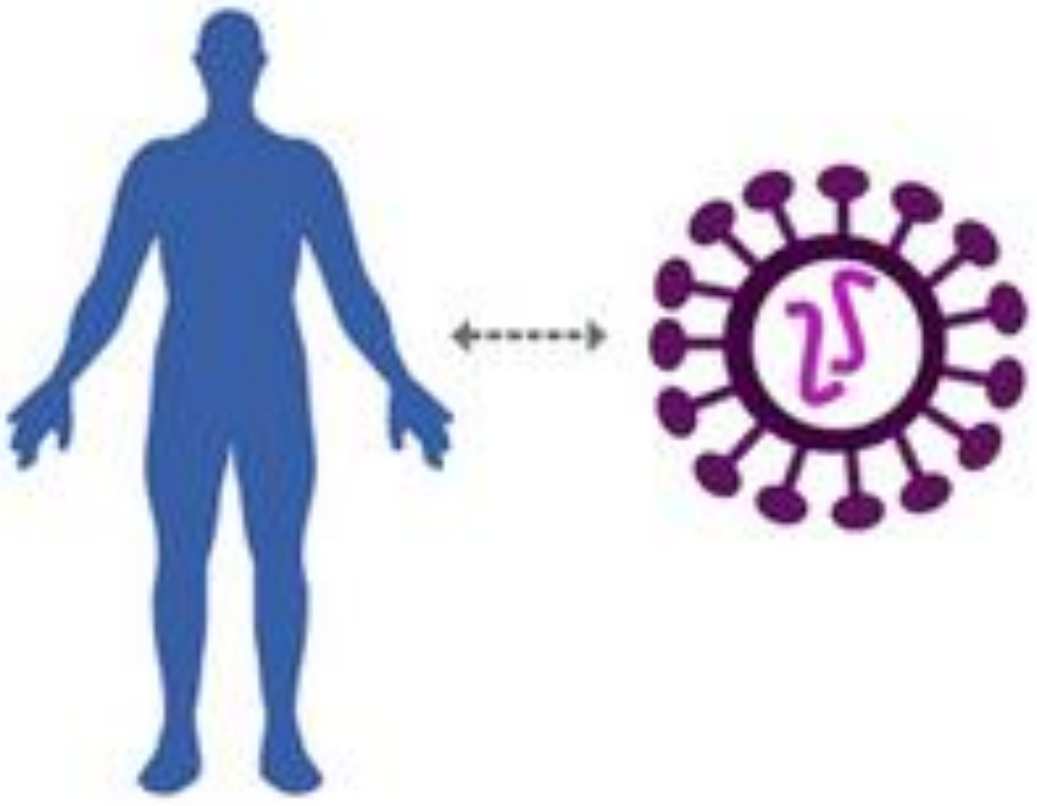
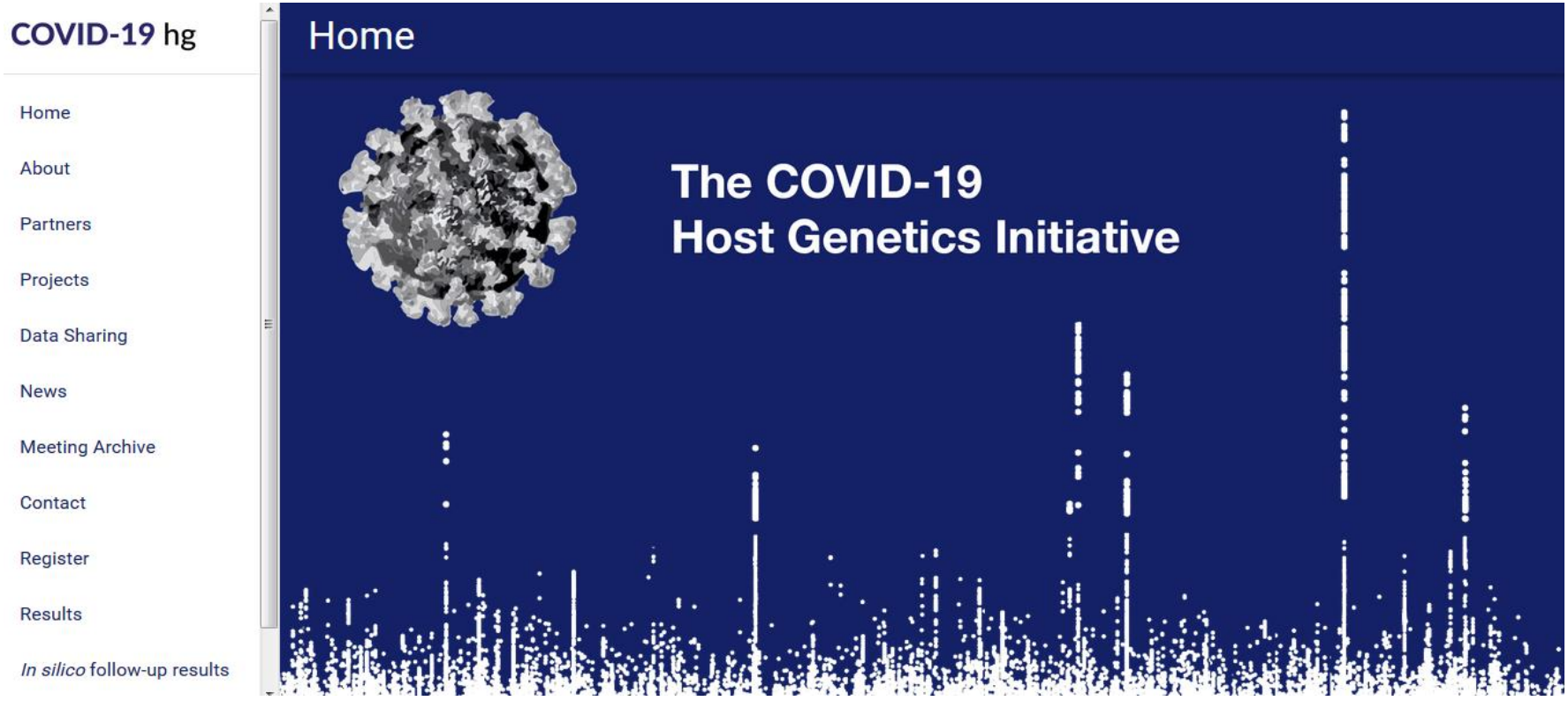


Figure 1: LamPORE is a simple and fast process comprising of amplification, library preparation and sequencing steps to identify whether the SARS-CoV-2 virus is present in a sample.

Main problems

- Collecting the large number of samples
- Handling in BSL2 laboratorios
- Data handling




COVID-19 hg
Home
About
Partners
Projects
Data Sharing
News
Meeting Archive
Contact
Register
Results
In silico follow-up results
Home
The COVID-19
Host Genetics Initiative

Covid-19 Host Genetics Initiative

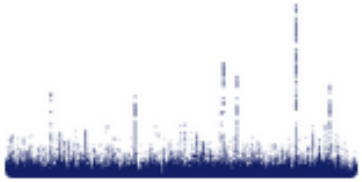
Mission

The COVID-19 host genetics initiative brings together the human genetics community to generate, share and analyze data to learn the genetic determinants of COVID-19 susceptibility, severity and outcomes. Such discoveries could help to generate hypotheses for drug repurposing, identify individuals at unusually high or low risk, and contribute to global knowledge of the biology of SARS-CoV-2 infection and disease.

Objectives

The COVID-19 host genetics initiative is a bottom-up collaborative effort that has three main aims:

Aim 1: Provide an environment to foster the sharing of resources to facilitate COVID-19 host genetics research (e.g. protocols, questionnaires).



Aim 2: Organize analytical activities across studies to identify genetic determinants of COVID-19 susceptibility and severity.

Aim 3: Provide a platform to share the results from meta-analytical activities to benefit the broader scientific community.

See our partners

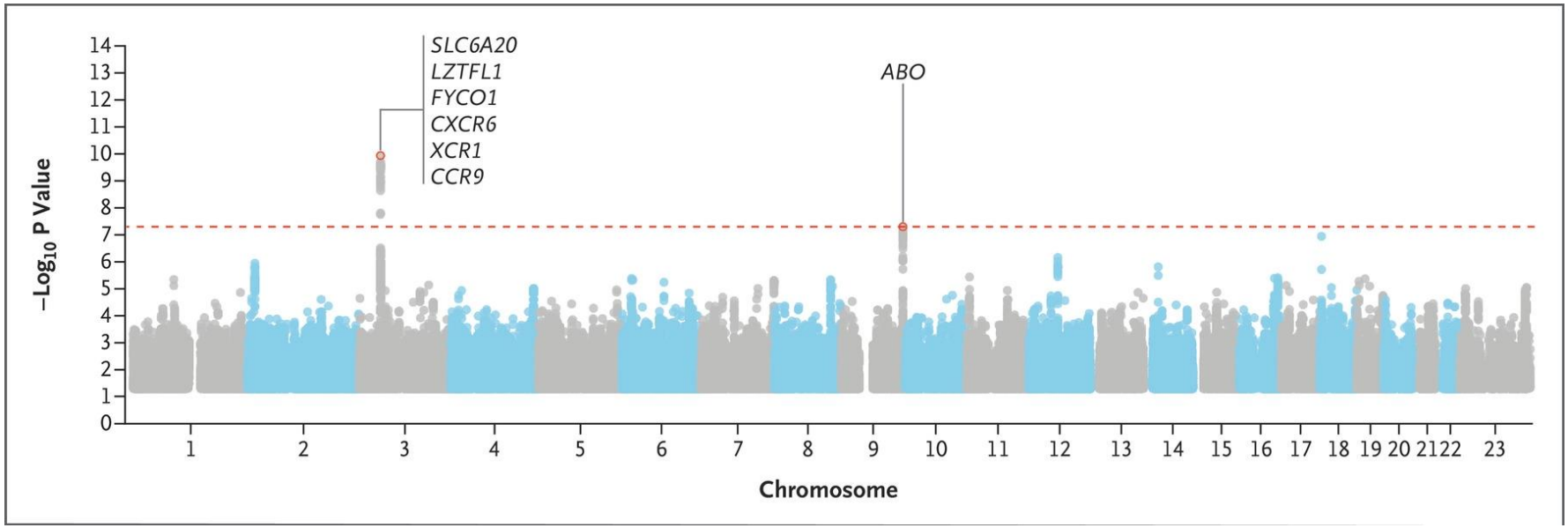
The NEW ENGLAND JOURNAL of MEDICINE

June 17, 2020
DOI: 10.1056/NEJMoa2020283

ORIGINAL ARTICLE

Genomewide Association Study of Severe Covid-19 with Respiratory Failure

David Ellinghaus, Ph.D., Frauke Degenhardt, M.Sc., Luis Bujanda, M.D., Ph.D., Maria Buti, M.D., Ph.D., Agustín Albillos, M.D., Ph.D., Pietro Invernizzi, M.D., Ph.D., Javier Fernández, M.D., Ph.D., Daniele Prati, M.D., Guido Baselli, Ph.D., Rosanna Asselta, Ph.D., Marit M. Grimsrud, M.D., Chiara Milani, Ph.D., et al., for The Severe Covid-19 GWAS Group*


SLC6A20
LZTFL1
FYCO1
CXCR6
XCR1
CCR9
ABO
−Log10 P Value
Chromosome
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

JAMA | Preliminary Communication

Presence of Genetic Variants Among Young Men With Severe COVID-19

Caspar I. van der Made, MD; Annet Simons, PhD; Janneke Schuurs-Hoeijmakers, MD, PhD; Guus van den Heuvel, MD; Tuomo Mantere, PhD; Simone Kersten, MSc; Rosanne C. van Deuren, MSc; Marloes Steehouwer, BSc; Simon V. van Reijmersdal, BSc; Martin Jaeger, PhD; Tom Hofste, BSc; Galuh Astuti, PhD; Jordi Corominas Galbany, PhD; Vyne van der Schoot, MD, PhD; Hans van der Hoeven, MD, PhD; Wanda Hagmolen of ten Have, MD, PhD; Eva Klijn, MD, PhD; Catrien van den Meer, MD; Jeroen Fiddelaers, MD; Quirijn de Mast, MD, PhD; Chantal P. Bleeker-Rovers, MD, PhD; Leo A. B. Joosten, PhD; Helger G. Yntema, PhD; Christian Gilissen, PhD; Marcel Nelen, PhD; Jos W. M. van der Meer, MD, PhD; Han G. Brunner, MD, PhD; Mihai G. Netea, MD, PhD; Frank L. van de Veerdonk, MD, PhD; Alexander Hoischen, PhD

IMPORTANCE Severe coronavirus disease 2019 (COVID-19) can occur in younger, predominantly male, patients without preexisting medical conditions. Some individuals may have primary immunodeficiencies that predispose to severe infections caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

OBJECTIVE To explore the presence of genetic variants associated with primary immunodeficiencies among young patients with COVID-19.

Editorial

Supplemental content

RESULTS The 4 male patients had a mean age of 26 years (range, 21-32), with no history of major chronic disease. They were previously well before developing respiratory insufficiency due to severe COVID-19, requiring mechanical ventilation in the ICU. The mean duration of ventilatory support was 10 days (range, 9-11); the mean duration of ICU stay was 13 days (range, 10-16). One patient died. Rapid clinical whole-exome sequencing of the patients and segregation in available family members identified loss-of-function variants of the X-chromosomal *TLR7*. In members of family 1, a maternally inherited 4-nucleotide deletion was identified (c.2129_2132del; p.[Gln710Argfs*18]); the affected members of family 2 carried a missense variant (c.2383G>T; p.[Val795Phe]). In primary peripheral blood mononuclear cells from the patients, downstream type I interferon (IFN) signaling was transcriptionally downregulated, as measured by significantly decreased mRNA expression of *IRF7*, *IFNB1*, and *ISG15* on stimulation with the TLR7 agonist imiquimod as compared with family members and controls. The production of IFN-γ, a type II IFN, was decreased in patients in response to stimulation with imiquimod.

CONCLUSIONS AND RELEVANCE In this case series of 4 young male patients with severe COVID-19, rare putative loss-of-function variants of X-chromosomal *TLR7* were identified that were associated with impaired type I and II IFN responses. These preliminary findings provide insights into the pathogenesis of COVID-19.

EASI Genomics

European Advanced infraStructure for Innovative Genomics

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824110

Extraordinary call for Covid-19 – 30. June 2020

- **Focus**
  - Host genetic factors
  - Severe cases <50 years old
  - No comorbidities

- **Selected studies**
  - 9 projects
  - >900 exomes
  - RNAseq
  - Longitudinal studies
  - Single cell

www.easi-genomics.eu


1+MillionGenomes
Member State Initiative

Declaration for delivering cross-border access to **genomic databases**

1 million **genomes accessible** in the EU by 2022

Linking access to existing and future genomic databases across the EU

Providing **proper scale** for research with clinical impact

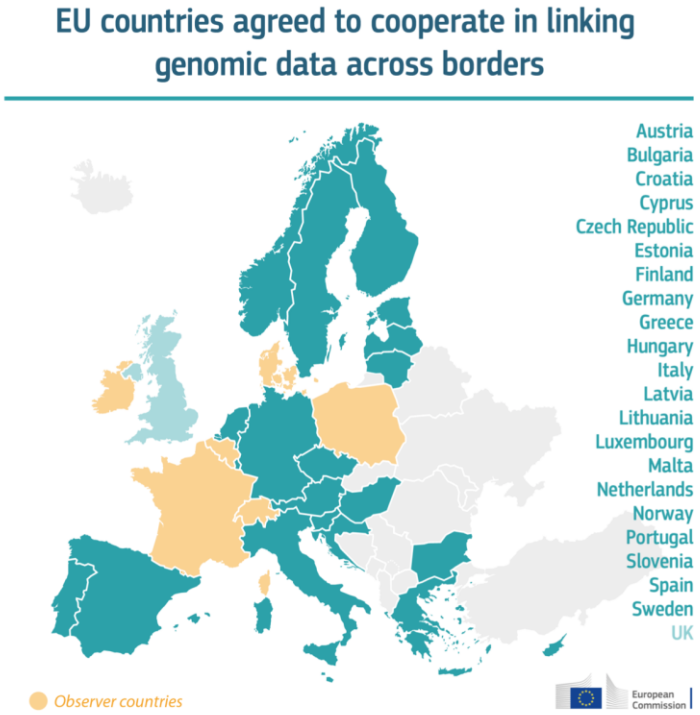
European Commission

1+MG Declaration of cooperation - April 2018

DECLARATION OF COOPERATION

Towards access to at least 1 million sequenced genomes in the European Union by 2022

https://ec.europa.eu/digital-single-market/en/european-1-million-genomes-initiative



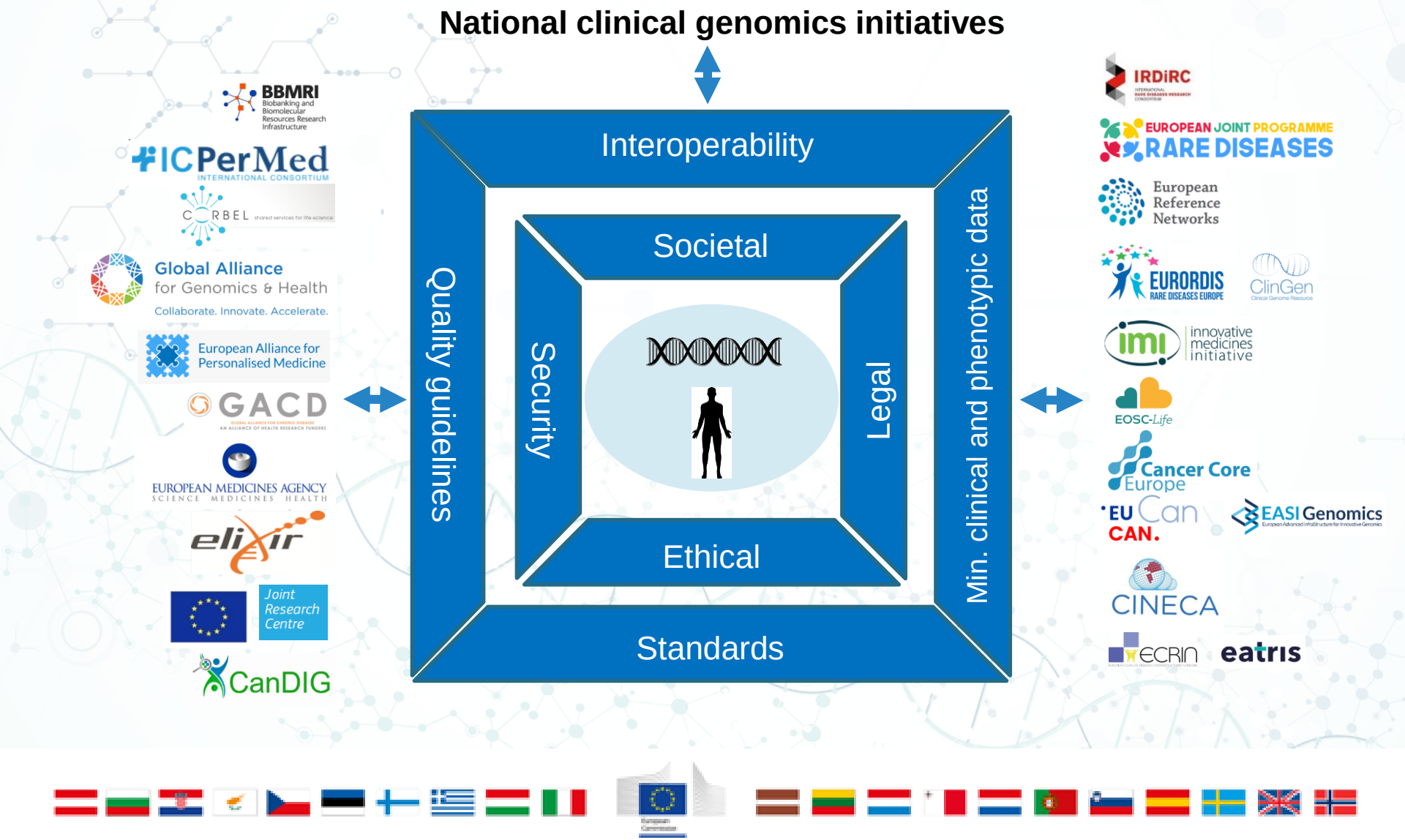
22 countries have now signed; 6 are observers

The 1+ Million Genome initiative

- Federated framework that would allow **secure** and **authorised** cross-border access to **genomic** and other **health data** across the EU, supporting **research, health care and prevention**.
- To allow users to search and access the data through a user-friendly and effective data governance structure **building on existing national and European initiatives**.
- To ensure that citizens, researchers and health systems in Europe can benefit from the full potential of genomics to **advance targeted health care interventions** leading to better **prevention, early diagnosis and treatment of diseases**

Member States of the European Union, the European Economic Area (EEA) and the European Free Trade Association (EFTA)

Join forces!


National clinical genomics initiatives
Interoperability
Societal
Security
Legal
Ethical
Quality guidelines
Min. clinical and phenotypic data
Standards
BBMRI
Biobanking and Biomolecular Resources Research Infrastructure
ICPerMed
INTERNATIONAL CONSORTIUM
CORBEL
shared services for life-science
Global Alliance for Genomics & Health
Collaborate. Innovate. Accelerate.
European Alliance for Personalised Medicine
GACD
EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH
elixir
Joint Research Centre
CanDIG
IRDiRC
EUROPEAN JOINT PROGRAMME RARE DISEASES
European Reference Networks
EURORDIS
RARE DISEASES EUROPE
ClinGen
Clinical Genome Resource
imi
innovative medicines initiative
EOSC-Life
Cancer Core Europe
EU Can CAN.
EASI Genomics
European Advanced Infrastructure for Innovative Genomics
CINECA
ECRIN
eatris

Creation of working groups

- WG1 Scope, stakeholders and governance
- WG2 Ethical, Legal, and Societal Issues (ELSI)
- WG3 Common standards and min. dataset for clinical and phenotypic data
- WG4 Good sequencing practice
- WG5 Federated, secure, interoperable and privacy-respecting framework and access governance
- WG6 Health economics and outcome research
- WG7 Involvement of the private sector
- WG8 Use case - Rare diseases
- WG9 Use case - Cancer
- WG10 Use case - Populations, Precision prevention, Pharmacogenomics
- WG11 Use case – Covid-19, Infectious diseases

In conclusion

- Interaction of the virus and the host is very important
- Virus needs to be handled in BSL2 laboratories
- Host can present many predisposing low-risk contributing genetic factors – polygenic risk scores
- Host can have rare variants that confer high risk – rare disease

- Sequencing of the virus and host is possible
- Data sharing accelerates discoveries
- This can provide insight into the potential outcomes and guide treatment decisions

Acknowledgements to funders

cnag
centre nacional d'anàlisi genòmica
centro nacional de análisis genómico

CRG
Centre
for Genomic
Regulation

cnag

baldiri reixac, 4
pcb - tower i, 2nd floor
08028 barcelona

t +34 93 4020542
f +34 93 4037279
www.cnag.eu

CRG Centre for Genomic Regulation

cnag