iScience, Volume *26*

Supplemental information

A GWAS in the pandemic epicenter highlights the severe COVID-19 risk locus introgressed by Neanderthals

Matteo Breno, Marina Noris, Nadia Rubis, Aneliya Ilieva Parvanova, Davide Martinetti, Sara Gamba, Lucia Liguori, Caterina Mele, Rossella Piras, Silvia Orisio, Elisabetta Valoti, Marta Alberti, Olimpia Diadei, Elena Bresin, Miriam Rigoldi, Silvia Prandini, Tiziano Gamba, Nadia Stucchi, Fabiola Carrara, Erica Daina, Ariela Benigni, Giuseppe Remuzzi, and on behalf of the ORIGIN study group

Supplementary Figures

Figure S1. Distribution of birth places of the 1195 analyzed subjects across the three groups, related to Table 1. Bergamo means born in the province of Bergamo; Lombardy means born in Lombardy, but outside of the province of Bergamo, Italy means born in Italy, but outside Lombardy; Other means born outside Italy.

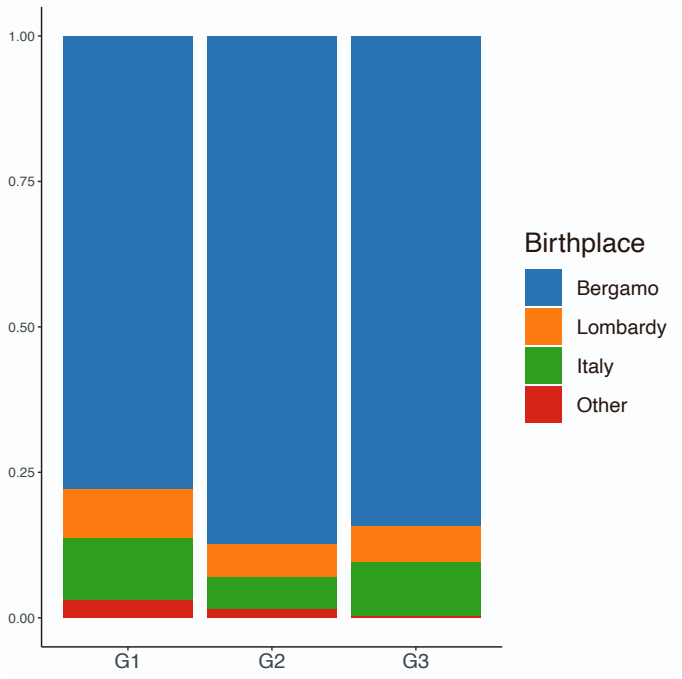

Figure S2. Distribution of the ageclasses of the 1195 analyzed subjects across the 3 groups, related to Table 1.

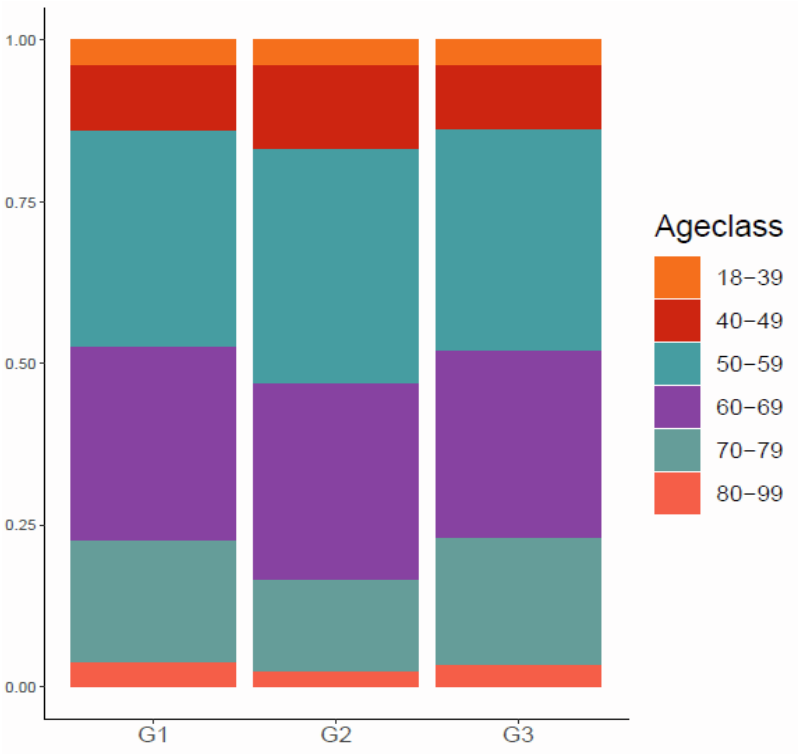

Figure S3: Distribution of risk factors of the 1195 analyzed subjects across the three groups, related to Table 1. Many subjects had more than one risk factor. They were classified according to the following: Cardiovascular (CV) > Diabetes (DB) > Hypertension (IP) > Overweight (SOV) > None (none).

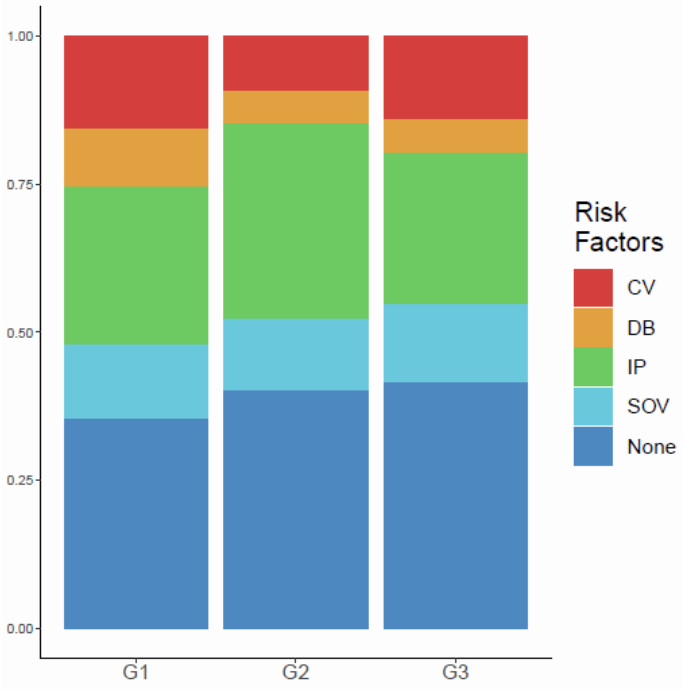

Figure S4. Conditional analysis of the Severity GWAS of the ORIGIN cohort, related to Figure 5. Each dot represents a marker. On the y-axis the minus log10 transformed p-values are shown. The top panel shows the results of the Severity analysis. The middle panel shows the same markers after conditioning on the lead variant. The bottom panel contains the gene annotations of the region. The horizontal blue bars represent the genome-wide significance threshold (P=5e-08) and the suggestive threshold (P=1e-5).

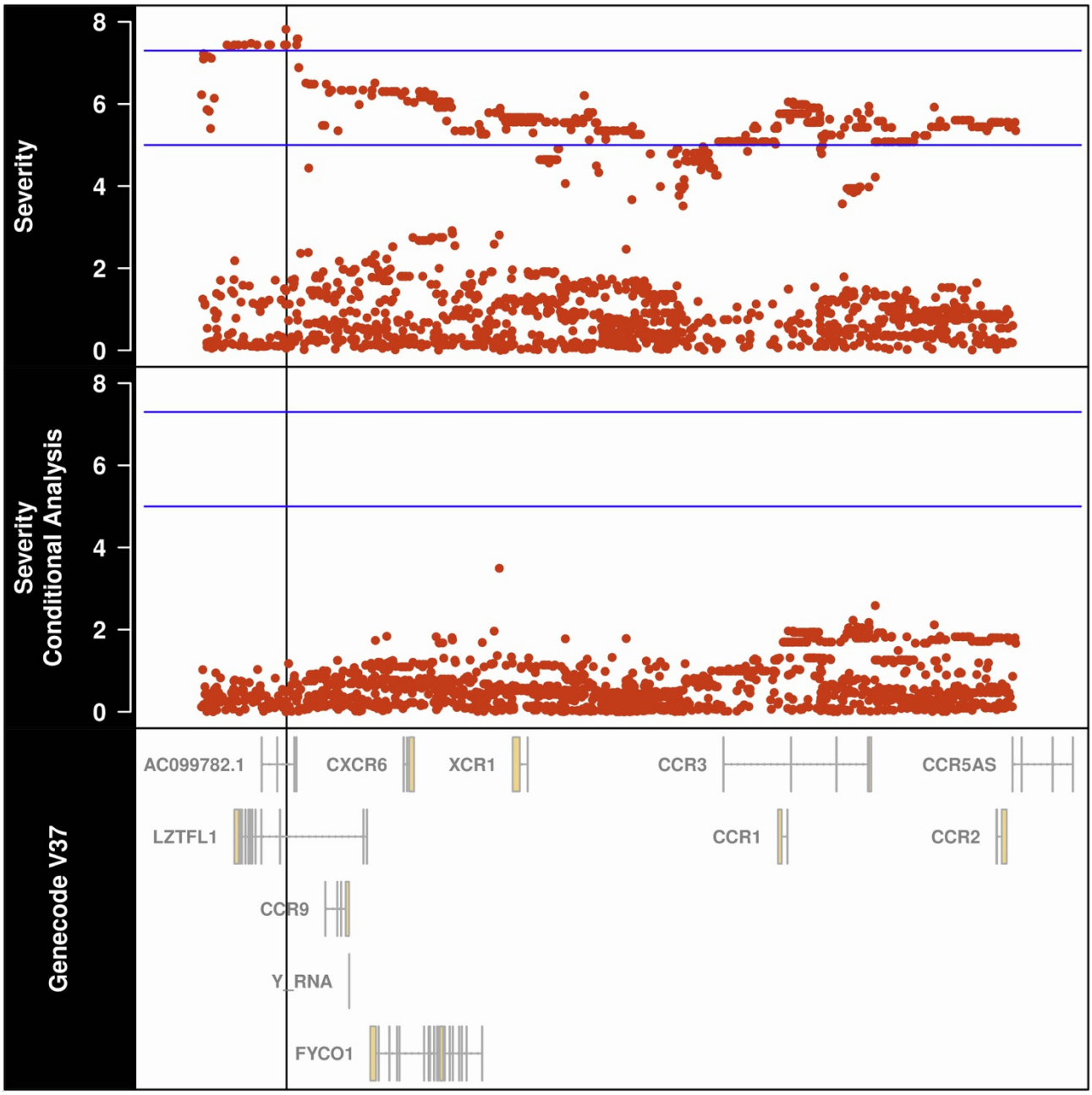

Figure S5. Distribution of the genotypes of the 3p21.31 lead variant in the severity analysis across age-classes and severity status, related to Figure 5. Individuals were divided among those younger and older than 60 years. Genotypes of chr3:45859142:G:C are represented compared to the reference genome (GRCh38): 0|0 is homozygous references (GG), 0|1(GC) is heterozygous and 1|1 is homozygous alternate allele (CC).

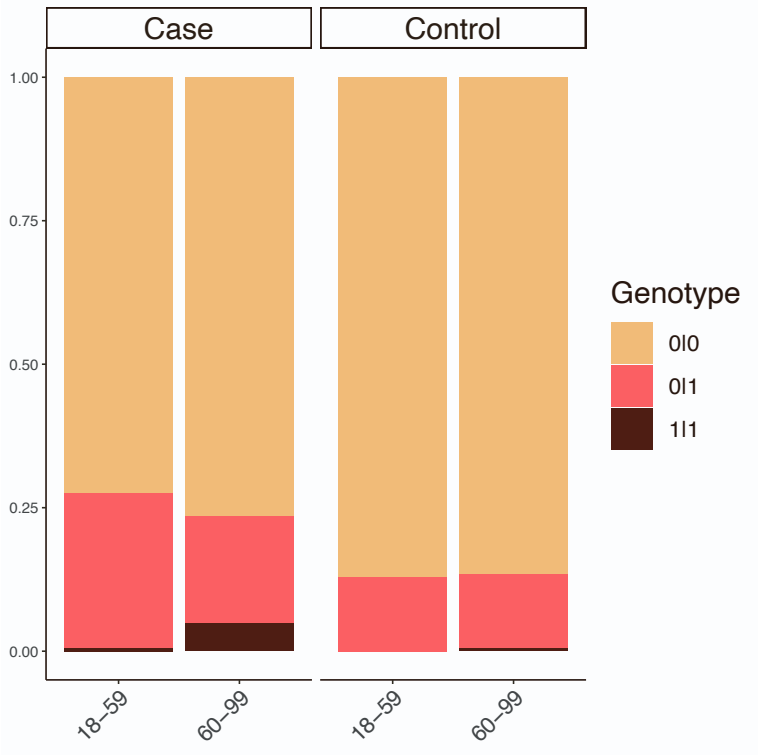

Figure S6. Distribution of the genotypes of the 3p21.31 lead variant in the severity analysis across gender and severity status, related to Figure 5. Genotypes of chr3:45859142:G:C are represented compared to the reference genome (GRCh38): 0|0 is homozygous references (GG), 0|1(GC) is heterozygous and 1|1 is homozygous alternate allele (CC).

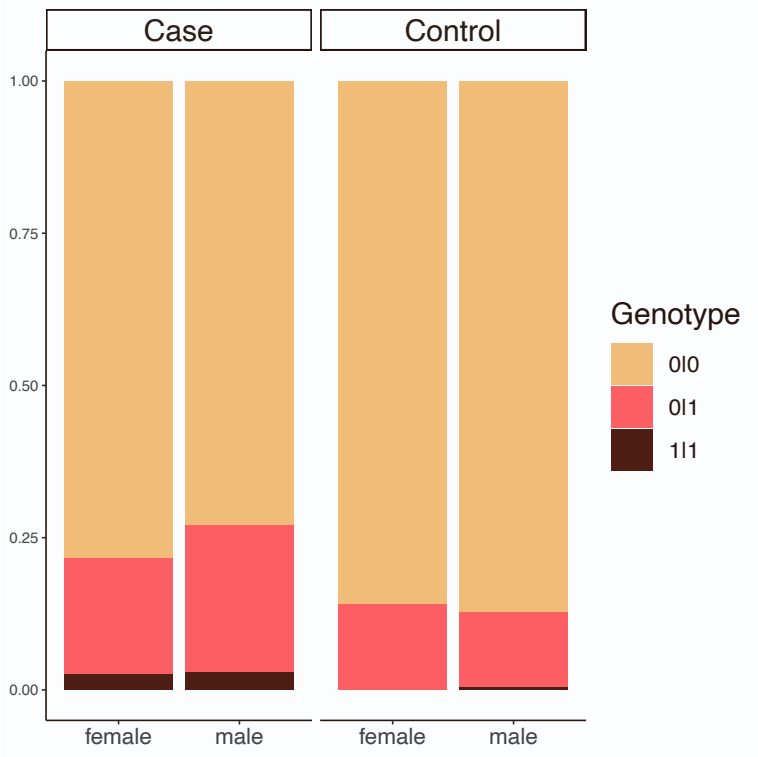

Figure S7. Comparison with COVID-19-HGI lead variants, related to Figure 6. Comparison of odds ratio (OR) between the ORIGIN cohort and the COVID-19 HGI lead variants for Severity (top panel) and Susceptibility (lower panel). Vertical lines depict the 95% confidence intervals of the ORIGIN estimates.