nature portfolio

Corresponding author(s): Peter A. Nigrovic

Last updated by author(s): June 08, 2024

# 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 | There is no use of commercial, open code, and custom code in this study. |
| Data analysis | SAS 9.4 was used for the extraction of our sequencing data. All statistical analysis was done using GraphPad software (Prism 9) using one-way analysis of variance (ANOVA) for comparison of multiple groups, or unpaired t-test (non-parametric Mann-Whitney test) for two groups. |

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

Source data are provided with this paper. All data supporting the findings of this study are available within the paper and its Supplementary Information. The Human reference genome utilized in this study was sourced from the genome assembly GRCh37 (https://www.ncbi.nlm.nih.gov/datasets/genome/

GCF_000001405.13/). The mass spectrometry proteomics data have been deposited to the ProteomeXchnage Consortium via the PRIDE[70] partner repository (https://www.ebi.ac.uk/pride/archive/projects/PXD048367) with the dataset identifier PXD048367 and 10.6019/PXD048367.

## 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 | This study involved 35 genotyped subjects, including 12 males and 23 females by self-declared sex. Sex was not considered in the recruitment of subjects or the choice of cell lines for in vitro manipulation. |
| Reporting on race, ethnicity, or other socially relevant groupings | In our research, we did not catergorize or analyze participant based on any socially constructed or socially relevant variables. |
| Population characteristics | Healthy subjects aged between 23 and 64 with 12 C/C, 11 C/G, and 12 G/G genotypes for SNP rs2297550 were recruited, excluding individuals who diagnosed with autoimmune and/or immune-mediated diseases. |
| Recruitment | 35 healthy participants were recruited from genotyped volunteer donors within the Mass General Brigham Biobank through the recruitment center of the Joint Biology Consortium (www.jbcwebportal.org). Each participant was compensated with a gift-card, and their sex was determined based on self-report. |
| Ethics oversight | This study was approved under Brigham and Women's Hospital IRB protocol 2019P003709 and the Boston Children's Hospital IRB protocol P000043108. All healthy participants provided informed consent. Blood samples from healthy controls were obtained from blood bank leukoreduction collars without written consent and were not employted for genomic studies. |

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 | • We did not perform a sample size calculation for SNP-seq. We chose three biological replicates based on their high correlation (⍰=0.89, P<1×10-15) observed between sample pairs and between PBMC and BL2 sample. This high correlation ensured reliability in our results, leading to the identification of 248 SNPs.<br>• We did not perform a sample size calculation for EMSA. We used three biological replicates based on their demonstrated successful reproducibility.<br>• We performed luciferase reporter assay in eight biological replicates without conducting a sample size calculation. To mitigate the risk of false positive errors, we applied statistical analysis using Mann-Whitney two-tailed U-test with two-stage step-up correction for multiple hypothesis testing.<br>• We compared RNA and protein levels among C/C (n=4), C/G (n=3), and G/G (n=3) based edited clones of Daudi without conducting a sample size calculation. The statistical difference of RNA levels was evaluated using one-way ANOVA with Tukey's multiple comparisons correction. Protein levels showed comparable trends to the RNA levels.<br>• To measure IKKe expression using RT-qPCR and flow cytometry, we collected a total of 35 samples, comprising 12 C/C, 11 C/G, and 12 G/G donors. With a sample size of 35, we estimated that we would have over 90 % statistical power for linear regression analysis between genotype groups, assuming a large effect size (https://bwhbioinfo.shinyapps.io/powerEQTL/). The sample size for each experiment was specified in the corresponding figure and method section.<br>• We performed pulldown and silver staining without conducting a sample size calculation. This experiment was repeated twice to ensure reproducibility.<br>• We performed EMSA-supershift without conducting a sample size calculation. This experiment was repeated twice to ensure reproducibility.<br>• We performed ChIP-qPCR in wild-type and CRISPR-HDR edited C/G Daudi cells without conducting a sample size calculation. We utilized three biological replicates that exhibited significant difference of Ikaros binding between groups.<br>• We performed pulldown and western blotting once without conducting a sample size calculation. This experiment result was confirmed using complementary approaches including EMSA-supershift, and ChIP-qPCR.<br>• We performed RT-qPCR, western blotting, and ChIP-qPCR in Ikaros knockout Daudi clones once without conducting a sample size calculation. We utilized three knockout clones and two wild-type clones that displayed significant difference of Ikaros, IKBKE/IKKe, and Ikaros binding. |
| Data exclusions | No data were excluded from the analyses. |
| Replication | • SNP-seq was performed using nuclear extracts from PBMCs of three different healthy individuals. The same assay conducted in parallel using nuclear extracts from BL2 B cells, each tested in triplicates. All replicates were successfully conducted.<br>• EMSA for 52 SNPs was confirmed with three replicates using distinct nuclear extracts. We selected SNP that were replicated in at least two replicates (supplementary figure 5). |

• Luciferase reporter assay was validated with eight replicates per allele of candidate SNPs in Daudi B cells, performed once.
• We successfully tested the differences in RNA and protein levels in base-edited Daudi clones, which included 4 C/C, 3 C/G, and 3 G/G clones, in a single experiment.
• Association analysis between SNP rs2297550 and IKKe expression by flow cytometry was successfully conducted once, using 35 healthy individuals with genotypes 12 C/C, 11 C/G, and 12 G/G.
• Pulldown and silver staining for eight SNPs was performed with two replicates using distinct nuclear extracts. Four of the SNPs showed differential binding.
• We repeated EMSA-supershift twice. Both replicates successfully demonstrated Ikaros binding to the G allele of rs2297550 in Daudi cells.
• We successfully identified Ikaros allele-specific binding to rs2297550 in wild-type and CRISPR-HDR edited C/G Daudi cells using ChIP-qPCR with three biological replicates, performed once.
• We successfully demonstrated Ikaros binding to the G allele of rs2297550 in Daudi using pulldown and western blotting, performed once.
• We successfully demonstrated altered expression of Ikaros protein, IKBKE RNA, and IKKe protein, as well as altered binding of Ikaros in thee Ikaros knockout clones and two wild-type controls, performed once.

Randomization

• For SNP-seq, EMSA, luciferase assay, Pulldown and silver staining/western blotting, ChP-qPCR, RT-qPCR, and Western blotting, the allocation of samples into experimental groups was based on their biological replicates rather than randomization. As these assays focused on biological replicates, the concept of randomization was not applicable. To control for potential covariates, we ensured that the biological replicates were handled and processed under identical conditions.
• Randomization was not applicable for the mass spectrometry analysis to identify proteins.
• CRISPR-HDR edited Daudi clones were separated into groups based on their modified genotype at rs2297550.
• To assess the association of genotype with IKKe expression, human samples were separated into groups based on their genotype for the SNP rs2297550.

Blinding

This study focused on assessing the impact of genetic variation at SNP rs2297550 on a pathway regulating IKKe expression. To achieve this, we recruited 35 healthy subjects based on their genotype from the MGB Biobank donors. The analysis of the experiment results involved grouping subjects by the genotype. Due to this specific approach, blinding was not applicable.

# 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 |

## Antibodies

Antibodies used

A list of antibodies for western blotting:
All antibodies were diluted at a ratio of 1:1000.
Anti-Ikaros antibody (Cell Signaling, clone D10E5, catalog number 9034S, validated by the manufacturer in Jurkat, Ramos, and Raji cells), anti-GAPDH antibody (Cell Signaling, clone 14C10, catalog number 2118S, validated by the manufacturer in NIH/3T3, C6, HUVEC, and L929 cells), anti-IKKε antibody (Cell Signaling, clone D20G4, catalog number 2905S, validated by the manufacturer in Ramos, and RL cells), anti-vinculin antibody (Bio-Rad, clone V284, catalog number MCA465GA, validated by the manufacturer in Hela cells), goat anti-Rabbit IgG (H+L) cross-adsorbed secondary antibody HRP (Thermo Scientific, catalog number G-21234), and goat anti-mouse IgG-HRP (Santa Cruz, catalog number sc-525408).

A list of antibodies for EMSA supershift:
10 µl of antibody was added to the 20 µl reaction.
anti-IKZF1 antibody (Cell Signaling, clone D10E5, catalog number 9034S, validated by the manufacturer in Jurkat, Ramos, and Raji cells), anti-E2F4 antibody (Cell signaling, clone E3G2G, catalog number 40291S, validated by the manufacturer in Jurkat, SEM, HCT15, HT1080, and COS-7 cells)

A list of antibodies for ChIP:
10 µl of antibody was incorporated into 200 µl ChIP reaction.
anti-Ikaros antibody (Cell Signaling, clone D10E5, catalog number 9034S, validated by the manufacturer in Jurkat cells)

A list of antibodies to measure IKKε expression in subsets of primary PBMCs:
All antibodies were diluted at a ratio of 1:200.
CD19-APC-e780 (eBioscience, clone HIB19, catalog number 47-0199-41, validated by the manufacturer in human PBMC); CD3-BV421 (Biolegend, clone UCHT1, catalog number 300433, validated by the manufacturer in human PBMC); CD4-PE-eFluor 610 (eBioscience, clone RPA-T4, catalog number 61-0049-41, validated by the manufacturer in human PBMC); CD8-BV786 (eBioscience, clone RPA-T8,

| | |
|---|---|
| | catalog number 417-0088-42, validated by the manufacturer in human PBMC); CD56-PE-Cy7 (eBioscience, clone NCAM, catalog number 25-0567-41, validated by the manufacturer in human PBMC); CD14-APC (eBioscience, clone 61D3, catalog number 17-0149-41); anti-IKKε-PE for intracellular staining. IKKε antibody (Cell Signaling, clone D20G4, catalog number 2905S) was labeled with PE (Biotium, catalog number 92299). IKKε-PE for intracellular staining was validated in this study by comparing it to background signal from an isotype control in human PBMCs. |
| Validation | Information on the application and validation of each primary antibody is described in the above section. |

## Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

| | |
|---|---|
| Cell line source(s) | BL2 was from DSMZ (catalog number ACC 625) and Daudi was provided by Dr. Baglaenko Yuriy. |
| Authentication | The cell lines used were not authenticated. |
| Mycoplasma contamination | The cell lines were not tested for mycoplasma contamination. |
| Commonly misidentified lines (See ICLAC register) | BL2 and Daudi are not commonly misidentified cell lines. |

## Plants

| | |
|---|---|
| Seed stocks | Not applicable |
| Novel plant genotypes | Not applicable |
| Authentication | Not applicable |

## Flow Cytometry

### Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

| | |
|---|---|
| Sample preparation | PBMCs were isolated from blood via Ficoll gradient. |
| Instrument | We used the BD LSRFortessa cell analyzer for flow cytometry. |
| Software | The flow cytometer was operated using the BD FACSDiva software and the flow cytometry data was analyzed using Flowjo software version 10.8.0. |
| Cell population abundance | Flow cytometry revealed that over 95% of the isolated PBMCs were classified within the lymphocyte gate. |
| Gating strategy | The positive region was identified by gating based on its fluorescence signal being greater than the background signal set by the negative isotype control.<br>We identified the major populations of PBMCs as follows: CD19+ B cells, CD3+CD4+ T cells, CD3+CD8+ T cells, CD3-CD56+ NK cells, and CD14+ monocytes. |

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.