Update on COVNET (Genetic Determinants of COVID-19 Outcomes & Susceptibility)

COVNET Team

Division of Cancer Epidemiology and Genetics

NIH NATIONAL CANCER INSTITUTE

February 10, 2021

Genetic Susceptibility to SARS-CoV2 Outcomes

Under Construction

Three NCI/NIH Genetic Efforts in Progress

1. **COVNET**- Large-scale GWAS and Whole Genome Sequencing (WGS) of COVID-19 infection
   - 40,000 GWAS & 4-5,000 WGS
2. **COVIDcode** Study- NIH Clinical Center IRB approved study of 2,500
   - Genetic Analyses
   - Extensive Immunologic assessment (NIAID)
3. **NCI COVID-19 in Cancer Patients Study** (NCCAPS): Prospective Study of COVID-19 in Cancer Patients
   - Credit for Clinical Trials Enrollment (& Support)
   - 24-month follow-up

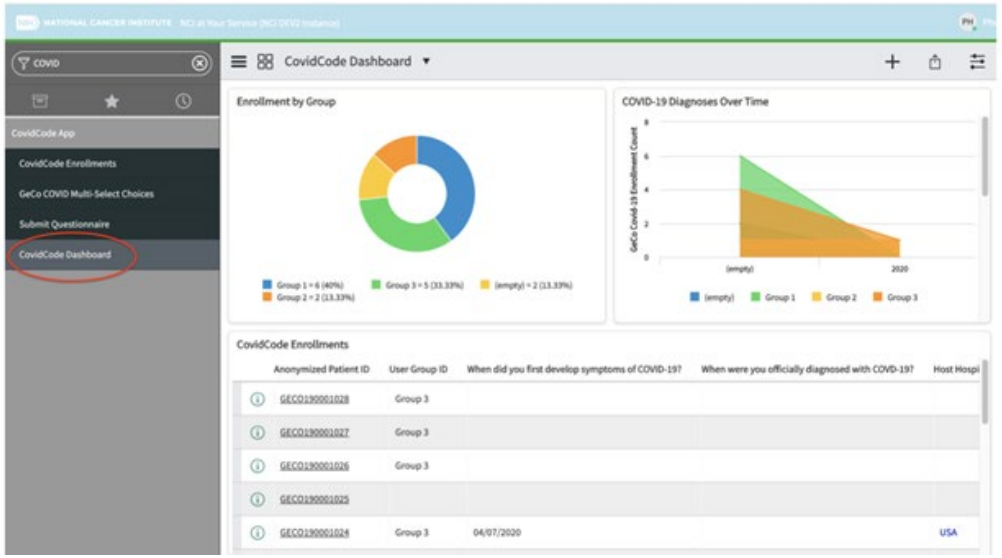
Principal Investigator: Les Biesecker, M.D. NHGRI
Associate Investigators: Stephen Chanock, M.D. NCI
Steven Holland, M.D. NIAID
Sharon Savage, M.D. NCI

- **Primary objectives**
  - To identify germline susceptibility variants that determine host responses to COVID-19 disease in case-case design.
  - Identify common and rare germline variants associated with host susceptibility to severe or fatal COVID-19 disease.
  - Deposit and share data quickly to enable community analyze according to NIH data sharing precepts.

- **Secondary objectives**
  - Exploratory analyses of epigenetic signatures, serologic immune markers and antibody profiles using whole blood collected.
  - Collect *(when possible & distributed carefully):*
    - Whole blood specimens for sera and DNA and RNA
    - Explore B and T cell repertoire
    - Serum or plasma to explore humoral response and soluble mediators
    - RNA sample tubes for transcriptomic analysis

COVIDcode – data collection

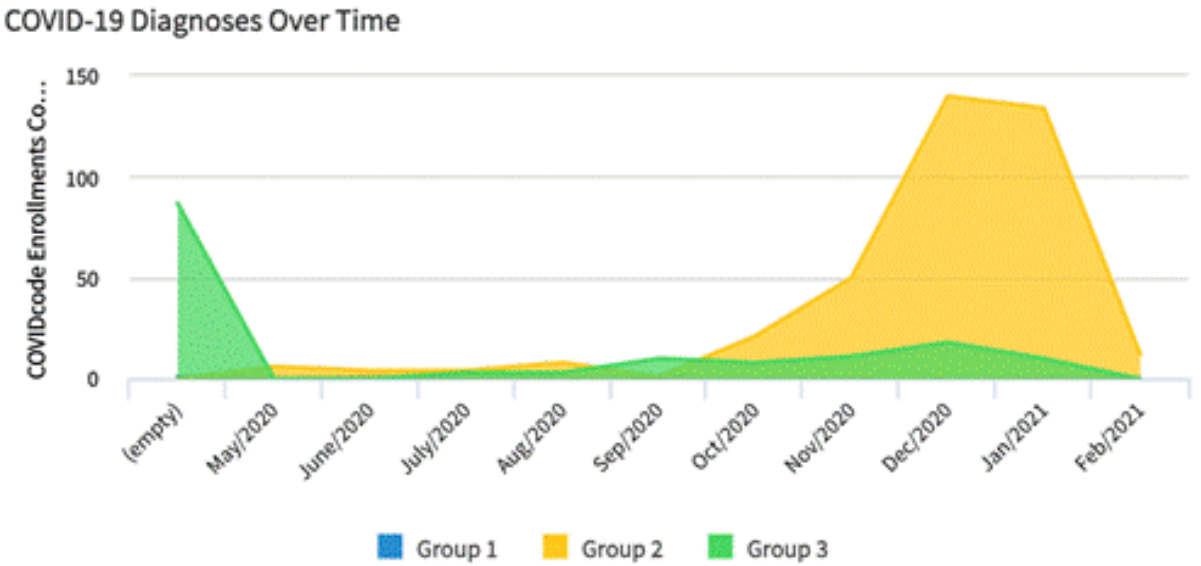
- Target recruitment= 2500
- Aligned with COVID-19 Host Genetics Initiative: https://www.covid19hg.org
- Uses CGB's CHARMS (communications hub and research management system)
- Minimum Questionnaire
  - ~30 questions
- Extended Questionnaire
  - ~150 questions
  - More clinical details



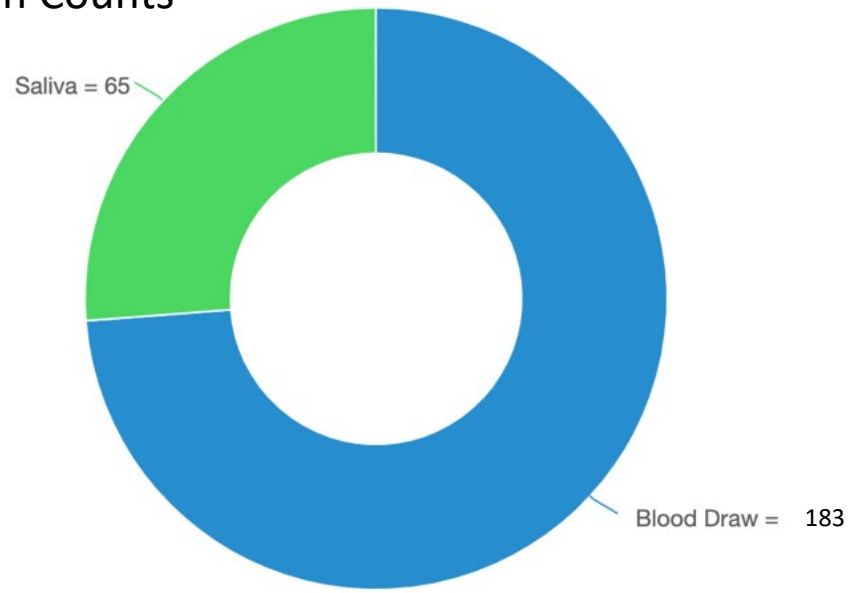
COVIDcode progress

Enrollment by Group

Group 2 = 397 (72.31%) Group 3 = 150 (27.32%) Group 1 = 2 (0.36%)



Specimen Counts



Blood Draw = 183 (73.79%) Saliva = 65 (26.21%)

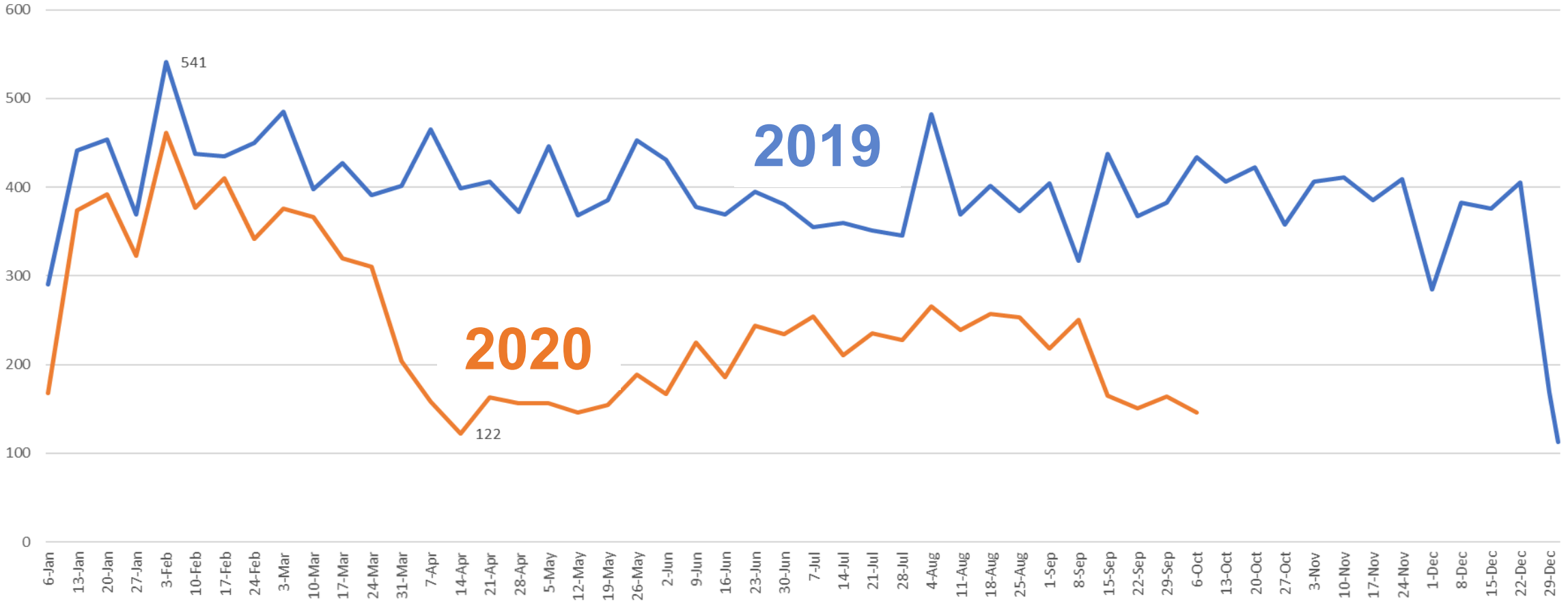
COVIDcode: Notable Features

- Approved by NIH IRB
  - Oral Consent
  - Participants or Next of Kin (if acutely ill)
- Capacity to recontact participants or family members
- Collection of risk factors Rapid oral/on-line questionnaire of 31 queries
  - “Harmonized” with COVID Host Genetics Initiative

- Protocol publicly posted and available for use

NCI Cancer Center Treatment Trial Accrual – 2019 and 2020

Weekly totals 1/6 to 12/29*



Data as of 11/4/2020. *2020 data through 10/6/2020. Cancer Centers report quarterly.
Cancer Center trials include Institutional/Investigator-Initiated trials without direct NCI grant funding.
Excludes NCTN, NCORP, etc. and Industrial trial accrual.

NCI COVID-19 in Cancer Patients Study (NCCAPS)

1. Cohort of cancer patients infected with COVID-19 comprising **all age groups** for collection of a comprehensive dataset on the cancers, treatments, medications, symptoms, course, and recovery, and co-morbidities with longitudinal follow-up until return to pre-morbid status;
2. Follow subset of pts for >1 yr to assess impact of COVID-19 on survivorship and cancer outcomes;
3. Collect blood samples at study entry and then every 2-3 months for 1 yr to estimate antibody and cellular immune response, genetic susceptibility, coagulation abnormalities, and for biomarker development;
4. Collect imaging and QOL data longitudinally;
5. Public database/biospecimens.

NCI COVID-19 in Cancer Patients Study (NCCAPS)

Critical Study Milestones

- **Study opened in late May 2020**
  - 6 weeks from idea to patient entry
- Enroll the first 500 patients within 3 months of trial activation
- Target accrual of 2,000 patients with biospecimens @ Nationwide
- Follow-up and survivorship evaluations by early 2022
- Begin biomarker studies on blood samples soon after initial 500 patients accrued—supported by new Congressional appropriation
- GWAS and NGS on all 2,000 to be done through CGR/DCEG

NCI COVID-19 in Cancer Patients Study (NCCAPS)

NIH NATIONAL CANCER INSTITUTE

Home > News & Events > Cancer Currents Blog

How Does COVID-19 Affect People with Cancer? NCCAPS Will Help Find Out

Subscribe

May 21, 2020, by James H. Doroshow, M.D.

With the sudden explosion of the COVID-19 pandemic, we are all living with a great deal of fear, uncertainty, and anxiety. As an oncologist and cancer researcher, I know that those feelings are heightened for many people with cancer.

People with cancer are already facing the shock of a cancer diagnosis, the tribulations that accompany treatment, or the stress of survivorship. On top of that, we're learning that people with cancer may be at higher risk of severe illness from COVID-19 because their cancer, or its treatment, has left them more vulnerable to complications.

NCI has launched a study called NCCAPS that will help scientists answer questions about COVID-19's impact on cancer patients and cancer's impact on the course of COVID-19.

Credit: iStock

850

TRIAL SITES ACTIVATED IN ALL STATES AND PUERTO RICO

778

PATIENTS SCREENED

687

ENROLLED

Figures as of January 5, 2021.

Germline genetics of COVID-19 susceptibility and manifestations

COVNET weblink: https://dceg.cancer.gov/research/how-we-study/genomic-studies/covnet

Manual for Conducting a Large-scale GWAS and Whole Genome Sequencing (WGS) of COVID-19 infection

Version 1.0 (June 12, 2020)

NIH Lead Investigator:
Stephen Chanock, M.D. (chanocks@mail.nih.gov)
Director, Division of Cancer Epidemiology and Genetics, NCI

NIH Investigators:
Sharon Savage, M.D. (savagesh@mail.nih.gov)
Lisa Mirabello, Ph.D. (mirabellol@mail.nih.gov)
Meredith Yeager, Ph.D. (yeagerm@mail.nih.gov)
Mitchell Machiela, Sc.D. (mitchell.machiela@nih.gov)
Joshua Sampson, Ph.D. (joshua.sampson@nih.gov)
Amy Hutchinson, M.S. (hutchiam@mail.nih.gov)
Belynda Hicks, M.S. (hicksbel@mail.nih.gov)
Division of Cancer Epidemiology and Genetics, NCI

Mary Carrington, Ph.D. (carringm@mail.nih.gov)
Center for Cancer Research, NCI

Leslie Biesecker, M.D. (lesb@mail.nih.gov)
Teri Manolio, M.D. (manoliot@mail.nih.gov)
National Human Genome Research Institute

Steven Holland, M.D. (sholland@niaid.nih.gov)
National Institute of Allergy and Infectious Diseases

Project Manager

Vibha Vij, M.S., MPH (vibha.vij@nih.gov)
Division of Cancer Epidemiology and Genetics, NCI

Status of COVNET- February 2021

- Over 150 programs/Geneticists in US approached
  - >100 Teleconferences
- MTAs (marker of commitment in US)
  - 17 Signed
  - 18 in process
- Sampled received
  - 3500 (12 studies)
  - 7500 expected in coming weeks
- Genotyped
  - 3500 (phenotypes received for ~1,000. **Over *2/3 awaiting phenotypes***)

COVNET-Current Collaborators

International Collaborators

Innsbruck Medical University, Austria

Seoul National University Hospital, Korea

Academy of Athens, Greece

Latin American countries (next slide)

COVNET Latin America

Large-scale genome-wide association study and whole genome sequencing of COVID-19 severity

January 2021

LATIN AMERICA - COLLABORATORS

Country	Investigator	Planned sample size	Current samples	MTA	Planned ship date (month)
Brazil	Eduardo Tarazona, Maria Cássia, Leandro Colli	1,300	300	In progress	**March**
Chile/Argentina/ Guatemala/Colombia	Luis A. Quiñones, Matias Olguin	1,000	500	Executed	**500 February/500 April**
Chile	Catterina Ferreccio, Vanessa Van de Wyngard	300	300	Final draft ready	**300 February**
Chile	Alvaro Cerda, Monica Aguilar	300		Pending IRB approval	**April**
Colombia/Venezuela	Bladimiro R Orozco	125	50	In progress	**50 February/75 April**
Peru	Meddly Santolalla	250	100	In progress	**March**

New countries:

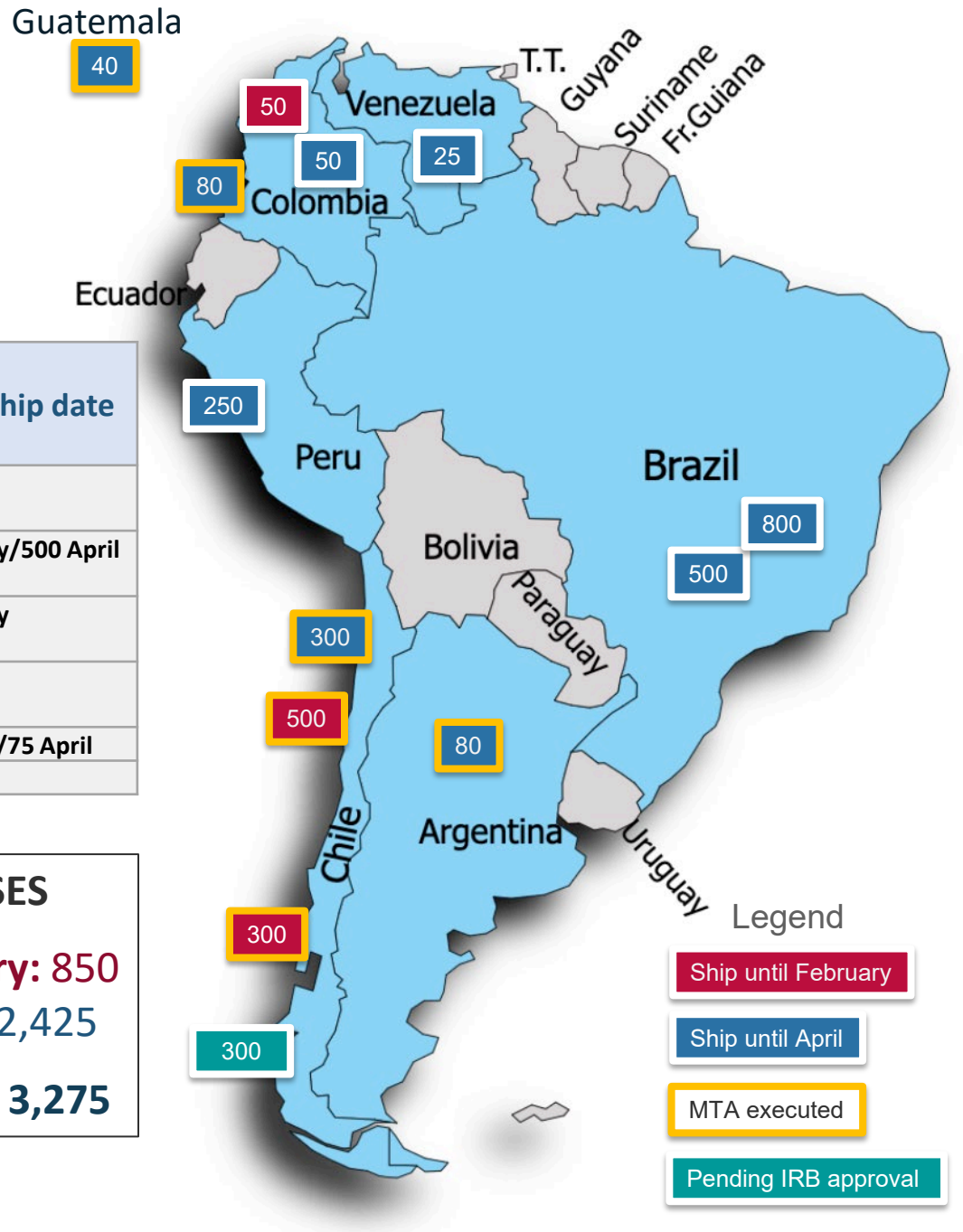
- Argentina
- Guatemala

CASES

February: 850

April: 2,425

TOTAL: 3,275



Distinctive Opportunities in COVNET

- First look at susceptibility to SARS-CoV2
  - CDC Studies x 2
  - Austrian Isghl Study
- Liquidator Cohort in Ukraine (post Chernobyl)
  - Long term consequences of radiation exposure
- Population Genetics of Latin America

Sample requirements and processing

Amy Hutchinson

Director of Operations, CGR

Sample Requirements

DNA

Mass: 1.0 – 1.5ug

Volume: ≥ 30ul

** Sample Kit use is encouraged. **

Blood

Whole Blood, Buffy Coat, or PBMCs

Volume: ≥ 150ul

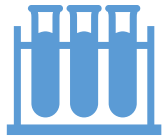
Buccal

Oragene, Mouthwash, or Saliva

Volume: ≥ 1000ul

Sample Processing

Samples are processed in a highly-automated, high-throughput laboratory environment following established SOPs



Laboratory

Tracking via a highly-integrated, customized Laboratory Information Management System (LIMS)

Extraction

KingFisher Flex

DNA QC

Volume verification

PicoGreen quantification

Genotyping

Standard input = 200ng

Minimum input = 50ng

1-3% QC replicates

Post-Genotyping Data QC

- Sample-level QC
  - Array processing: remove samples that fail to generate valid idat/gtc files
  - Completion rate: cutoff 0.8 for samples and 0.8 for loci, followed by 0.95 for samples and 0.95 for loci
  - Sample contamination: exclude samples with >10% contamination as predicted by VerifyIDintensity
  - Expected replicate removal: include replicate with higher call rate

- Subject-level QC
  - Sex verification: reported vs. observed sex based on chrX method-of-moments F coefficient. Cutoff at 0.5 (expected to be 0.0 for females and 1.0 for males)
  - Unexpected replicates: phenotypes are assessed before filtering unexpected replicates

Data sharing

Speaker: Stephen Chanock

Data Sharing Plan

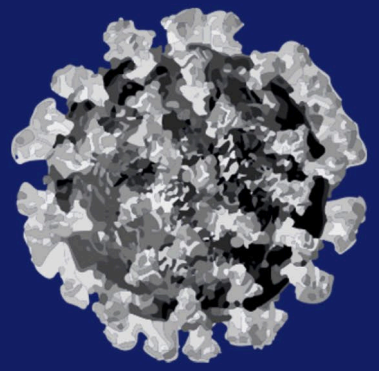
- Following primary data QC, a delivery package will be provided to each site via NIH secured transfer on the Box platform.
- Data access is restricted to you (and those you designate) and designated NCI staff only.
- Delivery package includes:
  - **QC Reports (.xls and .doc)** – These reports are generated from the genotyping batch of which your samples were a part. You will therefore notice additional samples included in both reports, this is expected and samples that are not your own can be ignored (no proprietary information is provided here).
  - **Called Genotypes** – Called genotypes for your subjects from CGR's pipeline in standard PLINK format.
  - **Sample ID Linkage File** – This is the master sample ID file and provides a listing of the various IDs contained in the reports and genotyping files for your samples only. This provides the link between IDs we generated internally as part of our laboratory pipelines and the IDs you provided with your samples. Any samples excluded from genotyping are listed at the top of the spreadsheet with a comment regarding the reason for exclusion.
  - **Illumina Infinium Genotyping_V3** – This is CGR's standard material and methods (M&M) document for Illumina genotyping and describes genotyping and data QC workflows currently in place. Specific details related to data QC of this project are included in the QC Report.

COVID-19 hg

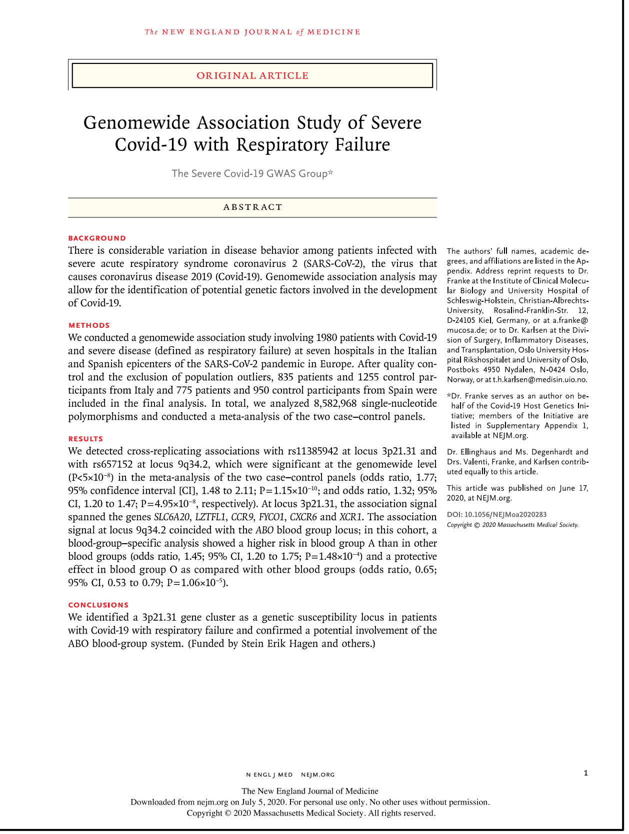
Home

- Home
- About
- Partners
- Projects
- Publications
- Data Sharing
- News
- Meeting Archive
- Contact
- Register
- Results
- *In silico* follow-up results
- Acknowledgements
- Media

Home



The COVID-19 Host Genetics Initiative


The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Genomewide Association Study of Severe Covid-19 with Respiratory Failure

The Severe Covid-19 GWAS Group*

ABSTRACT

BACKGROUND
There is considerable variation in disease behavior among patients infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes coronavirus disease 2019 (Covid-19). Genomewide association analysis may allow for the identification of potential genetic factors involved in the development of Covid-19.

METHODS
We conducted a genomewide association study involving 1980 patients with Covid-19 and severe disease (defined as respiratory failure) at seven hospitals in the Italian and Spanish epicenters of the SARS-CoV-2 pandemic in Europe. After quality control and the exclusion of population outliers, 835 patients and 1255 control participants from Italy and 775 patients and 950 control participants from Spain were included in the final analysis. In total, we analyzed 8,582,968 single-nucleotide polymorphisms and conducted a meta-analysis of the two case–control panels.

RESULTS
We detected cross-replicating associations with rs11385942 at locus 3p21.31 and with rs657152 at locus 9q34.2, which were significant at the genomewide level ($P<5\times10^{-8}$) in the meta-analysis of the two case–control panels (odds ratio, 1.77; 95% confidence interval [CI], 1.48 to 2.11; $P=1.15\times10^{-10}$; and odds ratio, 1.32; 95% CI, 1.20 to 1.47; $P=4.95\times10^{-8}$, respectively). At locus 3p21.31, the association signal spanned the genes *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6* and *XCR1*. The association signal at locus 9q34.2 coincided with the *ABO* blood group locus; in this cohort, a blood-group–specific analysis showed a higher risk in blood group A than in other blood groups (odds ratio, 1.45; 95% CI, 1.20 to 1.75; $P=1.48\times10^{-4}$) and a protective effect in blood group O as compared with other blood groups (odds ratio, 0.65; 95% CI, 0.53 to 0.79; $P=1.06\times10^{-5}$).

CONCLUSIONS
We identified a 3p21.31 gene cluster as a genetic susceptibility locus in patients with Covid-19 with respiratory failure and confirmed a potential involvement of the ABO blood-group system. (Funded by Stein Erik Hagen and others.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Franke at the Institute of Clinical Molecular Biology and University Hospital of Schleswig-Holstein, Christian-Albrechts-University, Rosalind-Franklin-Str. 12, D-24105 Kiel, Germany, or at a.franke@mucosa.de; or to Dr. Karlsen at the Division of Surgery, Inflammatory Diseases, and Transplantation, Oslo University Hospital Rikshospitalet and University of Oslo, Postboks 4950 Nydalen, N-0424 Oslo, Norway, or at t.h.karlsen@medisin.uio.no.

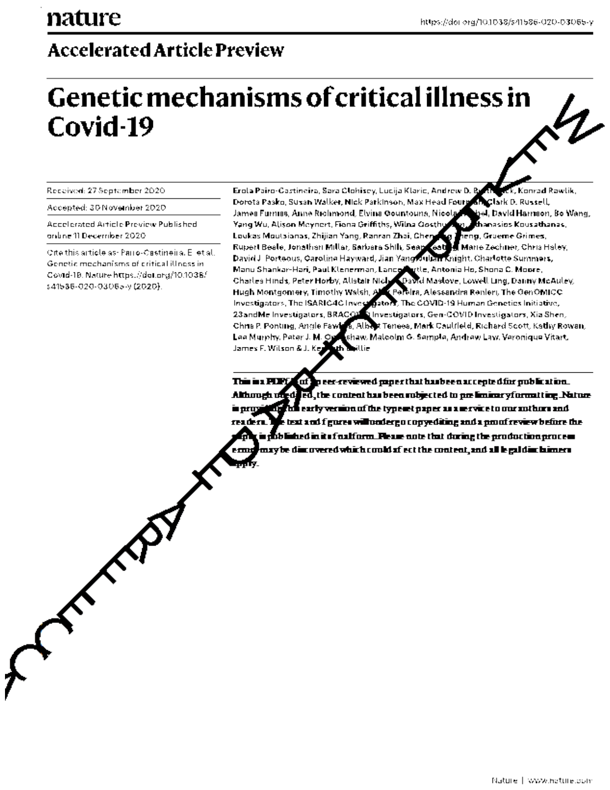
*Dr. Franke serves as an author on behalf of the Covid-19 Host Genetics Initiative; members of the Initiative are listed in Supplementary Appendix 1, available at NEJM.org.

Dr. Ellinghaus and Ms. Degenhardt and Drs. Valenti, Franke, and Karlsen contributed equally to this article.

This article was published on June 17, 2020, at NEJM.org.

DOI: 10.1056/NEJMoa2020283
Copyright © 2020 Massachusetts Medical Society.

The New England Journal of Medicine
Downloaded from nejm.org on July 5, 2020. For personal use only. No other uses without permission.
Copyright © 2020 Massachusetts Medical Society. All rights reserved.


nature https://doi.org/10.1038/s41586-020-03065-y

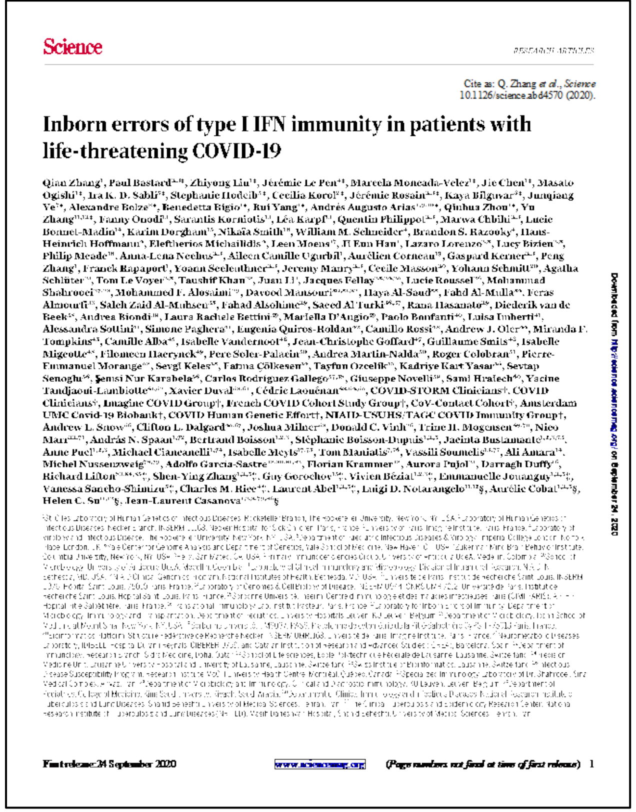
Accelerated Article Preview

Genetic mechanisms of critical illness in Covid-19

Received: 27 September 2020
Accepted: 30 November 2020
Accelerated Article Preview Published online 11 December 2020

Cite this article as: Pairo-Castineira, E. et al. Genetic mechanisms of critical illness in Covid-19. Nature https://doi.org/10.1038/s41586-020-03065-y (2020).

ACCELERATED ARTICLE PREVIEW


Science RESEARCH ARTICLES

Cite as: Q. Zhang et al., Science 10.1126/science.abd4570 (2020).

Inborn errors of type I IFN immunity in patients with life-threatening COVID-19

Downloaded from http://science.sciencemag.org/ on September 24, 2020

2 GWAS Loci
Chr 13
Chr 9- ABO Blood Group??

4 GWAS
Confirm Chr 13
New:
Chr 19 (*DPP9*)
Chr 12 (*OAS1-3*)
Chr 21 (*IFNAR2*)
But no ABO….

Set of Rare Variants in IFN genes

Other Large Scale Exome project
Have not replicated so far.

Challenge of Rare Variants….

Criteria for Whole Genome Sequencing (30X)

- Home team of COVNET will work with each study
- Highest priorities
  - Extreme phenotypes (mortality as well as survival with many comorbidities)
  - Population genetics
- Expect to do 15-20% of COVNET study set (all of NCCAPS & Covidcode)
- NGS at
  - American Genome Center at Uniformed Services University
  - HudsonAlpha Institute for Biotechnology with UAB

Analyses and preliminary findings: GWAS

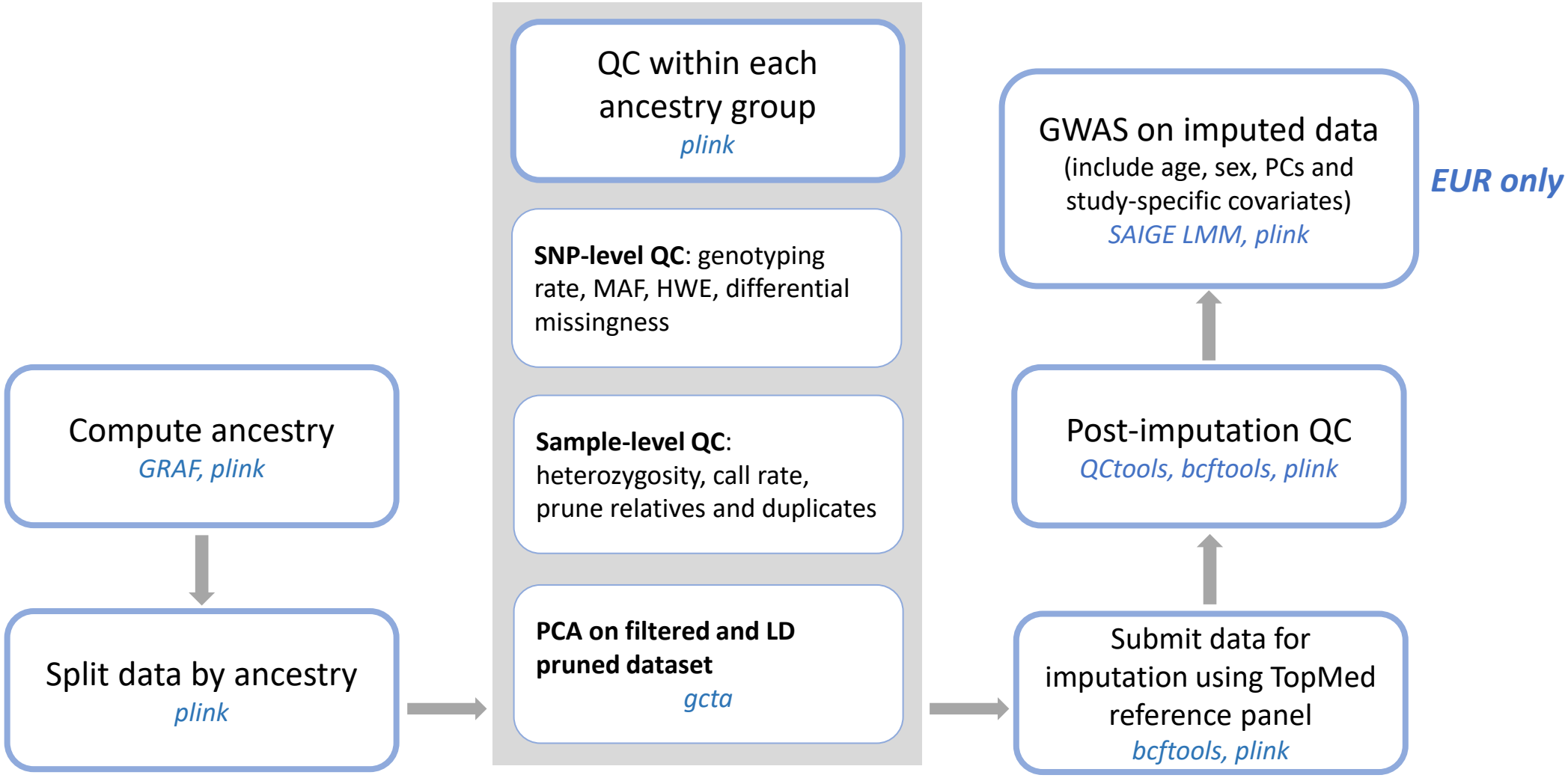
Meredith Yeager

Scientific Director, CGR

Lisa Mirabello

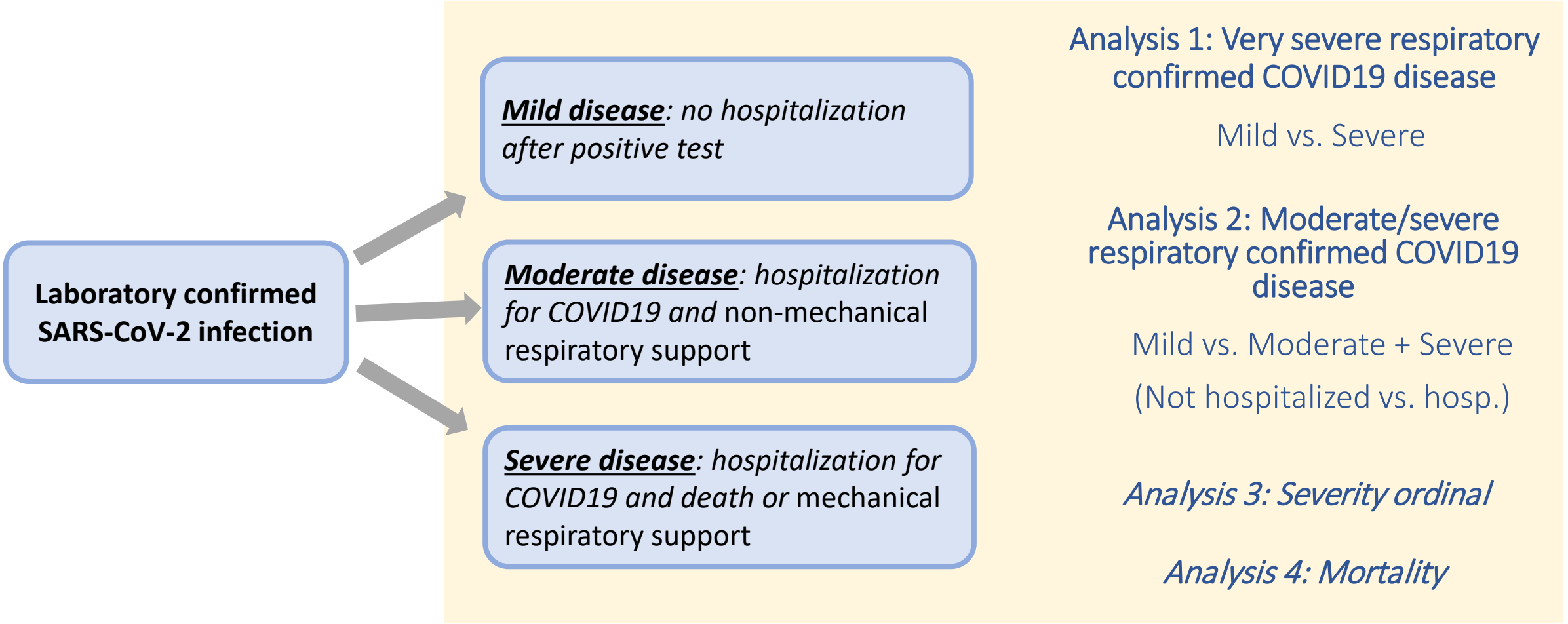
Senior Investigator, DCEG

GWAS Analysis plan



Adapted from Pairo-Castineira et al., Genetic mechanisms of critical illness in Covid-19. *Nature* 2020

Analysis Plan: Primary association analyses for disease severity



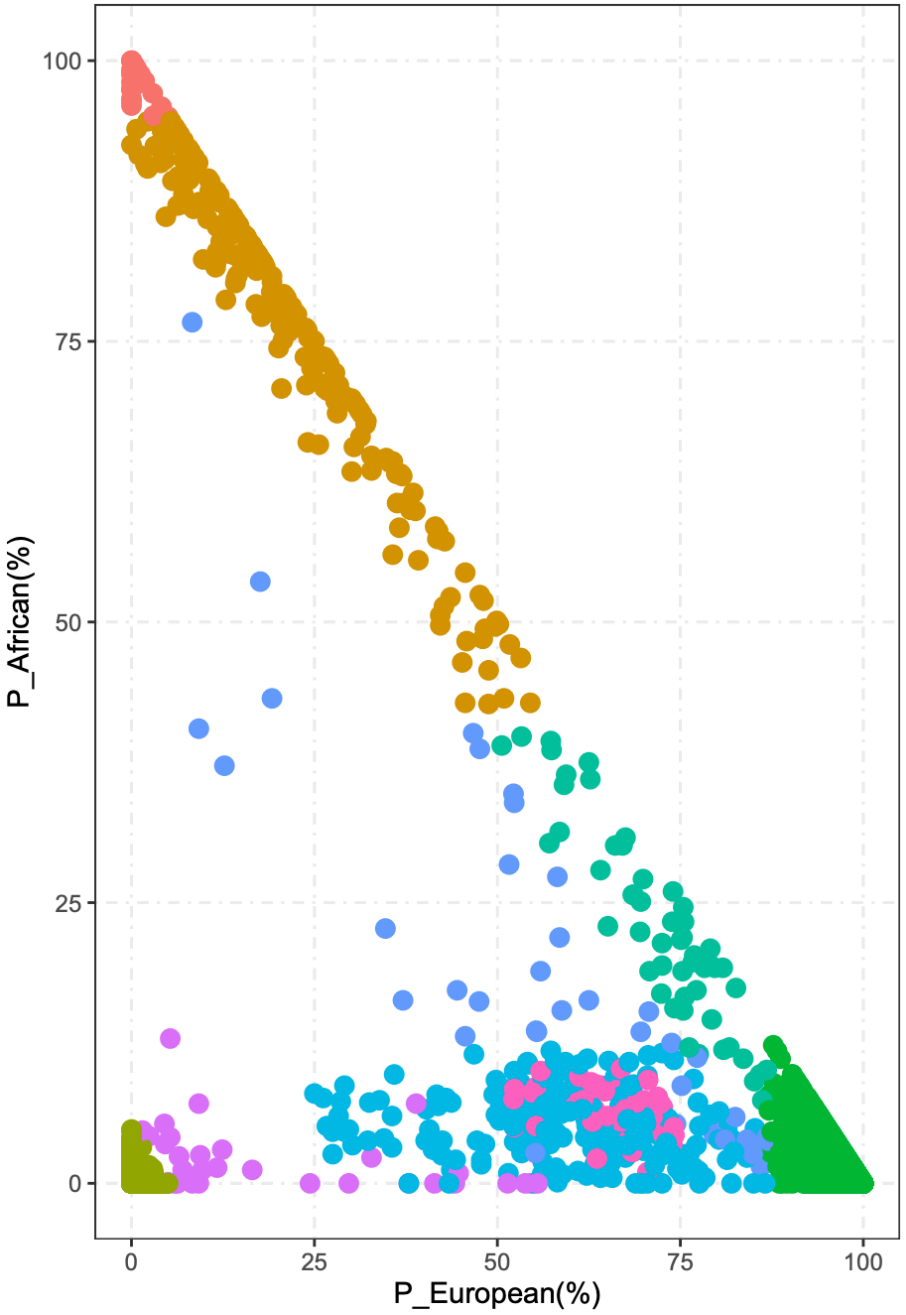
For EUR ancestry individuals: generalized linear mixed model used to evaluate imputed SNP associations for the specified outcomes; models adjusted for PCs, sex, age.

Samples received: Phenotype data required for analyses

Samples received	N	% of total
Total genotyped	2,405	
Total with phenotype	1,126	46.8%
Total phenotypes excluded or reclassified	188	16.7%
reason: no confirmed positive test	62	5.5%
reason: phenotype discordance/reclassified	66	5.9%
reason: not enough data to classify disease	60	5.3%

- Phenotype data needed using our provided template in order to harmonize disease severity classification across studies

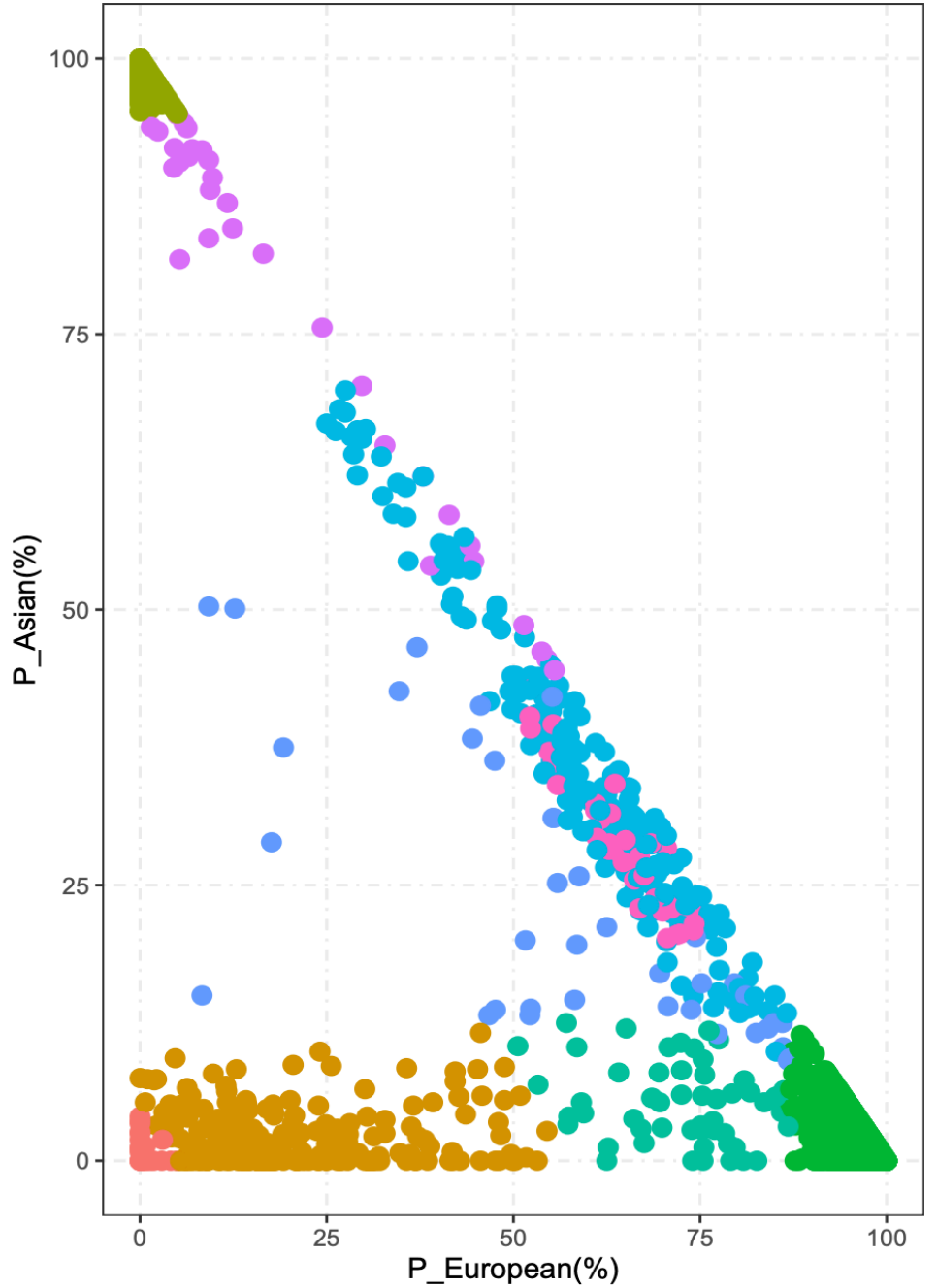
Ancestry computation with GRAF

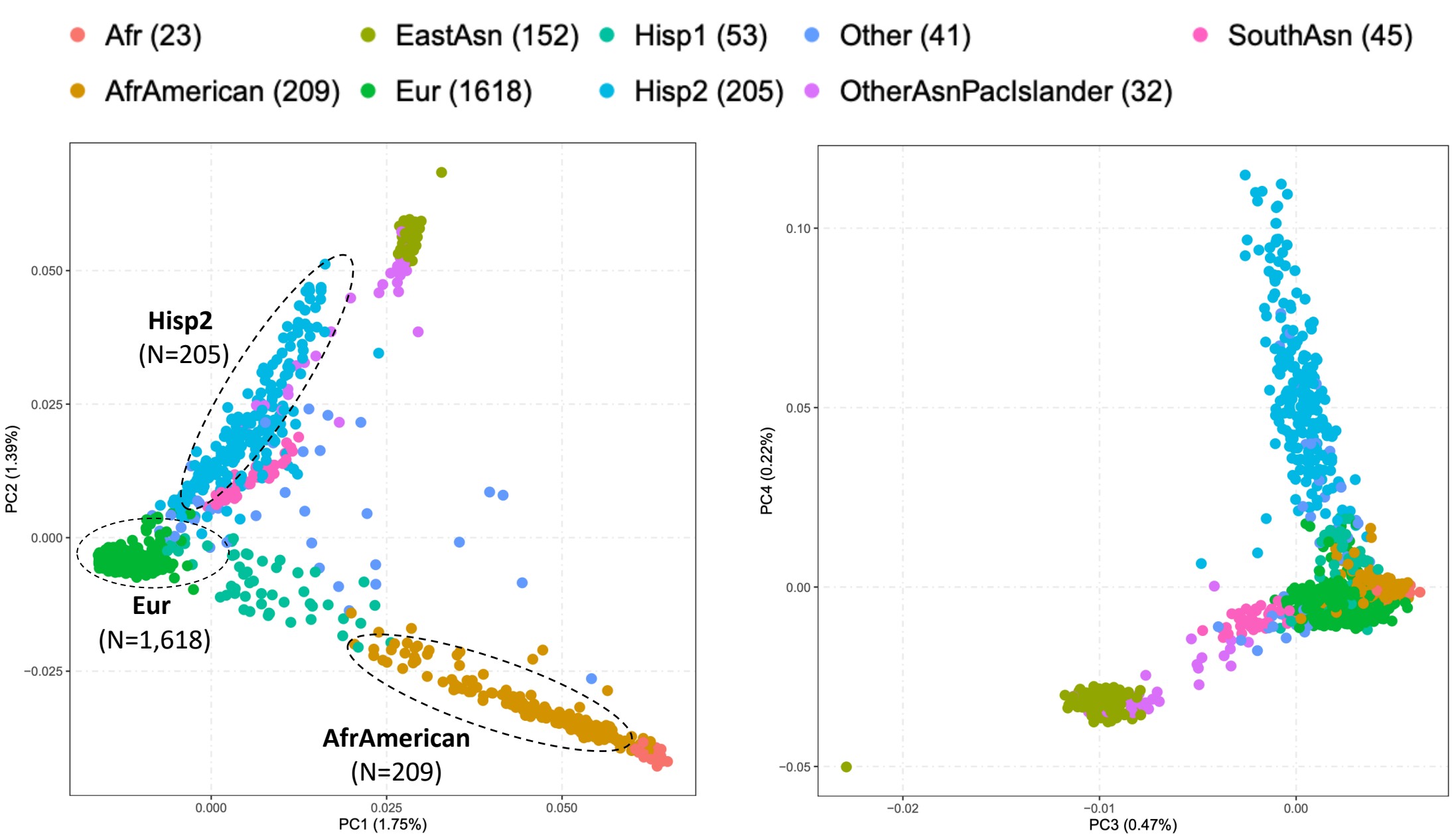


- Assumes that each subject is an admixture of three ancestries:
  - European
  - Asian
  - African
- Estimates proportions of the 3 ancestries and assigns each subject to a population

Computed ancestry

- African
- AfricanAmerican
- EastAsian
- European
- Hispanic1
- Hispanic2
- Other
- OtherAsianorPacificIslander
- SouthAsian

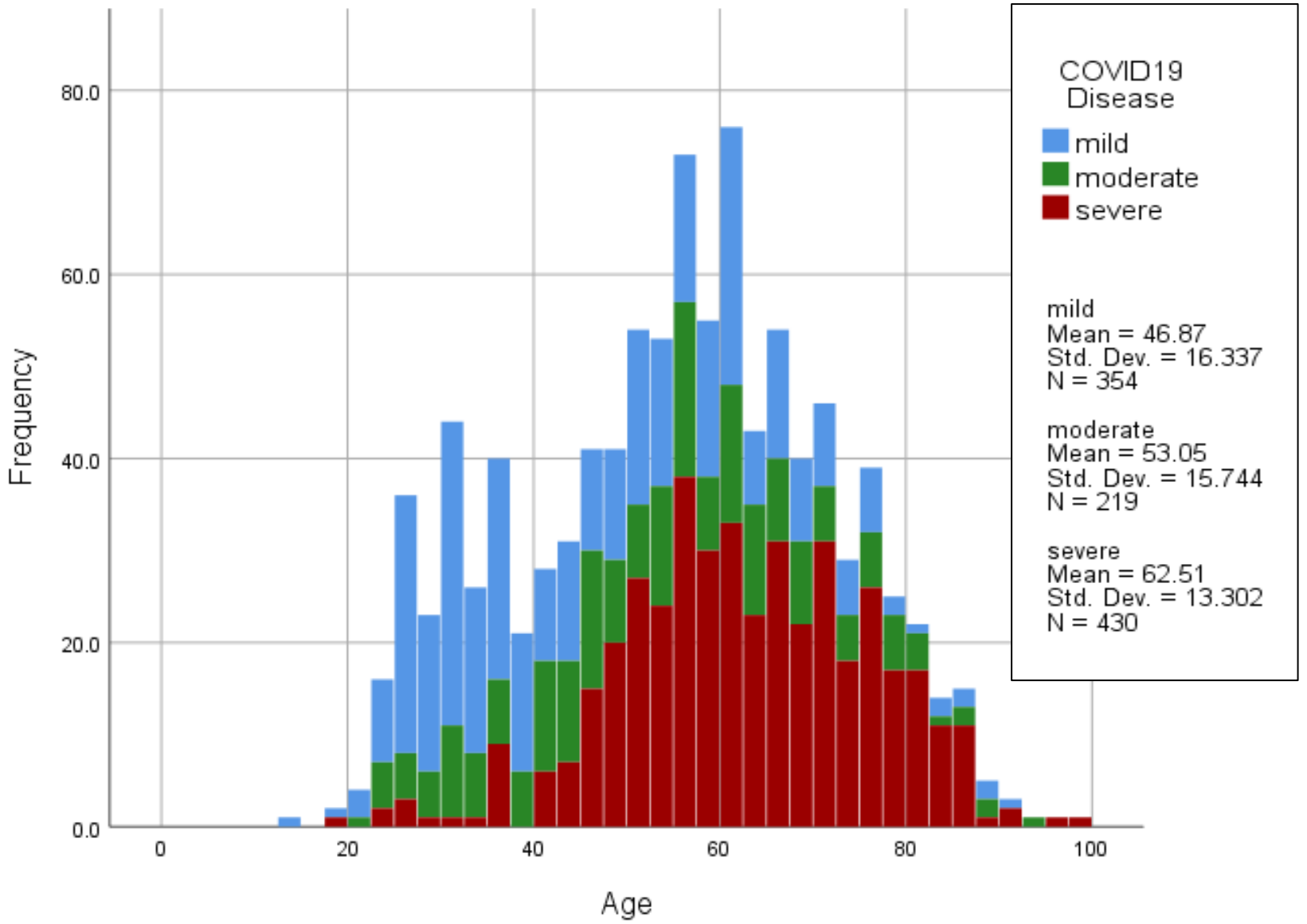




- PCA performed on filtered (MAF>0.1, HWE p-value>10^{-6}, missingness<1%) and LD-pruned SNPs (1000 markers, a step size of 80 markers and an r2 threshold of 0.1)

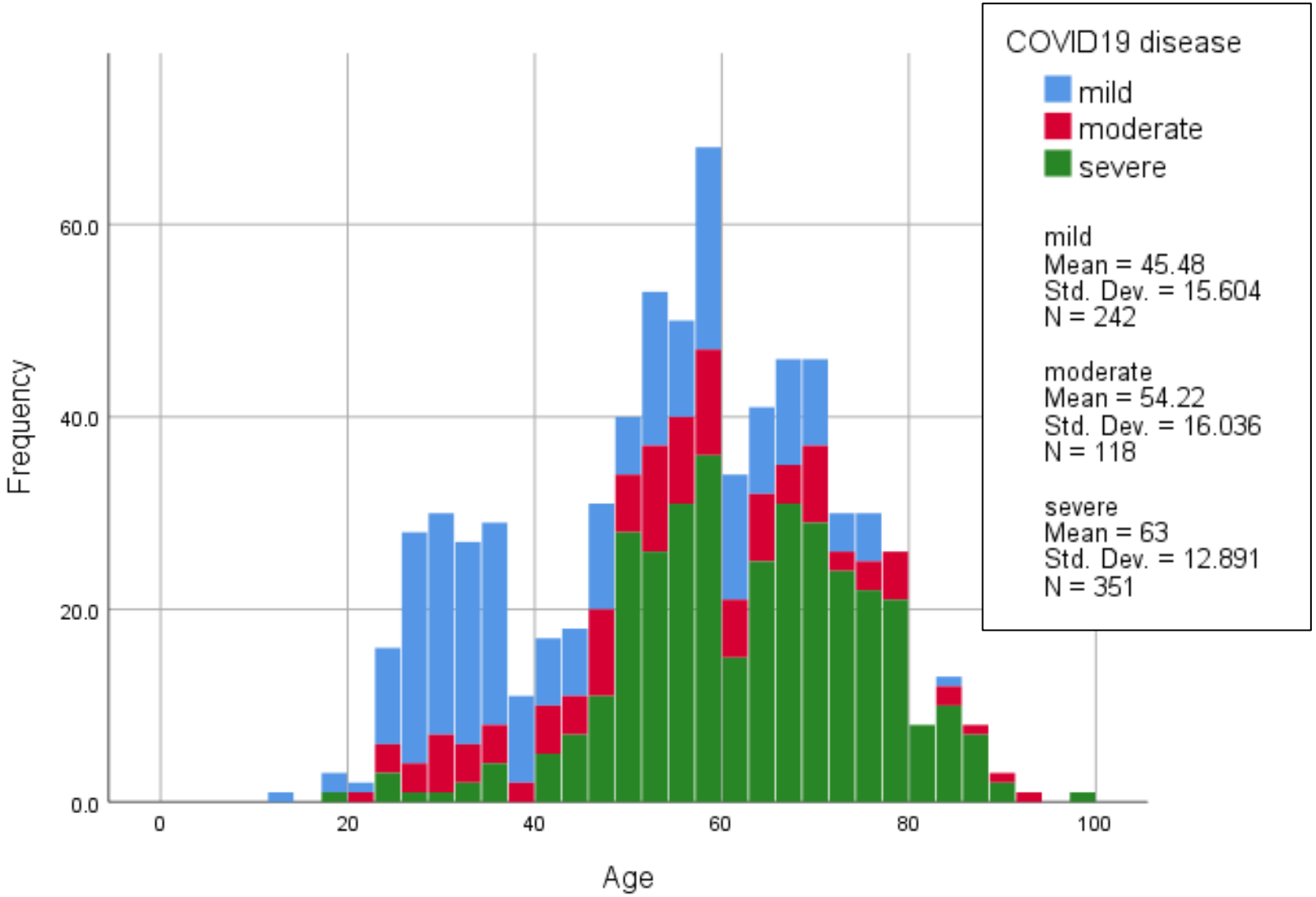
1,003 samples with genotype and phenotype data

COVID patients		N	Column %
Disease			
	mild	354	35.3%
	moderate	219	21.8%
	severe	430	42.9%
Sex			
	male	553	55.1%
	female	450	44.9%
Vital status			
	alive	616	88.0%
	deceased	84	12.0%
Smoking			
	current	18	4.6%
	former	95	24.2%
	never	280	71.2%



EUR samples with genotype and phenotype data

COVID patients		N	Column %
Disease			
	mild	242	34.0%
	moderate	118	16.6%
	severe	351	49.4%
Sex			
	male	396	55.7%
	female	315	44.3%
Vital status			
	alive	435	85.6%
	deceased	73	14.4%
Smoking			
	current	4	2.1%
	former	50	26.7%
	never	133	71.1%



Analysis of EUR samples with COVID disease severity classification available (N=702 unique subjects)

Sample-level QC

- 15 pairs of relatives (up to 3rd degree) and 23 pairs of unexpected duplicates detected
  - Retain the sample with a higher call rate
- 8 samples excluded due to heterozygosity > 3 s.d. from mean
- 659 samples remain

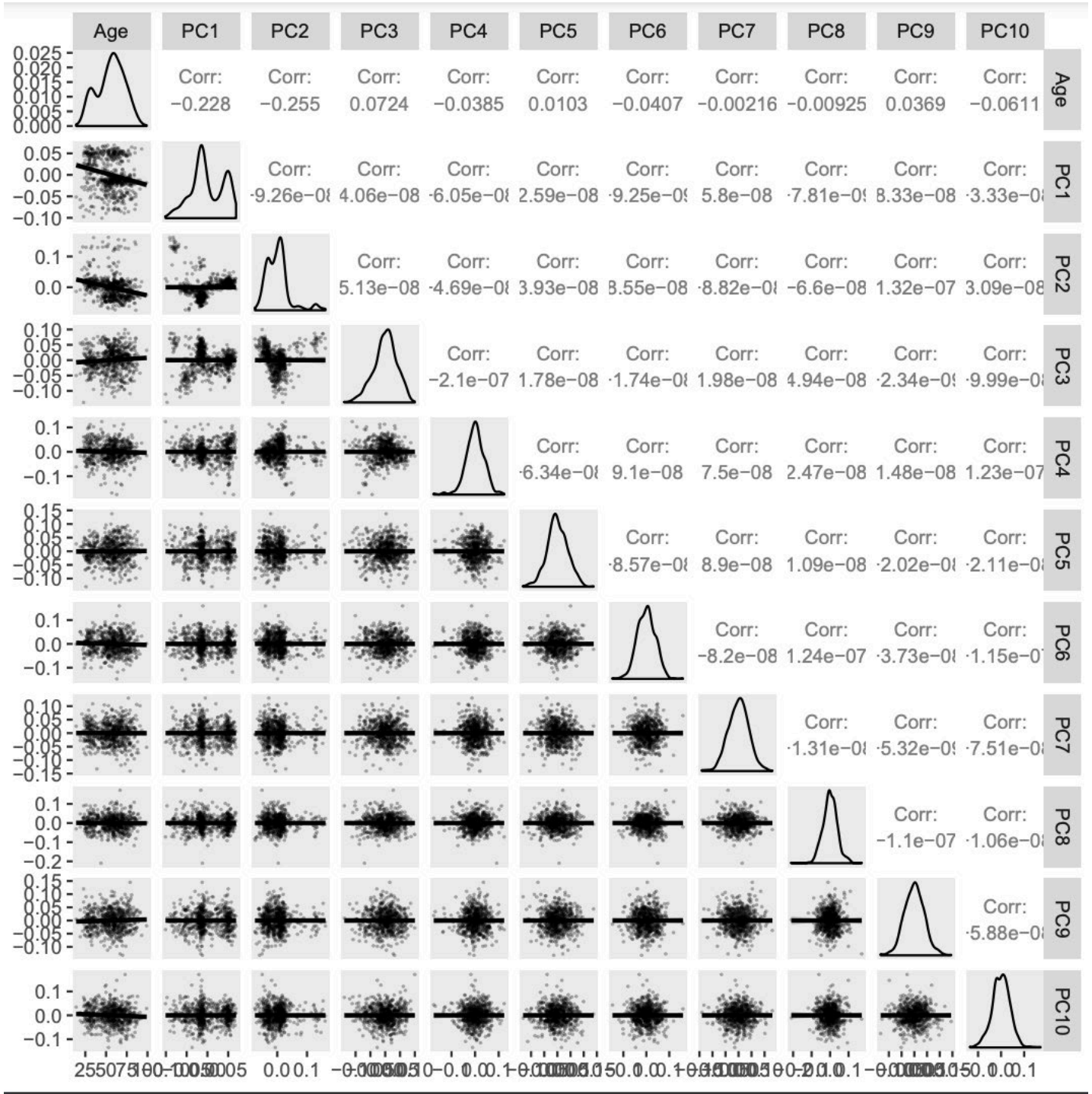
SNP-level QC

- Select variants with call rate > 99%, minor allele frequency (MAF) > 1%, Hardy-Weinberg equilibrium (HWE) p-value > 10^{-6}
- Exclude SNPs with significant differential missingness (P<0.001) across disease states

- Compute PCs with LD-pruned SNPs using gcta1.9
- Submit QC'd dataset for imputation using TopMed reference panel
- Download and QC imputed data (imputation quality score > 0.9, ~6.8M SNPs remain) and convert to BGEN format

Multicollinearity and regression on disease status and covariates

TEST	OR	P
Sex	0.68	0.09*
Age	1.04	3.6E-10*
PC1	4.0E-11	1.4E-17*
PC2	3.6E-14	9.9E-19*
PC3	3.1E+04	0.002*
PC4	0.06	0.28
PC5	39.37	0.21
PC6	1.84	0.83
PC7	357	0.05*
PC8	1.87	0.82
PC9	3.88	0.64
PC10	2.45	0.76



The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Genomewide Association Study of Severe Covid-19 with Respiratory Failure

The Severe Covid-19 GWAS Group*

ABSTRACT

BACKGROUND
There is considerable variation in disease behavior among patients infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes coronavirus disease 2019 (Covid-19). Genomewide association analysis may allow for the identification of potential genetic factors involved in the development of Covid-19.

METHODS
We conducted a genomewide association study involving 1980 patients with Covid-19 and severe disease (defined as respiratory failure) at seven hospitals in the Italian and Spanish epicenters of the SARS-CoV-2 pandemic in Europe. After quality control and the exclusion of population outliers, 835 patients and 1255 control participants from Italy and 775 patients and 950 control participants from Spain were included in the final analysis. In total, we analyzed 8,582,968 single-nucleotide polymorphisms and conducted a meta-analysis of the two case–control panels.

RESULTS
We detected cross-replicating associations with rs11385942 at locus 3p21.31 and with rs657152 at locus 9q34.2, which were significant at the genomewide level ($P<5\times10^{-8}$) in the meta-analysis of the two case–control panels (odds ratio, 1.77; 95% confidence interval [CI], 1.48 to 2.11; $P=1.15\times10^{-10}$; and odds ratio, 1.32; 95% CI, 1.20 to 1.47; $P=4.95\times10^{-8}$, respectively). At locus 3p21.31, the association signal spanned the genes *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6* and *XCR1*. The association signal at locus 9q34.2 coincided with the *ABO* blood group locus; in this cohort, a blood-group–specific analysis showed a higher risk in blood group A than in other blood groups (odds ratio, 1.45; 95% CI, 1.20 to 1.75; $P=1.48\times10^{-4}$) and a protective effect in blood group O as compared with other blood groups (odds ratio, 0.65; 95% CI, 0.53 to 0.79; $P=1.06\times10^{-5}$).

CONCLUSIONS
We identified a 3p21.31 gene cluster as a genetic susceptibility locus in patients with Covid-19 with respiratory failure and confirmed a potential involvement of the ABO blood-group system. (Funded by Stein Erik Hagen and others.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Franke at the Institute of Clinical Molecular Biology and University Hospital of Schleswig-Holstein, Christian-Albrechts-University, Rosalind-Franklin-Str. 12, D-24105 Kiel, Germany, or at a.franke@mucosa.de; or to Dr. Karlsen at the Division of Surgery, Inflammatory Diseases, and Transplantation, Oslo University Hospital Rikshospitalet and University of Oslo, Postboks 4950 Nydalen, N-0424 Oslo, Norway, or at t.h.karlsen@medisin.uio.no.

*Dr. Franke serves as an author on behalf of the Covid-19 Host Genetics Initiative; members of the Initiative are listed in Supplementary Appendix 1, available at NEJM.org.

Dr. Ellinghaus and Ms. Degenhardt and Drs. Valenti, Franke, and Karlsen contributed equally to this article.

This article was published on June 17, 2020, at NEJM.org.

DOI: 10.1056/NEJMoa2020283
Copyright © 2020 Massachusetts Medical Society.

The New England Journal of Medicine
Downloaded from nejm.org on July 5, 2020. For personal use only. No other uses without permission.
Copyright © 2020 Massachusetts Medical Society. All rights reserved.

2 GWAS Loci
Chr 13
Chr 9- ABO Blood Group??

nature https://doi.org/10.1038/s41586-020-03065-y

Accelerated Article Preview

Genetic mechanisms of critical illness in Covid-19

Received: 27 September 2020
Accepted: 30 November 2020
Accelerated Article Preview Published online 11 December 2020
Cite this article as: Pairo-Castineira, E. et al. Genetic mechanisms of critical illness in Covid-19. Nature https://doi.org/10.1038/s41586-020-03065-y (2020).

Erola Pairo-Castineira, Sara Clohisey, Lucija Klaric, Andrew D. Bretherick, Konrad Rawlik, Dorota Pasko, Susan Walker, Nick Parkinson, Max Head Fourman, Clark D. Russell, James Furniss, Anne Richmond, Elvina Gountouna, Nicola Wrobel, David Harrison, Bo Wang, Yang Wu, Alison Meynert, Fiona Griffiths, Wilna Oosthuyzen, Athanasios Kousathanas, Loukas Moutsianas, Zhijian Yang, Ranran Zhai, Chenqing Zheng, Graeme Grimes, Rupert Beale, Jonathan Millar, Barbara Shih, Sean Keating, Marie Zechner, Chris Haley, David J. Porteous, Caroline Hayward, Jian Yang, Julian Knight, Charlotte Summers, Manu Shankar-Hari, Paul Klenerman, Lance Turtle, Antonia Ho, Shona C. Moore, Charles Hinds, Peter Horby, Alistair Nichol, David Maslove, Lowell Ling, Danny McAuley, Hugh Montgomery, Timothy Walsh, Alex Pereira, Alessandra Renieri, The GenOMICC Investigators, The ISARIC4C Investigators, The COVID-19 Human Genetics Initiative, 23andMe Investigators, BRACOVID Investigators, Gen-COVID Investigators, Xia Shen, Chris P. Ponting, Angie Fawkes, Albert Tenesa, Mark Caulfield, Richard Scott, Kathy Rowan, Lee Murphy, Peter J. M. Openshaw, Malcolm G. Semple, Andrew Law, Veronique Vitart, James F. Wilson & J. Kenneth Baillie

This is a PDF file of a peer-reviewed paper that has been accepted for publication. Although unedited, the content has been subjected to preliminary formatting. Nature is providing this early version of the typeset paper as a service to our authors and readers. The text and figures will undergo copyediting and a proof review before the paper is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply.

ACCELERATED ARTICLE PREVIEW

4 GWAS
Confirm Chr 13
New:
Chr 19 (*DPP9*)
Chr 12 (*OAS1-3*)
Chr 21 (*IFNAR2*)
But no ABO....

Science

Cite as: Q. Zhang *et al.*, *Science* 10.1126/science.abd4570 (2020).

Inborn errors of type I IFN immunity in patients with life-threatening COVID-19

Qian Zhang, Paul Bastard, Zhiyong Liu, Jérémie Le Pen, Marcela Moncada-Velez, Jie Chen, Masato Ogishi, Ira K. D. Sabli, Stephanie Hodeib, Cecilia Korol, Jérémie Rosain, Kaya Bilguvar, Junqiang Ye, Alexandre Bolze, Benedetta Bigio, Rui Yang, Andrés Augusto Arias, Qinhua Zhou, Yu Zhang, Fanny Onodi, Sarantis Korniotis, Léa Karpf, Quentin Philippot, Marwa Chbihi, Lucie Bonnet-Madin, Karim Dorgham, Nikaïa Smith, William M. Schneider, Brandon S. Razooky, Hans-Heinrich Hoffmann, Eleftherios Michailidis, Leen Moens, Ji Eun Han, Lazaro Lorenzo, Lucy Bizien, Philip Meade, Anna-Lena Neehus, Aileen Camille Ugurbil, Aurélien Corneau, Gaspard Kerner, Peng Zhang, Franck Rapaport, Yoann Seeleuthner, Jeremy Manry, Cécile Masson, Yohann Schmitt, Agatha Schlüter, Tom Le Voyer, Taushif Khan, Juan Li, Jacques Fellay, Lucie Roussel, Mohammad Shahrooei, Mohammed F. Alosaimi, Davood Mansouri, Haya Al-Saud, Fahd Al-Mulla, Feras Almourfi, Saleh Zaid Al-Muhsen, Fahad Alsohime, Saeed Al Turki, Rana Hasanato, Diederik van de Beek, Andrea Biondi, Laura Rachele Bettini, Mariella D'Angio, Paolo Bonfanti, Luisa Imberti, Alessandra Sottini, Simone Paghera, Eugenia Quiros-Roldan, Camillo Rossi, Andrew J. Oler, Miranda F. Tompkins, Camille Alba, Isabelle Vandernoot, Jean-Christophe Goffard, Guillaume Smits, Isabelle Migeotte, Filomeen Haerynck, Pere Soler-Palacin, Andrea Martin-Nalda, Roger Colobran, Pierre-Emmanuel Morange, Sevgi Keles, Fatma Çölkesen, Tayfun Ozcelik, Kadriye Kart Yasar, Sevtap Senoglu, Şemsi Nur Karabela, Carlos Rodríguez Gallego, Giuseppe Novelli, Sami Hraiech, Yacine Tandjaoui-Lambiotte, Xavier Duval, Cédric Laouénan, COVID-STORM Clinicians†, COVID Clinicians†, Imagine COVID Group†, French COVID Cohort Study Group†, CoV-Contact Cohort†, Amsterdam UMC Covid-19 Biobank†, COVID Human Genetic Effort†, NIAID-USUHS/TAGC COVID Immunity Group†, Andrew L. Snow, Clifton L. Dalgard, Joshua Milner, Donald C. Vinh, Trine H. Mogensen, Nico Marr, András N. Spaan, Bertrand Boisson, Stéphanie Boisson-Dupuis, Jacinta Bustamante, Anne Puel, Michael Ciancanelli, Isabelle Meyts, Tom Maniatis, Vassili Soumelis, Ali Amara, Michel Nussenzweig, Adolfo García-Sastre, Florian Krammer, Aurora Pujol, Darragh Duffy, Richard Lifton, Shen-Ying Zhang‡, Guy Gorochov‡, Vivien Béziat‡, Emmanuelle Jouanguy‡, Vanessa Sancho-Shimizu‡, Charles M. Rice‡, Laurent Abel‡, Luigi D. Notarangelo§, Aurélie Cobat§, Helen C. Su§, Jean-Laurent Casanova§

Downloaded from http://science.sciencemag.org/ on September 24, 2020

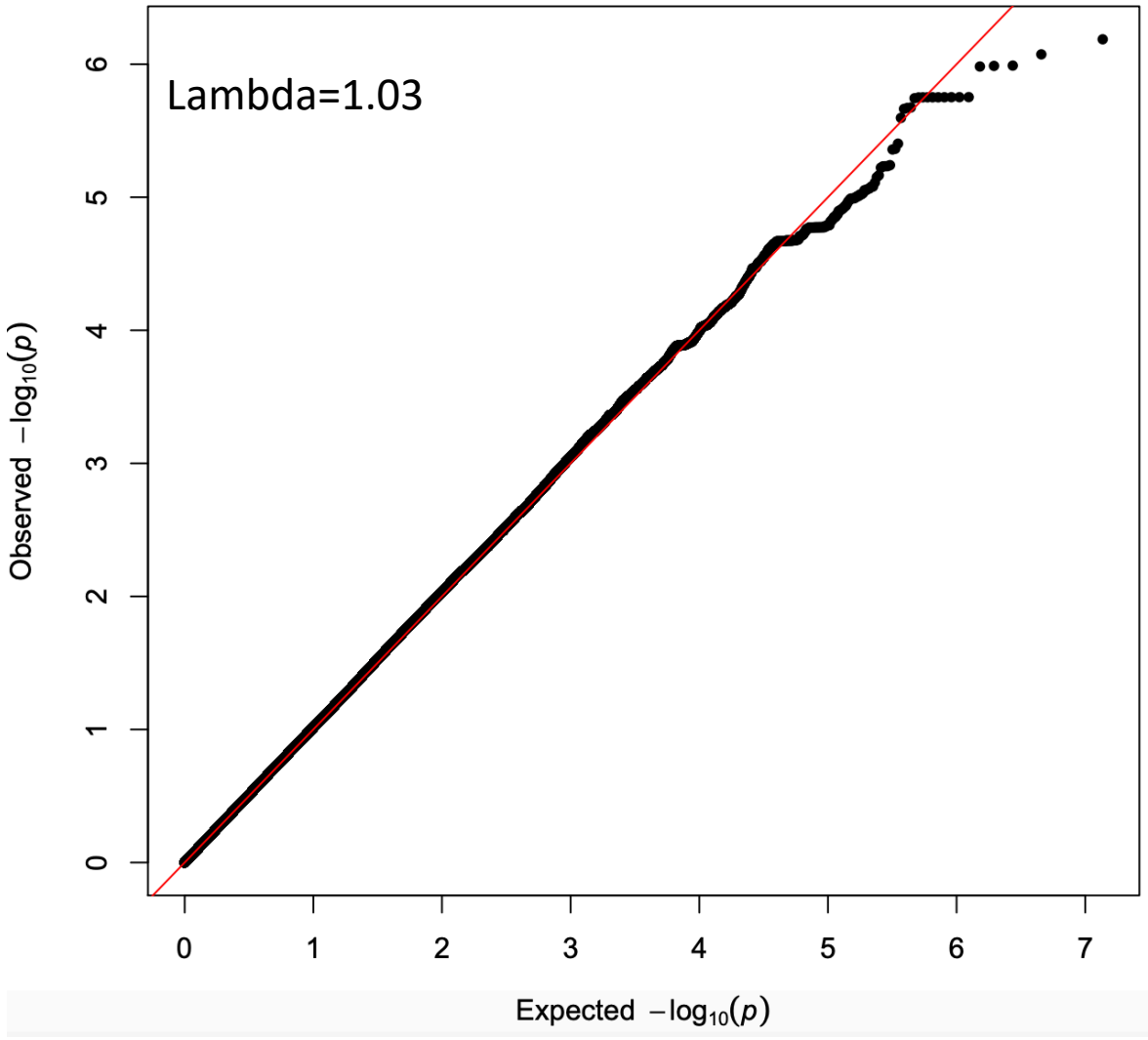
Set of Rare Variants in IFN genes

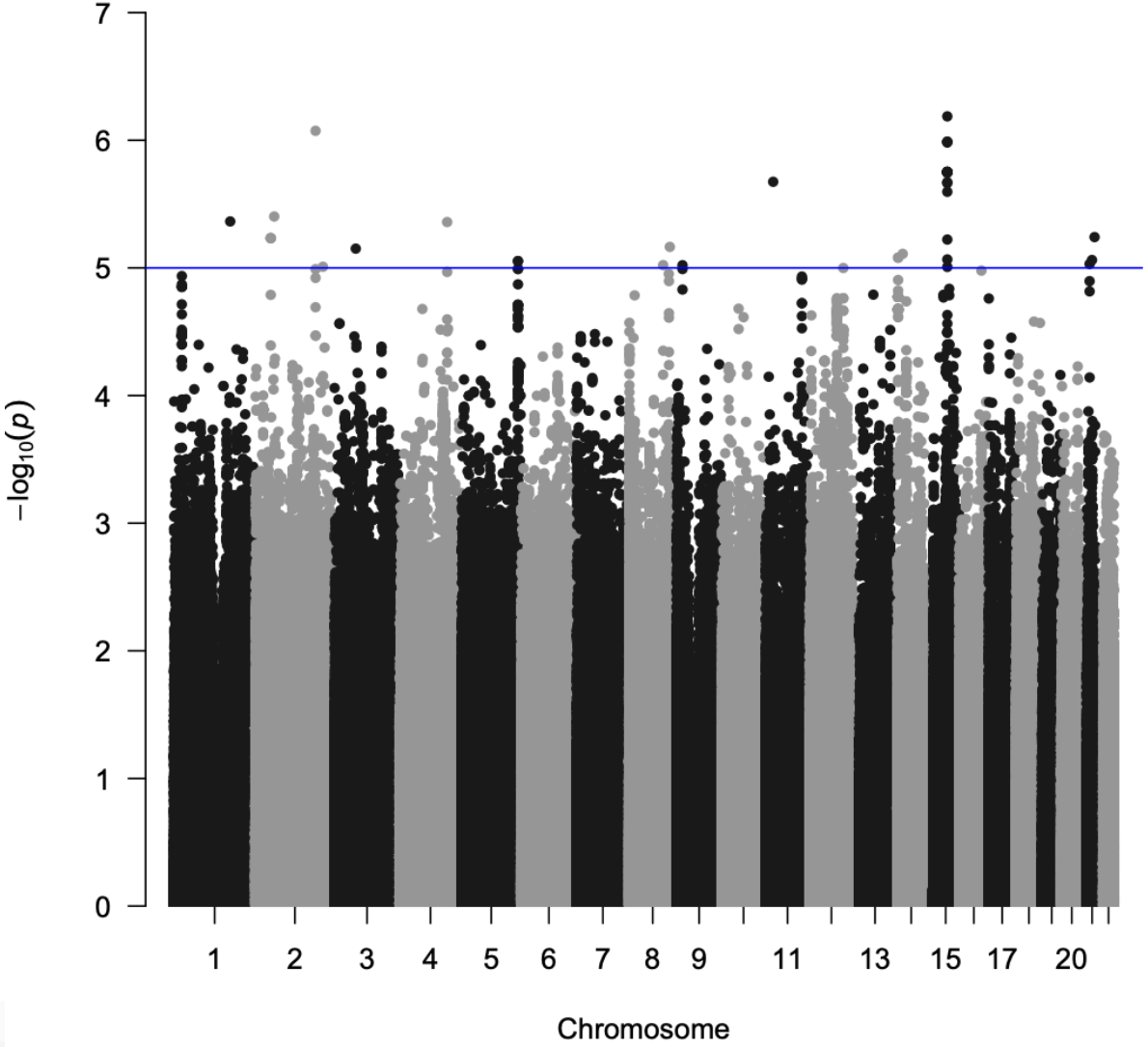
Other Large Scale Exome project
Have not replicated so far.

Challenge of Rare Variants....

GWAS: EUR COVID-19 Mild (N=223) vs. Severe (N=326) after QC of imputed data

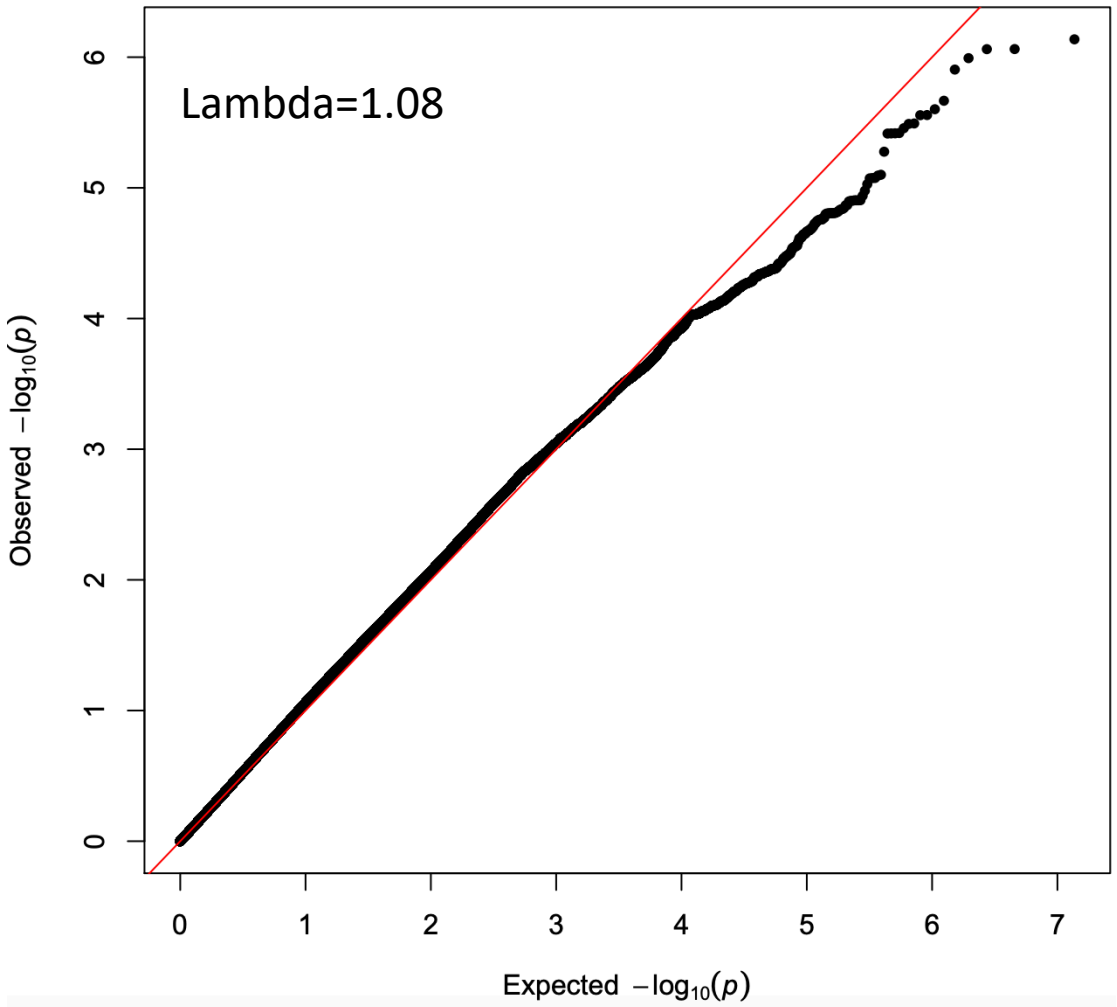
- Results generated with Plink. GWAS with SAIGE LMM in progress.
- **Covariates**: PCs1-10, Age, Mean-centered Age^2, Sex
- **Top SNP**: chr15:65913160:A:G (b38); p-value = 6.5e-07

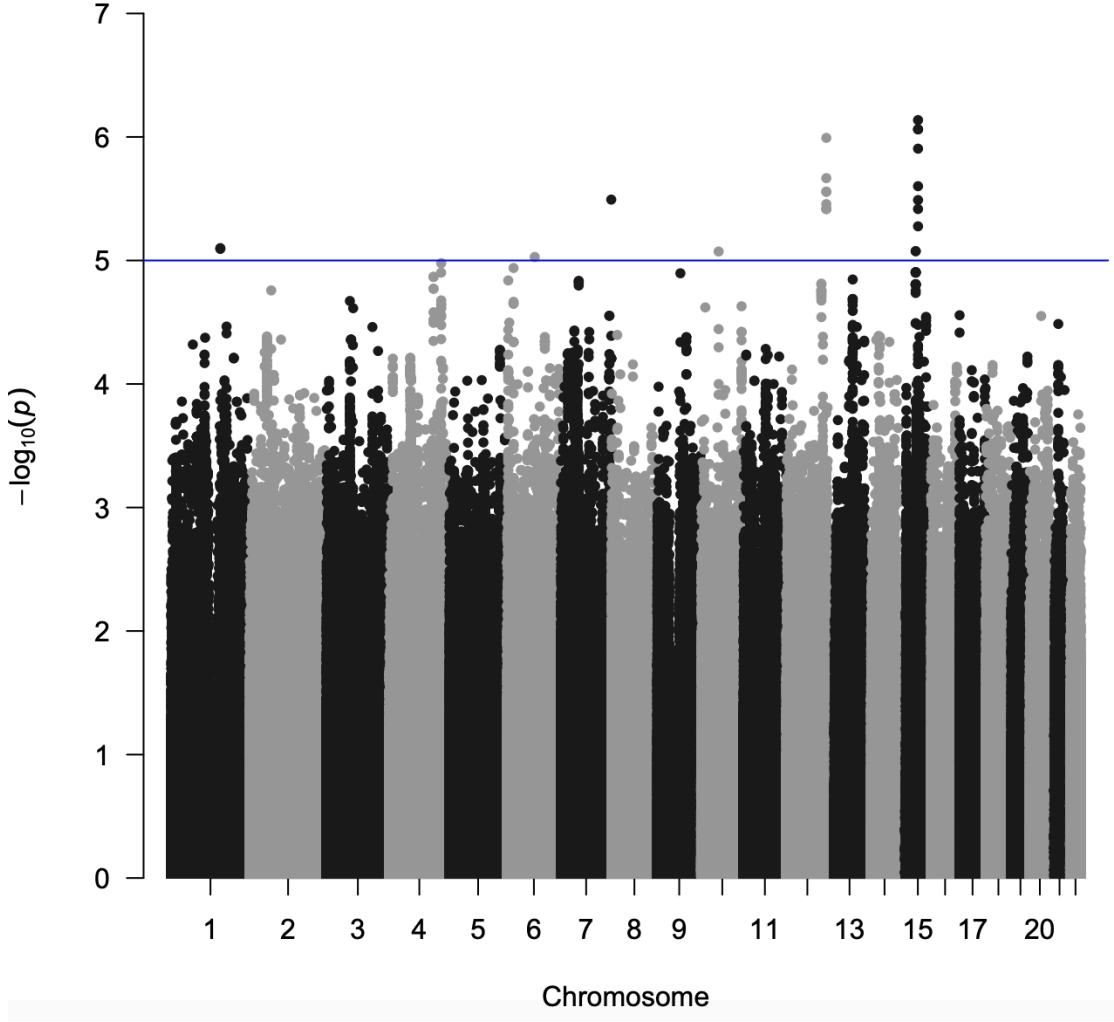




GWAS: EUR COVID-19 Mild (N=223) vs. Moderate + Severe (N=436) after QC of imputed data

- Results generated with Plink. GWAS with SAIGE LMM in progress.
- **Covariates**: PCs1-10, Age, Mean-centered Age^2, Sex
- **Top SNP**: chr15:61507130:G:A (b38); p-value = 7. 3e-07

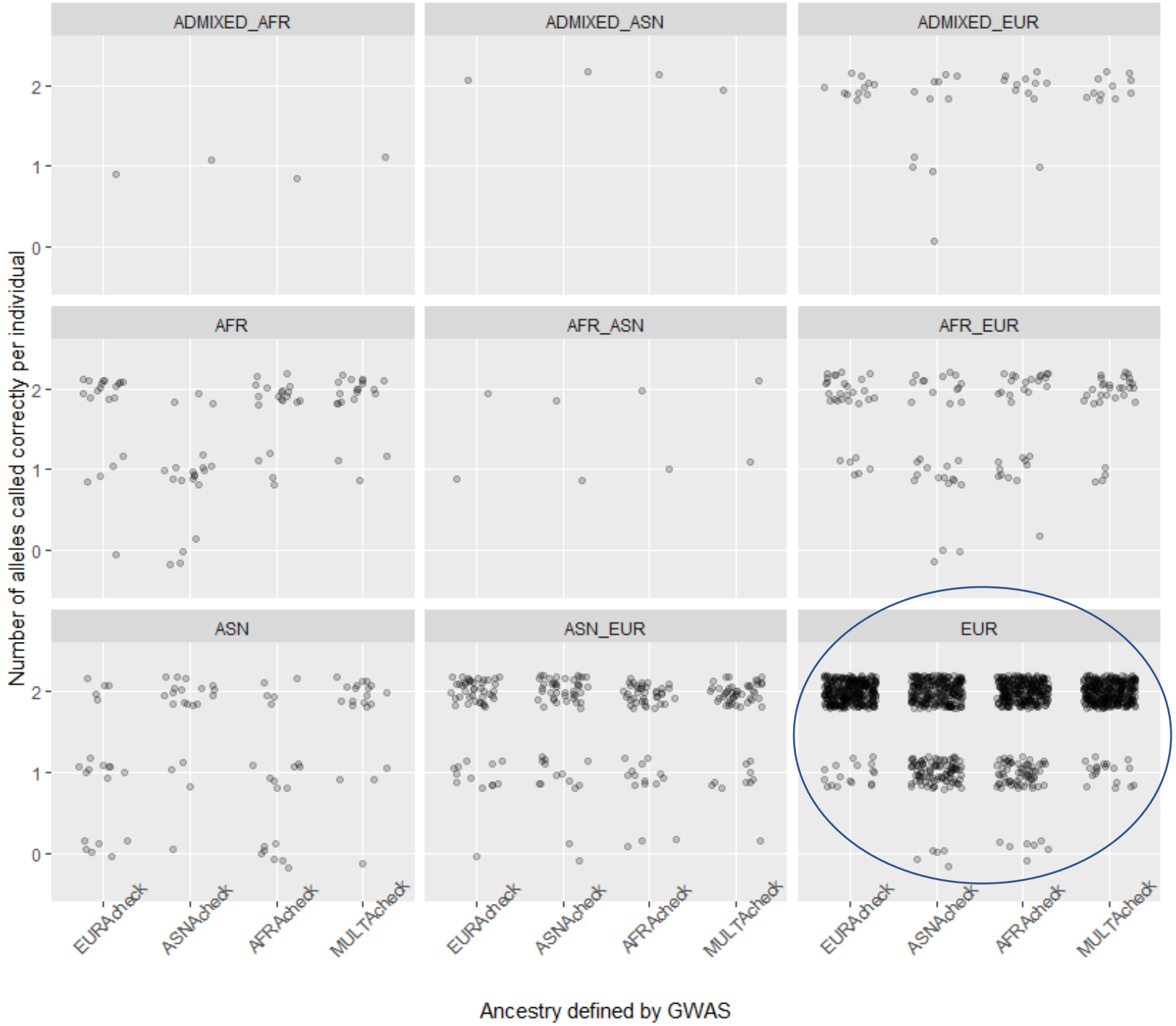




HLA: A locus
Imputation accuracy check

~500 subjects

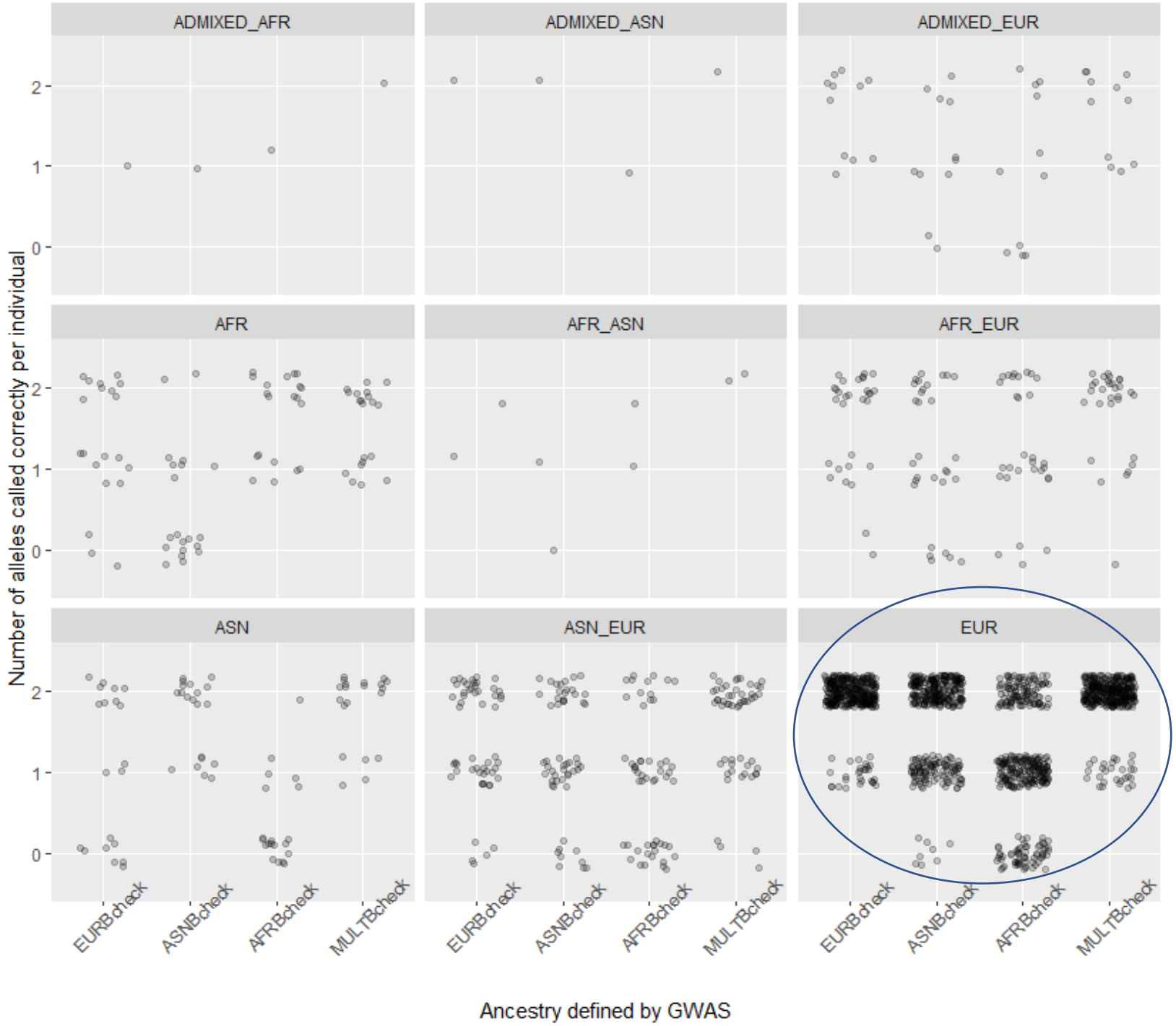
- HLA typing
- GSA genotyping
- Inferred ancestry
- Imputation accuracy



Courtesy: M Carrington lab

HLA: B locus
Imputation accuracy check

- B locus is more polymorphic
  - Decrease in imputation accuracy
- Imputation probability score is not predictive
- HLA typing may be important



Courtesy: M Carrington lab

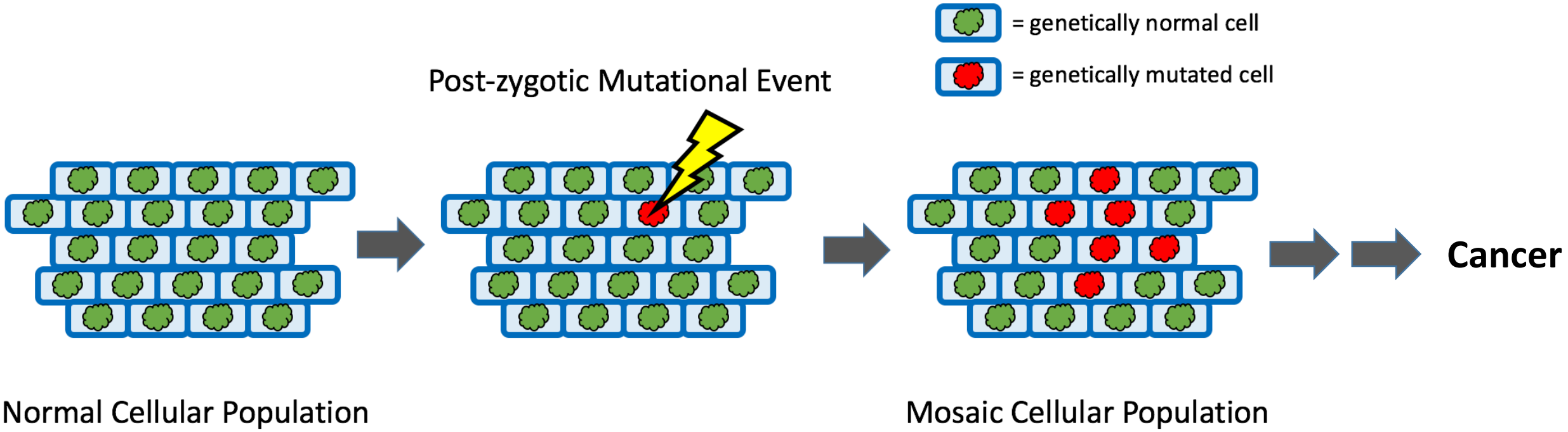
Analyses and preliminary findings: Clonal mosaicism

Mitch Machiela

Earl Stadtman Investigator, DCEG

Clonal Mosaicism

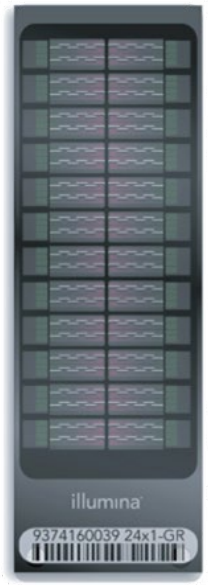
- The presence of an acquired somatic mutation(s) in a clonal subset of cells that differs from the inherited germline genome



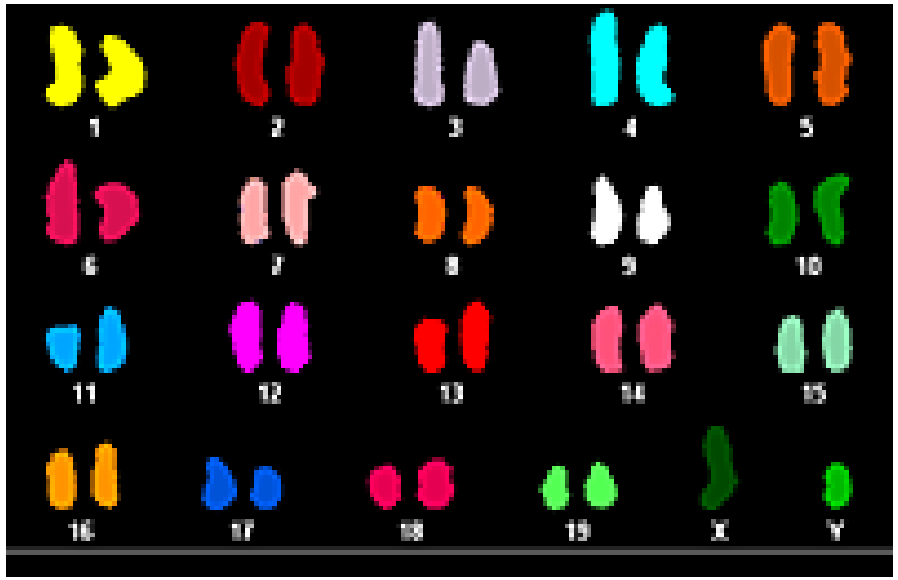
Machiela *Curr Opin Genet Dev* 2017

Detecting Mosaic Chromosomal Alterations

Create “virtual” karyotypes from existing SNP genotype array intensity data



High-density genotyping array

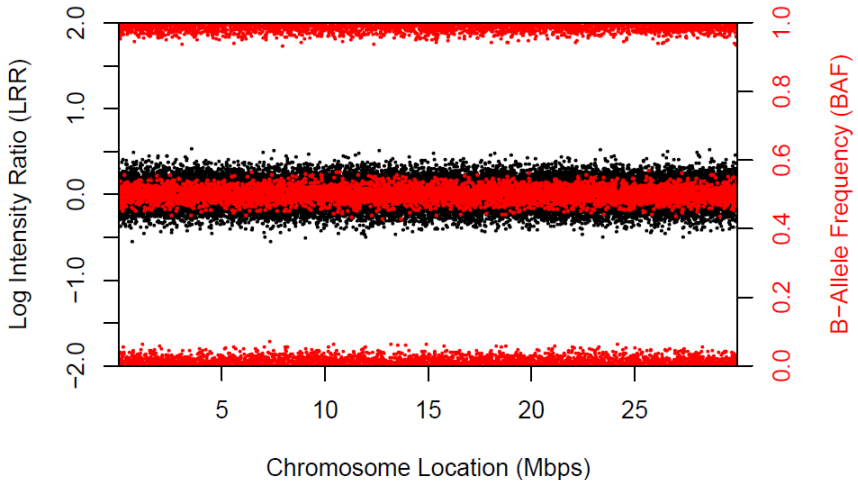


SKY/FISH Karyotype

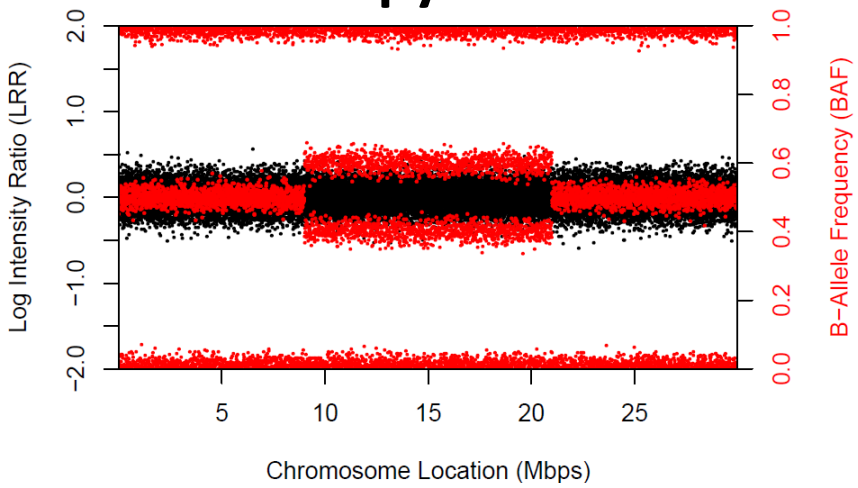
Rodríguez-Santiago *AJHG* 2010, Vattathil *AJHG* 2016, Loh *Nature 2019*

Detecting Mosaic Chromosomal Alterations

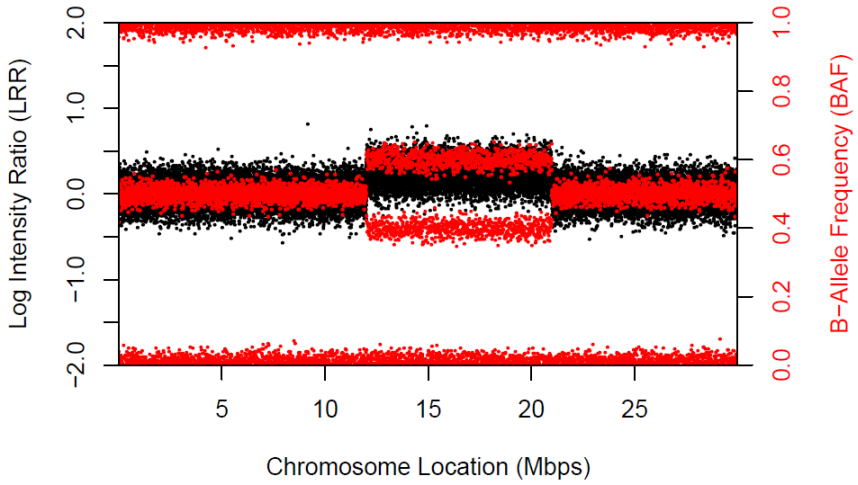
Normal



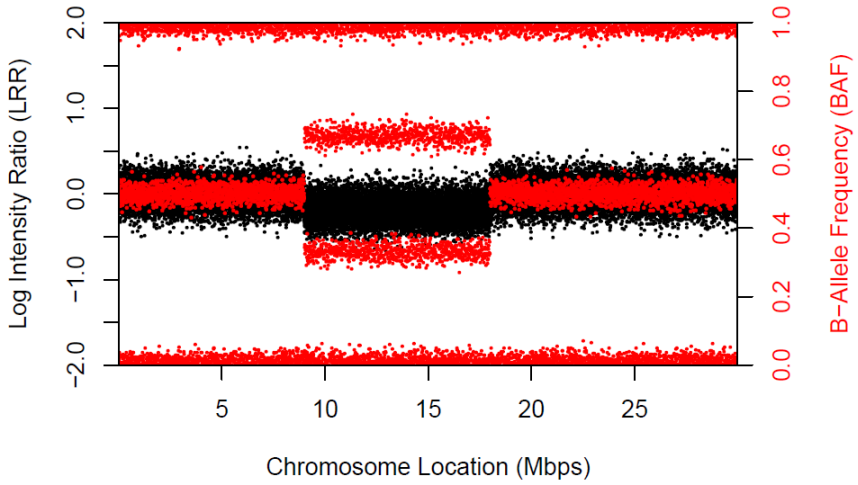
Mosaic Copy Neutral LOH



Mosaic Gain



Mosaic Loss



Mosaic Chromosomal Alterations in COVNET

We scanned all three current genotyped batches using MoChA

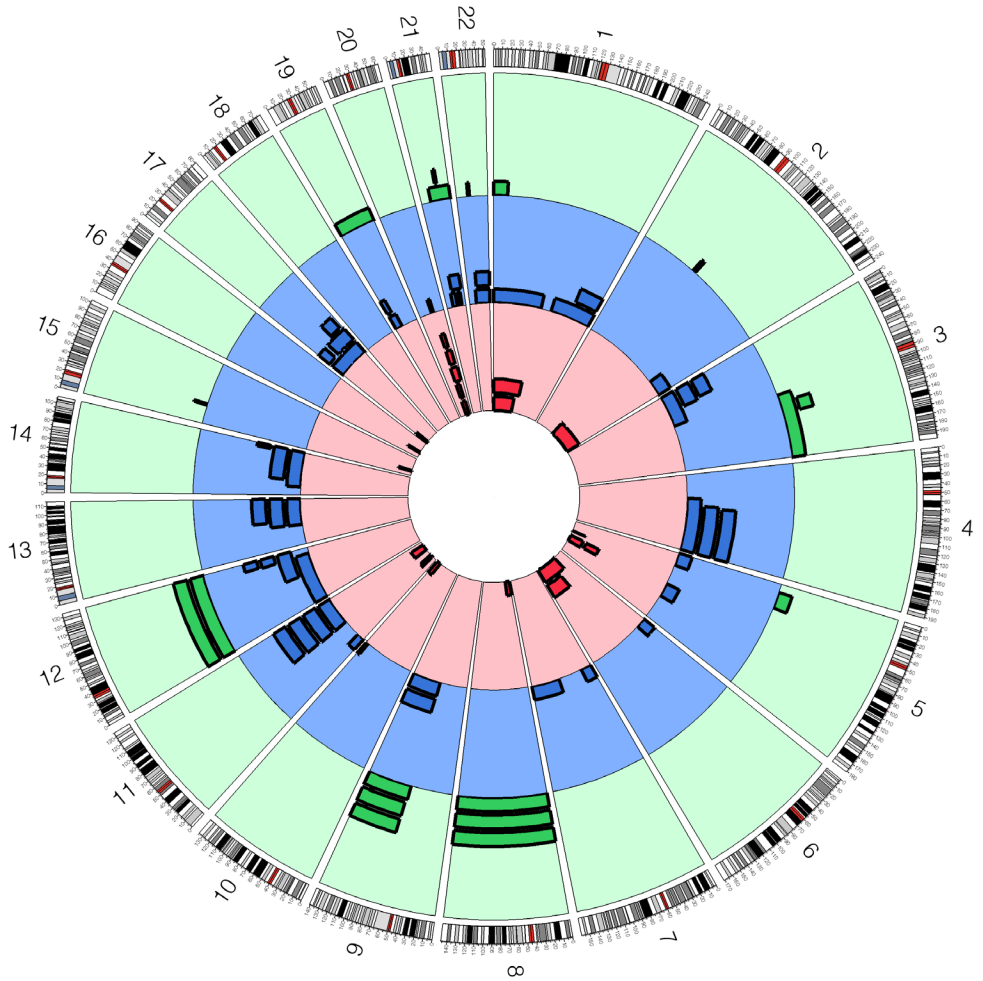
Analytic subject counts:

- N = 2,108 subjects in autosomal denominator file
- N = 1,209 subjects in female chromosome X denominator file
- N = 899 subjects in male chromosome Y PAR denominator file

Event counts:

- 84 autosomal events with >=2MB from 69 subjects (69/2108=3.3%)
- 30 female chrX events with >=2MB from 29 subjects (29/1209=2.4%)
- 68 male Y PAR events from 68 subjects (68/899=7.6%)

Detected COVNET Autosomal mCAs (N=84)



Detected COVNET Sex Chromosome mCAs

Female X Chromosome

Type	Event counts
GAIN	9
LOSS	15
NEUTRAL	5
Undetermined	1
Total	**30**

Male Y Chromosome

Type	Event counts
GAIN	19
LOSS	49
Total	**68**

Next steps

Create Working Groups

- GWAS Analyses
- NGS- Rare Variant Analyses
- Population Genetics & Admixture
- Susceptibility in well-studied settings (CDC & Ishgl)
- Mosaicism/Clonal Hematopoiesis

Next Major Steps

- Increase Accrual- target 40,000 for GWAS
- Select NGS/WGS for American Genome Center (USUHS)
  - 4-5,000 (COVNET)
    - Selection- extreme phenotypes, URM and population genetics
  - 2,000 (NCCAPS)
- Data sharing with dbGap/AnVil
- COVID Human Genetics Initiative

Important Take Home Points

- Critical need for providing *minimal phenotype data*- focus on acute COVID19 outcomes
  - Major impediment to analyses
- Potential for studying 'Long-COVID19'
  - Challenge of defining phenotype
  - NIH Strategic Initiative
- Large capacity for GWAS and NGS
  - Targets
    - 40,000 GWAS
    - 5,000 NGS/WGS
- Time to engage in different groups

QUESTIONS?

NATIONAL CANCER INSTITUTE

https://dceg.cancer.gov/research/how-we-study/genomic-studies/covnet

www.cancer.gov

www.cancer.gov/espanol