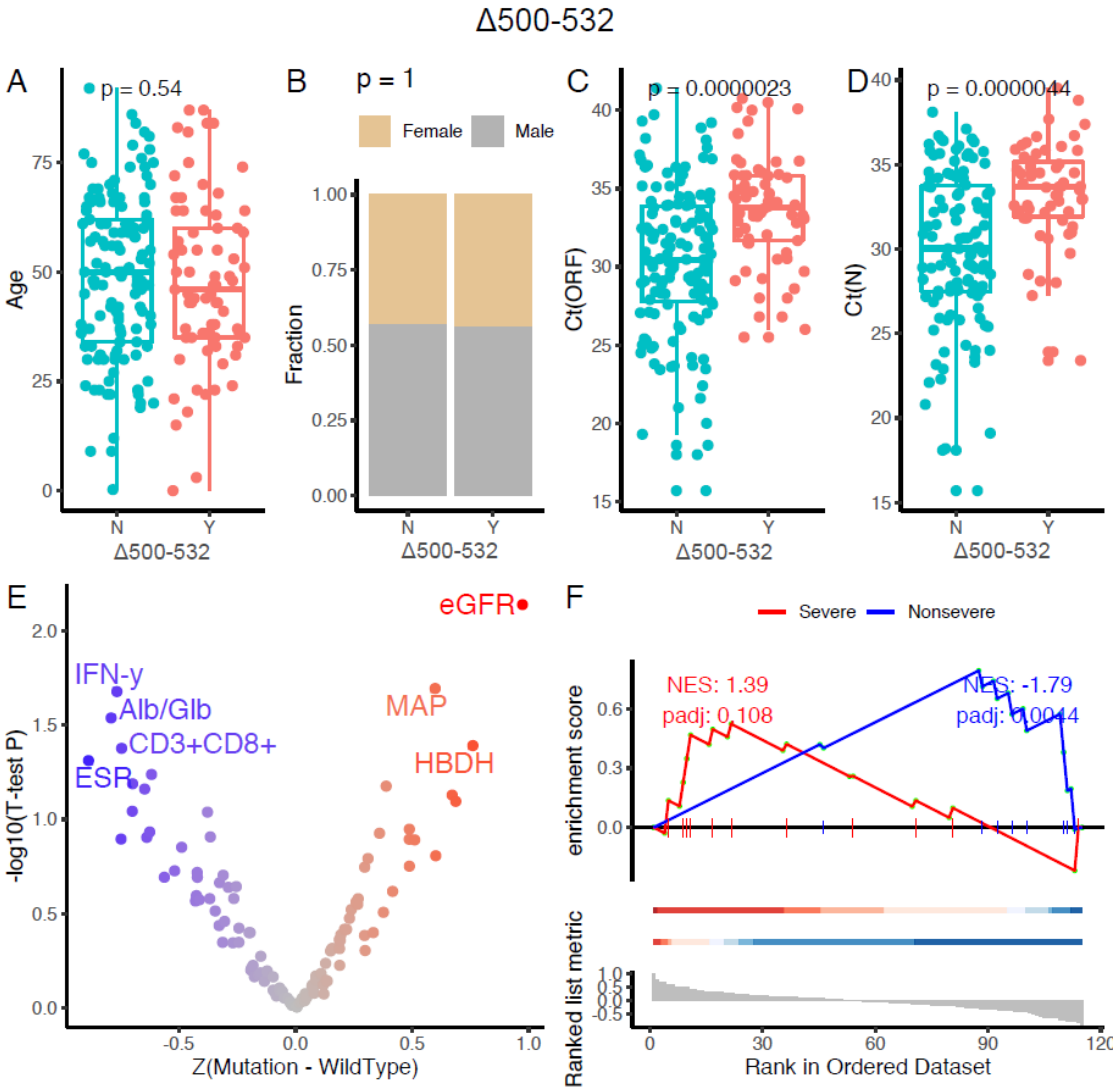
Data S2. Comparison between groups with or without a virus variant (related to Figure 3)

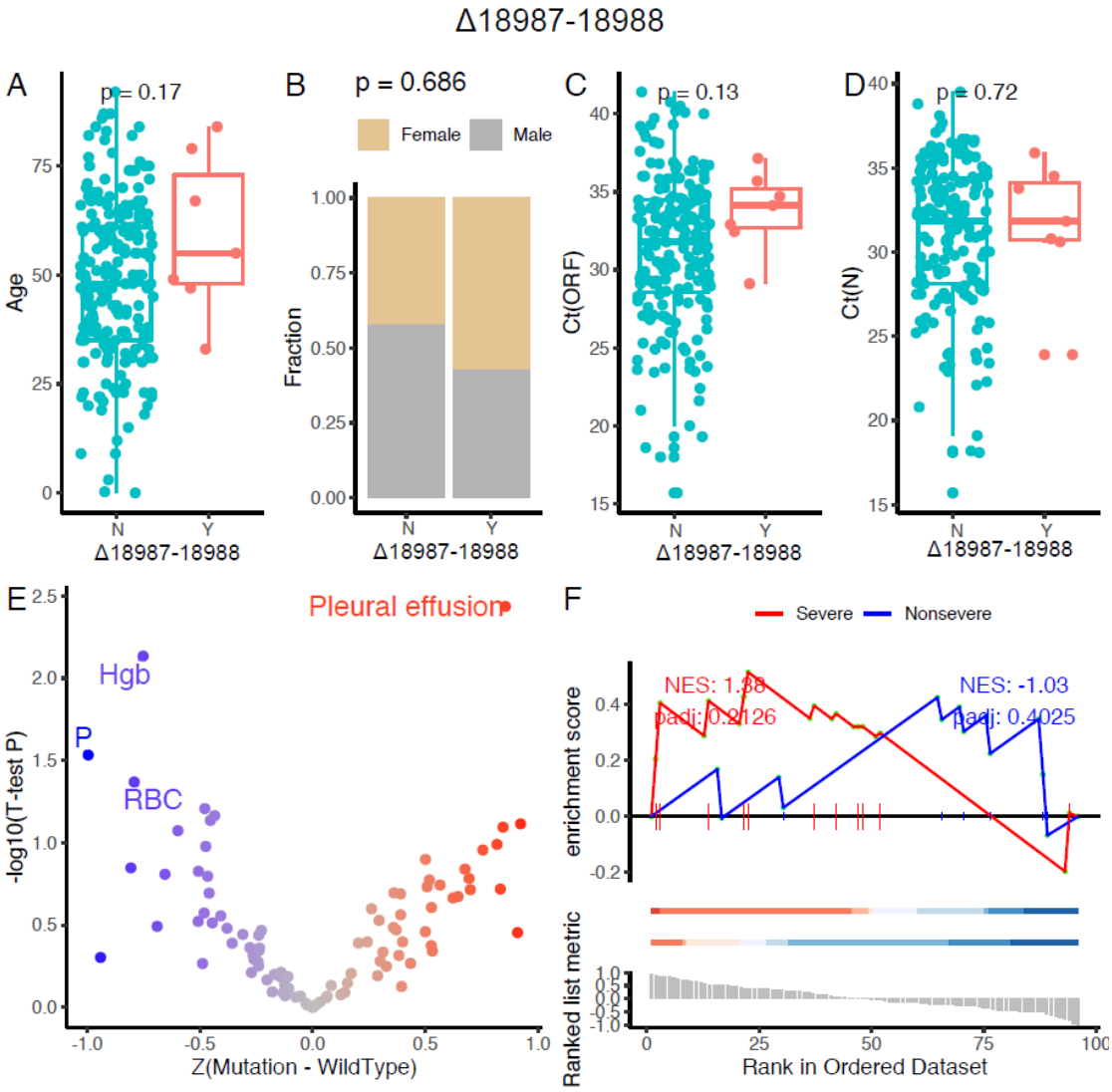
Contents

1. Δ500-532 in Nsp1



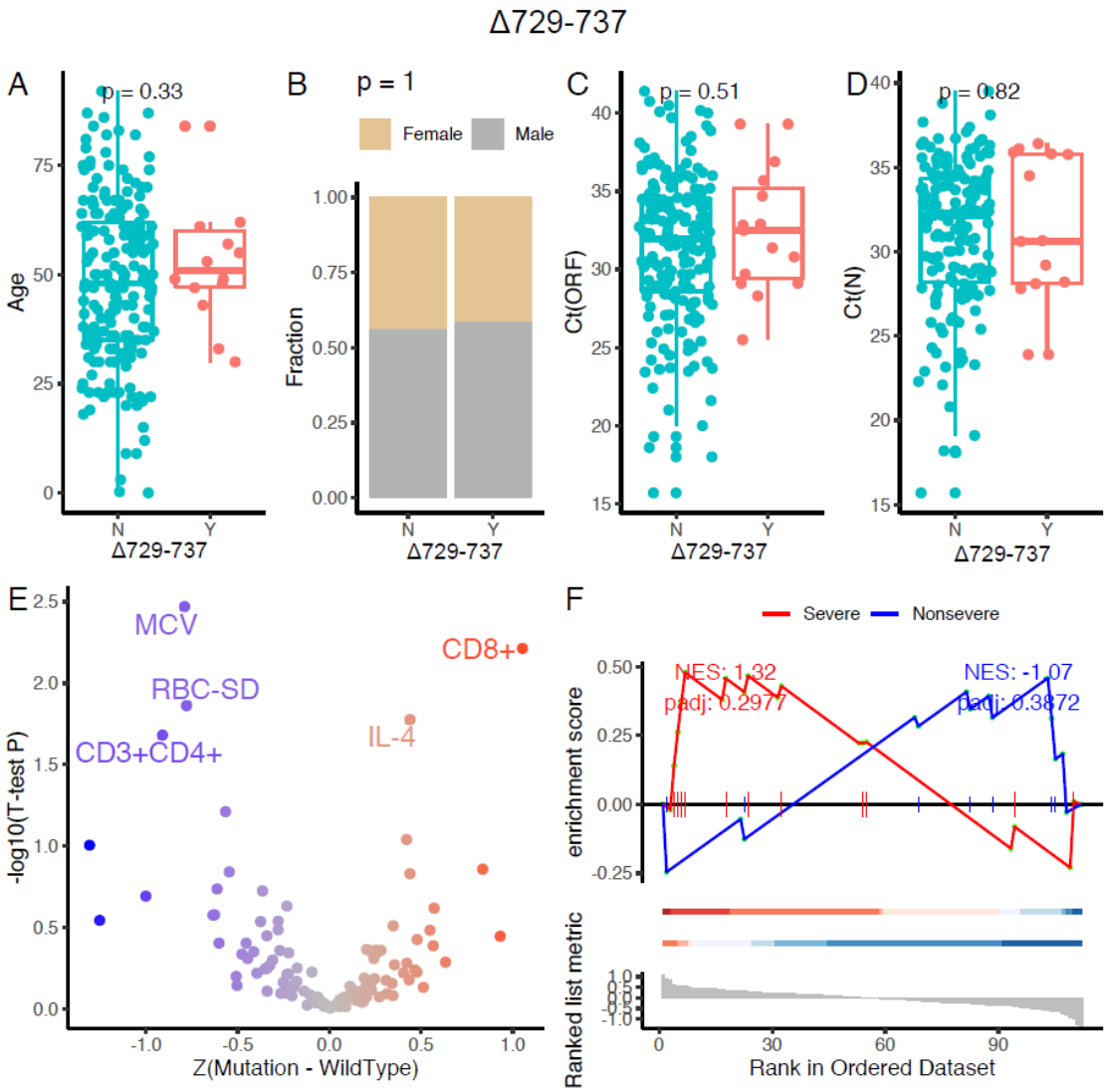
Δ500-532 is an in-frame deletion of Nsp1 (**A**$_{79}$PHGHVMVELV$_{89}$, predicted to have a moderate impact on the protein), which has been found in 23 samples (26.00%) in Sichuan and 39 samples (24.00%) in Wuhan, Hubei. The group with or without this deletion showed no difference in age (**A**, Wilcox-test, p = 0.54) and gender (**B**, chi-square test, p = 1). When comparing the Ct values of ORF and N genes based on qPCR, the group with Δ500-532 has a significant increase of Ct value in ORF (**C**, Wilcox-test, p = 2.3×10^{-6}) and N gene (**D**, Wilcox-test, p = 4.4×10^{-6}). We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of eGFR, MAP and HBDH, and has lower levels of ESR, IFN-γ and $CD3^+$ $CD8^+$ T cell counts in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the Δ500-532 has a –1.79 NES in non-severe traits (padj = 4.4×10^{-3}) and a 1.39 NES in severe traits (padj = 0.108, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

2. Δ18987-18988 in Nsp14



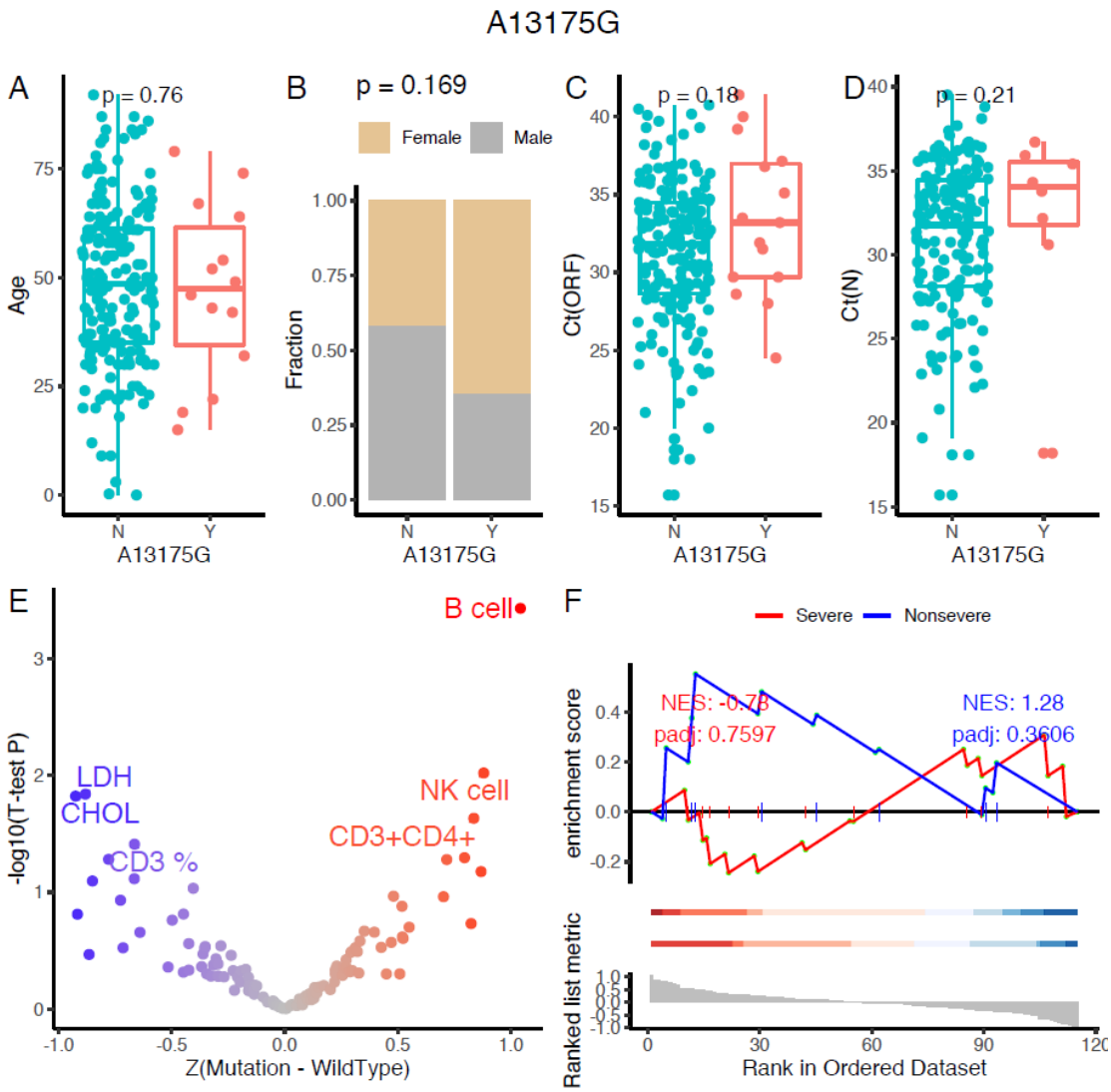
Δ18987-18988 is a frameshift variant of Nsp14 (QHD43415.1:p.6241-6242VV>VX, predicted to have high impact on the protein), which has been found in 4 samples (5.00%) in Sichuan. The group with or without this deletion showed no difference in age (**A**, Wilcox-test, p = 0.17), gender (**B**, chi-square test, p = 0.686), Ct values of ORF (**C**, Wilcox-test, p = 0.13) and N genes (**D**, Wilcox-test, p = 0.72) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group without the deletion has higher levels of Hgb, P and RBC in valcano plot, while the group with the variant has a higher level of pleural effusion (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the Δ18987-18988 has a –1.03 NES in non-severe traits (padj = 0.4025) and a 1.38 NES in severe traits (padj = 0.2126, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

3. Δ729-737 in Nsp1



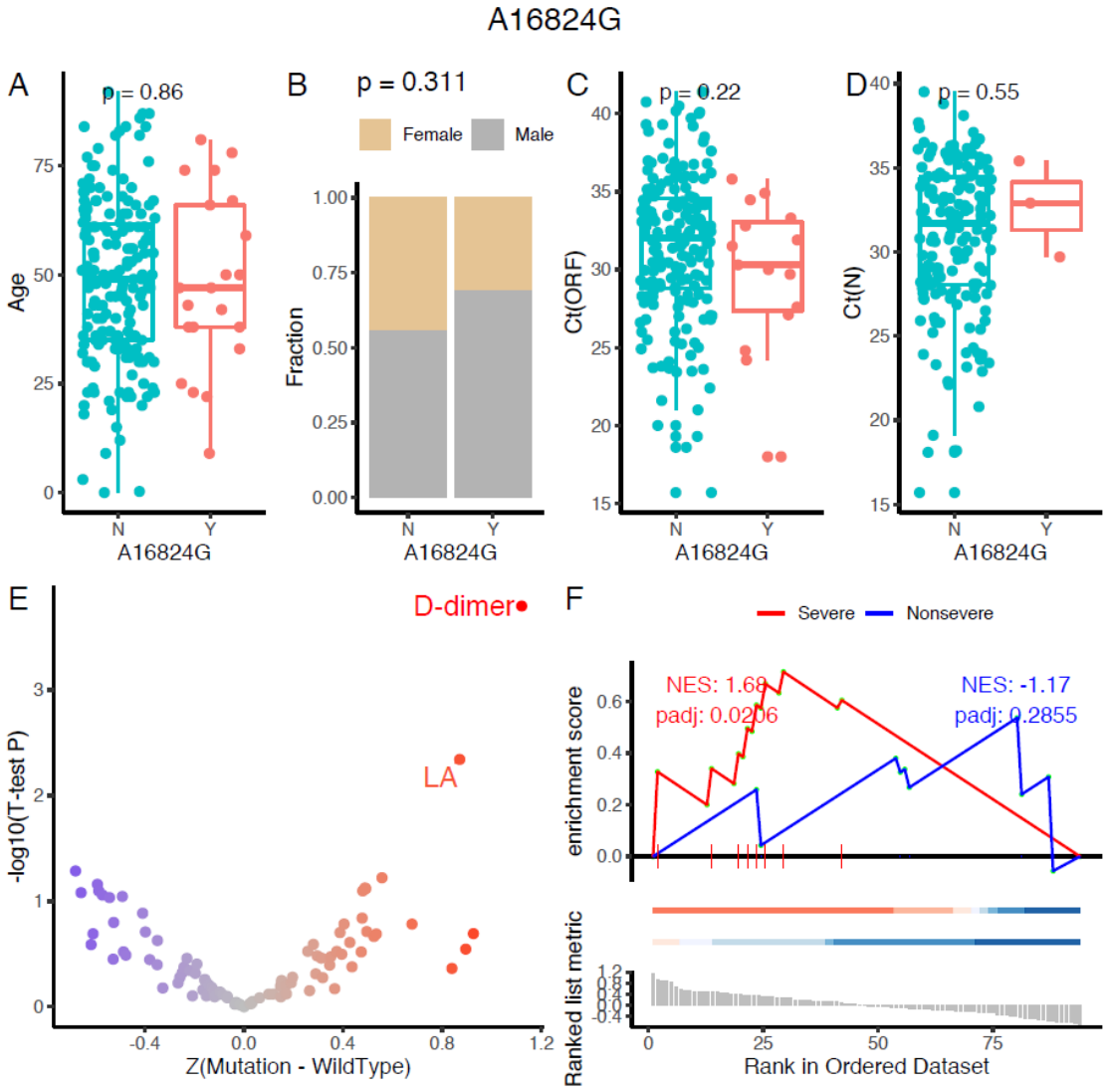
Δ729-737 is an in-frame deletion of Nsp1 (QHD43415.1:p.155-158EDFQ>E, predicted to have a moderate impact on the protein), which has been found in 2 samples (2.00%) in Sichuan and 4 samples (3.00%) in Hubei. The group with or without this deletion showed no difference in age (**A**, Wilcox-test, p = 0.33) and gender (**B**, chi-square test, p = 1), Ct values of ORF (**C**, Wilcox-test, p = 0.51) and N genes (**D**, Wilcox-test, p = 0.82) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of IL-4 and $CD8^+$ T cell counts, and has lower levels of MCV, RBC-SD and $CD3^+$ and $CD4^+$ T cell counts in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the Δ729-737 has a –1.07 NES in non-severe traits (padj = 0.3872) and a 1.32 NES in severe traits (padj = 0.2977, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

4. A13175G in Nsp10



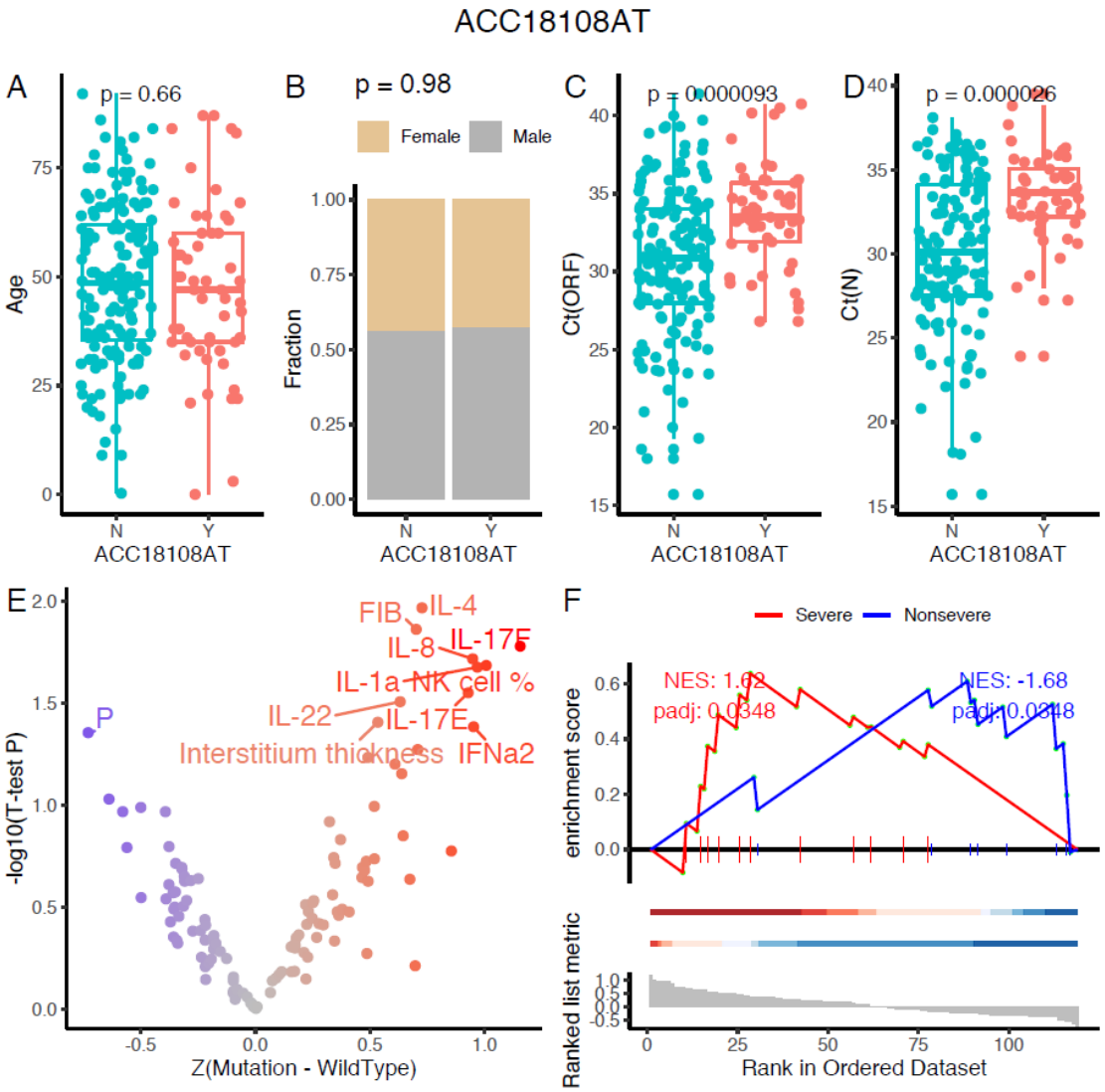
A13175G is a missense variant of Nsp10 (QHD43415.1:p.4304T>A, predicted to have a moderate impact on the protein), which has been found in 4 samples (5.00%) in Sichuan and 1 additional sample (0.02%) in Asia. The group with or without this deletion showed no difference in age (**A**, Wilcox-test, p = 0.76), gender (**B**, chi-square test, p = 0.169), Ct values of ORF (**C**, Wilcox-test, p = 0.18) and N genes (**D**, Wilcox-test, p = 0.21) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of B cell, NK cell, and $CD3^+$ and $CD4^+$ T cell counts, and has lower levels of LDH, CHOL and $CD3^+$ T cell percentage in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the A13175G has a 1.28 NES in non-severe traits (padj = 0.3606) and a –0.78 NES in severe traits (padj = 0.7597, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

5. A16824G in Nsp13



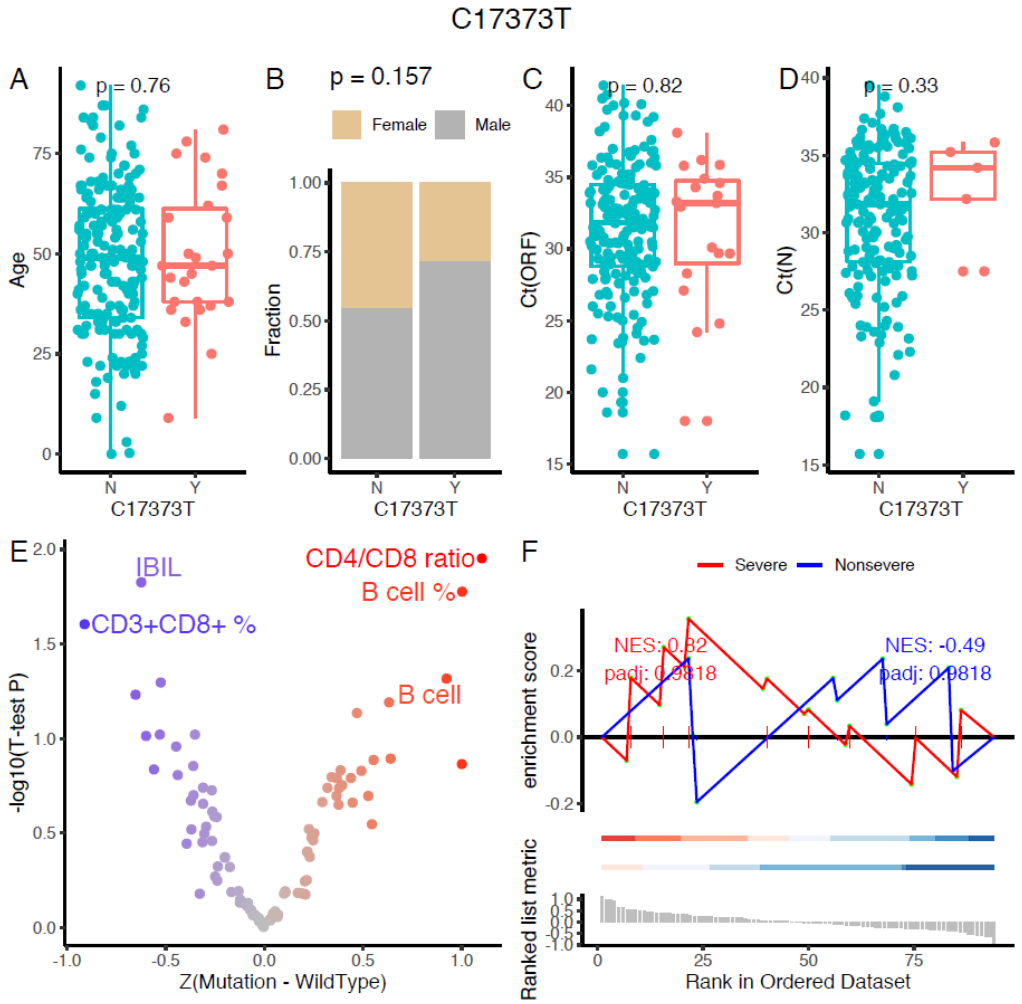
A16824G is a synonymous variant of Nsp13 (QHD43415.1:p.5520G, predicted to have a low impact on the protein), which has been found in 6 samples (7.00%) in Sichuan. The group with or without this deletion showed no difference in age (**A**, Wilcox-test, p = 0.86), gender (**B**, chi-square test, p = 0.311), Ct values of ORF (**C**, Wilcox-test, p = 0.22) and N genes (**D**, Wilcox-test, p = 0.55) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of LA and D-dimer in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the A16824G has a –1.17 NES in non-severe traits (padj = 0.2855) and a 1.69 NES in severe traits (padj = 0.0206, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

6. ACC18108AT in Nsp14



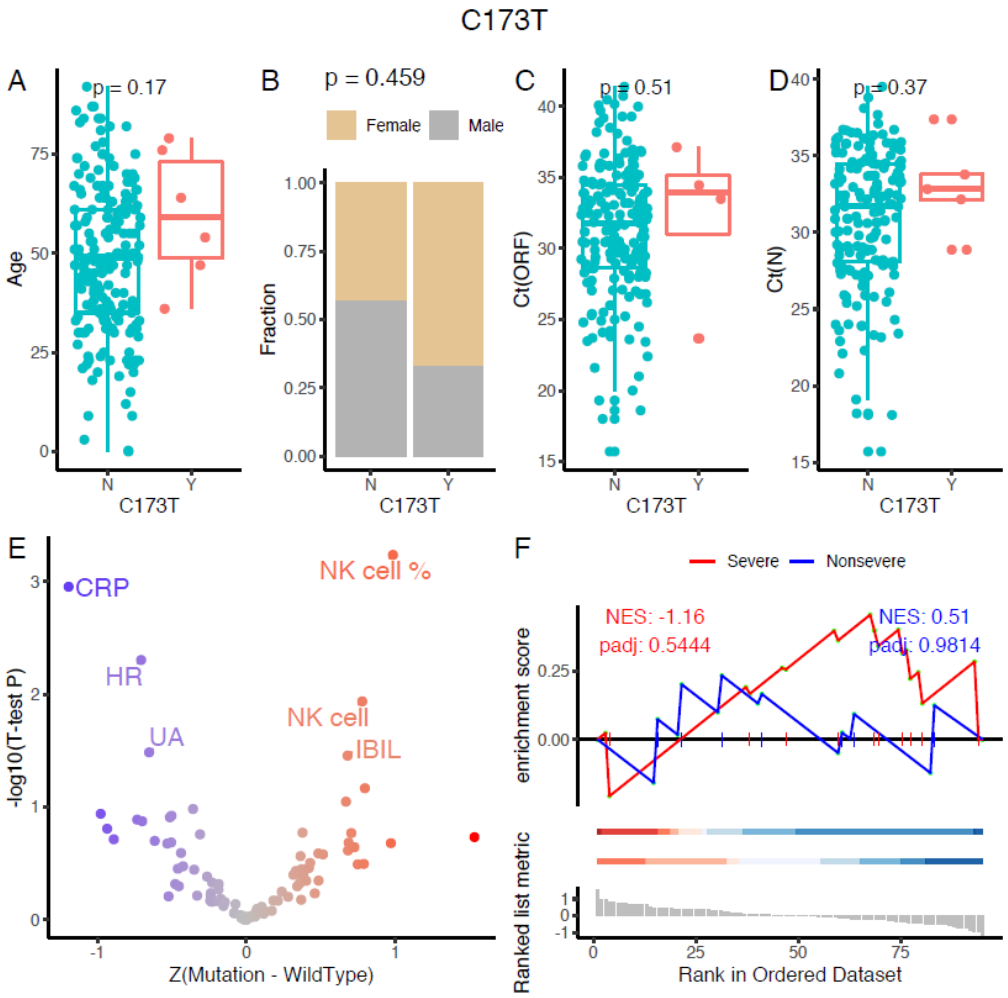
ACC18108AT is a frameshift variant of Nsp14 (QHD43415.1:p.5949P>X, predicted to have high impact on the protein), which has been found in 12 samples (14.00%) in Sichuan and 12 samples (8.00%) in Hubei. The group with or without this deletion showed no difference in age (**A**, Wilcox-test, p = 0.66) and gender (**B**, chi-square test, p = 0.98). When comparing the Ct values of ORF and N genes based on qPCR, the group with ACC18108AT has a significant increase of Ct value in ORF (**C**, Wilcox-test, $p = 9.3\times10^{-5}$) and N gene (**D**, Wilcox-test, $p = 2.6\times10^{-5}$). We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of IL-4, IL-8, IL-22, IL-1a, IL-17E, IL-17F, FIB, interstitium thickness and NK cell percentage, and has a lower level of P in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the ACC18108AT has a –1.68 NES in non-severe traits (padj = 0.0348) and a 1.62 NES in severe traits (padj = 0.0348, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

7. C17373T in Nsp13



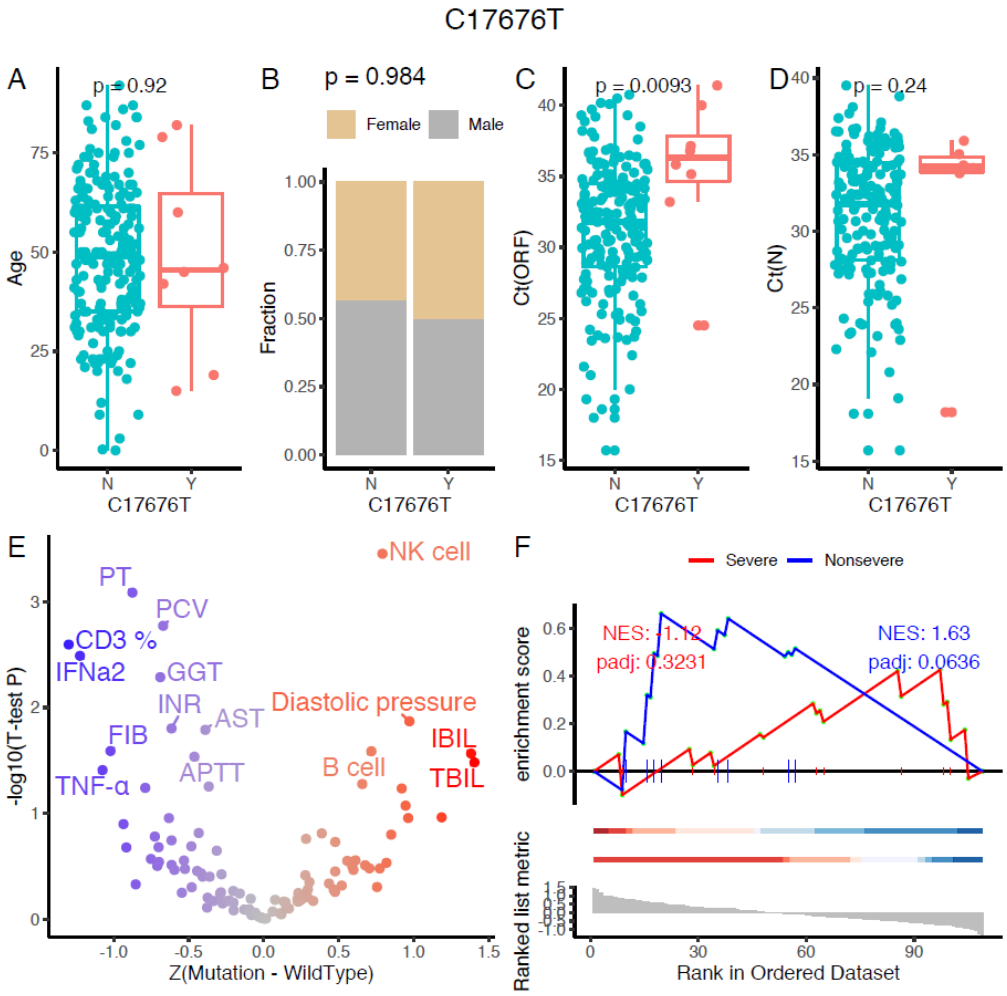
C17373T is a synonymous variant of Nsp13 (QHD43415.1:p.5703A, predicted to have a low impact on the protein), which has been found in 7 samples (8.00%) in Sichuan, 17 samples (0.29%) in Asia, 14 samples (0.03%) in Europe, 1 sample (0.01%) in Oceania and 2 samples (0.01%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.76), gender (**B**, chi-square test, p = 0.157), Ct values of ORF (**C**, Wilcox-test, p = 0.82) and N genes (**D**, Wilcox-test, p = 0.33) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of CD4/CD8 ratio, B cell numbers and percentage, and has lower levels of IBIL, $CD3^+$ and $CD8^+$ T cell percentage in volcano plot in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C17373T has a –0.49 NES in non-severe traits (padj = 0.9818) and a 0.82 NES in severe traits (padj = 0.9818, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

8. C173T in upstream region



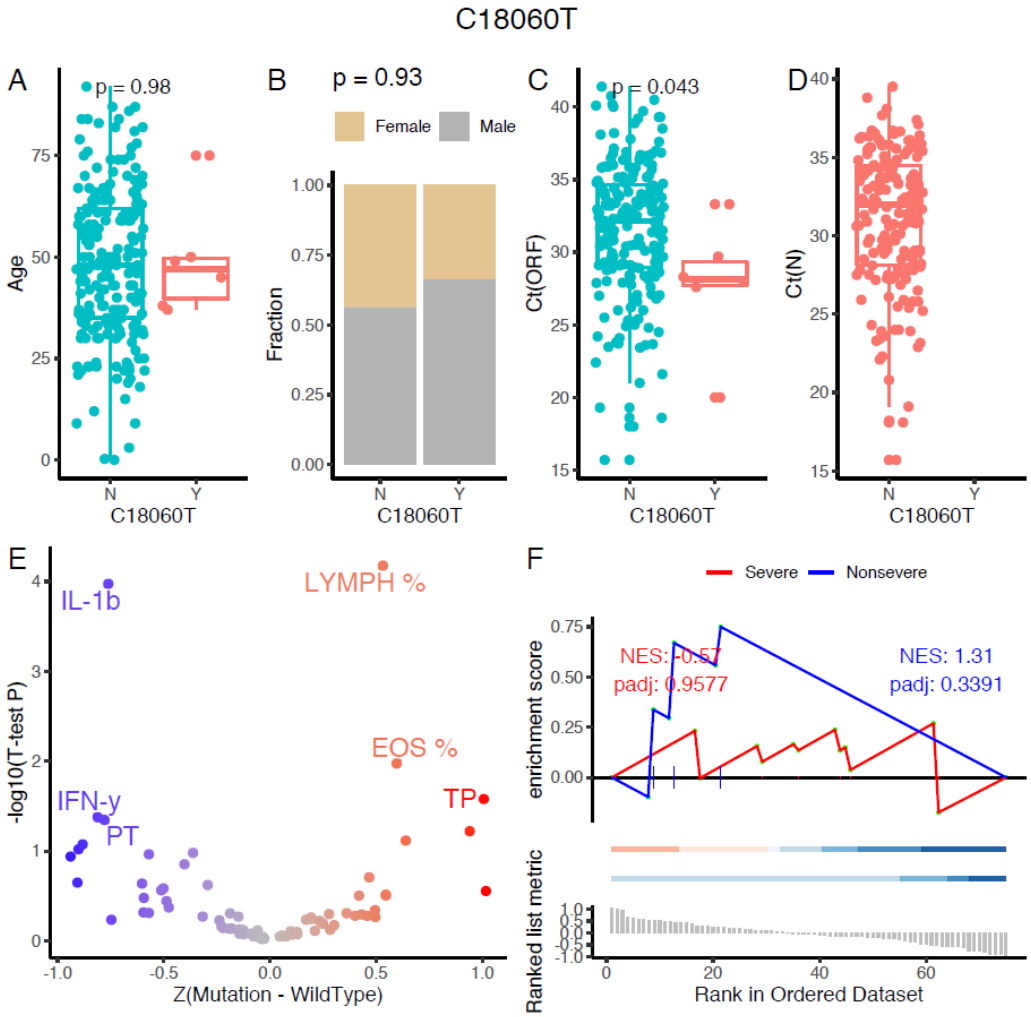
C173T is an upstream variant of 5'UTR (QHD43415.1, predicted to have "modifier"), which has been found in 2 samples (2.00%) in Sichuan, 2 samples (0.00%) in Europe, 1 sample (0.01%) in Oceania and 2 samples (0.01%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.17), gender (**B**, chi-square test, p = 0.459), Ct values of ORF (**C**, Wilcox-test, p = 0.51) and N genes (**D**, Wilcox-test, p = 0.37) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of IBIL, NK cell numbers and percentage, and have lower levels of CPR, HR and UA in volcano plot in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C173T has a 0.51 NES in non-severe traits (padj = 0.9814) and a –1.16 NES in severe traits (padj = 0.5444, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

9. C17676T in Nsp13



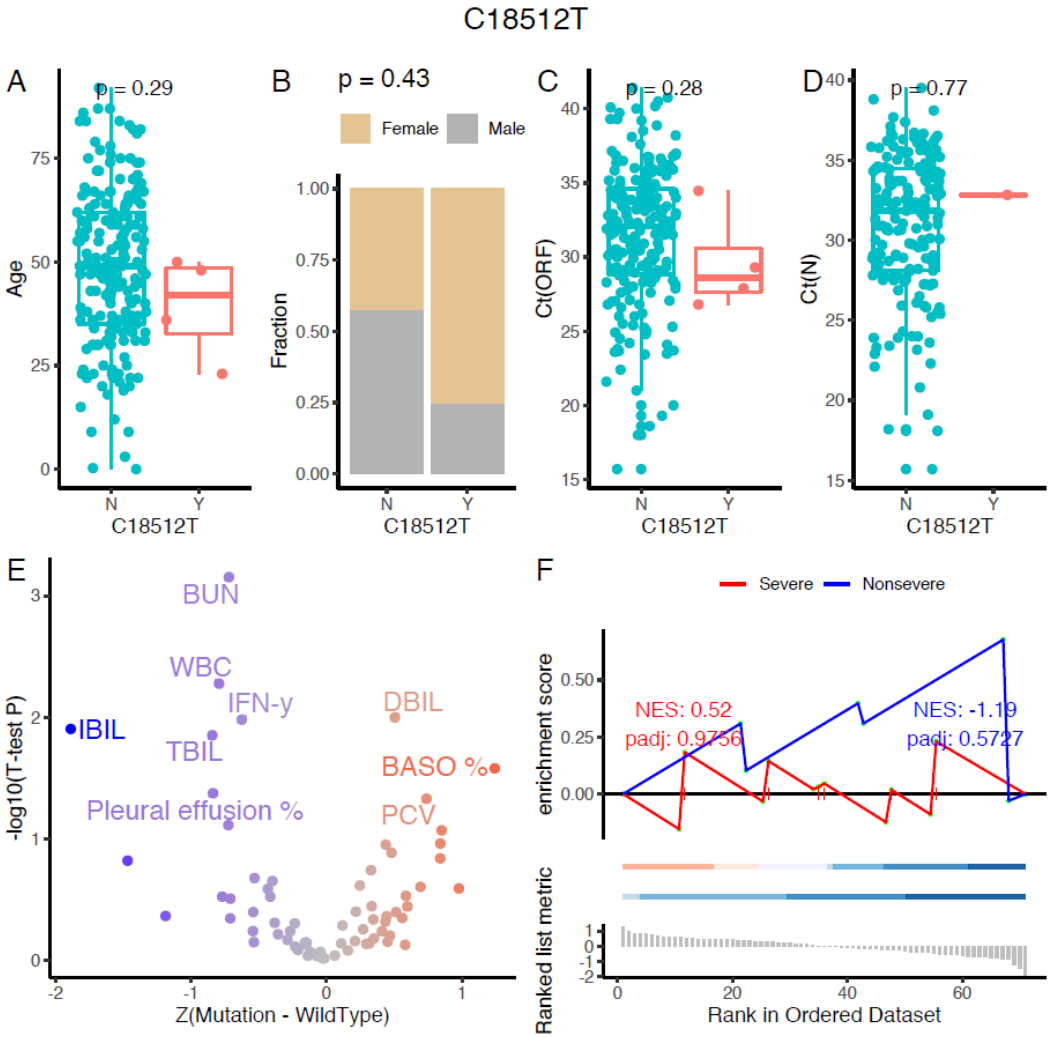
C17676T is a synonymous variant of Nsp13 (QHD43415.1: p.5804I, predicted to have a low impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 3 samples (0.22%) in Africa, 29 samples (0.06%) in Europe, 5 samples (0.09%) in Asia, 3 samples (0.32%) in South America and 36 samples (0.19%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.92), gender (**B**, chi-square test, p = 0.984) and N genes (**D**, Wilcox-test, p = 0.24) based on qPCR. When comparing the Ct values of ORF based on qPCR, the group with C17676T has a significant increase of Ct value in ORF (**C**, Wilcox-test, $p = 9.3\times10^{-3}$). We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of IBIL, NK cell numbers, diastolic pressure, TBIL and B cell, and has lower levels of PT, PCV, CD3 T cell percentage, GGT, IFNa2, INR, FIB, AST, TNF- and APTT in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C17676T has a 1.63 NES in non-severe traits (padj = 0.0636) and a –1.12 NES in severe traits (padj = 0.3231, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

10. C18060T in Nsp14



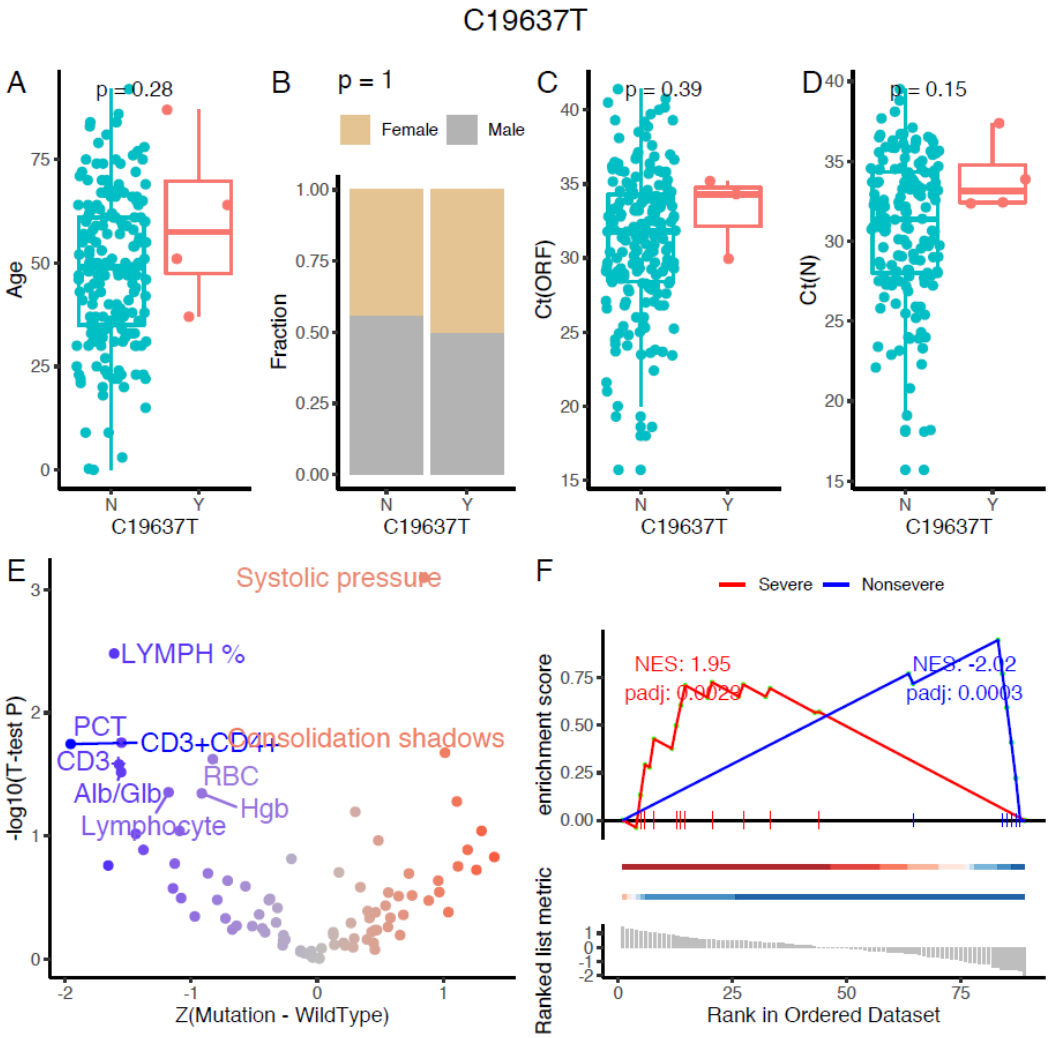
C18060T is a synonymous variant of Nsp14 (QHD43415.1:p.5932L, predicted to have a low impact on the protein), which has been found in 3 samples (3.00%) in Sichuan, 32 samples (0.07%) in Europe, 18 samples (0.31%) in Asia, 27 samples (3.00%) in South America, 71 samples (1.00%) in Oceania and 1563 samples (8.00%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.98), gender (**B**, chi-square test, p = 0.93). When comparing the Ct values of ORF based on qPCR, the group with C18060T has a decrease of Ct value in ORF (**C**, Wilcox-test, p = 0.043). We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of TP, LYMPH and EOS percentage, and has lower levels of IL-1b, PT and IFN-γ concentration in serum in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C18060T has a 1.31 NES in non-severe traits (padj = 0.3391) and a –0.57 NES in severe traits (padj = 0.9577, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

11. C18512T in Nsp14



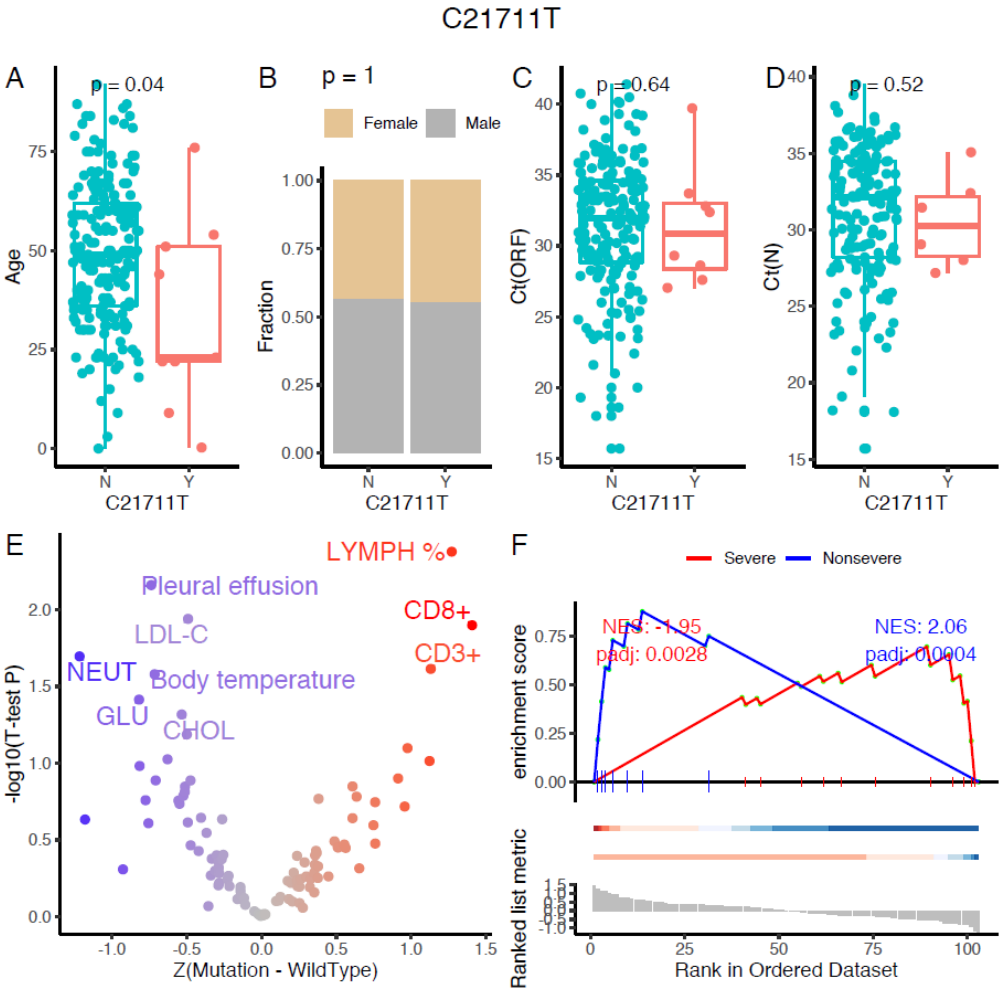
C18512T is a missense variant of Nsp14 (QHD43415.1:p.6083P>L, predicted to have a moderate impact on the protein), which has been found in 3 samples (3.00%) in Sichuan, 2 samples (0.00%) in Europe and 1 sample (0.02%) in Asia. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.29), gender (**B**, chi-square test, p = 0.43), Ct values of ORF (**C**, Wilcox-test, p = 0.28) and N genes (**D**, Wilcox-test, p = 0.77) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of DBIL, PCV and BASO percentage, and has lower levels of IBIL, BUN, WBC, TBIL, IFN-γ concentration in serum and pleural effusion percentage in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C18512T has a –1.19 NES in non-severe traits (padj = 0.5727) and a 0.52 NES in severe traits (padj = 0.9756, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

12. C19637T in Nsp15



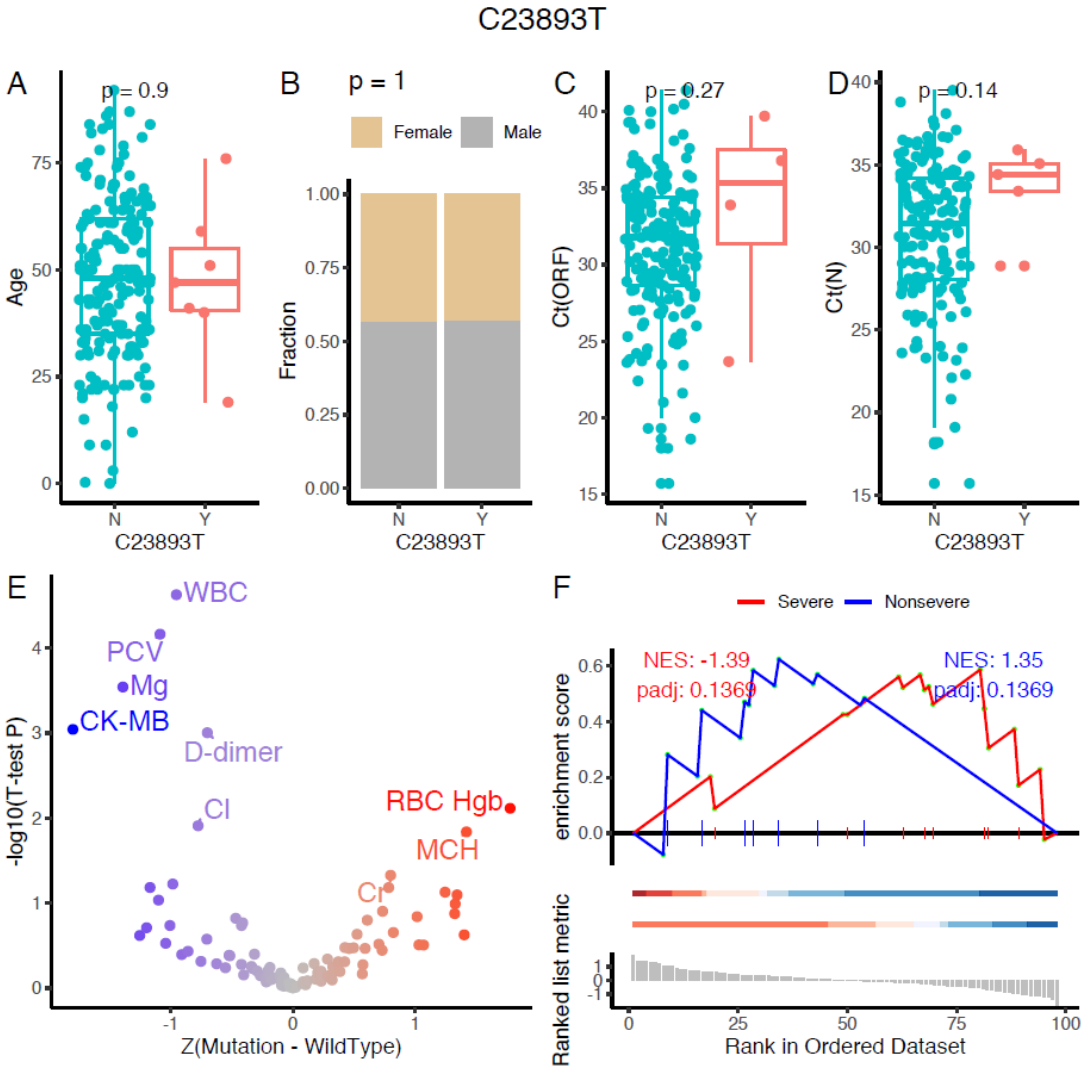
C19637T is a missense variant of Nsp15 (QHD43415.1:p.6458A>V, predicted to have a moderate impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 1 sample (0.00%) in Europe and 1 sample (0.02%) in Asia. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.28), gender (**B**, chi-square test, p = 1), Ct values of ORF (**C**, Wilcox-test, p = 0.39) and N genes (**D**, Wilcox-test, p = 0.15) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of systolic pressure and consolidation shadows, and has lower levels of LYMPH percentage, PCT, RBC, $CD3^+$ T cell counts, Alb, Hgb, Glb, lymphocyte, $CD3^+$ and $CD4^+$ T cell counts in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C19637T has a –2.02 NES in non-severe traits (padj = 3×10^{-4}) and a 1.95 NES in severe traits (padj = 2.9×10^{-3}, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

13. C21711T in S



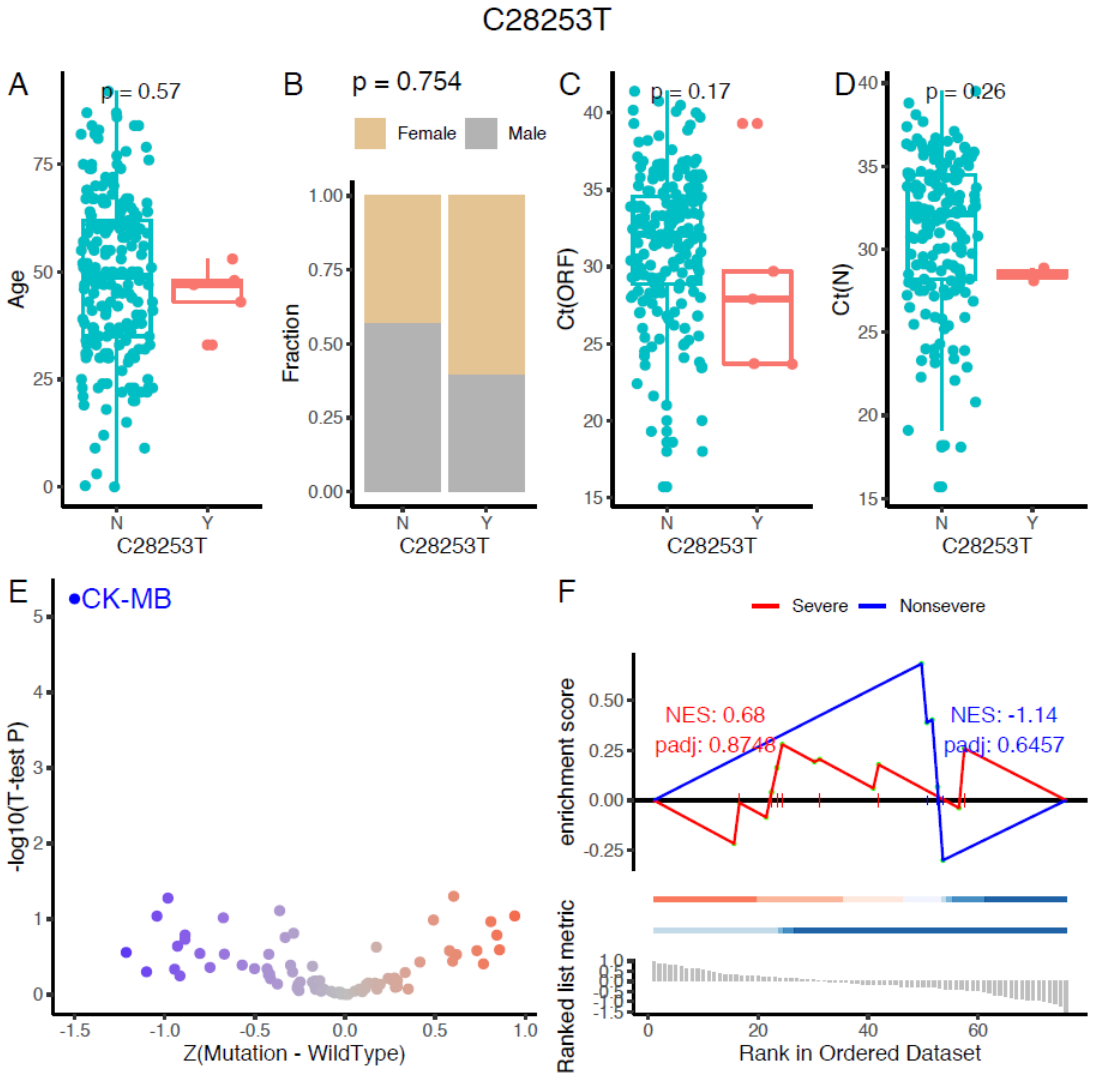
C21711T is a missense variant of S gene (QHD43416.1:p.50S>L, predicted to have a moderate impact on the protein), which has been found in 3 samples (3.00%) in Sichuan, 1 sample (0.62%) in Hubei, 2 samples (0.00%) in Europe, 4 samples (0.07%) in Asia, 1 sample (0.01%) in Oceania and 8 samples (0.04%) in North America. The group with or without this missense variant showed no difference in gender (**B**, chi-square test, p = 1.00), Ct values in ORF (**C**, Wilcox-test, p = 0.64) and N genes (**D**, Wilcox-test, p = 0.52) based on qPCR, but showed a significant difference in age (**A**, Wilcox-test, p = 0.04). We further compared the 117 clinical phenotypes between the two groups, the group with the C21711T has higher levels of LYMPH percentage, $CD3^+$ and $CD8^+$ T cell counts, and has lower levels of GLU, CHOL, NEUT, LDL-C, Pleural effusion and Body temperature in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C21711T has a 2.06 NES in non-severe traits (padj = 4.00×10^{-4}) and a –1.95 NES in severe traits (padj = 2.8×10^{-3}, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

14. C23893T in S



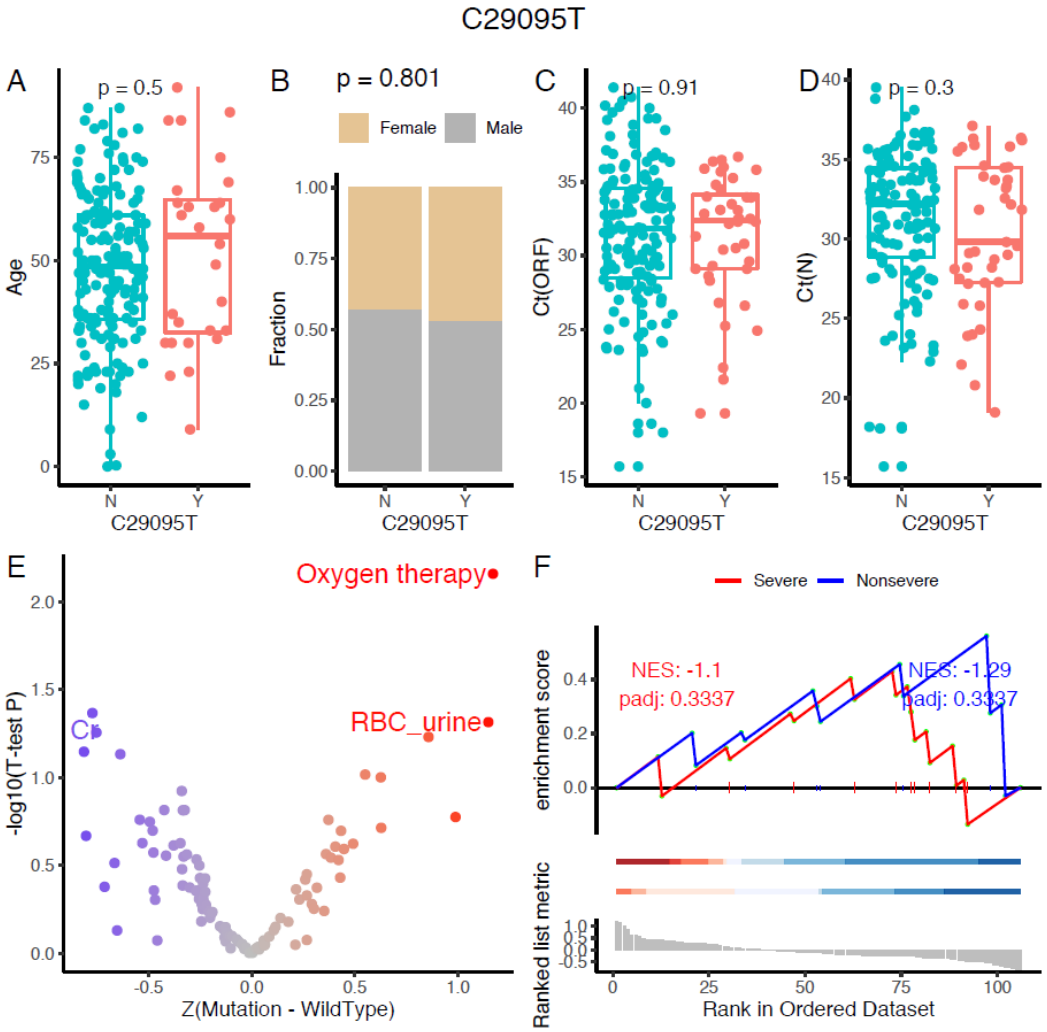
C23893T is a synonymous variant in S gene (QHD43416.1:p.777N, predicted to have a low impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 13 samples (0.03%) in Europe and 1 sample (0.01%) in North America. The group with or without this synonymous variant showed no difference in age (**A**, Wilcox-test, p = 0.90), gender (**B**, chi-square test, p = 1.00) and Ct values in ORF (**C**, Wilcox-test, p = 0.27) and N genes (**D**, Wilcox-test, p = 0.14) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the C23893T has higher levels of RBC, Hgb , Cr and MCH, and has lower levels of CK-MB, PVC, WBC, D-dimer, Mg and Cl in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C23893T has a 1.35 NES in non-severe traits (padj = 0.1369) and a –1.39 NES in severe traits (padj = 0.1369, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

15. C28253T in ORF8



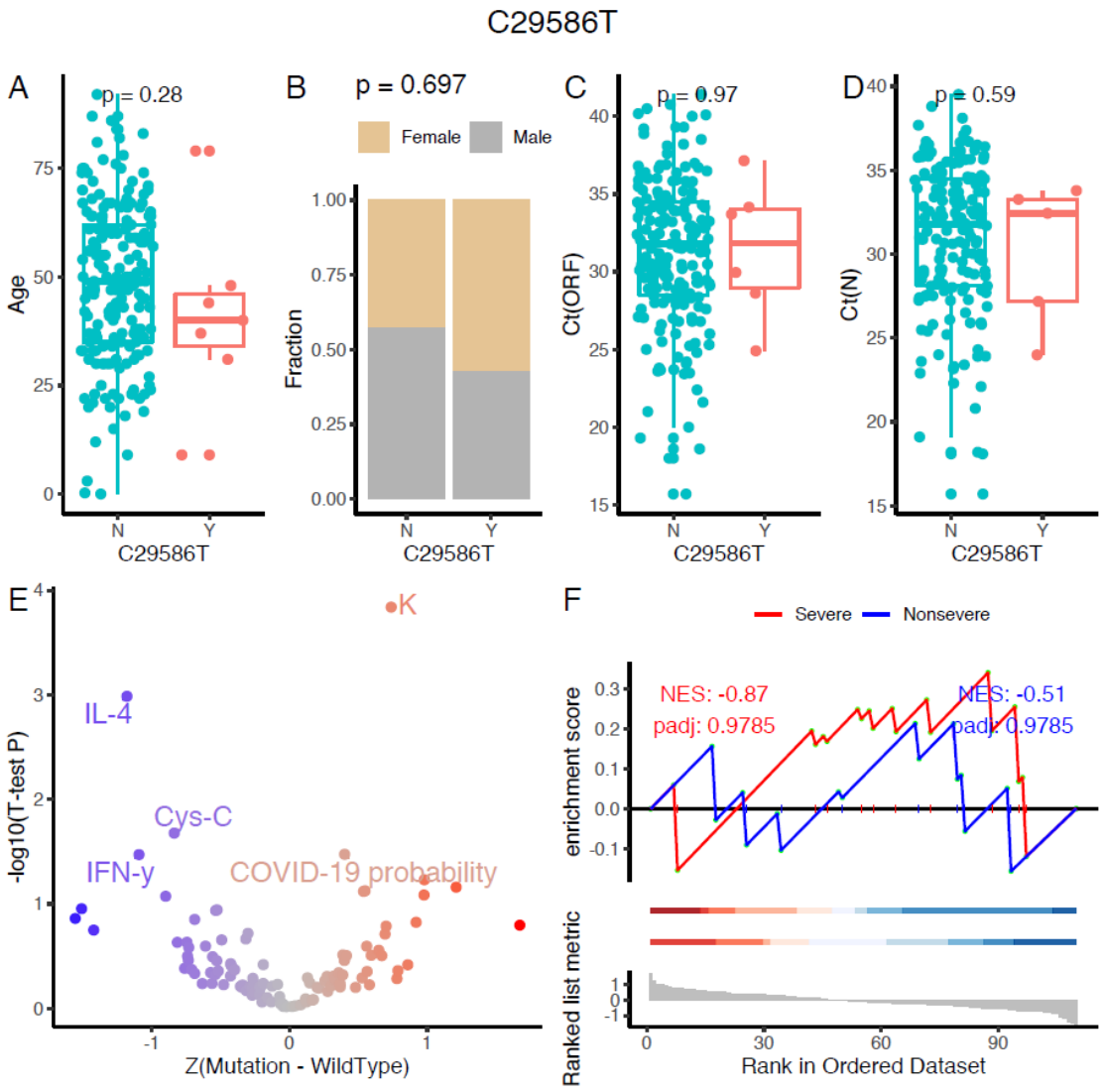
C28253T is a synonymous variantin in ORF8 (QHD43422.1:p.120F, predicted to have a low impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 17 samples (1.00%) in Africa, 187 samples (0.40%) in Europe, 20 samples (0.34%) in Asia, 27 samples (3.00%) in South America, 20 samples (0.29%) in Oceania and 39 samples (0.20%) in North America. The group with or without this synonymous variantin showed no difference in age (**A**, Wilcox-test, p = 0.57), gender (**B**, chi-square test, p = 0.745). and Ct values in ORF (**C**, Wilcox-test, p = 0.17) and N genes (**D**, Wilcox-test, p = 0.26) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the C28253T has lower level of CK-MB in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C28253T has a –1.14 NES in non-severe traits (padj = 0.6457) and a 0.68 NES in severe traits (padj = 0.8746, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

16. C29095T in N



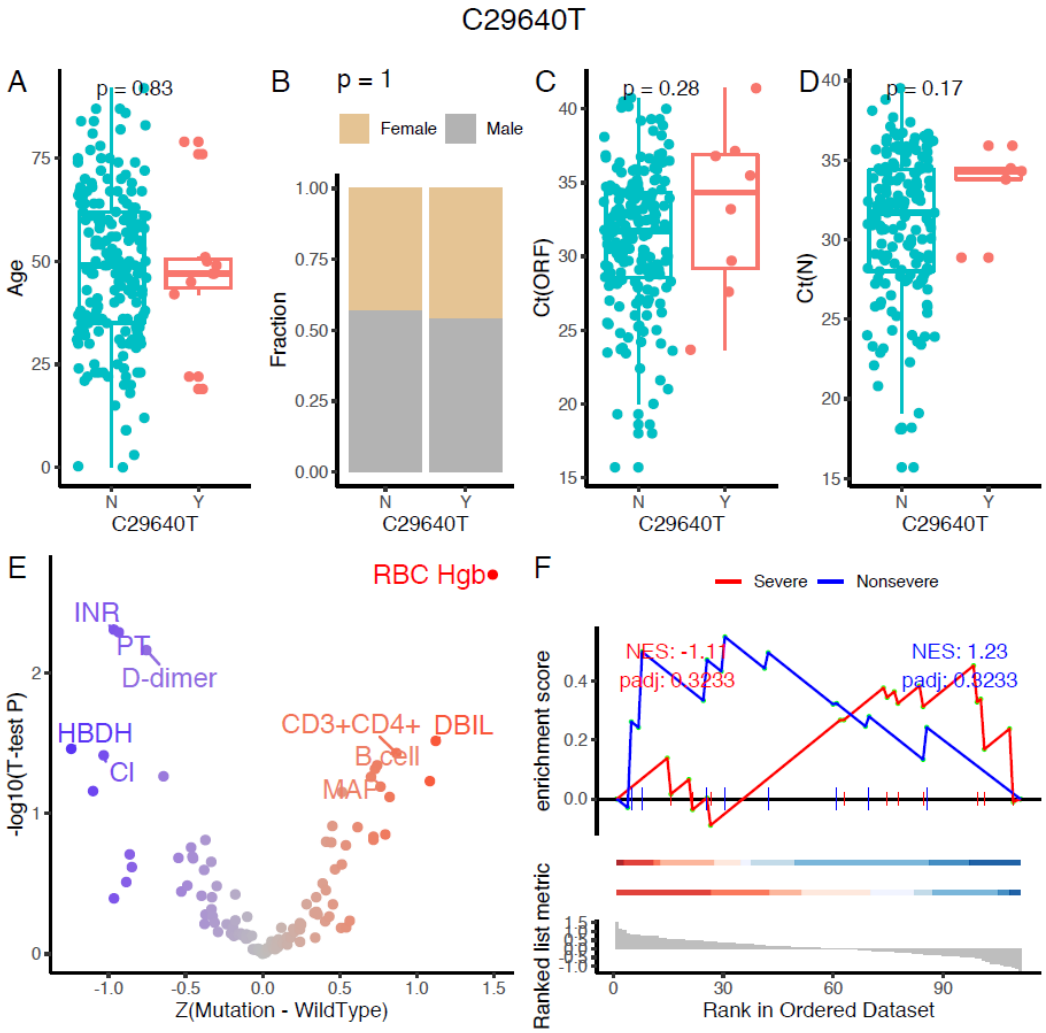
C29095T is a synonymous variant in N gene (QHD43423.2: p.274F, predicted to have a low impact on the protein), which has been found in 3 samples (3.00%) in Sichuan, 8 samples (5.00%) in Hubei, 3 samples (0.22%) in Africa, 503 samples (1.00%) in Europe, 51 samples (0.87%) in Asia, 1 sample (0.11%) in South America, 4 samples (0.06%) in Oceania and 24 samples (0.12%) in North America. The group with or without this synonymous variant showed no difference in age (**A**, Wilcox-test, p = 0.50), gender (**B**, chi-square test, p = 0.801) and Ct values in ORF (**C**, Wilcox-test, p = 0.91) and N genes (**D**, Wilcox-test, p = 0.30) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the C29095T has higher levels of Oxygen therapy and RBC_urine, and has lower level of Cr in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C29095T has a –1.29 NES in non-severe traits (padj = 0.3337) and a –1.1 NES in severe traits (padj = 0.3337, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

17. C29586T in ORF10



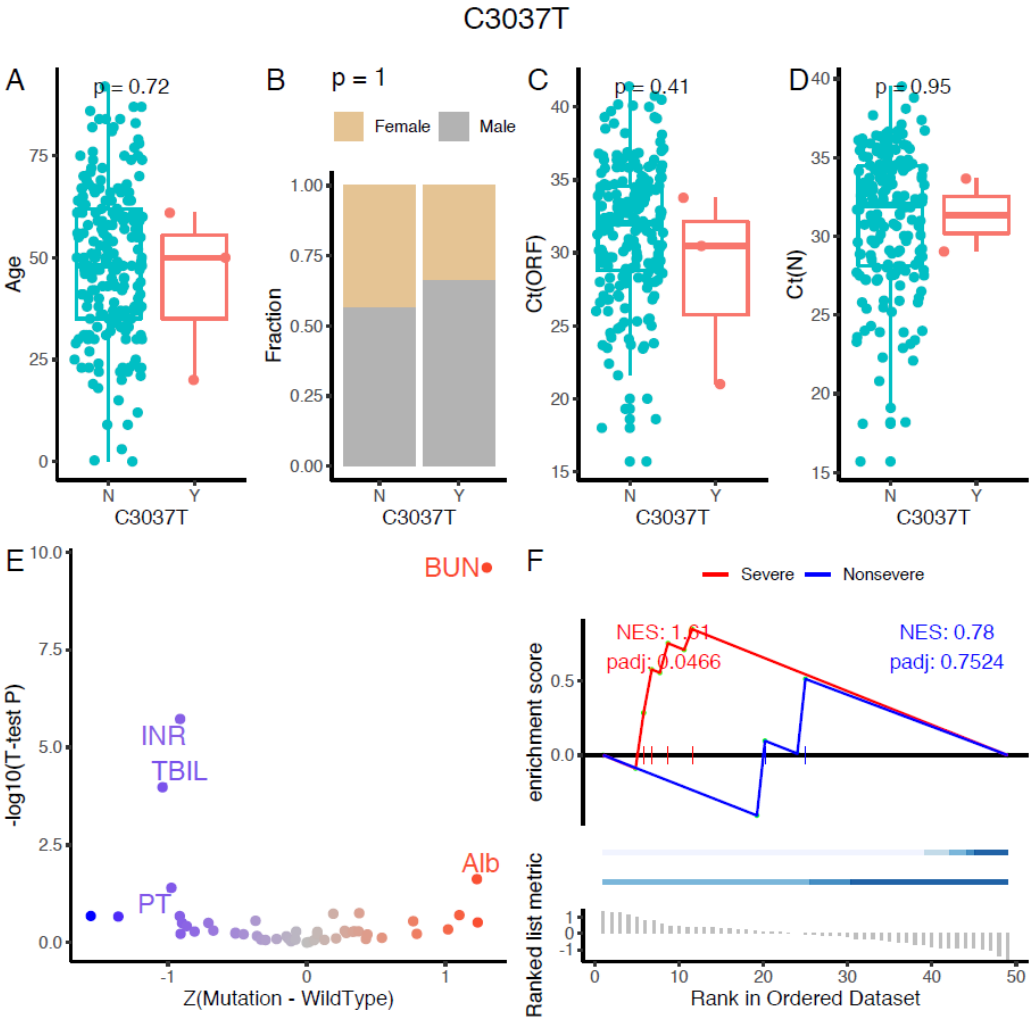
C29586T is a missense variant in ORF10 (QHI42199.1:p.10P>L, predicted to have a moderate impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 3 samples (0.01%) in Europe and 2 samples (0.03%) in Oceania. The group with or without this missense variant showed no difference in age (**A**, Wilcox-test, p = 0.28), gender (**B**, chi-square test, p = 0.697) and Ct values in ORF (**C**, Wilcox-test, p = 0.97) and N genes (**D**, Wilcox-test, p = 0.59) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the C29586T has higher levels of K and COVID−19 probability, and has lower levels of IL-4, Cys-C and IFN-γ concentration in serum in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C29586T has a –0.51 NES in non-severe traits (padj = 0.9785) and a –0.87 NES in severe traits (padj = 0.9785, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

18. C29640T in ORF10



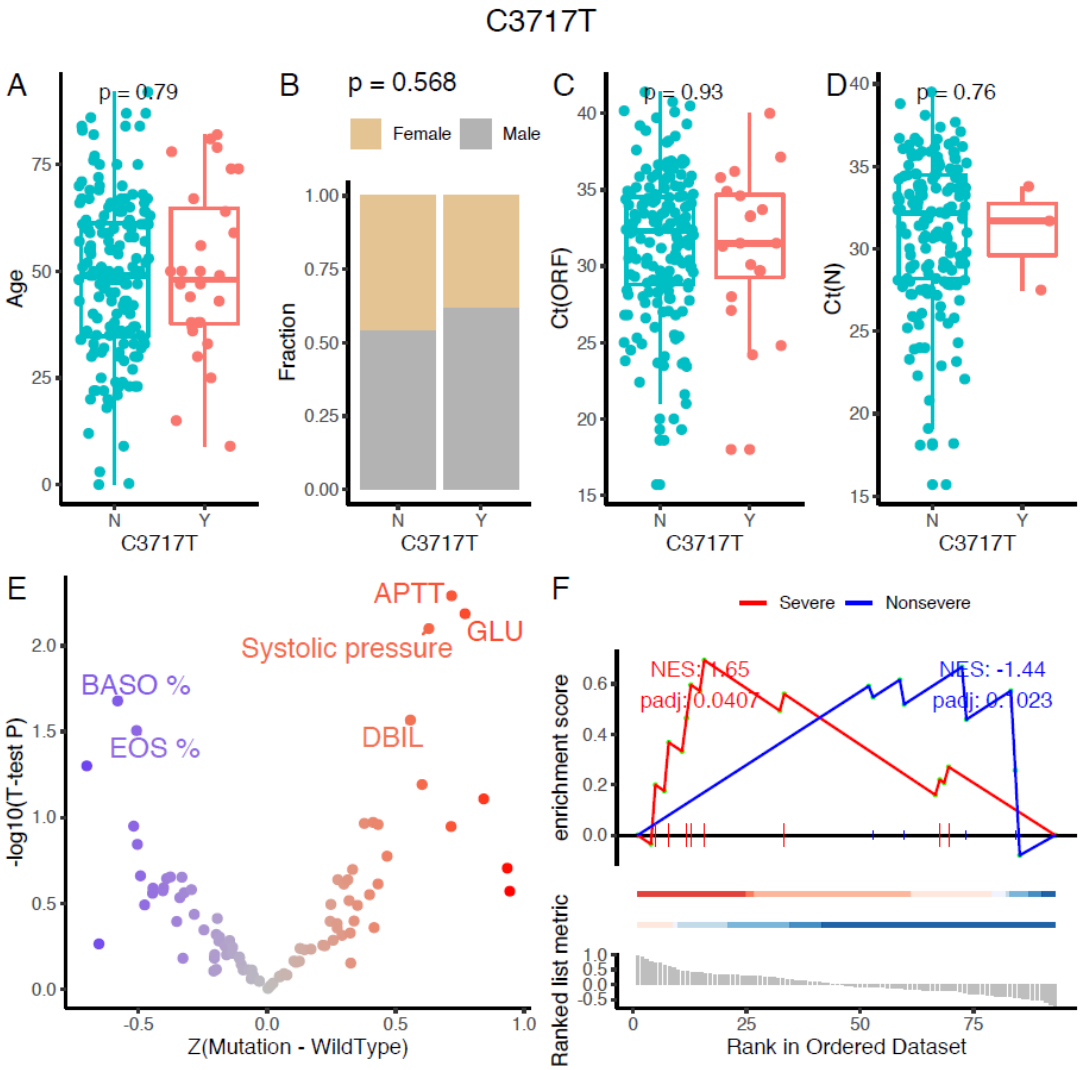
C29640T is a missense variant in ORF10 (QHI42199.1:p.28A>V, predicted to have a moderate impact on the protein), which has been found in 4 samples (5.00%) in Sichuan, 20 samples (0.04%) in Europe, 4 samples (0.07%) in Asia, 1 sample (0.01%) in Oceania and 9 samples (0.05%) in North America. The group with or without this missense variant showed no difference in age (**A**, Wilcox-test, p = 0.83), gender (**B**, chi-square test, p = 1.00) and Ct values in ORF (**C**, Wilcox-test, p = 0.28) and N genes (**D**, Wilcox-test, p = 0.17) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the C29640T has higher levels of RBC, Hgb, MAP, the $CD3^+$ $CD4^+$ T cell counts and B cell in the peripheral bloods, and has lower levels of PT, INR, HBDH, D-dimer and Cl in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C29640T has a 1.23 NES in non-severe traits (padj = 0.3233) and a –1.11 NES in severe traits (padj = 0.3233, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

19. C3037T in Nsp3



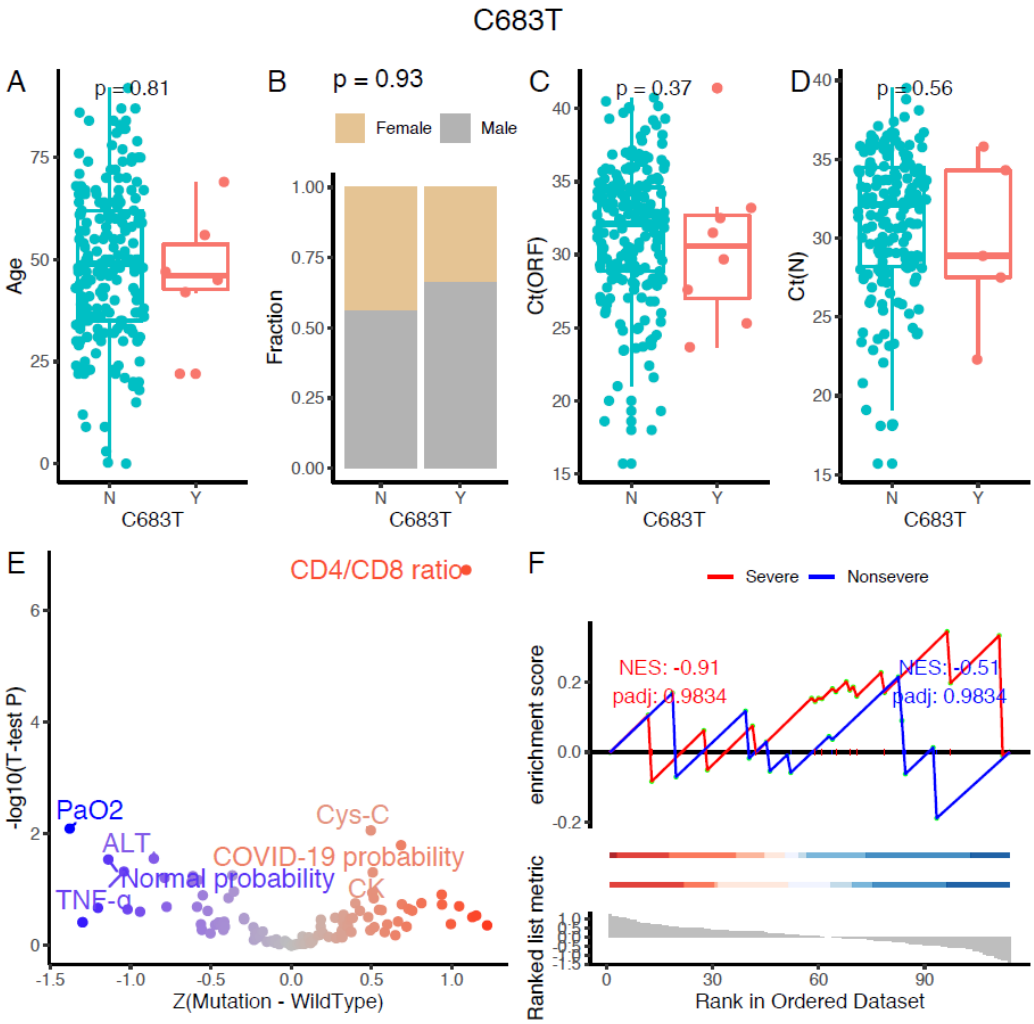
C3037T is a synonymous variant of Nsp3 (QHD43415.1:p.924F, predicted to have a low impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 1314 samples (96.00%) in Africa, 44726 samples (95.00%) in Europe, 3976 samples (70.00%) in Asia, 883 samples (95.00%) in South America, 6349 samples (93.00%) in Oceania and 16596 samples (85.00%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.72), gender (**B**, chi-square test, p = 1.00), Ct values in ORF (**C**, Wilcox-test, p = 0.41) and N genes (**D**, Wilcox-test, p = 0.95) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of BUN and Alb, and has lower levels of INR, TBIL and PT in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C3037T has a 0.78 NES in non-severe traits (padj = 0.7524) and a 1.61 NES in severe traits (padj = 0.0466, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

20. C3717T in Nsp3



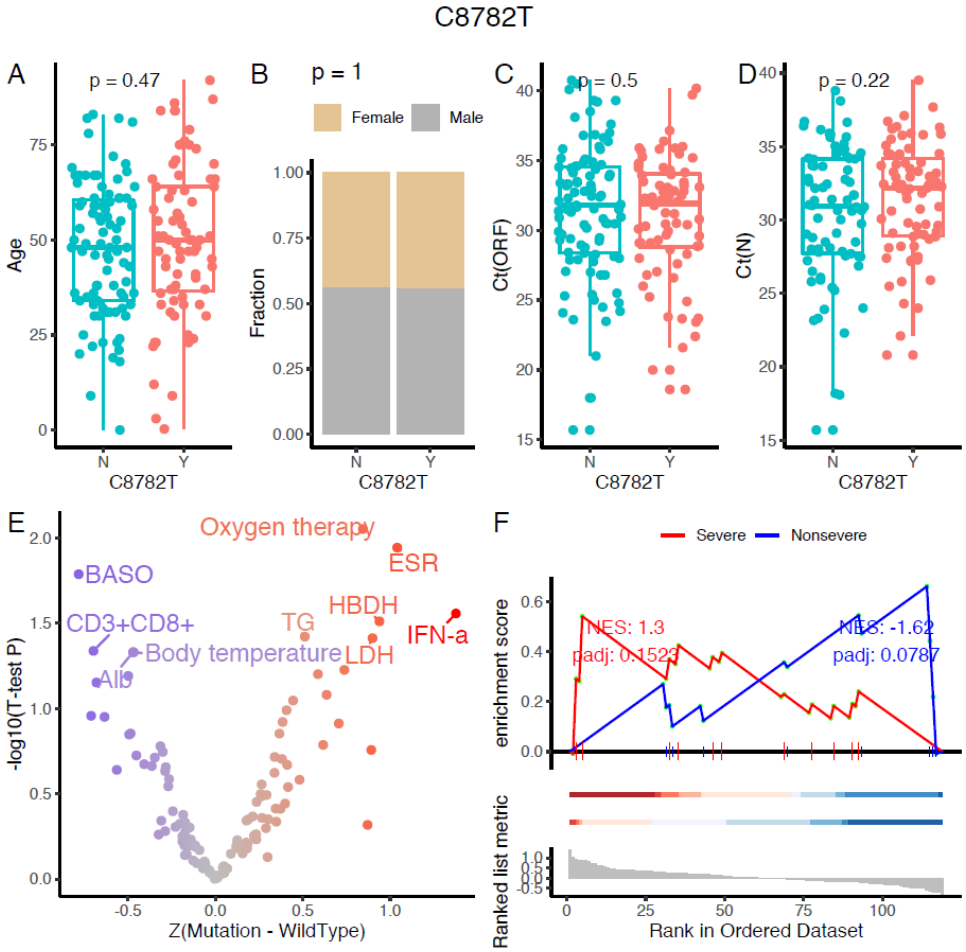
C3717T is a missense variant of Nsp3 (QHD43415.1:p.1151A>V, predicted to have a moderate impact on the protein), which has been found in 3 samples (3.00%) in Sichuan, 1 sample (0.62%) in Hubei, 11 samples (0.02%) in Europe and 4 samples (0.02%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.79), gender (**B**, chi-square test, p = 0.568), Ct values in ORF (**C**, Wilcox-test, p = 0.93) and N genes (**D**, Wilcox-test, p = 0.76) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of APTT, GLU, DBIL and systolic pressure, and has lower levels of BASO and EOS percentage in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C3717T has a –1.44 NES in non-severe traits (padj = 0.1023) and a 1.65 NES in severe traits (padj = 0.0407, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

21. C683T in Nsp1



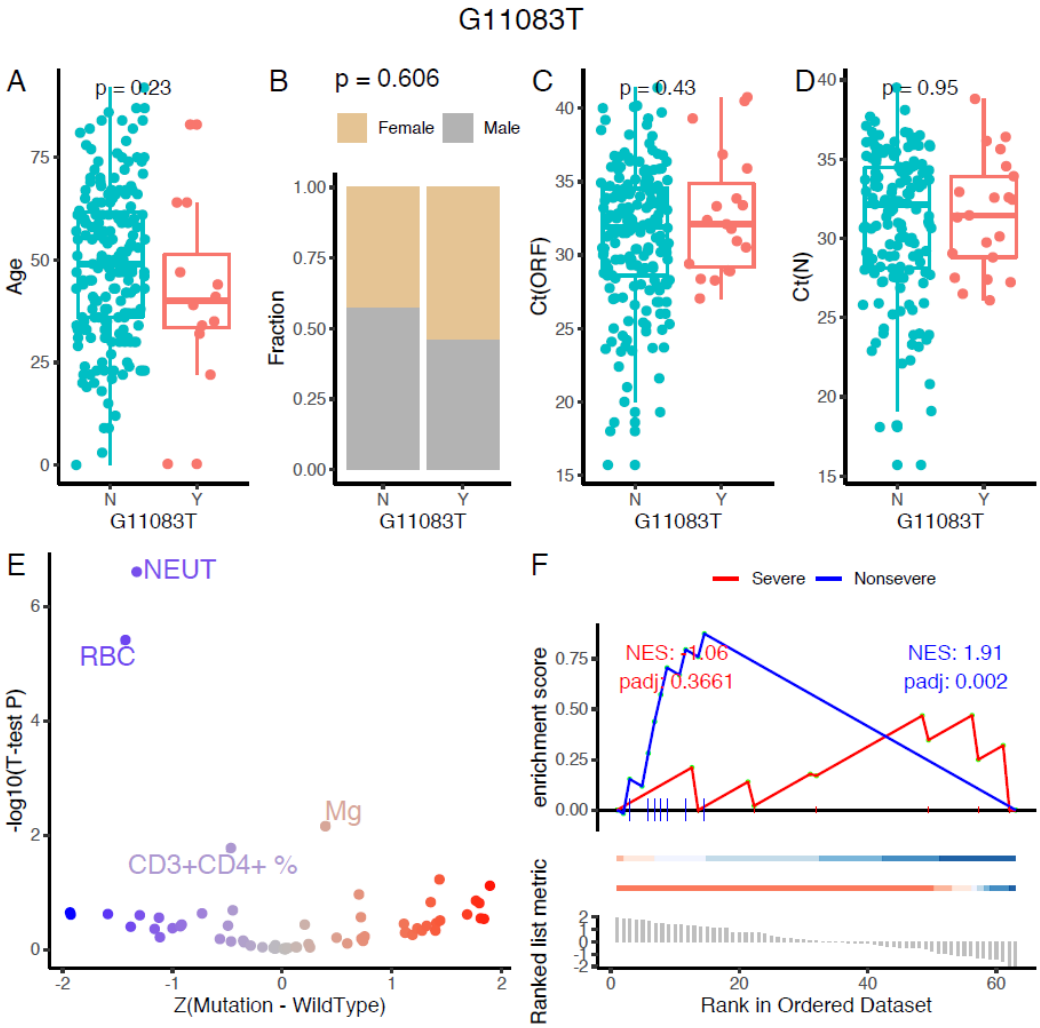
C683T is a synonymous variant of Nsp1 (QHD43415.1:p.140L, predicted to have a low impact on the protein), which has been found in 3 samples (3.00%) in Sichuan, 68 samples (0.14%) in Europe, 24 samples (0.41%) in Asia, 1 sample (0.11%) in South America, 5 samples (0.07%) in Oceania and 8 samples (0.04%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.81), gender (**B**, chi-square test, p = 0.93), Ct values in ORF (**C**, Wilcox-test, p = 0.37) and N genes (**D**, Wilcox-test, p = 0.56) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of CD4/CD8 ratio, Cys-C, CK and COVID-19 probability, and has lower levels of PaO2, ALT, TNF and normal probability in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C683T has a –0.51 NES in non-severe traits (padj = 0.9834) and a ‑0.91 NES in severe traits (padj = 0.9834, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

22. C8782T in Nsp4



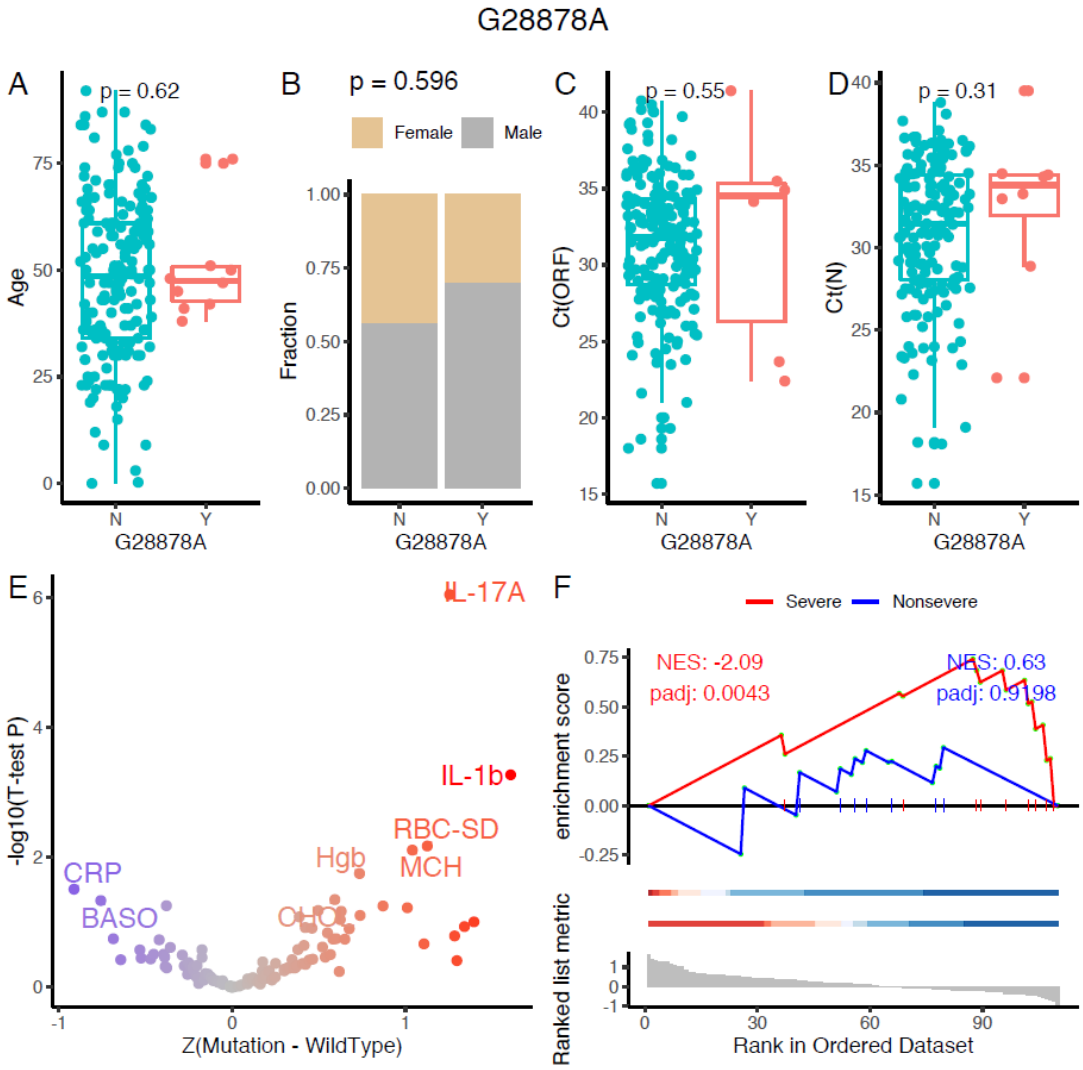
C8782T is a synonymous variant of Nsp4 (QHD43415.1:p.2839S, predicted to have a low impact on the protein), which has been found in 10 samples (11.00%) in Sichuan, 7 samples (4.00%) in Hubei, 36 samples (3.00%) in Africa, 420 samples (0.89%) in Europe, 537 samples (10.00%) in Asia, 28 samples (3.00%) in South America, 247 samples (4.00%) in Oceania and 2024 samples (10.00%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.47), gender (**B**, chi-square test, p = 1.00), Ct values in ORF (**C**, Wilcox-test, p = 0.50) and N genes (**D**, Wilcox-test, p = 0.22) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has higher levels of oxygen therapy, ESR, HBDH, TG, LDH and IFN-α concentration in serum, and has lower levels of body temperature, BASO, Alb, $CD3^+$ $CD8^+$ T cell counts in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the C8782T has a –1.62 NES in non-severe traits (padj = 0.0787) and a 1.3 NES in severe traits (padj = 0.1523, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

23. G11083T in Nsp6



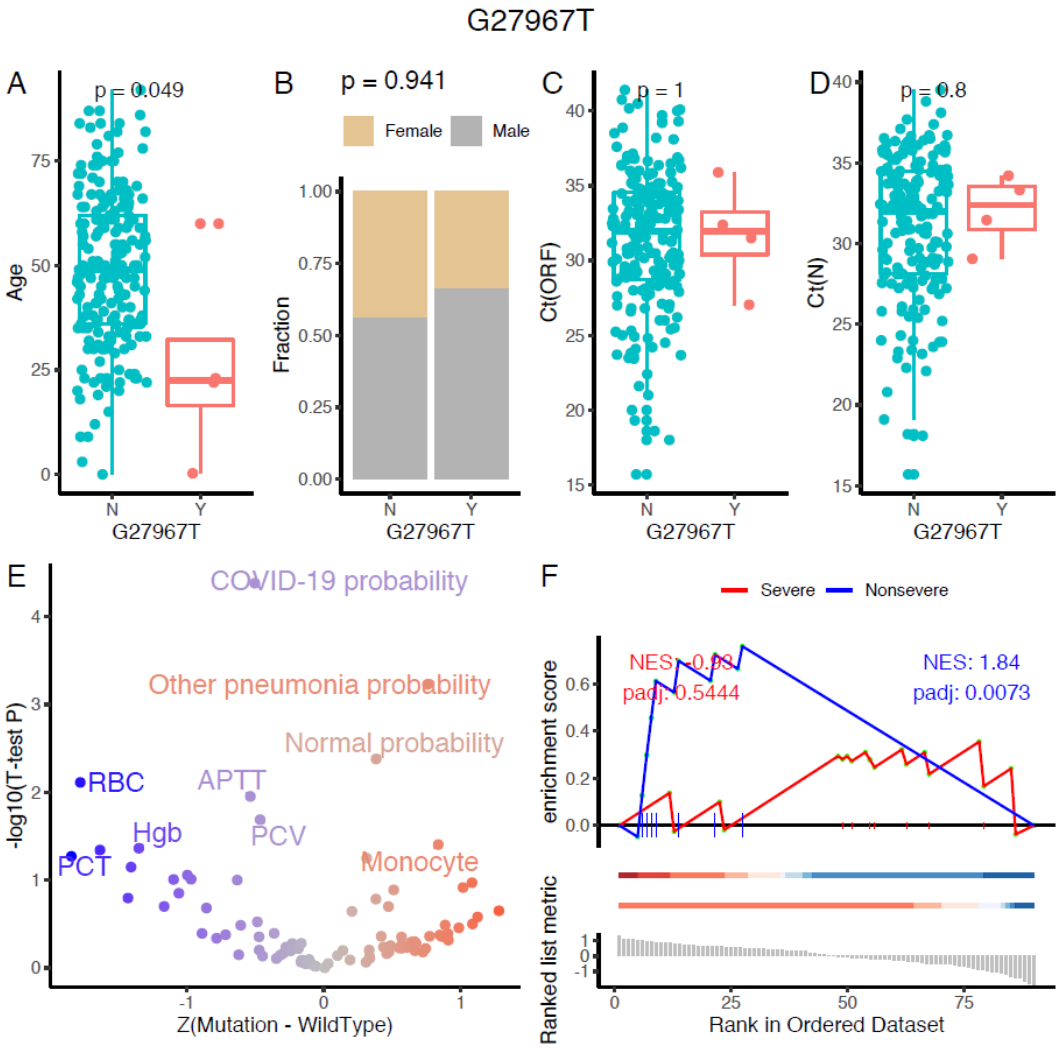
G11083T is a missense variant of Nsp6 (QHD43415.1:p.3606L>F, predicted to have a moderate impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 11 samples (7.00%) in Hubei, 65 samples (5.00%) in Africa, 3193 samples (7.00%) in Europe, 864 samples (15.00%) in Asia, 24 samples (3.00%) in South America, 261 samples (4.00%) in Oceania and 809 samples (4.00%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.23), gender (**B**, chi-square test, p = 0.61), Ct values in ORF (**C**, Wilcox-test, p = 0.43) and N genes (**D**, Wilcox-test, p = 0.95) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the variant has a higher level of Mg, and has lower levels of NEUT, RBC, $CD3^+$ and $CD4^+$ T cell percentage in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the G11083T has a 1.91 NES in non-severe traits (padj = 2×10^{-3}) and a $-$1.06 NES in severe traits (padj = 0.3661, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

24. G28878A in N



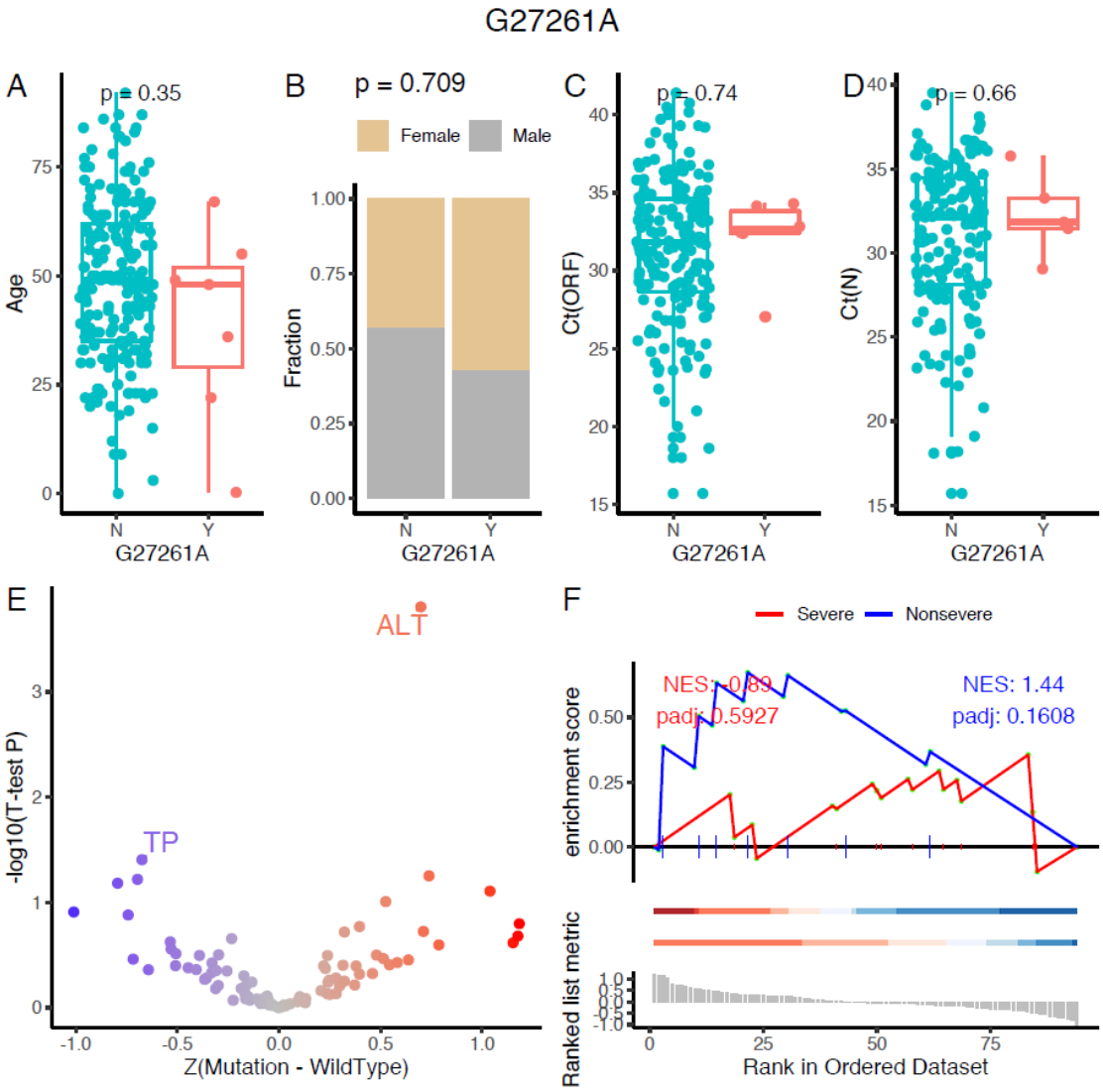
G28878A is a missense variant of N (QHD43423.2:p.202S>N, predicted to have a moderate impact on the protein), which has been found in 3 samples (3.00%) in Sichuan, 36 samples (3.00%) in Africa, 106 samples (0.23%) in Europe, 191 samples (3.00%) in Asia, 16 samples (0.23%) in Oceania and 61 samples (0.31%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.62), gender (**B**, chi-square test, p = 0.596), Ct values of ORF (**C**, Wilcox-test, p = 0.55) and N genes (**D**, Wilcox-test, p = 0.31) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the missense variant has higher levels of IL-17A, IL-1b, RBC-SD, MCH, CHOL and Hgb, and has lower levels of CRP and BASO in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the G28878A has a 0.63 NES in non-severe traits (padj = 0.9198) and a –2.09 NES in severe traits (padj = 4.3×10^{-3}, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

25. G27967T in ORF8



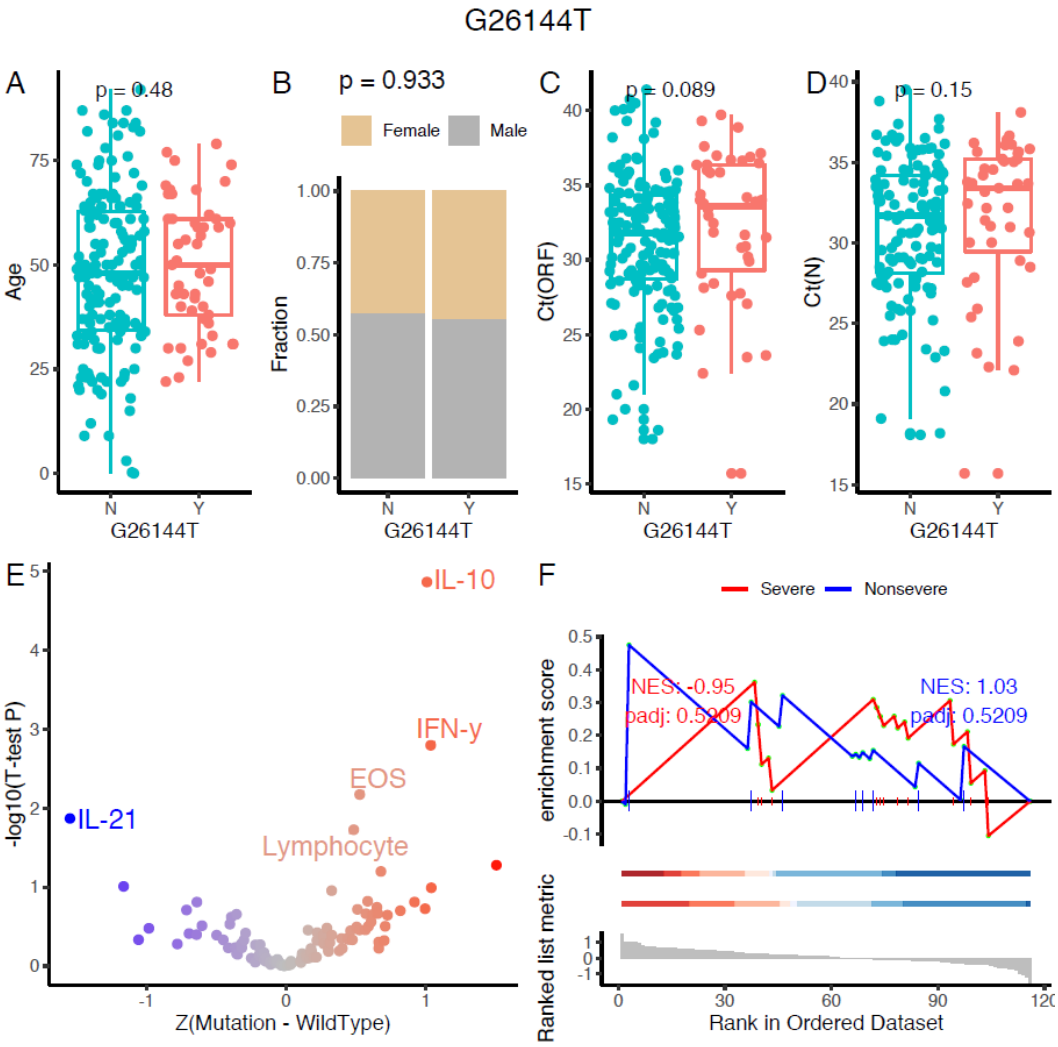
G27967T is a missense variant of ORF8 (QHD43422.1:p.25C>F, predicted to have a moderate impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 1 sample (0.62%) in Hubei, 2 samples (0.15%) in Africa and 72 samples (0.15%) in Europe. The group with or without this variant showed no difference in gender (**B**, chi-square test, p = 0.941), Ct values of ORF (**C**, Wilcox-test, p = 1) and N genes (**D**, Wilcox-test, p = 0.8) based on qPCR, but has a significant difference in age (**A**, Wilcox-test, p = 0.049). We further compared the 117 clinical phenotypes between the two groups, the group with the missense variant has higher levels of monocyte, other pneumonia probability and normal probability, and has lower levels of RBC, PCT, Hgb, APTT, PCV and COVID-19 probability in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the G27967T has a 1.84 NES in non-severe traits (padj =7.3×10^{-3}) and a –0.93 NES in severe traits (padj = 0.5444, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

26. G27261A in ORF6



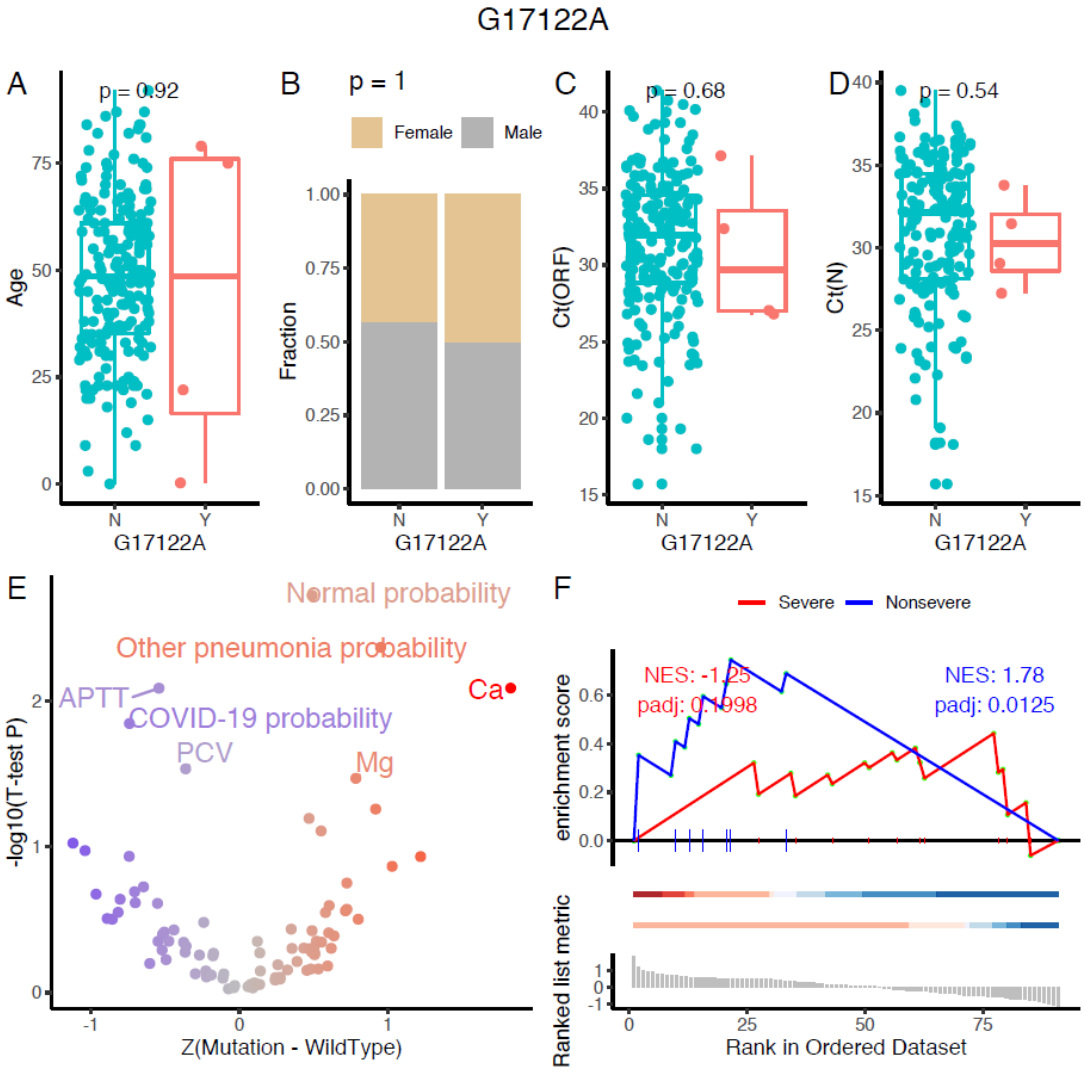
G27261A is a synonymous variant of ORF6 (QHD43420.1:p.20R, predicted to have a low impact on the protein), which has been found in 3 samples (3.00%) in Sichuan, 1 sample (0.07%) in Africa, 1 sample (0.00%) in Europe and 1 sample (0.02%) in Asia. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.35), gender (**B**, chi-square test, p = 0.709), Ct values of ORF (**C**, Wilcox-test, p = 0.74) and N genes (**D**, Wilcox-test, p = 0.66) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the synonymous variant has a higher level of ALT, and has a lower level of TP in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the G27261A has a 1.44 NES in non-severe traits (padj = 0.1608) and a –0.89 NES in severe traits (padj = 0.5927, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

27. G26144T in ORF3a



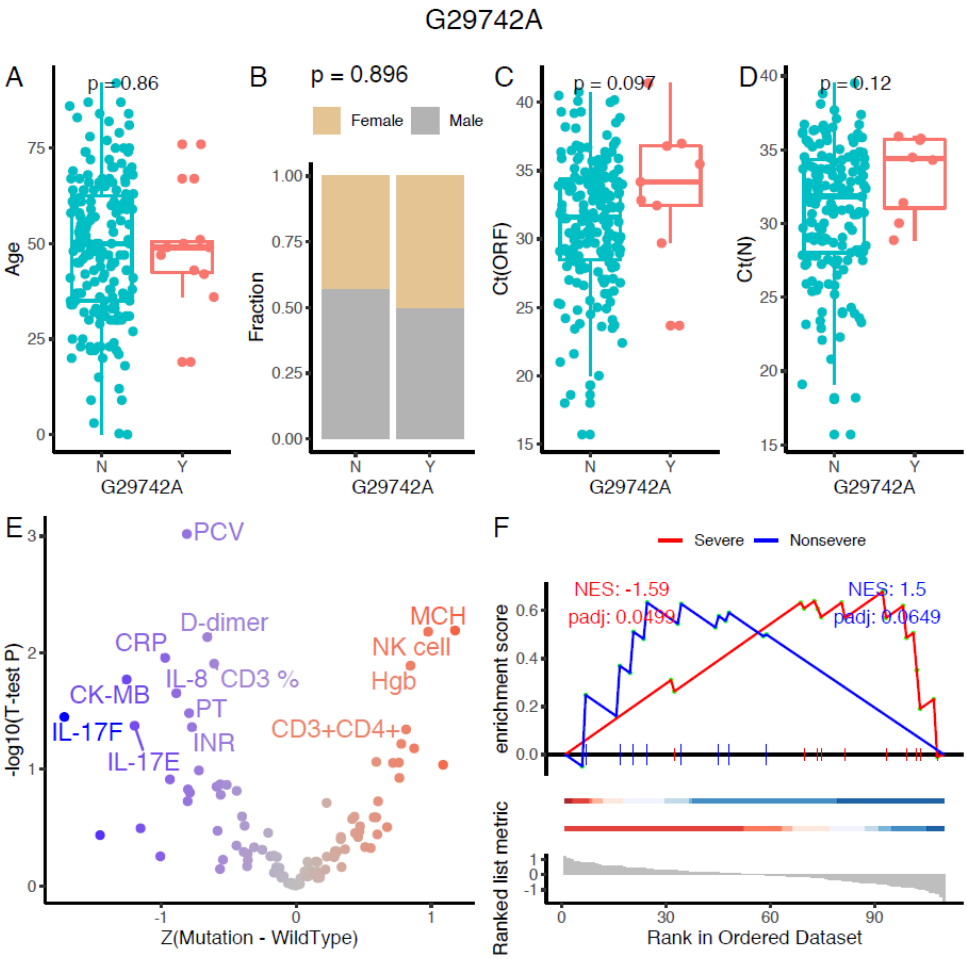
G26144T is a missense variant of ORF3a (QHD43417.1:p.251G>V, predicted to have a moderate impact on the protein), which has been found in 4 samples (5.00%) in Sichuan, 15 samples (9.00%) in Hubei, 7 samples (0.51%) in Africa, 946 samples (2.00%) in Europe, 230 samples (4.00%) in Asia, 10 samples (1.00%) in South America, 114 samples (2.00%) in Oceania and 287 samples (1.00%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.48), gender (**B**, chi-square test, p = 0.933), Ct values of ORF (**C**, Wilcox-test, p = 0.089) and N genes (**D**, Wilcox-test, p = 0.15) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the missense variant has higher levels of IL-10, IFN-γ concentration in serum, EOS and Lymphocyte, and has a lower level of IL-21 in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the G26144T has a 1.03 NES in non-severe traits (padj = 0.5209) and a –0.95 NES in severe traits (padj = 0.5209, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

28. G17122A in Nsp13



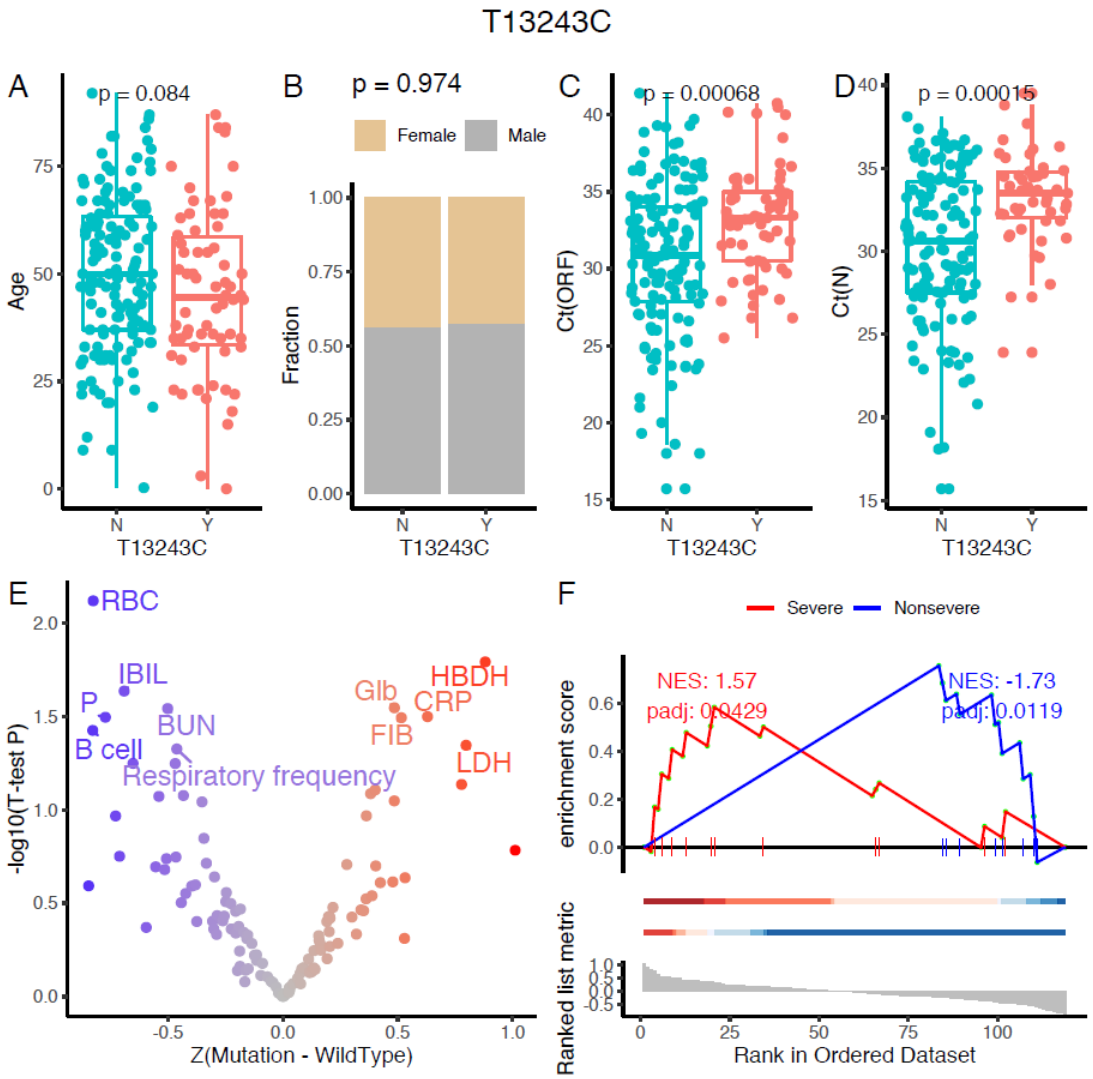
G17122A is a missense variant of Nsp13 (QHD43415.1:p.5620A>T, predicted to have a moderate impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 11 samples (0.02%) in Europe and 5 samples (0.09%) in Asia. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.92), gender (**B**, chi-square test, p = 1), Ct values of ORF (**C**, Wilcox-test, p = 0.68) and N genes (**D**, Wilcox-test, p = 0.54) based on qPCR. We further compared the 117 clinical phenotypes between the two groups, the group with the missense variant has higher levels of Ca, Mg, Normal probability and Other pneumonia probability, and has lower levels of PCV, COVID-19 probability and APTT in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the G17122A has a 1.78 NES in non-severe traits (padj = 0.0125) and a –1.25 NES in severe traits (padj = 0.1998, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

29. G29742A in downstream region



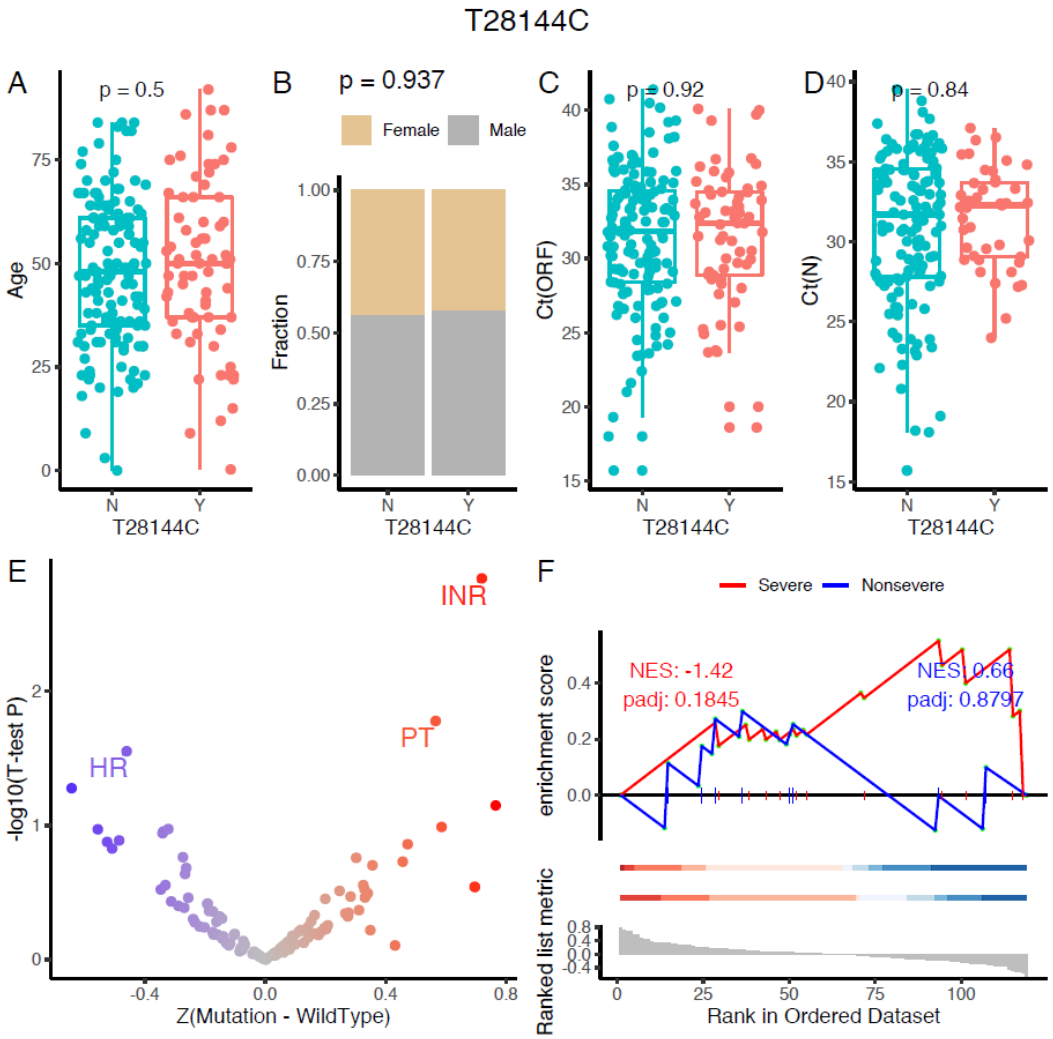
G29742A is a downstream variant in 3'UTR (QHI42199.1, predicted to be “modifier”), which has been found in 4 samples (5.00%) in Sichuan, 32 samples (2.00%) in Africa, 115 samples (0.24%) in Europe, 184 samples (3.00%) in Asia, 6 samples (0.09%) in Oceania and 82 samples (0.42%) in North America. The group with or without this deletion showed no difference in age (**A**, Wilcox-test, p=0.86), gender (**B**, chi-square test, p = 0.896) and the Ct values of ORF (**C**, Wilcox-test, p = 0.097) and N gene (**D**, Wilcox-test, p = 0.12). When comparing the clinical phenotypes between the two groups, the group with the variant has higher levels of MCH, NK cell, Hgb, $CD3^+$ $CD4^+$ T cell counts in the peripheral bloods, and has lower levels of PCV, D-dimer, CRP, CK-MB, PT, INR, IL-17F, IL-17E and CD3 T cell percentage in the peripheral bloods in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the G29742A has a 1.5 NES in non-severe traits (padj = 0.0649) and a ‑1.59 NES in severe traits (padj =0.0499, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

30. T13243C in Nsp10



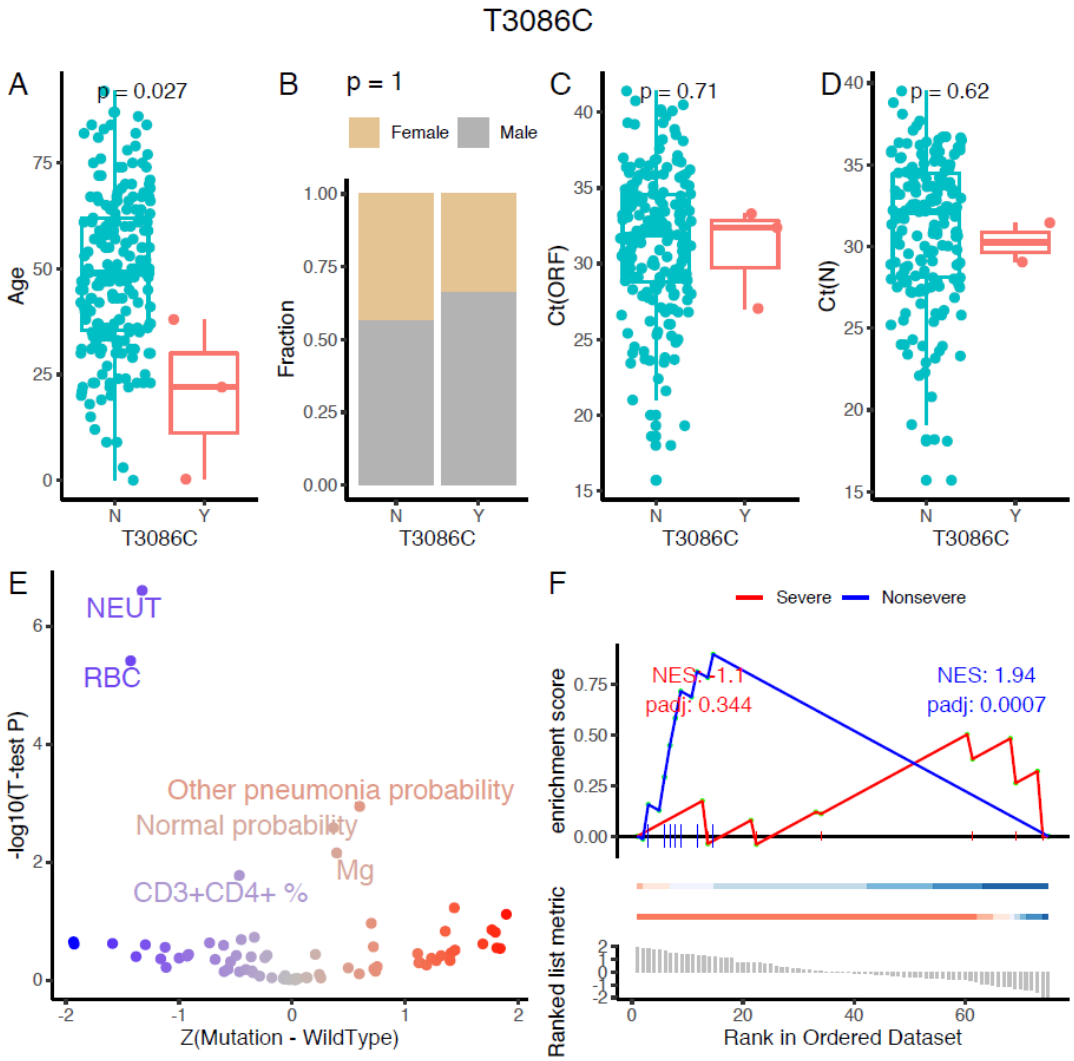
T13243C is a synonymous variant in Nsp10 (QHD43415.1:p.4326C, predicted to have a low impact on the protein), which has only been found in China, with 9 samples (10.00%) in Sichuan and 24 samples (15.00%) in Hubei. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.084), gender (**B**, chi-square test, p = 0.974), while significant higher Ct values of ORF (**C**, Wilcox-test, p = 6.8×10^{-4}) and N genes (**D**, Wilcox-test, p = 1.5×10^{-4}) based on qPCR were observed. When comparing the clinical phenotypes between the two groups, the group with the variant has higher levels of HBDH, CRP, Glb and LDH, and has lower levels of RBC, IBIL, P, BUN, B cell counts and respiratory frequency in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, the T13243C has a −1.73 NES in non-severe traits (padj =0.0119) and a 1.57 NES in severe traits (padj = 0.0429, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

31. T28144C in ORF8



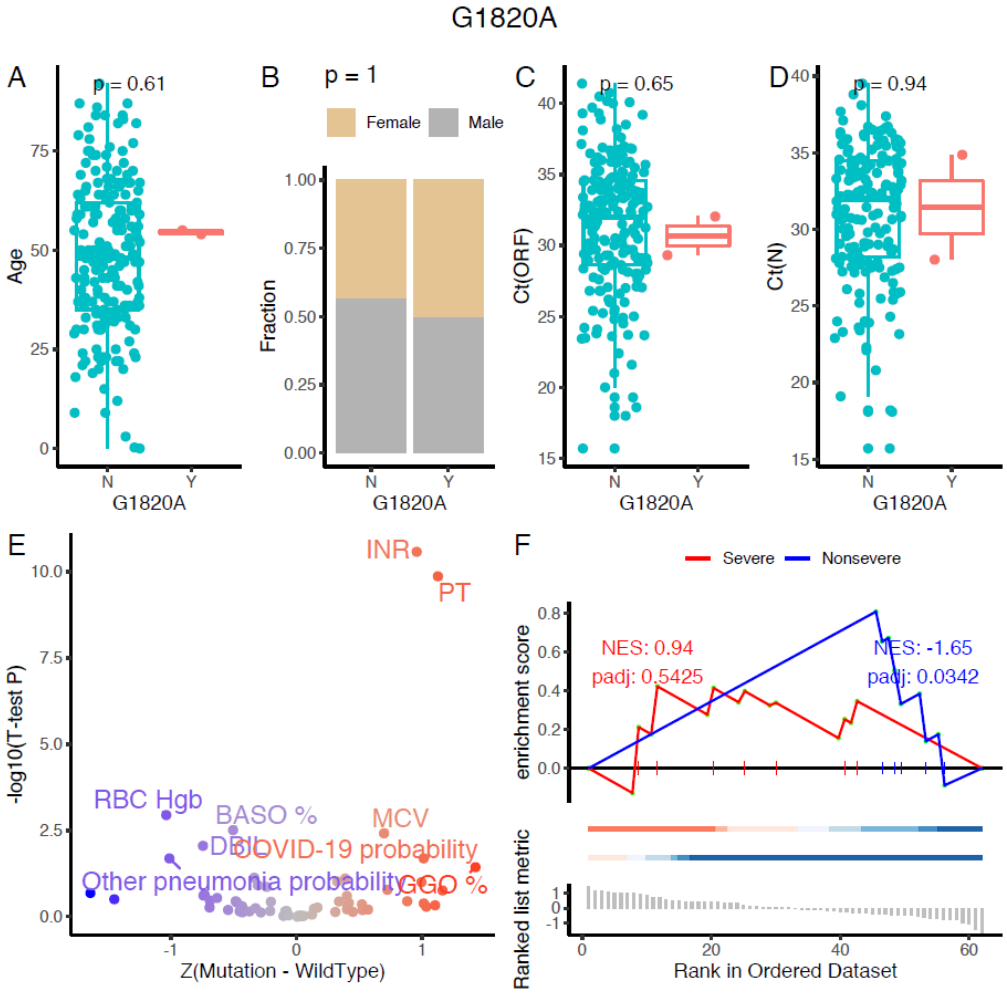
T28144C is a missense variant in ORF8 (QHD43422.1:p.84L>S, predicted to have a moderate impact on the protein), which has been found in 29 samples (33.00%) in Sichuan, 25 samples (16.00%) in Hubei, 35 samples (3.00%) in Africa, 410 samples (0.87%) in Europe, 547 samples (9.00%) in Asia, 28 samples (3.00%) in South America, 247 samples (4.00%) in Oceania and 2022 samples (10.00%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.5), gender (**B**, chi-square test, p = 0.94) and the Ct values of ORF (**C**, Wilcox-test, p =0.92) and N gene (**D**, Wilcox-test, p = 0.84). When comparing the clinical phenotypes between the two groups, the group with the variant has higher levels of INR and PT, and has lower levels of HR in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, has a 0.66 NES in non-severe traits (padj = 0.8797) and a −1.42 NES in severe traits (padj = 0.1845, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

32. T3086C in Nsp3



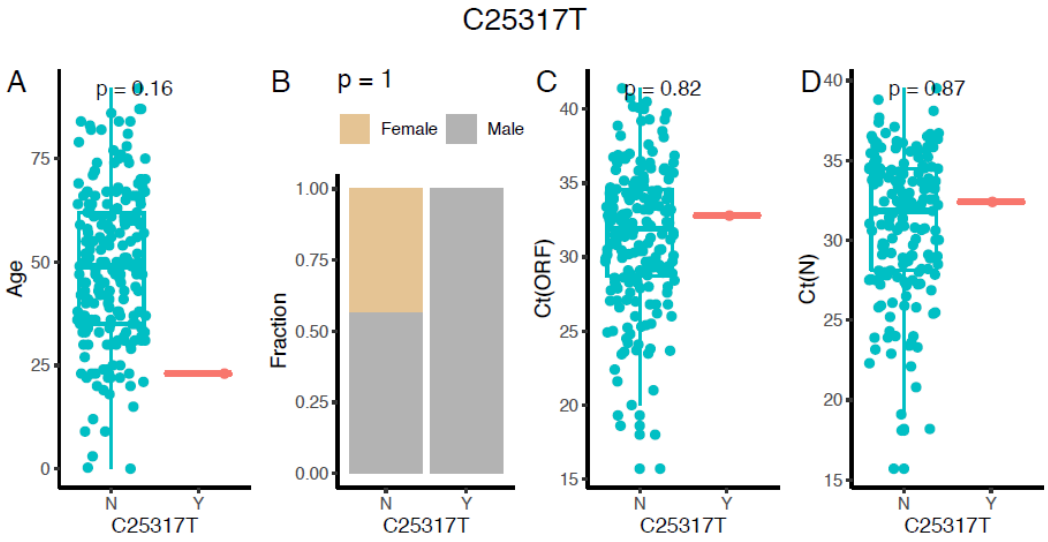
T3086C is a missense variant in Nsp3 (QHD43415.1:p.941F>L, predicted to have a moderate impact on the protein), which has been found in 2 samples (2.00%) in Sichuan and 5 samples (0.09%) in Asia. The group with or without this variant showed a significant difference in age (**A**, Wilcox-test, p = 0.027) and no difference in gender (**B**, chi-square test, p = 1) and the Ct values of ORF (**C**, Wilcox-test, p = 0.71) and N gene (**D**, Wilcox-test, p = 0.62). When comparing the clinical phenotypes between the two groups, the group with the variant has higher levels of other pneumonia probability, normal probability and Mg, and has lower levels of NEUT, RBC, $CD3^+$ and $CD4^+$ T cell percentage in the peripheral bloods in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, T3086C has a 1.94 NES in non-severe traits (padj = 7×10^{-4}) and a −1.11 NES in severe traits (padj = 0.344, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

33. G1820A in Nsp2



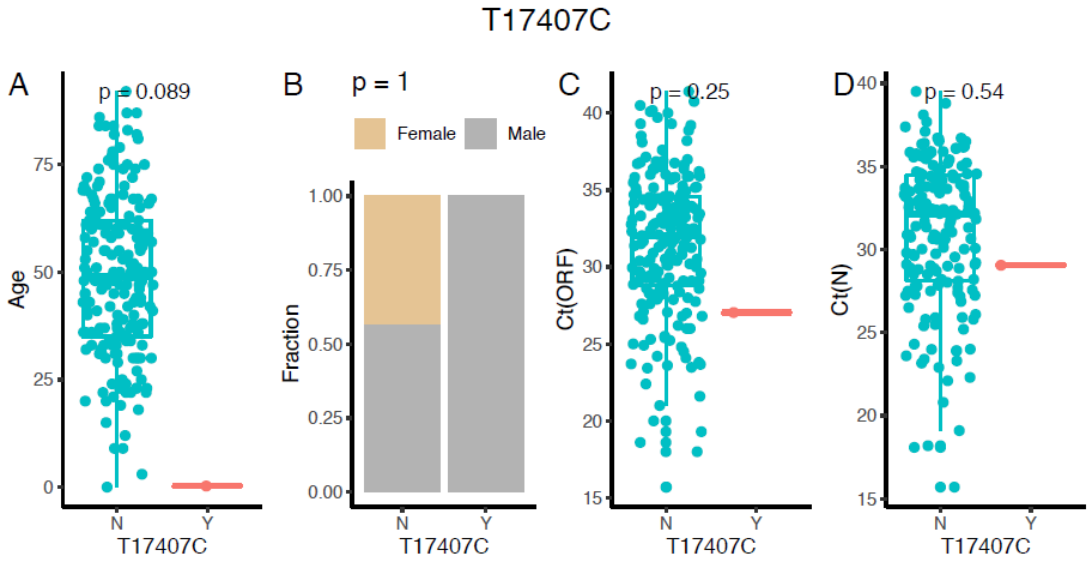
G1820A is a missense variant in Nsp2 (QHD43415.1:p.519G>S, predicted to have a moderate impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 4 samples (0.29%) in Africa, 156 samples (0.33%) in Europe, 32 samples (0.55%) in Asia, 4 samples (0.06%) in Oceania and 28 samples (0.14%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.61), gender (**B**, chi-square test, p = 1) and the Ct values of ORF (**C**, Wilcox-test, p = 0.65) and N gene (**D**, Wilcox-test, p = 0.94) based on qPCR. When comparing the clinical phenotypes between the two groups, the group with the variant has higher levels of INR, PT, MCV GGO% and COVID-19 probability, and has lower levels of RBC, Hgb, BASO percentage, DBIL and other pneumonia probability in volcano plot (**E**). The Y axis is the –log10 (P value) of traits, the traits with P value less than 0.05 were shown. The X axis is the difference of Z score between the two groups. In the enrichment analysis, G1820A has a −1.65 NES in non-severe traits (padj = 0.0342) and a 0.94 NES in severe traits (padj = 0.5425, **F**). NES, normalized enrichment score. P value were adjusted using permutation.

34. C25317T in S



C25317T is a missense variant of S gene (QHD43416.1:p.1252S>F, predicted to have a moderate impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 7 samples (0.01%) in Europe, 2 samples (0.22%) in South America and 3 samples (0.02%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.16), gender (**B**, chi-square test, p = 1) and the Ct values of ORF (**C**, Wilcox-test, p = 0.82) and N gene (**D**, Wilcox-test, p = 0.87) based on qPCR. Due to the only one sample with clinical information, the comparison of clinical phenotypes and enrichment analysis were not performed.

35. T17407C in Nsp13



T17407C is a synonymous variant in Nsp13 (QHD43415.1:p.5715L, predicted to have a low impact on the protein), which has been found in 2 samples (2.00%) in Sichuan, 1 sample (0.00%) in Europe, and 1 sample (0.01%) in North America. The group with or without this variant showed no difference in age (**A**, Wilcox-test, p = 0.089), gender (**B**, chi-square test, p = 1) and the Ct values of ORF (**C**, Wilcox-test, p = 0.25) and N gene (**D**, Wilcox-test, p = 0.54) based on qPCR. Due to the only one sample with clinical information, the comparison of clinical phenotypes and enrichment analysis were not performed.