nature portfolio

Corresponding author(s): DAPR, COMMSBIO-24-7574B

Last updated by author(s): Jun 16, 2025

# Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed |
|---|---|
| ☐ | ☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ A description of all covariates tested |
| ☐ | ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | No software was used for data collection. The study utilized next-generation sequencing (NGS) data, including both DNA methylation and RNA-seq data, which were independently generated in-house. DNA methylation data were generated using the MGI DNBSEQ-T7RS platform, and RNA-seq data were produced using the Illumina NovaSeq 6000 platform. |
| Data analysis | Data analysis was performed using the following software and tools: Fastp (v0.23.1), FastQC (v0.11.9), Bismark (v0.23.1), methylKit R package (v1.20.0), annotatr R package (v1.20.0), STAR (v2.7.10b), RSEM (v1.3.3), sva R package (v3.54.0), DESeq2 R package (v1.38.3), and ShinyGO (v0.82). The FASTQ files were mapped to the human reference genome (GRCh38.p13) provided by GENCODE version 42. Custom scripts developed for data processing and statistical analysis have been made publicly available on GitHub and are referenced in the 'Code Availability' section of the manuscript. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All next-generation sequencing (NGS),including both DNA methylation and RNA-seq data,used in this study were generated in-house. Details regarding data availability and access are provided in the 'Data Availability' section of the manuscript. Raw sequencing data and materials used in the study are available from the corresponding author upon request.

## Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

| | |
|---|---|
| Reporting on sex and gender | Human participants of all sexes were included in this project. The self-reported sex of the subjects, recorded as either male or female, is provided in the baseline characteristics. No subjects refused to disclose their gender identity. Our analysis assumes the self-reported sex information corresponds to the biological sex of the participants. |
| Reporting on race, ethnicity, or other socially relevant groupings | This study was conducted in the Republic of Korea and included only participants of Korean nationality. Variables related to race, ethnicity, or other socially defined groupings were not collected, as they were not applicable in the study context. However, sampling bias was inevitable due to the geographical settings and the socio-cultural diversity inherent to the population. |
| Population characteristics | Human participants of the current project are self-reported as descendants of a single ancestry across all age groups. |
| Recruitment | Participants were recruited through informed consent obtained during hospitalization. All individuals were enrolled after providing written consent in accordance with the approved institutional ethics protocols. |
| Ethics oversight | The study complies with the ethical guidelines and regulations set forth by the Institutional Review Board (IRB), as detailed in the Materials and methods section in manuscript. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No statistical tests were performed to pre-determine the sample size during the study design. Instead, all samples that met quality control requirements were included in downstream analyses. A total of 46 longitudinal whole blood samples from hospitalized COVID-19 patients (across different infection phases), 90 convalescent samples, and 344 healthy control samples were used in this study. To mitigate potential bias due to the relatively small sample size in the hospitalized group, a leave-one-out (LOO) approach was applied in downstream analyses to ensure robust and precise biomarker identification. The number of human individuals included in each analysis is clearly indicated in the Online Methods and figure legends. |
| Data exclusions | All samples that failed to pass quality check requirements were excluded. |
| Replication | Due to the nature and scope of the study, no independent replication analysis was conducted. However, internal validation was performed using cross-validation and leave-one-out approaches to support the robustness of the findings. |
| Randomization | Randomization was not applicable in this observational study using clinical samples. Instead, covariates such as age and sex were included in the statistical models to adjust for potential confounding effects. |
| Blinding | Blinding was not performed in this study, as sample group information (e.g., severity classification) was required during data processing and analysis. However, all analyses were conducted following predefined criteria to minimize bias. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology and archaeology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |
| ☒ | ☐ Plants |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Plants

| | |
|---|---|
| Seed stocks | N/A |
| Novel plant genotypes | N/A |
| Authentication | N/A |