nature portfolio

Corresponding author(s): Abdel-Wahab, Bolton

Last updated by author(s): Dec 22, 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 *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 *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	NovaSeq6000, NovaSeqX Plus, FACS Diva Software flow cytometry sorting (BD Biosciences), Microsoft Excel 2008.
Data analysis	ArCH v2.2.0, R v4.3.2, Mutect2 v4.2.0.0, VarDictJava v1.8.2, LoFreq2 v2.1.5, bcftools v1.15, 10x Genomics CellRanger v8.0.0, Seurat v5.1.0, scDblFinder v1.19.0, Cellmarker 2.0, ChatGPT4o, SCpubr v2.0.2 The custom analysis code used for the scRNA can be found here: https://github.com/pangxueyu233/CDK4-6-inhibition-mitigates-clonal-hematopoiesis The custom analysis code used for the Figures in this manuscript can be found here: https://github.com/kbolton-lab/CDK46_CH

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

Data were deposited and are available through Gene Expression Omnibus (GEO) accession number GSE285875 and Sequence Read Archive (SRA) access number PRJNA1270003

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	The findings in this paper apply to both biological sex. Patient recruitment was randomized to account for both sex where applicable. Sex was self-reported. Additionally all statistical analyses were adjusted for sex where information was provided.
Reporting on race, ethnicity, or other socially relevant groupings	Race as well as ethnicity were both used as potentially confounding covariates in the analysis of this manuscript due to the potential differences in clonal hematopoiesis acquisition rates between different ethnic and racial groups. All reported race and ethnicity were self-reported.
Population characteristics	Age as well as smoking status were both used as potentially confounding covariates in the analysis of this manuscript due to clonal hematopoiesis being an age-related disease and further confounded by previous smoking histories. All reported ages and smoking status were self-reported by the participants of this study. Additionally, all statistical analyses were adjusted for age as well as smoking where information was provided and available.
Recruitment	Subjects were enrolled on four clinical trials of trilaciclib which are referenced in the manuscript. Written and informed consent was obtained from each patient or participant before initiation of all study procedures.
Ethics oversight	The original clinical studies patient data and study was designed and conducted in compliance with the principles of the Declaration of Helsinki and the Good Clinical Practice guidelines of the International Council for Harmonisation. The study protocol and all study subjects underwent informed consent through local investigational site institutional review boards. For the participants in the Healthy Controls and Panel of Normal samples, the study protocol was approved by the IRB of Washington University School of Medicine in St. Louis (IRB:201109148) and the IRB of BJC WEBB (IRB:202012143) respectively.

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	With 10 mice in each group, we have 80% power to detect a 20% difference in cell populations assuming an effect size ($\Delta\mu/\sigma$) of 2.0 between groups using a Wilcoxon rank-sum test at the 0.05 significance level. There were a total of 224 patients in total split between the four clinical trials (Small Cell Lung Cancer N=65, Colorectal Cancer N=125, Triple Negative Breast Cancer N=34). Untreated patients totaled 176 to match the same number of treated patients. No prior sample size calculation was made due to data availability being limited to the total number of participants recruited for each clinical trial. As many patients as possible were sampled to yield the highest coverage where possible.
Data exclusions	No data exclusions.
Replication	In vivo assays were done in biological triplicate to assess reproducibility with 10 mice per group in each arm of each experiment. Clinical data analysis was performed for four clinical trials with similar treatment regiments and windows to ensure associated signals were concurrent between different cancer types.
Randomization	Mice were randomized into different treatment groups used for all animal studies.

Blinding	Investigators performing in vivo treatments were blinded to treatment type/group (i.e. control versus drug treatment). Investigators performing the genetic analyses were blinded to the patient information including treatment status, demographic data, and identifiable characteristics throughout the duration of data collection and data processing until the statistical analyses.

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	The following anti-mouse antibodies were used for flow cytometry: APC-Cy7-CD45.1 (A20), PE-CD45.2 (104), FITC-CD11b (M1/70), APC-Gr-1 (R86-8C5), PerCP5.5-B220 (RA3-6B2), and BV711-CD3 (17A2). Stem/progenitor cells were identified with the following markers: pacific blue-Lineage (a cocktail of CD3e, Ly6G/Ly6C, CD11b, CD45R/B220, TER119), FITC-CD34 (RAM34), BUV395-CD117(2B8), AF700-Sca-1 (D7), BV711-CD48 (hm48-1), APC-CD150 (Q38-480), PE-CD16/32 (2.4G2) and APC-Cy7-CD45.1 (A20), and BV605-CD45.2 (104). The immunophenotypic cell surface antigens are: LT-HSC (Lin-cKIT+Sca1+CD48- CD150+); ST-HSC (Lin- cKIT+ Sca1+ CD48- CD150-); multipotent progenitors (MPP; Lin-cKIT+ Sca1+ CD48+ CD150-); LSK (Lin- cKIT+ Sca1+); granulocyte macrophage progenitors (GMP; Lin-cKIT+ Sca1- CD16/32+ CD34+); common myeloid progenitors (CMP; Lin- cKIT+ Sca1- CD16/32- CD34+); megakaryocyte-erythroid progenitors (MEP; Lin- cKIT+ Sca1- CD16/32- CD34-) following the manufacturer's protocol. DAPI (BioLegend, 422801, 1:1000) was used to exclude dead cells.
Validation	For commercially available antibodies, validation has been performed by the manufacturer and corresponding certificates of analysis are available at the manufacturer's website. Reactivity of primary antibodies used in flow cytometry assays was validated by the manufacturer in immunocytochemistry and frozen immunohistochemistry.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	8-10 week-old CD45.1+ mice s well as Vav-cre p53R172H/WT mice and Vav-cre p53 wild-type control mice (all of which were purchased from Jackson Lab).
Wild animals	Not applicable.
Reporting on sex	Female B6.SJL-Ptprca Pepcb/BoyJ mice were used as recipients.
Field-collected samples	Not applicable.
Ethics oversight	All mouse procedures were completed in accordance with the Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committees at MSKCC. All mouse experiments were performed in accordance with a protocol approved by the MSKCC Institutional Animal Care and Use Committees (13-04-003).

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

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*

Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Plants

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

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	Peripheral blood and bone marrow samples collected from transgenic mice and littermate controls were first lysed with ACK lysis buffer to remove red blood cells and washed with ice-cold phosphate-buffered saline. Cells were stained with monoclonal antibodies against cell surface markers in phosphate-buffered saline/2% bovine serum albumin for 30 minutes on ice.
Instrument	Fortessa II
Software	FlowJo version 10 software
Cell population abundance	Mouse cells were gated on by FSC/SSC followed by gating on DAPI- CD45+ cells to determine percentage of live CD45+ cells.
Gating strategy	Gating strategies for analysis of murine immune cell subsets in Figure 3c.

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