nature research

Corresponding author(s): Maxim N. Artyomov

Last updated by author(s): Aug 28, 2020

Reporting Summary

Nature Research 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 Research policies, see Authors & Referees 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	No software was used
Data analysis	ULI-ChIP-seq processing pipeline is available on GitHub: https://github.com/JetBrains-Research/washu Custom tools (SPAN and JBR browser), models and tutorials are available online at https://artyomovlab.wustl.edu/aging/tools.html. The R statistical language with various Bioconductor packages as noted in Methods. STAR (v2.6.1b) FastQC (v0.11.3) Picard tools (v2.18.4) HTSeq (v0.9.1) Methpipe (v3.4.3) Bedtools2 (v2.25.0) Bowtie (v1.1.1) Limma (v3.34.5) Pheatmap (1.0.12) DeSeq2 (v1.24.0) fgsea (v1.10.0) BisRNA (v0.2.2) ChIPseeker (v1.20.0) IGV (v2.3.72) lumi (v2.26.4) sva (v3.22.0) Macs2 (v2.1.1) SICER (v1.1) Diffbind (v2.4.8) diffReps (v1.55.6)

ChIPDiff (version from March 27, 2008)
Details and references can be found in the relevant Methods sections.

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Raw human sequencing data (Figures 2, 3, and 4) are deposited to Synapse repository (https://www.synapse.org/#!Synapse:syn22020090/wiki/602603) and available for download upon request. Processed sequencing data, as well as raw and processed metabolic and proteomics data are available at the dedicated online portal at https://artyomovlab.wustl.edu/aging/.

Following publicly available datasets have been used in present study:
• GSE56045 MESA transcriptomic dataset – expression table and annotation are provided. (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE56045)
• GSE56046 MESA DNA methylation dataset – table of M values and annotation are provided. (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE56046)
• EGAD00001002523 Blueprint dataset – bigwig files with methylation signal were downloaded. Experiment ids are EGAX00001097775, EGAX00001097772, EGAX00001086967, EGAX00001086968, EGAX00001086970, EGAX00001097774.
• GSE31263 WGBS dataset – BED files with methylation signal were used. (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31263)
• ENCODE ChIP-Seq samples were used: GSM1102782, GSM1102785, GSM1102788, GSM1102793, GSM1102797.
• ENCODE ChromHMM annotation ENCSR907LCD was used. (https://www.encodeproject.org/annotations/ENCSR907LCD/)
• ReMap 2018 database was downloaded for hg19. (http://pedagogix-tagc.univ-mrs.fr/remap/download/remap2018/hg19/MACS/remap2018_TF_archive_nr_macs2_hg19_v1_2.tar.gz)
• For a list of datasets used in Figure 6 and Extended Data 9 please see “Integration with public data” section of Methods.
• Neal lab analysis of UK Biobank was downloaded from http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank

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	Sample size was not determined a priori. We did not focus on a specific effect-size and performed a discovery study, including all available samples that passed QC into analysis. To our knowledge, a comparative epigenetic human study of this scale has not been previously undertaken. We discuss potential limitations of our sample size in Results and Discussion sections. Specifically, we show that large cohorts are required for detection of low-degree changes in expression levels (Extended Data 4B).
Data exclusions	1 outlier was excluded from analysis of metabolic data based on its separate location on PCA plot. 3 outliers were excluded from DNA methylation analysis based on PCA plots and dendogram (Extended Data 5H and 5I) as well as irregular distribution of their methylation percentages (Extended Data 5J). Outlier detection based on exploratory data analysis (PCA etc.) was performed as a part of a planned data analysis process. Data from 191 out of 200 ULI-ChIP-seq experiments passed initial QC. See corresponding sections in Methods and Supplementary Information for details.
Replication	Selected targets from proteomics profile were confirmed by ELISA (Extended Data 3D). All proteins besides one were validated. OMD has not been validated by ELISA, possibly, due to high technical variation. All human samples were used in analysis, additional validation cohort was not recruited. Metabolomic profiling was not additionally validated due to literature support for the most major findings. RNA-seq, RRBS and ChIP-seq data was confirmed via comparison to publicly available datasets from GEO, Blueprint, ENCODE.
Randomization	For high-throughput data generation samples were randomized between batches to account for possible batch-effect. Stringent inclusion criteria were set to account for other possible confounding variables. Reproducibility of the identified trends was shown using publicly available data from cohorts of different demographics, showing independence of the identified signatures from potential confounding variables.
Blinding	All data acquisition was performed blinded to types of individual samples.

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
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

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

Antibodies

Antibodies used

H3K27me3 (Millipore 07-449, LOT 2736613, concentration used: 0.3 µg); H3K27Ac (Abcam ab4729, LOT GR312651-1, concentration used: 0.05 µg); H3K4me3 (Abcam ab8580, LOT GR3201240-1, concentration used: 0.03 µg); H3K4me1 (Abcam ab8895, LOT GR271478-1 and GR283603-1, concentration used: 0.2 µg); H3K36me3 (Abcam ab9050, LOT GR260274-1, concentration used: 0.1 µg);
V450 CD16 (BD Biosciences #561310, Clone B73.1, LOT 6074532); FITC CD14 (BD Biosciences #347493, Clone MΦP9, LOT 6022603);

Validation

All antibodies are commercially available, and validation data are provided by the manufacturer.

We additionally validated ChIP-seq antobodies: using public data for CD14+CD16- monocytes in Roadmap Epigenomics Consortium and data pilot runs we developed a collection of positive and negative control primers. These primers were used to do an additional check of signal-to-noise ratio using real-time PCR (SYBR-green) before sequencing step (see Table S15 for sequences and locations).

Examples of previous publications validating the antibodies:
1. H3K27me3 – Prickaerts P, et al., 2016, Epigenetics & Chromatin, 9-46
2. H3K27Ac – Durbin A, et al., 2018, Nature Genetics 50, 1240-1246
3. H3K4me3 – Shah R, et al., 2018, Molecular Cell 72, 162-177
4. H3K4me1 – Shah R, et al., 2018, Molecular Cell 72, 162-177
5. H3K36me3 – Huang Y, et al., 2018, Journal of Biological Chemistry 293 (20), 7811-7823
6. CD14 – Badia R, et al., 2018, Nature Communications 9, 2739
7. CD16 – Maekawa T, et al., 2017, Leukemia 31, 2709-2716

Human research participants

Policy information about studies involving human research participants

Population characteristics

Healthy, Caucasian, non-obese (BMI under 30) males were enrolled in the study in two groups (Figure 1A). Young donors between 24-30 years (n=20) and old non-frail donors between 57-70 years (n=20) of age were included. Additional cohort recruited for monocyte proteomic analysis exactly matched all criteria described for the main cohort.

Recruitment

Healthy volunteers matching inclusion criteria described above were recruited from St. Louis area, MO, USA. Using a screening questionnaire, subjects were asked about lifestyle and health issues. All subjects self-identified themselves as healthy. Subjects with any previous history of cancer, inflammatory conditions (rheumatoid arthritis, Crohn's disease, colitis, dermatitis, fibromyalgia or lupus) or infections (HIV, hepatitis B, C) were excluded. Smokers were also excluded. Blood (~100ml) was collected by venous puncture in the morning (8 – 10 a.m.) after overnight fasting. Some of the older donors self-reported blood pressure alterations/medication, which was not considered as a reason for exclusion.

Ethics oversight

All human studies were approved by the Washington University in St. Louis School of Medicine Institutional Review Board (IRB-201604138). Written informed consent was obtained from all participants.

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

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

Raw human sequencing data are deposited to Synapse repository (https://www.synapse.org/#!Synapse:syn22020090/wiki/602603) and available for download upon request. Processed sequencing data are available at the dedicated online portal at https://artyomovlab.wustl.edu/aging/ .

Files in database submission	Please see https://artyomovlab.wustl.edu/aging/download_data.html
Genome browser session (e.g. UCSC)	Files for UCSC can be uploaded from https://artyomovlab.wustl.edu/aging/explore_data.html

Methodology

Replicates	5 histone modifications (H3K4me3, H3K4me1, H3K27ac, H3K27me3, H3K36me3) were profiled in classical monocytes from 20 young and 20 old donors. Consistency between samples is illustrated in Extended Data 6.
Sequencing depth	See Supplementary Table 13 for details. 51 bp single-end reads were generated. Average sequencing depth was ~52 million reads. Average number of uniquely aligned reads was ~44 million reads.
Antibodies	H3K27me3 (Millipore 07-449, LOT 2736613, concentration used: 0.3 µg); H3K27Ac (Abcam ab4729, LOT GR312651-1, concentration used: 0.05 µg); H3K4me3 (Abcam ab8580, LOT GR3201240-1, concentration used: 0.03 µg); H3K4me1 (Abcam ab8895, LOT GR271478-1 and GR283603-1, concentration used: 0.2 µg); H3K36me3 (Abcam ab9050, LOT GR260274-1, concentration used: 0.1 µg).
Peak calling parameters	Peak calling parameters were optimized for each sample using a novel semi-supervised approach. Please see Methods and Supplementary Information for details. Optimized parameters are listed in Supplementary Table 16.
Data quality	Please see Methods and Supplementary Information for details. Peak calling summary and optimized parameters are listed in Supplementary Table 16.
Software	ULI-ChIP-seq processing pipeline is available on GitHub: https://github.com/JetBrains-Research/washu Custom tools (SPAN and JBR browser), models and tutorials are available online at https://artyomovlab.wustl.edu/aging/tools.html. Bowtie (v1.1.1)