nature portfolio

Corresponding author(s): Ash A. Alizadeh

Last updated by author(s): 10/18/2023

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

DNA and bulk RNA sequencing libraries were prepared as described in the Methods section and sequenced on the Illumina HiSeq4000/X-ten or NovaSeq6000 platform using 2x150 bp paired end reads. Processing of sequencing data was performed using fastp (version 0.20.0), BWA (0.7.17-dev), samtools (versions 0.1.18, 1.3.1 and 1.8) and bedtools (versions 2.20.1 and 2.26.0). Digital error suppression and variant calling were performed using custom scripts as previously described (Newman, Nat Biotechnol 2016). Phased variant calling and monitoring was done as described in Kurtz et al., Nat Biotechnol 2021.

scGEX:
Illumina NovaSeq 26-10-10-90 (R1-I7-I5-R2) configuration, 150bp paired end reads

Data analysis

Statistical analyses were performed using Python (2.7.5) and R (version 3.6.1 or 4.2.1). The function numpy.histogram in Python was first used to convert TLEN data extracted from BAM files using samtools into histogram distributions. These distributions were then read using R, and converted into cfDNA fragmentomic features of interest. The following R libraries are needed for the analysis: : 'reshape2' (1.4.4), 'data.table' (1.14.0), 'MASS' (7.3.51.6), 'zoo' (1.8.8), 'DescTools' (0.99.36), 'gtools' (3.8.2), 'matrixStats' (0.58.0), 'glmnet' (4.0.2), 'survival' (3.2.3), 'ggplot2' (3.3.2), 'ComplexHeatmap' (2.4.2), and 'optparse' (1.6.6), 'discover' (0.9.4), 'DESeq2' (1.36.0), 'fgsea' (1.22.0), 'textmineR' (3.0.5) and 'MCMCpack' (1.6.3).
The language model analysis (LDA) of the cHL clustering analysis was done using the following R packages: 'textmineR' (3.0.5) and 'MCMCpack' (1.6.3).
The germline-free SNV classifier was done using the GradientBoostingClassifier from sklearn.ensemble. For the autoencoder component of the classifier, the Python library torch (2.0.1) was used.

FlowJo (V10.1) was used for flow data analysis and visualization.

scGEX analysis:
CellRanger v6.0.1 (10X Genomics)
R statistical computing environment v4.0.2
Seurat v4.0.3
patchwork v1.1.1
dplyr v.1.0.7

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

Sequencing data is available in the database of Genotypes and Phenotypes dbGaP (study accession phs003435.v1.p1) in compliance with European General Data Protection Regulation (GDPR) and US Health Insurance Portability and Accountability Act (HIPAA) as applicable. In addition, an online tool to assign new samples to the genetic subtypes H1 and H2 is available at https://hodgkin.stanford.edu.
The hg19 reference genome dataset (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.13/) was used for sequence alignments.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	The goal of this study was to provide a comprehensive landscape of genomic variants in classic Hodgkin lymphoma using ctDNA profiling. There was no restriction of sex or gender in the study enrollment. Sex was recorded as part of the clinical care since male sex is a known risk factor for Hodgkin lymphoma. Sex is therefore reported as part of our analyses, but no finding was restricted to either sex or gender. scGEX analyses: samples were obtained from an institutional biorepository, which contains deidentified patient samples; thus, sex is not a reportable variable from the data we are privy to for these samples.
Population characteristics	Study participants include patients with classic Hodgkin Lymphoma of all ages. Clinical and demographic details of the study cohort are provided within the Supplementary Note Tables of the paper. We included pediatric and adult patients diagnosed with Classic Hodgkin Lymphoma (cHL) with available plasma and/or tumor specimens in this study. All patients were treated at cancer centers across Europe and North America between 2011 and 2020. Patients treated in or outside of clinical trials at 3 cancer centers, namely Stanford (CA, USA), Bellinzona (CH) and Leuven (BE) were eligible. Patient characteristics and treatment regimens administered are summarized in Supplementary Note Tables 1-3 of the paper. Additionally, clinical trial collectives were included in this study. Pediatric patients enrolled in the Phase III trials AHOD1331 trial (NCT02166463) and cHOD17 (NCT03755804) were included. Pediatric patients were not considered for associations with outcome, yet included in genotyping analyses to characterize the full spectrum of the disease. Further, patients enrolled in the Phase III AHL2011 (NCT01358747), the Phase II BREACH trial (NCT02292979), a pilot study evaluating concurrent Pembrolizumab-AVD (pembrolizumab and doxorubicin, vinblastine, and dacarbazine, NCT03331341) as well as a Phase II study evaluating PVAB (Prednisone, Vinblastine, Doxorubicin and Bendamustine, NCT02414568) in elderly patients were included. Patient characteristics and treatment regimens administered are summarized in Supplementary Note Tables 1-3. scGEX analyses: lymph node excisional biopsy samples obtained at St. Jude Children's Research Hospital with cryopreserved cell suspension(s) available in an institutional biorepository were identified.
Recruitment	We included pediatric and adult patients diagnosed with cHL with available plasma and/or tumor specimens All patients were treated at cancer centers across Europe and North America between 2011 and 2020. Patients treated in or outside of clinical trials at 3 cancer centers, namely Stanford (CA, USA), Bellinzona (CH) and Leuven (BE) were eligible. Additionally, clinical trial collectives were included in this study. Pediatric patients enrolled in the Phase III trials AHOD1331 trial (NCT02166463) and cHOD17 (NCT03755804) were included. Further, patients enrolled in the Phase III AHL2011 (NCT01358747), the Phase II BREACH trial (NCT02292979), a pilot study evaluating concurrent Pembrolizumab-AVD (pembrolizumab and doxorubicin, vinblastine, and dacarbazine, NCT03331341) as well as a Phase II study evaluating PVAB (Prednisone, Vinblastine, Doxorubicin and Bendamustine, NCT02414568) in elderly patients were included. The sole Inclusion criteria were diagnosis of cHL and available plasma or tissue specimens. Based on these criteria, we do not expect the data to be affected by major biases such as self-selection bias.
Ethics oversight	Patients were consented either for observational studies approved by the local institutional review board (IRB) at Stanford University (Stanford, CA), the University of Bellinzona (CH), UC Leuven (BE) and St. Jude Children's Research Hospital (Memphis, TN), or IRB-approved protocols (i.e. NCT02166463, NCT03755804, NCT01358747, NCT02292979, NCT03331341, NCT02414568).

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 performed sample size analysis for our WES cohort. Based on a background mutation rate of 2.8/Mbp, we profiled n=119 cases to detect recurrent mutations with at least 10% frequency in the population with sufficient power (80%).
Data exclusions	Samples deemed to be technical failures during DNA isolation or library preparation were excluded.
Replication	We processed samples in batches over time, and validated our findings from early samples in the late samples (e.g., correlation between TMTV and ctDNA levels). No replicates were generated for primary patient specimens. The number of replicates generated for in vitro studies are provided in the respective Figure Legends.
Randomization	Randomization of individuals to different groups was not applicable in this study as there was no intervention.
Blinding	As there was no therapeutic intervention in this study, no blinding was performed.

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

Methods

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

Antibodies

Antibodies used

Human IL-4R alpha Antibody
Catalog #: MAB230, R&D Systems
Monoclonal Mouse IgG2A Clone # 25463
Dilution 1:100
'Detects human IL-4 R alpha in direct ELISAs and Western blots. In direct ELISAs, no cross-reactivity with recombinant human (rh) IL-9 R, recombinant mouse IL-4 R, rhIL-5 R alpha, rhIL-5 R beta, rhIL-13 R alpha 1, or rhIL-13 R alpha 2 is observed.'

P-Stat6 (Y641) (C11A12) Rabbit mAb
Catalog #: 9364S, Cell Signaling Technology
Dilution 1:1000
'Phospho-Stat6 (Tyr641) (C11A12) Rabbit mAb detects endogenous levels of Stat6 only when phosphorylated at Tyr641. This antibody does not cross-react with the corresponding phospho-tyrosine residues of other Stat proteins. Cross-reactivity has been observed with receptor tyrosine kinases including phosphorylated PDGF receptor that can lead to inapproriate membrane staining. Species reactivity is determined by testing in at least one approved application (e.g., western blot).'

Stat6 (D3H4) Rabbit mAb
Catolog #: 5397S, Cell Signaling Technology
Dilution 1:1000
'Stat6 (D3H4) Rabbit mAb recognizes endogenous levels of total Stat6 protein.Species reactivity is determined by testing in at least one approved application (e.g., western blot).'

GAPDH (14C10) Rabbit mAb
Catalog #: 2118S, Cell Signaling Technology
Dilution 1:20000
'GAPDH (14C10) Rabbit mAb detects endogenous levels of total GAPDH protein. Species reactivity is determined by testing in at least one approved application (e.g., western blot).'

IL-4Rα Antibody (H-4) Mouse mAb
Catalog #: sc-28361, Santa Cruz Biotechnology
Dilution 1:200
'Epitope Species
human
Product Reactivity/Detection
IL-4Rα
Species Reactivity/Detection
mouse, rat, human'

CD124 (IL-4Rα) PE/Cy7, clone G077F6
Catalog #: 355008, BioLegend
Dilution 1:25
'Verified Reactivity – Human
FC – Quality tested'

Stat6 (pY641) AF647, clone 18/P-Stat6
Catalog #: 562079 (612601 is the 50 test catalog number of the same clone), BD Biosciences
Dilution 1:25
'Human (QC Testing). Intracellular staining (flow cytometry) (Routinely Tested)'

Human Fc Block
Catalog #: 14-9161-73, Invitrogen
Dilution 1:100
'Species Reactivity – Human.'

Anti-Rabbit IgG (H+L), HRP Conjugate
Catalog #: W4011, Promega
Dilution 1:10000
'Applications – Western Blotting'

Anti-Mouse IgG (H+L), HRP Conjugate
Catalog #: W4021, Promega
Dilution 1:10000
'Applications – Western Blotting'

scGEX FACS:
Unlabled antibodies:
CD2 clone RPA2.10; catalog# 300202 BioLegend
CD54 clone HA58; catalog# 353102 BioLegend
CD58 clone TS2/9; catalog# 330902 BioLegend
CD11a/CD18 clone m24; catalog# 363402, BioLegend

Labeled antibodies:
CD31 APC/Cy7 clone WM59; catalog# 303119, BioLegend
CD326 APC/Cy7 clone 9C4; catalog# 324245, BioLegend
CD45 BV785 clone HI30; catalog# 304048,BioLegend
CD90(Thy1) PerCP/Cy 5.5 clone 5E10; catalog# 328117, BioLegend
CD49e PE clone NKI-SAM-1; catalog# 328009, BioLegend
CD20 AF700 clone 2H7; catalog# 3020322, BioLegend
CD15(SSEA-1) FITC clone HI98; catalog# 301904, BioLegend
CD30 PE/Cy7 clone BY88; catalog# 333918, BioLegend
CD40 PE-Dazzle 594 clone 5C3; catalog# 334342, BioLegend
CD95(Fas) APC clone DX2; catalog# 305612, BioLegend
CD64 BV421 Clone 10.1; catalog# 305020, BioLegend and 562872, BD Biosciences
CD5 BV605 Clone L17F12; catalog# 364019, BioLegend

Viability:
Tonbo Ghost Dye Violet 510; catalog# 13-0870-T100, Cytek Biosciences

Validation

All antibodies used in this study are commercially available. Antibodies used in a specific species or application have been appropriately validated by manufacturers for that application and this information is provided on their website and product information datasheets.

For those antibodies used in immunoblotting, specificity was validated by detection of the correct molecular weight and non-specific detection was minimized through optimizing concentration and blocking reagent (milk or BSA in 0.5% TBST). Specific validation statements by the vendor are included in the list of antibodies immediately above.

Biolegend Flow Cytometry Reagents: Specificity testing of 1-3 target cell types with either single- or multi-color analysis (including positive and negative cell types) is performed. Once specificity is confirmed, each new lot must perform with similar intensity to the in-date reference lot. Brightness (MFI) is evaluated from both positive and negative populations. Each lot product is validated by QC testing with a series of titration dilutions. Lot specific Certificates of Analysis are provided via search on the company's website.
BD Biosciences Flow Cytometry Reagents: The specificity is confirmed using multiple methodologies that may include a combination of flow cytometry, immunofluorescence, immunohistochemistry or western blot to test staining on a combination of primary cells, cell lines or transfectant models.
All flow cytometry reagents are titrated on the relevant positive or negative cells. To save time and cell samples for researchers, test size reagents are bottled at an optimal concentration with the best signal-to-noise ratio or relevant models during the product development. To ensure consistent performance from lot-to-lot, each reagent is bottled to match the previous lot MFI. Certificate of Analysis and the concentration of test-size human reagents from specific lots are available for look up via the Concentration Lookup page or BD regulatory documents. Lot specific Certificates of Analysis are available for lookup on the company's website.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	293FT Cell Line: Thermo Fisher, R70007, https://www.thermofisher.com/order/catalog/product/R70007#:~:text=The%20293FT%20Cell%20line%20is,the%20SV40%20large%20T%20antigen The DEV cell line was a gift from Dr. Diepstra (University Medical Center Groningen, Groningen, The Netherlands). KM-H2 cell line: DSMZ, ACC 8 https://www.dsmz.de/collection/catalogue/details/culture/ACC-8
Authentication	293FT, DEV and KM-H2 cell lines have been authenticated by STR profiling (The Centre for Applied Genomics, Toronto, Canada).
Mycoplasma contamination	The cell lines reported herein have been tested for mycoplasma contamination in-house using the VenorGeM Mycoplasma Detection kit (PCR based, Sigma-Aldrich), and have consistently tested negative.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

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	IL4R experiments: Samples were prepared as previously described in Viganò, E., Gunawardana, J., Mottok, A., et al. (2018). Somatic IL4R mutations in primary mediastinal large B-cell lymphoma lead to constitutive JAK-STAT signaling activation. Blood, 131(18), 2036–2046. https://doi.org/10.1182/blood-2017-09-808907. ScGEX FACS: LN excisional biopsies were initially processed to release cells by cutting the LN and scraping the cut edge to release cells in a Petri Dish. After cell collection, they were cryopreserved. Cryopreserved cell suspensions were then prepared as previously described in Reichel et al. Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. Blood (2015) 125(7):1061-1072. https://doi.org/10.1182/blood-2014-11-610436, with minor modifications.
Instrument	IL4R experiments: BD LSR Fortessa II Cell Analyzer https://www.bd.com/en-us/products-and-solutions/products/product-brands/lsrfortessa . scGEX FACS: USing BD FACSDiva (v9.0.1), labeled cells were then sorted using a FACSAria Fusion special order research product using a 130uM nozzle at 10p.s.i.
Software	IL4R experiments: BD FACSDiva software for sample collection https://www.bdbiosciences.com/en-ca/products/software/instrument-software/bd-facsdiva-software FlowJo software (V10.1) for data visualization and analysis https://www.flowjo.com/. scGEX FACS: FACSDIVA Software v9.0.1
Cell population abundance	IL4R experiments: Detailed in 'In-vitro cell line studies to functionally characterize IL4R mutations'. Briefly, transduced DEV and KM-H2 cells were FACS sorted for the GFP+/DAPI- populations. After sufficient cell expansion, sorted cells treated with 2-4µg/ml puromycin for 3-7 days.

scGEX FACS: absolute cell numbers of sorted populations were collected using FACSDIVA software.

Gating strategy

IL4R experiments:Gating strategy is defined as previously described in Viganò, E., Gunawardana, J., Mottok, A., et al. (2018). Somatic IL4R mutations in primary mediastinal large B-cell lymphoma lead to constitutive JAK-STAT signaling activation. Blood, 131(18), 2036–2046. https://doi.org/10.1182/blood-2017-09-808907.

scGEX FACS: Gating strategy is as described in Reichel et al. Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. Blood (2015) 125(7):1061-1072. https://doi.org/10.1182/blood-2014-11-610436, with some modifications made. We adjusted sorted populations over the course of the study based on total cell number and viability of individual samples to maximize yield of HRS cells being able to be processed appropriately for single cell RNA sequencing.

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