

COVID-19 risk alleles – susceptibility or severity?

Matti Pirinen

20.1.2022

Motivation (1/2)

- › Consider two GWAS on host phenotypes related to corona virus infection
- › **Infection** GWAS compares infected vs. population controls
 - Lab confirmed SARS-Cov-2 infection OR
 - Physician reported SARS-Cov-2 infection OR
 - Self reported SARS-Cov-2 infection
- › **Hospitalization** GWAS compares hospitalized for COVID-19 vs. population controls
 - Lab confirmed SARS-Cov-2 infection AND
 - Hospitalized for COVID-19

Motivation (2/2)

- › **Infection** GWAS compares infected vs. population controls
- › **Hospitalization** GWAS compares hospitalized for COVID-19 vs. population controls

ABO locus (rs505922-C)

DPP9 locus (rs2109069-A)

	OR	P-value	OR	P-value
Infection	1.10 (1.09-1.11)	4e-83	1.04 (1.03-1.05)	2e-13
Hospitalization	1.11 (1.09-1.13)	9e-23	1.12 (1.09-1.14)	2e-21

Is a variant associated with **susceptibility to infection** or **severity of disease**?

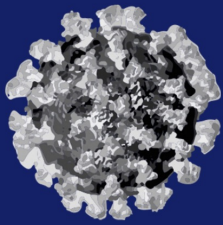
In this talk

ABO locus (rs505922-T)

DPP9 locus (rs2109069-A)

	OR	P-value	OR	P-value
Infection	0.91 (0.90-0.92)	4e-83	1.04 (1.03-1.05)	2e-13
Hospitalization	0.90 (0.88-0.92)	9e-23	1.12 (1.09-1.14)	2e-21

- › Why do we conclude that *ABO* is a *susceptibility* locus and *DPP9* is a *severity* locus?
- › How we can quantitatively evaluate to which group a locus belongs from GWAS data?

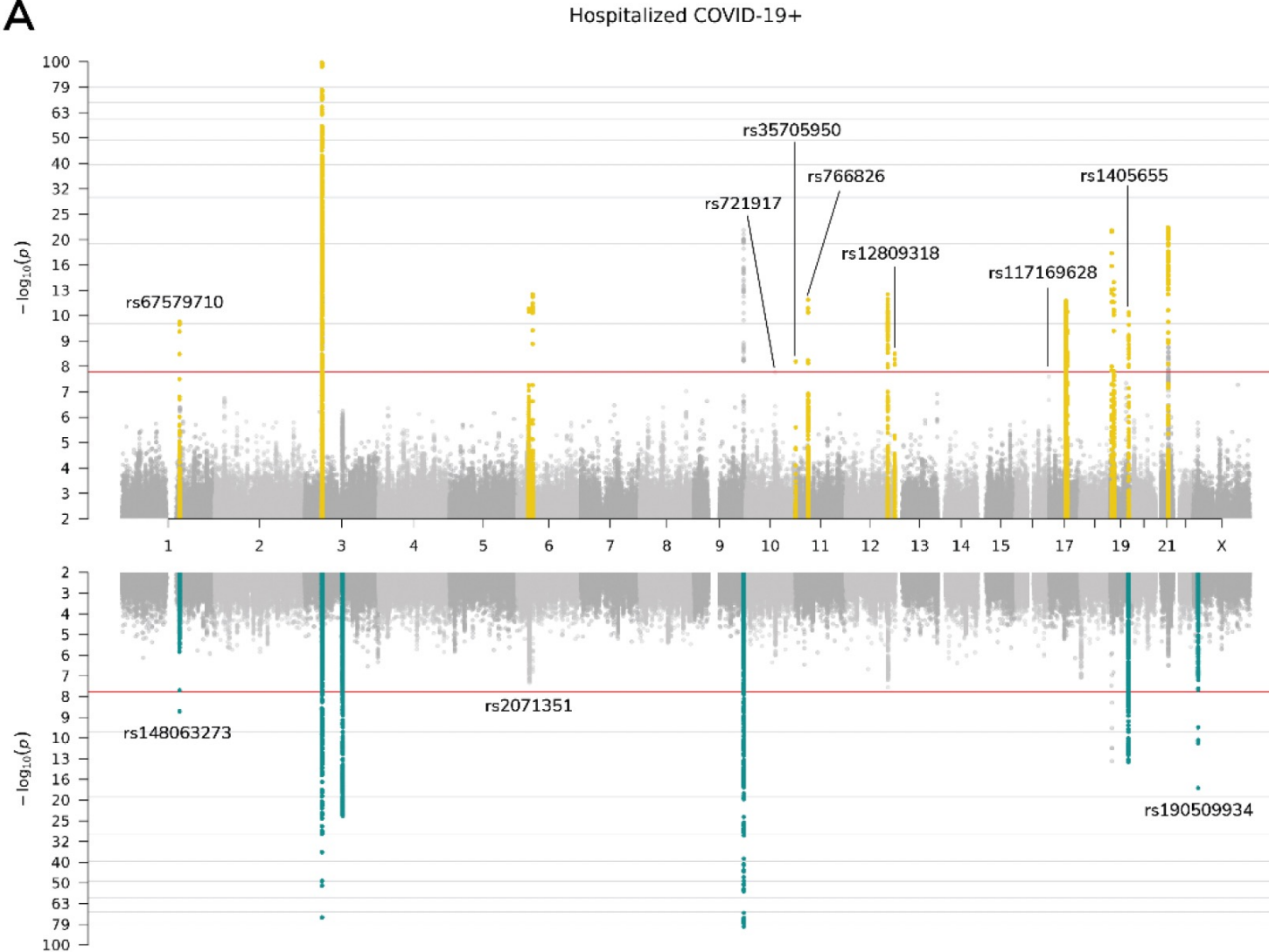


The COVID-19 Host Genetics Initiative

- › Generate, share, and analyze data to learn the genetic determinants of COVID-19 susceptibility, severity, and outcomes
- › Initiated by Andrea Ganna, Mark Daly and Ben Neale in March 2020
- › <https://www.covid19hg.org/>
- › Release 6 (June 2021)
 - 64 studies from 24 countries with over 3,300 researchers
 - Hospitalization GWAS of 23,988 hospitalized for COVID-19 vs. population controls (2,834,885)
 - Infection GWAS of 114,516 SARS-Cov-2 infected vs. population controls (2,138,237)
 - Manuscript in medRxiv

Release 6 results (medRxiv)

A



**Hospitalization
GWAS**

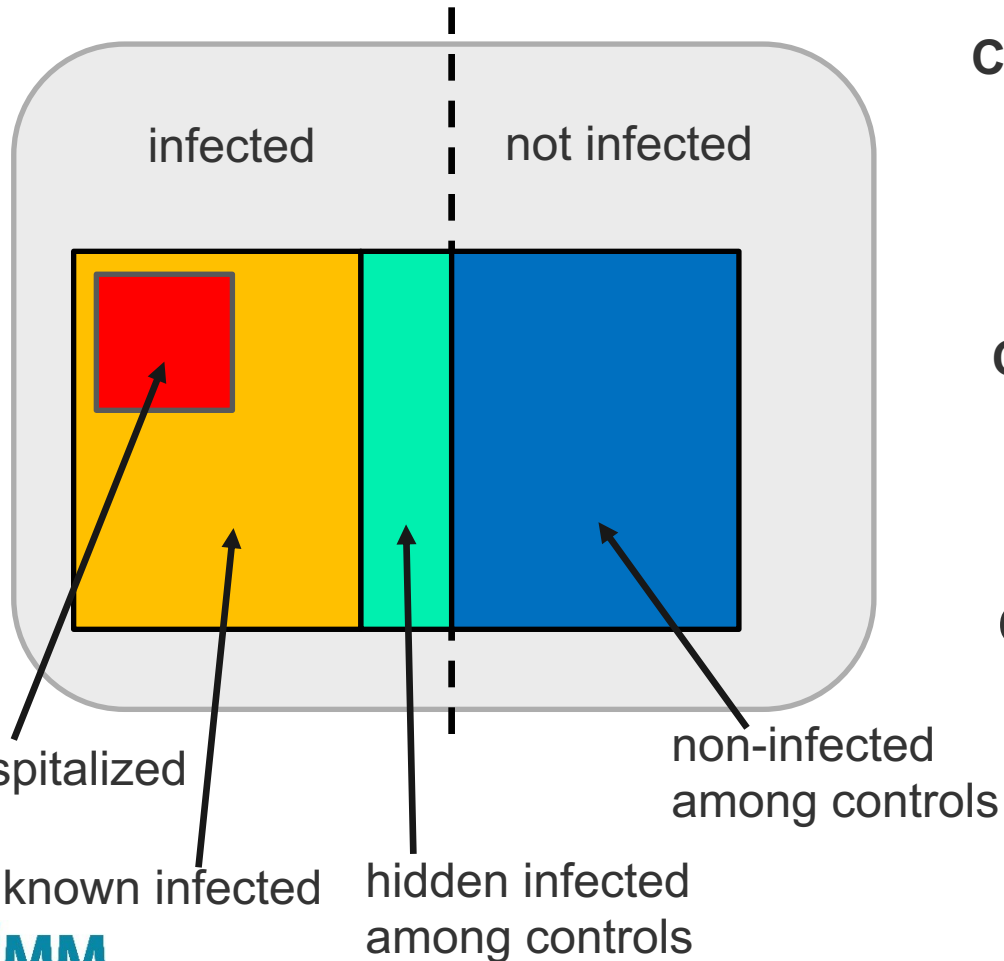
Together 23 loci
($P < 5e-8$)

**Infection
GWAS**

SARS-CoV-2 reported infection

Cases and controls

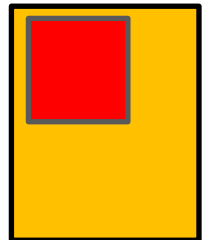
- Case groups are nested, controls are shared between the two GWAS



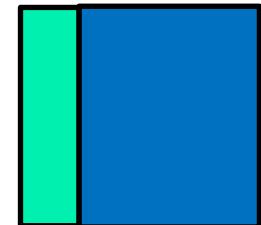
Cases in hospitalization GWAS



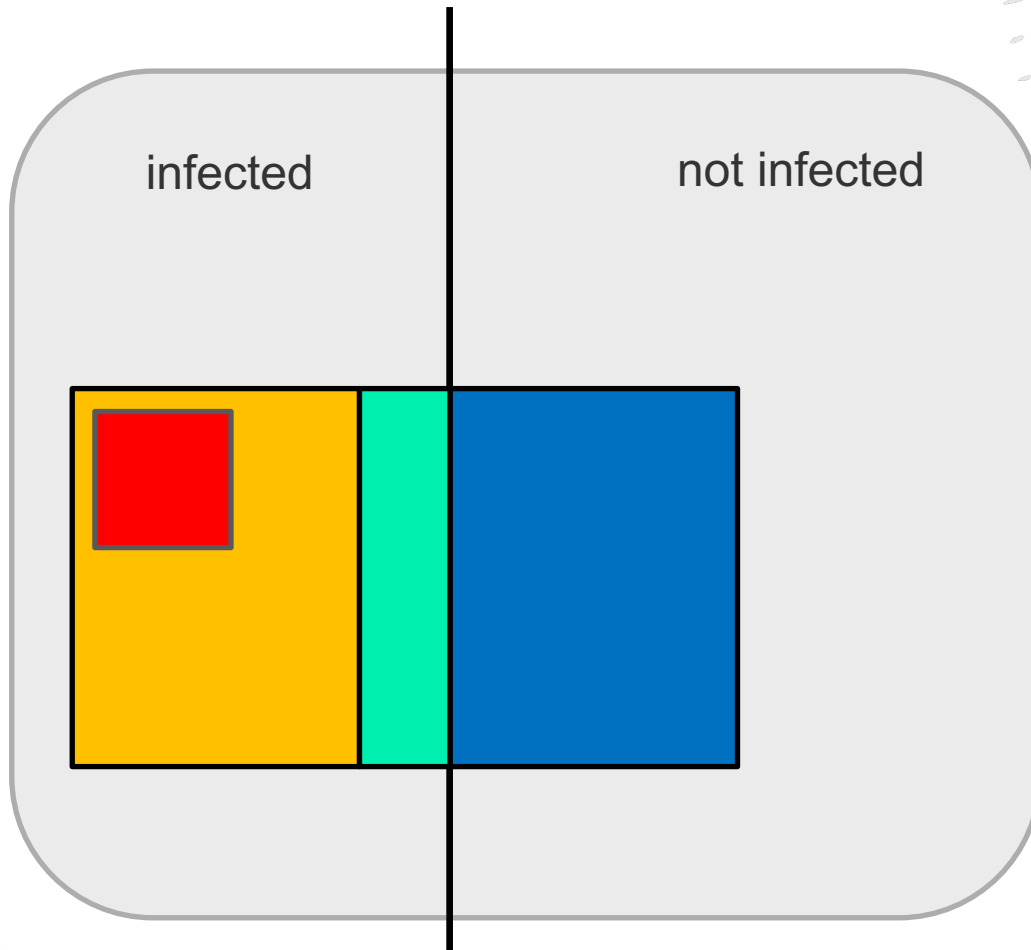
Cases in infection GWAS



Controls in both GWASes



Susceptibility variant

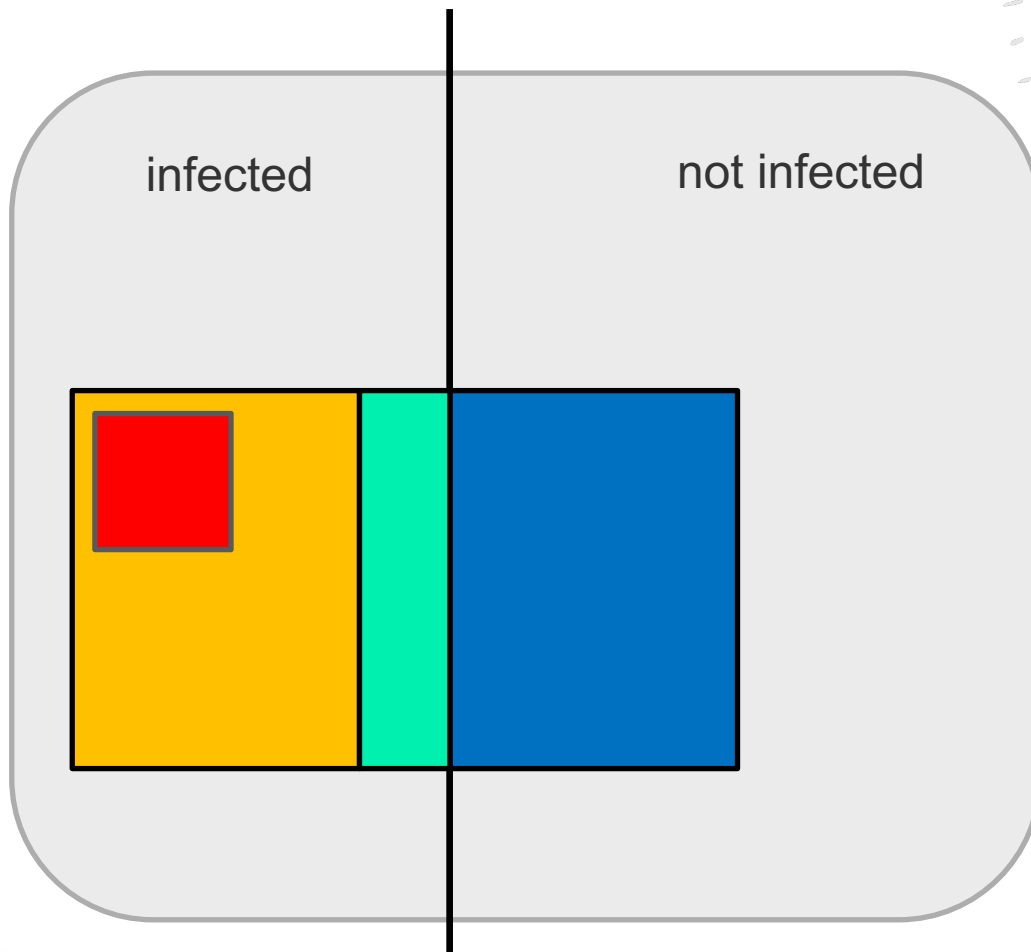


Higher frequency in infected than in non-infected but same in hospitalized cases as in all infected

$$\text{Red square} = \text{Yellow square} = \text{Cyan line} > \text{Blue square}$$

$$\beta_{INF} = \beta_{HOS} > 0$$

Severity variant (1/2)



Higher frequency among hospitalized than among other infected.

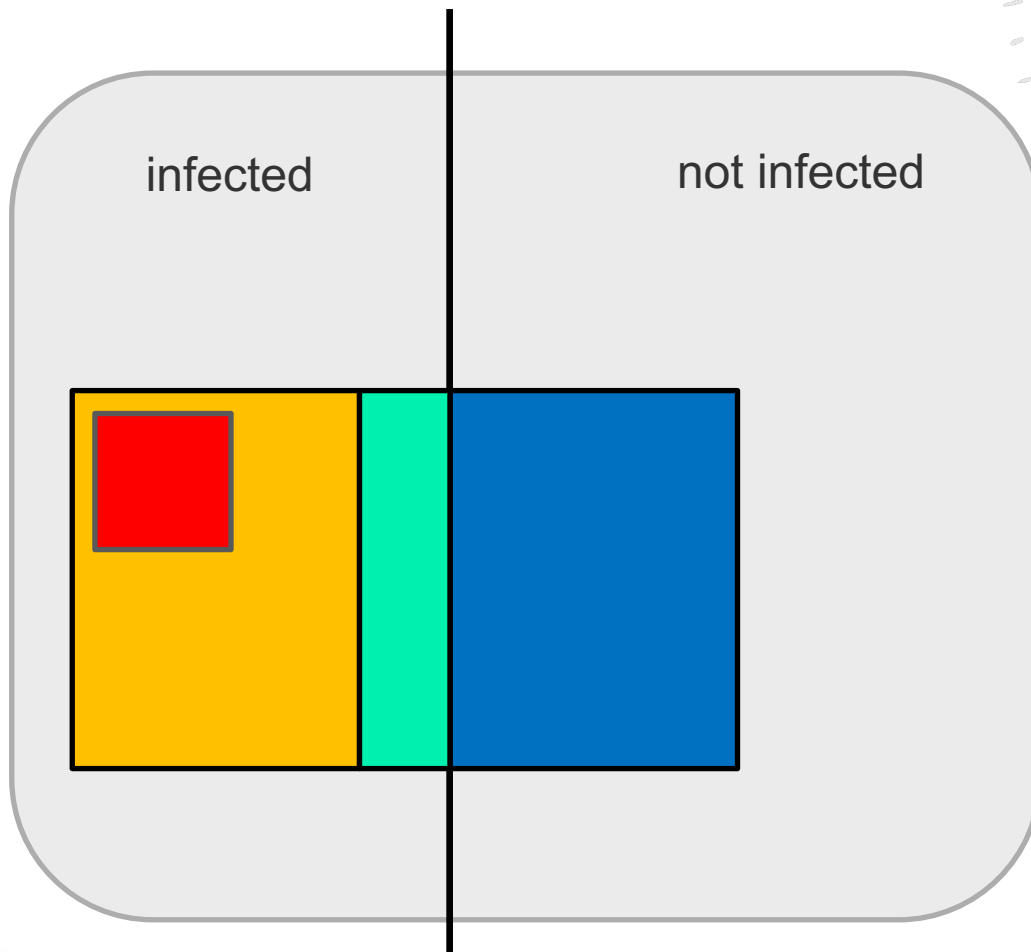
Same frequency among infected as among non-infected.

(Since hospitalized is a small part of all infected, non-hospitalized infected have approx. same frequency than non-infected.)

$$\text{Red square} > \text{Yellow square} \approx \text{Cyan square} \approx \text{Blue square}$$

$$\beta_{HOS} > \beta_{INF}$$

Severity variant (2/2)



If cases of infection GWAS were a random set of all infected, then, for a severity variant

$$\beta_{INF} = 0.$$

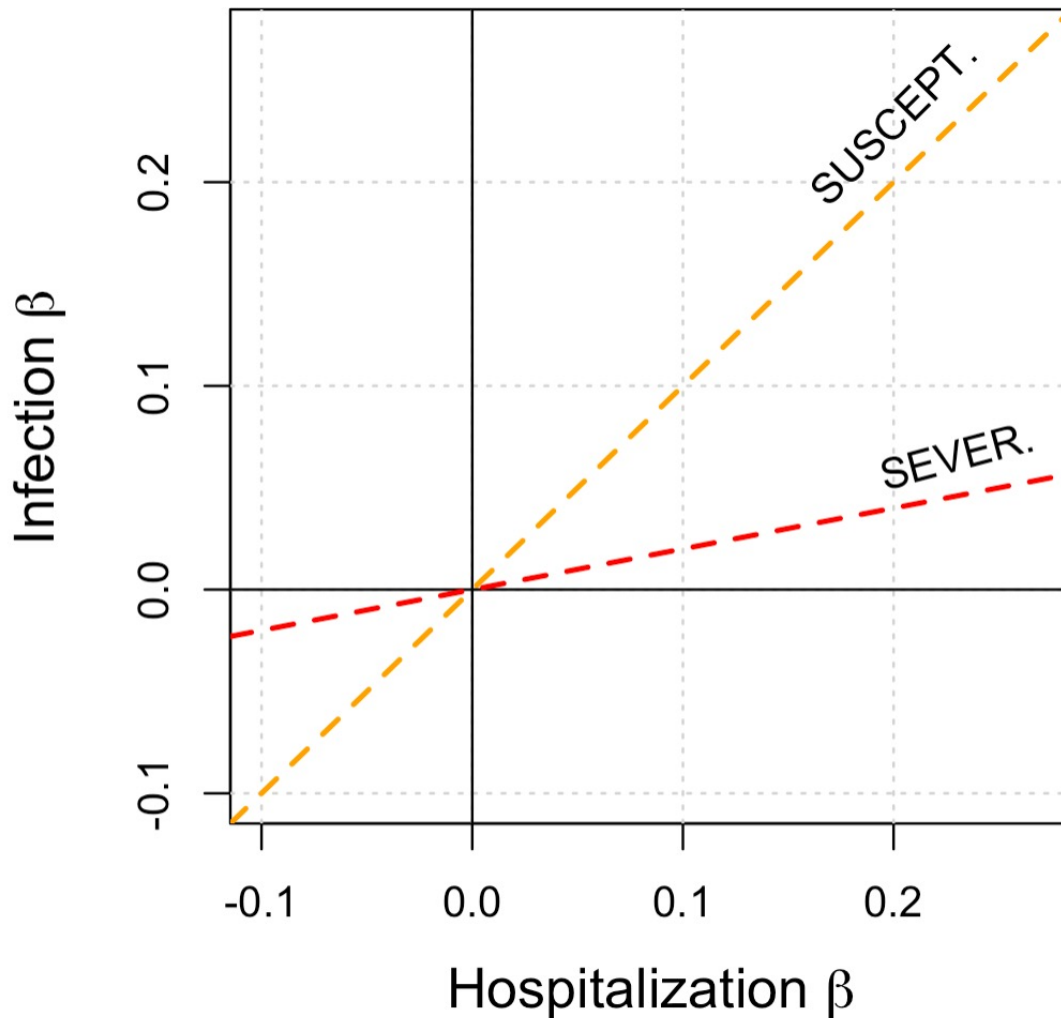
However, cases of infection GWAS is strongly enriched for hospitalized cases and therefore

$$\beta_{INF} > 0.$$

The weight of hospitalization GWAS in infection GWAS is 0.20.

$$\beta_{INF} \approx 0.2 \beta_{HOS} > 0$$

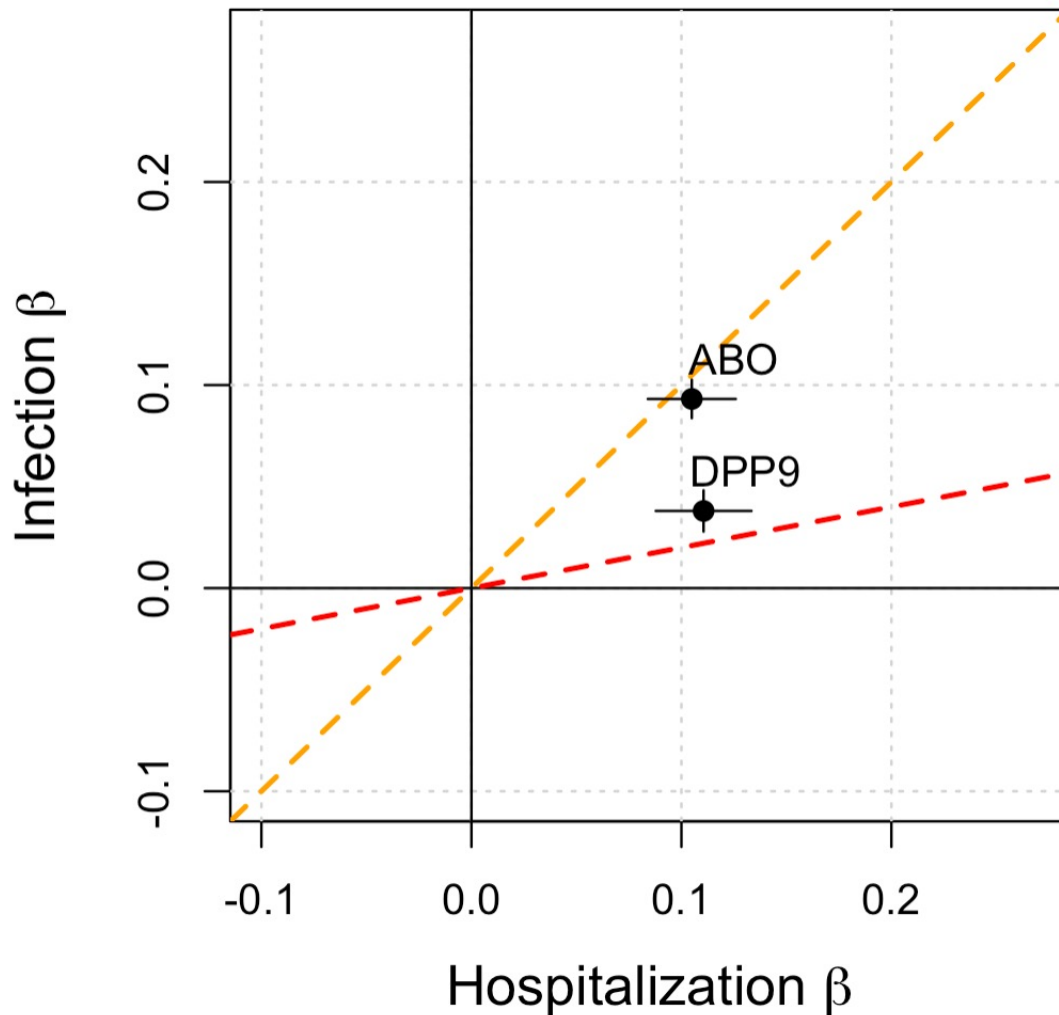
Expected GWAS effects



$$\beta_{INF} = \beta_{HOS} > 0$$

$$\beta_{INF} \approx 0.2 \beta_{HOS} > 0$$

Observed GWAS effects



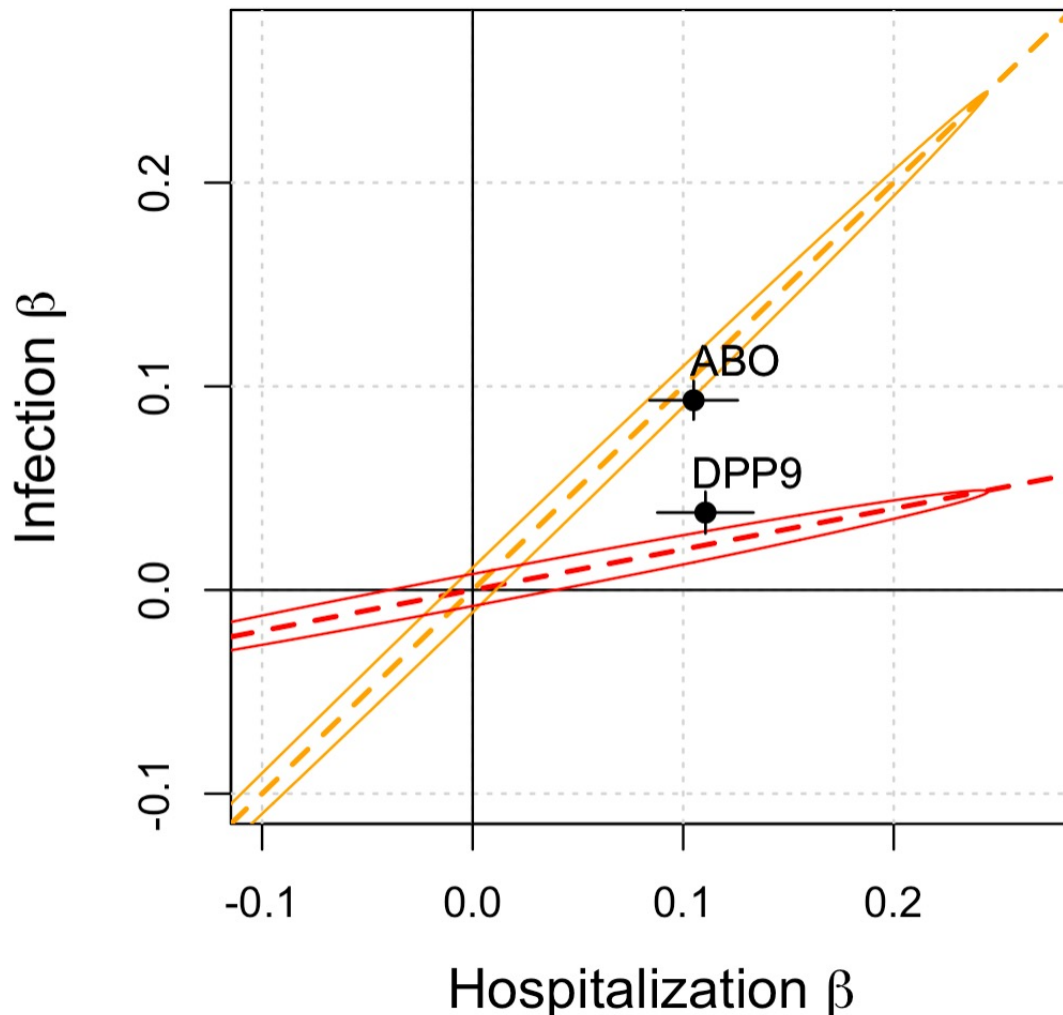
Susceptibility variants

$$\beta_{INF} = \beta_{HOS} > 0$$

Severity variants

$$\beta_{INF} \approx 0.2 \beta_{HOS} > 0$$

Models for GWAS effects



Susceptibility variants

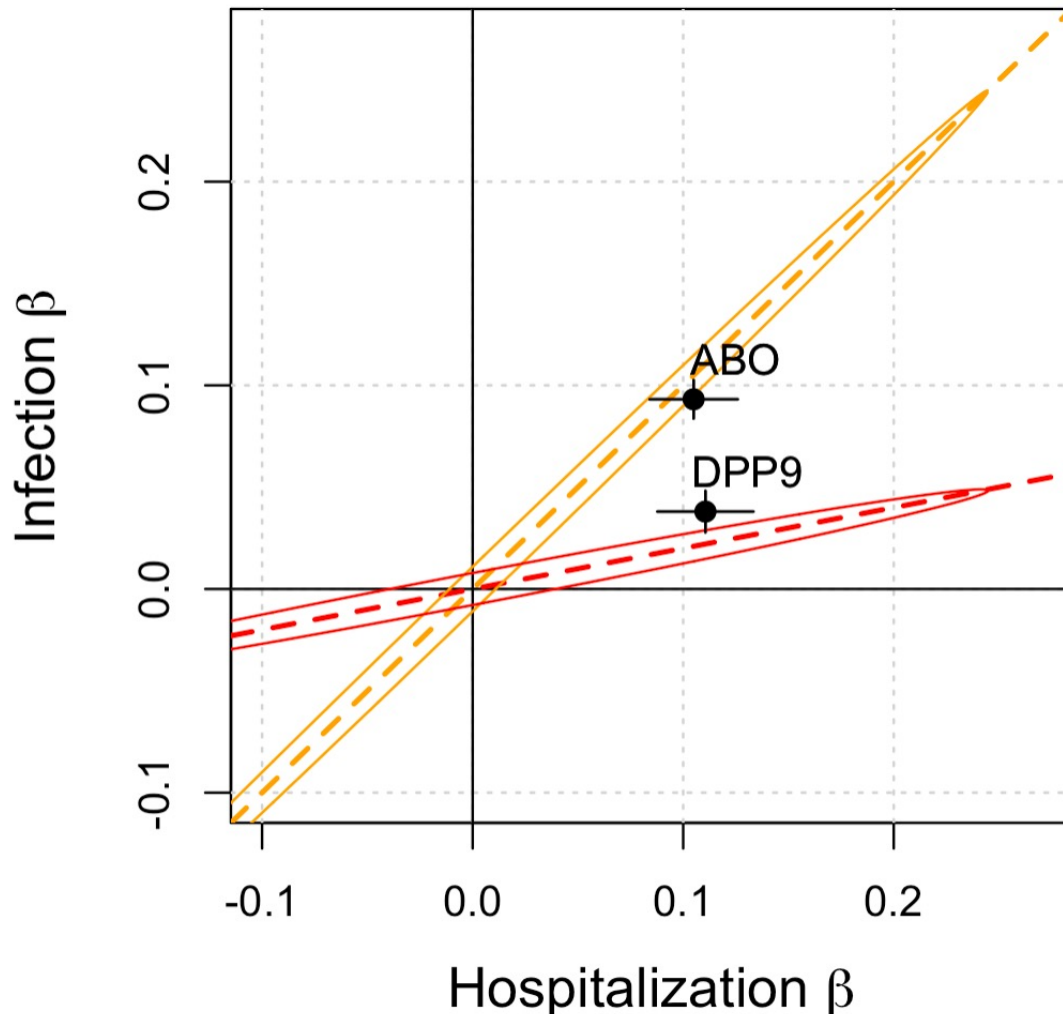
$$\beta_{INF} = \beta_{HOS} > 0$$

Severity variants

$$\beta_{INF} \approx 0.2 \beta_{HOS} > 0$$

We allow some deviation from theoretical relationships because of biases and approximations.

Which model is more probable?

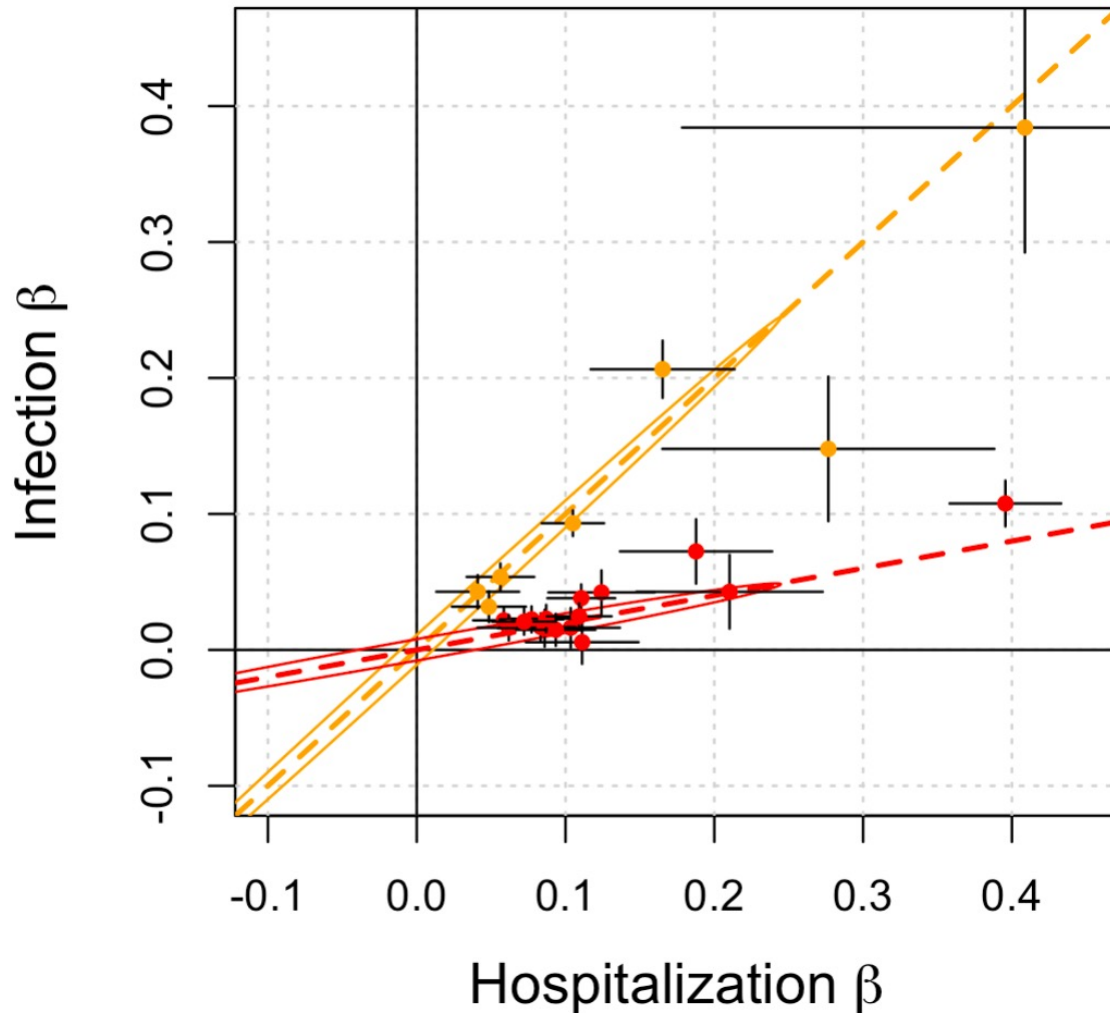


Bayes formula (with equal prior probabilities of models):

$$\frac{\Pr(\text{Sever.} | \hat{\beta})}{\Pr(\text{Suscep.} | \hat{\beta})} = \frac{\Pr(\hat{\beta} | \text{Sever.})}{\Pr(\hat{\beta} | \text{Suscep.})}$$

	Sever.	Suscep.
<i>ABO</i>	1e-38	1
<i>DPP9</i>	0.999999994	6e-8

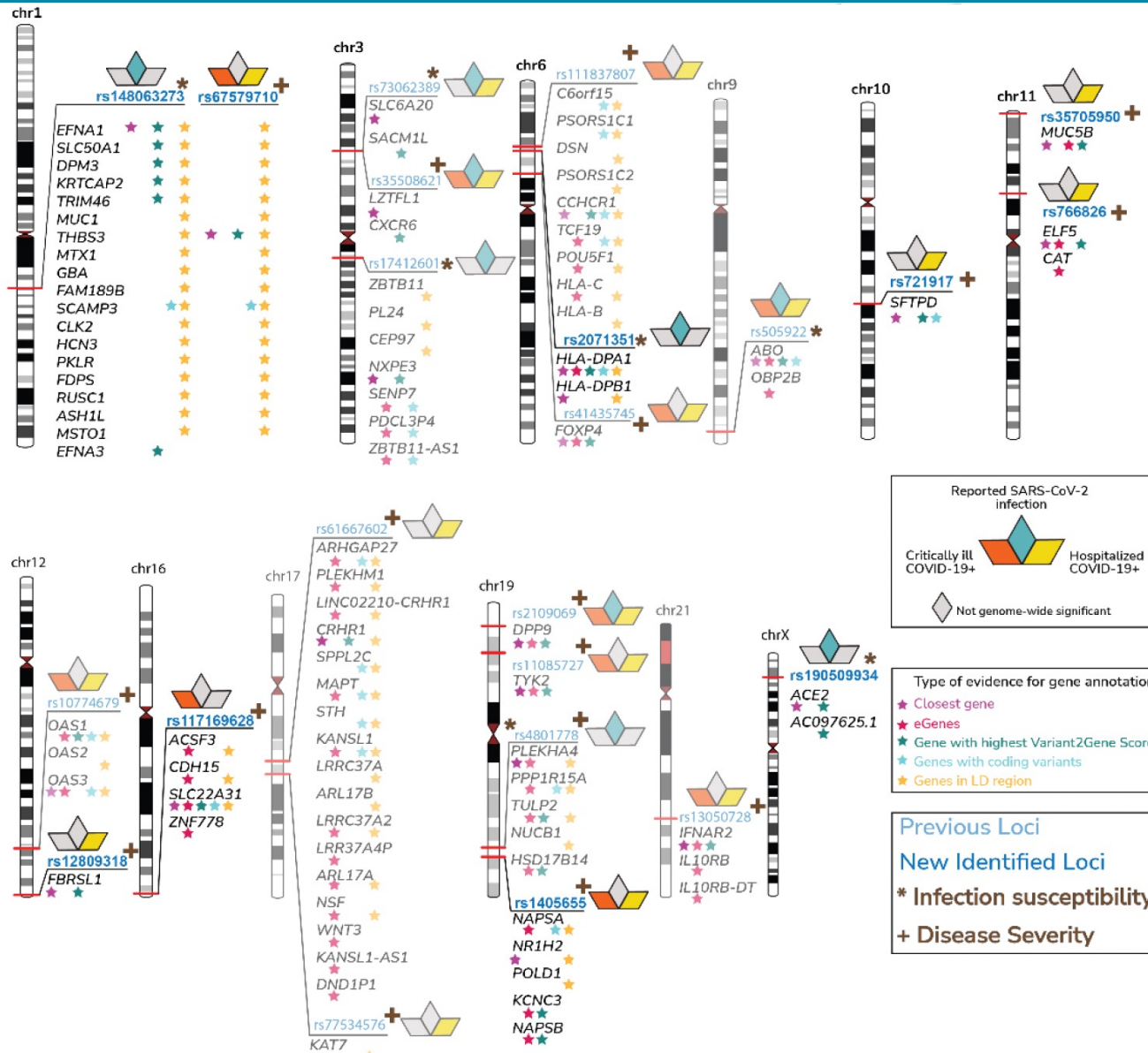
Which model is more probable?



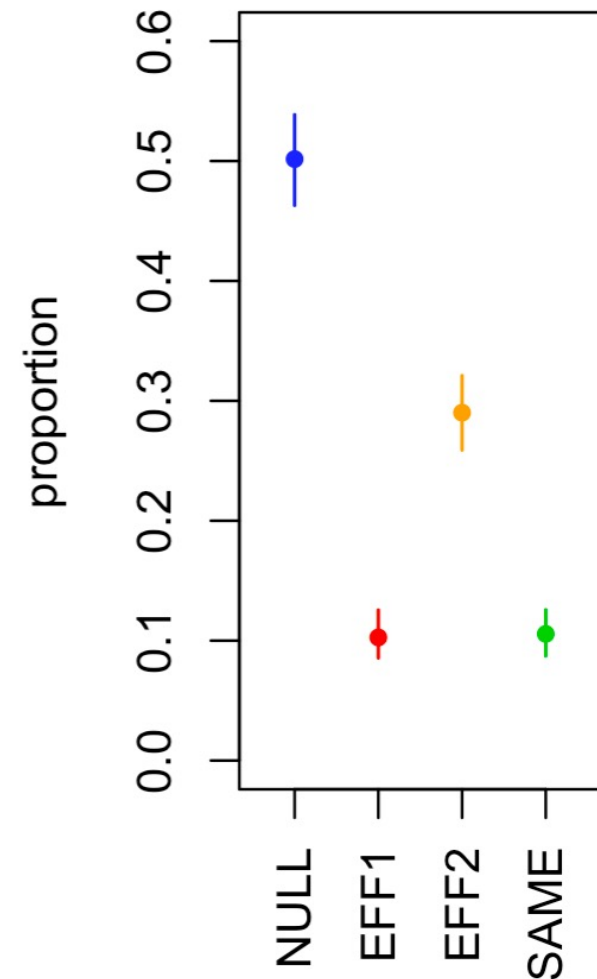
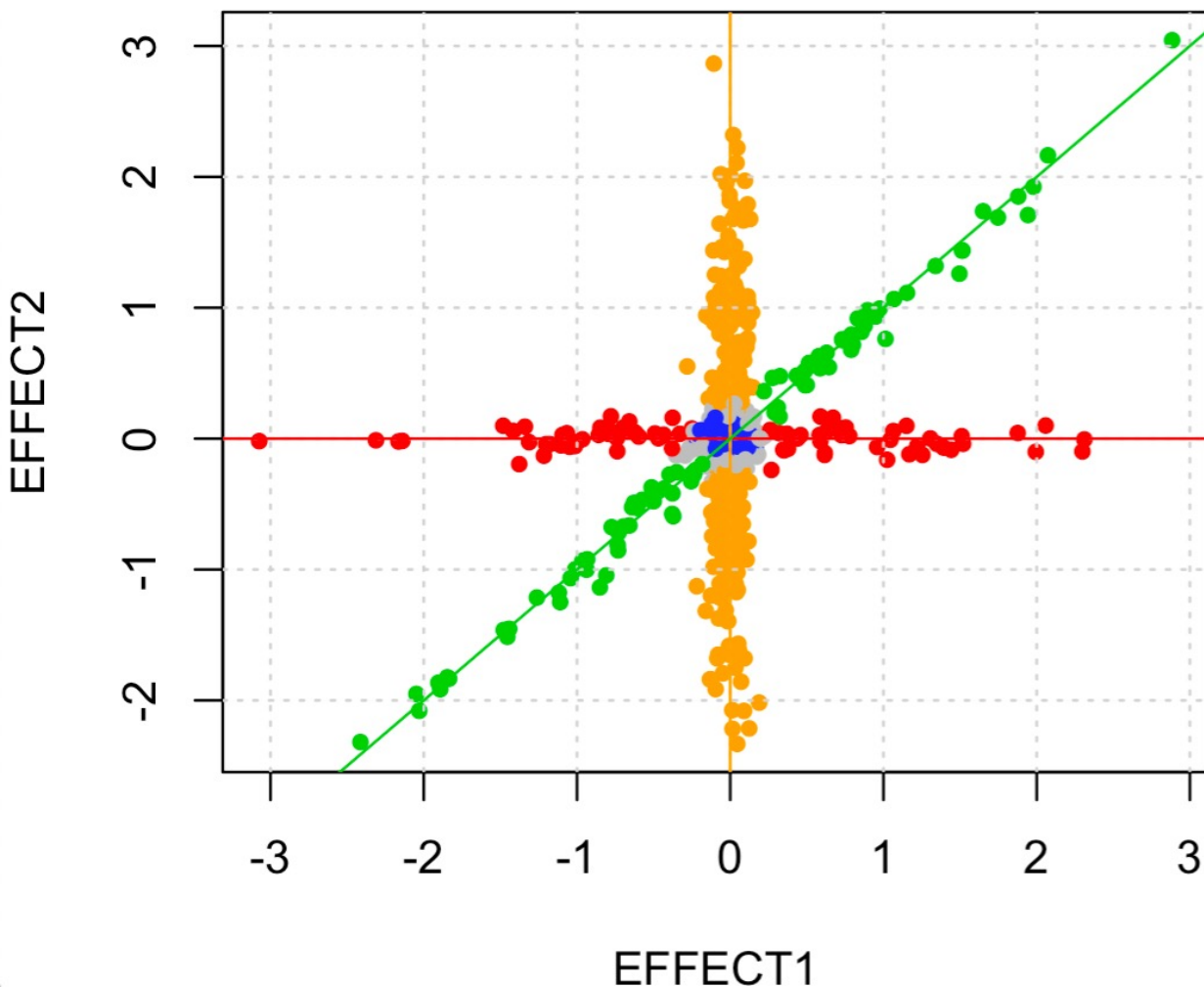
7 susceptibility loci
16 severity loci
(probability > 99%)

Code and documentation
in github / mjpipinen

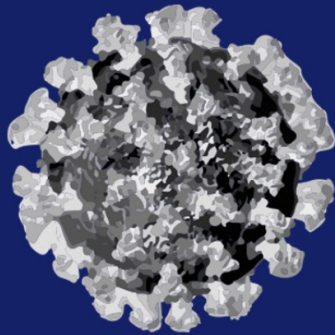
Release 6 results (medRxiv)



"Line Models" R function soon



Acknowledgements



The COVID-19 Host Genetics Initiative

Mark Daly
Andrea Ganna
Ben Neale
Mattia Cordioli
Gita Pathak
Juha Karjalainen