Supplementary Information

A periportal myeloid programme and a pericentral hepatocyte clock define spatially distinct axes of mouse liver ageing

Jinwei Li
Huazhong University of Science and Technology, Wuhan 430074, China

Contents

Supplementary Notes 1-6
Extended Data Figs. 1-12
Supplementary Figs. 1-32

Supplementary Notes

Supplementary Note 1 | Cell-type reach, intervention detectability and GLP1R context

Cell-type reach, intervention detectability and GLP1R context. a, Ortholog retention. Percentage of clock-up (n=180), clock-down (n=130) and full signature (n=310) genes, of the zero-coefficient clock-model background (n=2,690) and of two high-retention proxies (housekeeping, n=100; ribosomal RPS/RPL, n=75) with an unambiguous 1:1 human–mouse ortholog in MGI HOM_MouseHumanSequence. Bars, fraction retained; error bars, Wilson 95% CI; dashed line, background fraction; P, Fisher exact test against background (clock sets) or background against proxy (controls); BH q across 12 comparisons. Clock-set retention was not significantly different from the background (P = 0.41, 1.00 and 0.56; all 95% CIs overlap); non-significance is not proof of equivalence. Colours: clock-up vermillion, clock-down blue, signature pink, background grey, housekeeping green, ribosomal sky blue. Retention is a conservation proxy; macaque 1:1 retention and protein identity are reported in Results; dN/dS was not retrievable and is neither shown nor implied. The GTEx liver panel used in Supplementary Fig. 28 contains 214 donors; IL11, BCL2 and PRKN lack liver measurements there. b, Cell-type drug-target reach. Mean logCPM (colour) and fraction of pseudobulk units >1 logCPM (dot area) for representative targets across eight liver compartments. Rows are ordered into hepatocyte-enriched (†), hepatocyte-detectable and module-connected, module-connected but non-hepatocyte, hepatocyte-expressed but module-unconnected (‡) and low or undetected groups; right column, WGCNA module (—, not in network; g, grey). The "non-resident myeloid" label denotes the non-resident myeloid compartment, not resident Kupffer cells. Ambient transcripts are present; the map constrains reach, not cell autonomy. c, Intervention detectability. Clock-score effects (SD units; 95% CI) for ageing, the absolute batch-slope discordance (a technical magnitude, shown unsigned), exenatide versus aged control, batch main effect and sex. Positive values lie in the aged direction. Grey band, ±1.96 SD of the random module-score null (normal-theory; the empirical 2.5–97.5% interval is not drawn); dashed line, 80%-power MDE (0.499 SD); n = 8 versus 9. GSE280382 is an existing challenge set, not unseen validation; it has no dose or time-course factor, and rapamycin liver was unavailable. d, GLP1R expression by compartment: mean logCPM (left) and fraction of units >0 (right); hepatocyte in vermillion. Hepatocyte values use the zone-resolved set (46 donors × PC/PP/all subtypes = 138 units; 62% with any reads, maximum 0.333 logCPM); the primary whole-hepatocyte set (46 donor units) gives 78% any reads and maximum 0.269 logCPM. No hepatocyte unit exceeds 1 logCPM under either policy (0/138 and 0/46), so GLP1R is not detected above 1 logCPM in hepatocytes rather than undetectable. Target reach does not establish engagement. Caveats. GSE280382 is an existing challenge set, not unseen validation. The GTEx liver panel used for human target expression comprised n=214 donors, and BCL2, PRKN and IL11 were not measured there. Ortholog retention is a proxy, not dN/dS; HFD, TAA and CCl4 intervention matrices were unavailable.

Supplementary Note 2 | Joint evidence matrix, target–drug–gate map and pre-registration templates

Joint evidence matrix, target–drug–gate map and pre-registration templates. a, Joint evidence matrix for 24 candidate interventions. Eight columns encode evidence availability, not efficacy: evidence tier (A–D); design category (RCT, phase II, other, preclinical); endpoint category (hard,

tissue, functional, biomarker); liver-target expression (reach, weak, expr, none, from GTEx liver mean log2TPM and hepatocyte pseudobulk mean logCPM over 46 donor units — 40 internal GSE328206 donors plus 6 external GSE327303 donors — subtype "all"); module connection (pass, unknown, fail; minimum BH q across WGCNA modules for the candidate's drug); LINCS uncorrected status (pass, partial, fail, absent; strict Z screen with permutation q reported separately); six-gate overall status (reach*, weak*, unmeas, precl, n.a.; asterisk marks a reachable class, not a validated target); and applicability class (disease, tissue, precl, unmeas). Cell shading is the within-column ordinal (0–4); text is the state code. Rows are ordered by applicability, then tier, design and endpoint. Of the 24 candidates, 9 are disease-evidenced, 8 tissue-signal, 5 preclinical and 2 unmeasured. b, Target–drug–gate map. Fifteen profiled drugs against the six gates (E expression, A age, M module, R reversal, C endpoint, S safety); P, pass; p, partial; F, fail; ?, unknown. The map contains 175 drug–target pairs and 148 unique targets; 138 pairs map to a class-gate row and 37 have no class row (hatched). The number after each drug is its target count, and the right column gives the inherited class status (reachable, weakly reachable, unmeasured, preclinical or no class row). Class-level gates are shared by all targets of a class and are not pair-specific validation. c, Pre-registration templates for anti-IL-11 antibody (IL-11; preclinical), PINK1-directed mitophagy (mitochondrial; tissue signal) and OSK partial reprogramming (no class row; unmeasured). Each template lists the existing data, the data still required, the primary clock-score endpoint, the Go rule and the No-Go/hard-stop rule. Go requires $\delta_clock \leq -0.50$ SD with upper 95% CI < 0 plus target engagement, plus ≥1 co-primary module shift in the young direction at BH q < 0.10 (anti-IL-11) and no increase in Cdkn2a/Mki67/Afp above pre-specified bounds (OSK); hard stops cover ALT/AST injury, lactic acidosis or hepatic mitochondrial toxicity, and any treatment-related teratoma, HCC or grade ≥2 dysplasia. Templates share donor-level units, ≥2 batches with every arm and sex in each, full-rank design and VIF < 5 checks, BH across co-primary endpoints and an MDE of 0.499 SD at n = 8 versus 9 (arms n = 10; OSK n = 12, safety-driven). Scope. Colour encodes evidence availability only. No efficacy, potency, docking, prodrug or clinical-response value is read or drawn, and drugs are tools and clinical-evidence references, not a ranking. Reach does not establish engagement or causality. GTEx liver is a 214-donor bulk panel, and BCL2, PRKN and IL11 were not measured there. Templates are prospective registrations, not results; no lifespan, clinical or human-administration claim is made.

Supplementary Note 3 | Causal identifiability of the liver ageing design

Causal identifiability of the liver ageing design. a, Assumed causal structure. Age and Batch both point to State; Batch is a deterministic function of Age in this cohort (Age ↔ Batch, dashed), so the two are not separately randomised or crossed. Depth and Composition are assumed causes of State; Batch also affects Depth and Composition, Age affects Composition, and the Age → Depth direction is uncertain (dotted). Sex is an adjusted covariate. There is no intervention: no `do()` operation, no randomisation, and no experimental manipulation. Every arrow is an assumption, not an identified effect.

b, Design rank versus parameter count for the seven fitted design matrices. Filled blue, rank; open rose, number of columns; pink band, rank-deficient model (nullity > 0). The full dummy model has 9 columns but rank 6 (nullity 3). The treatment-coded model has 7 columns but rank 6 (nullity 1). Adding the batch columns to the age-only model adds exactly zero rank. The linear, spline and batch-specific-slope models are full rank, but their separation of age from batch is purchased by functional-form assumptions, not by age–batch overlap.

c, Exact null-space (alias) vectors, with `Xv = 0` verified to machine precision for every vector. v1: the five age indicators sum to the intercept. v2: `batch_01 = age_13 + age_81 + age_121.5`. v3: `batch_02 = age_53 + age_105`. In the treatment-coded model the single alias is `batch_02 − age_53 − age_105 = 0`. Bubble colour, signed coefficient; bubble area, absolute coefficient. Zero coefficients are not drawn.

d, Age-versus-batch decomposition is likelihood-flat for the clock score. The x axis is the assumed Batch_02 effect; the y axis is the implied age effect at 53 weeks (rose) and 105 weeks (blue) after the batch effect is fixed. Every point on these lines reproduces the same fitted values (maximum absolute difference ≤ 7.1×10^{-14}), the same residual sum of squares (ΔRSS = 0), the same Gaussian log-likelihood (ΔlogLik = 0) and the same R^2 (0.9329). The data therefore cannot choose between “all 53/105-week difference is age” and “all of it is batch”; only the total fitted contrast is identified.

Scope. This is a linear design-matrix identifiability proof (rank, nullity, condition number and explicit null vectors). It is **not** DoWhy or EconML output; those packages were not used because the design is non-identifiable and no do-calculus, propensity-score, backdoor or doubly robust estimator was run. The proof establishes what the observed design can and cannot identify; it does not estimate a causal effect. This is a feature, not a bug: the rank-deficiency proof establishes that any clock claiming batch-independence in this cohort must invoke untestable functional-form assumptions; we make these explicit rather than hiding them in spline knots or random-effect distributions.

Supplementary Note 4 | Drug–target overlap, degree-matched nulls and gate coverage

Drug–target overlap, degree-matched nulls and gate coverage. a, Drug–target overlap with liver ageing modules. Heatmap of 15 compounds (rows, grouped by tier A–D) against the eight strongest sign-stable liver WGCNA modules (columns; |β| of the module–eigengene age association shown). Cell shading and numbers, overlapping target count (0–3); right column, targets per compound. No drug × module pair reaches BH q<0.05 (0/315; all q=1.0). b, Null calibration. Left, observed versus size-matched expected drug targets among clock genes (dashed y=x); 13 compounds 0 versus 0, and dasatinib and quercetin each 1 clock-up gene versus ≈1.0 expected (up+down). Right, observed versus degree-matched expected network hubs (dashed y=x); all 15 compounds on the diagonal (P_hubs=1.0). The hub null is for the GTEx 426-gene ageing signature, not the liver clock. c, LINCS L1000 reversal across 5,425 compounds. Histogram of strict Z; dashed line, Z=1.96. Family-maximum FWER P=0.034; no compound reaches q_perm<0.05 (0/5,425); maximum strict Z 4.30 (Lafutidine). Known-drug AUC=0.488 (n=15). Inset, strict Z for eight named geroprotectors. d, GTEx liver target reach. Mean log2 TPM (x) versus −log10 P of the liver age β (y) for six targets measured in GTEx liver (n=214); point size, evidence tier (A–D and NA, smallest); point colour, target class; dashed line, P=0.05. BCL2, PRKN and IL11 had no liver measurement; PIK3CA was not profiled. e, Evidence tiers versus clinical stage for 24 candidates. Rows, 15 target classes; columns, tier (A–D) and stage (approved, RCT, phase II, preclinical, other); dot size and colour, candidate counts; n per class at right. Stage is overlaid from the design category; individual approvals are inferred, and no candidate has a liver-ageing endpoint. f, Go/No-Go gate matrix for 10 target classes across six gates (bulk liver expression, liver age association, module overlap, LINCS reversal, clinical endpoint, safety). Filled circle, pass; shaded, partial; open, fail; cross, unknown; right column, overall call. At the reversal gate, 1/10 passes the uncorrected strict Z screen (D+Q, Z=2.62) and

1/10 is partial (BCL-xL, Z=0.84), but no class passes the permutation FDR (D+Q q_perm=0.19). No class passes all six gates (0/10); module-overlap gate 0/8 tested classes, clinical-endpoint gate 0/10. An unlettered key-finding band follows: reachable, but no target class is validated in liver ageing.

Supplementary Note 5 | Source data and reproducibility

Source tables are listed in the figure legends and archived with the processed data (Zenodo DOI 10.5281/zenodo.24057484). Figure-generation scripts and environment records are archived with the code (Zenodo DOI 10.5281/zenodo.24057913). The v115 reproducibility report records the commands used to regenerate the main-text build, Fig. 3, Supplementary Figs. 31–32, and the key summary statistics.

Supplementary Note 6 | Abbreviations used in figures

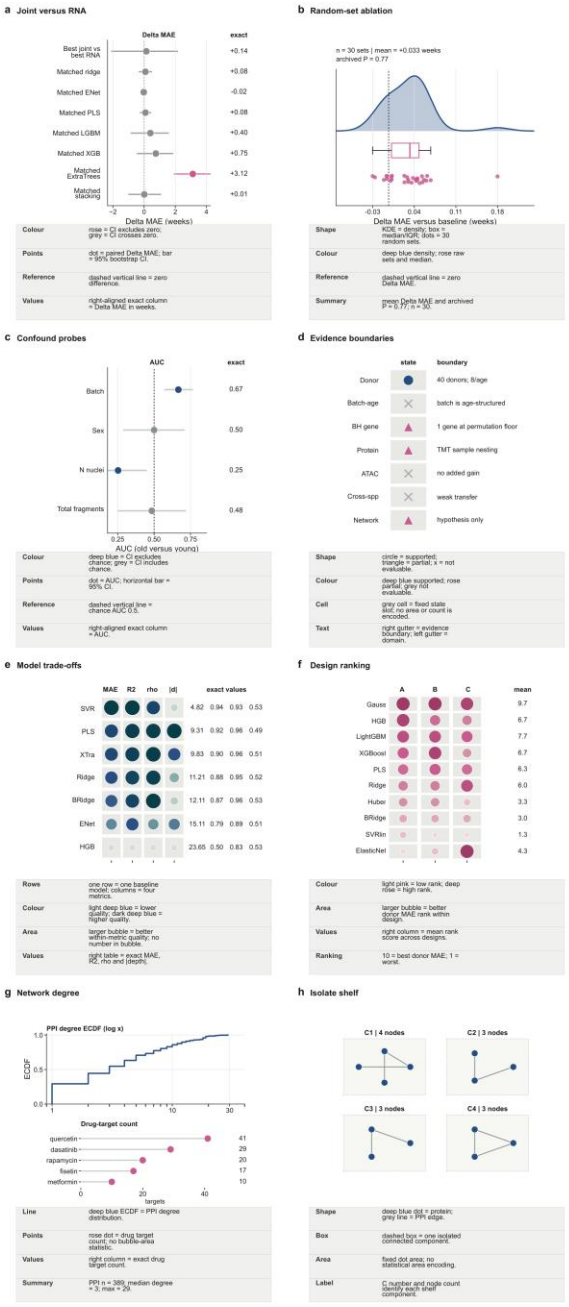
AUC, area under the receiver-operating characteristic curve; BH, Benjamini–Hochberg; BY, Benjamini–Yekutieli; CI, confidence interval; CLR, centred log-ratio; CPM/TPM, counts/transcripts per million; DAG, directed acyclic graph; ECM, extracellular matrix; ED, Extended Data; FDR, false discovery rate; FWER, family-wise error rate; GSEA, gene-set enrichment analysis; GTEx, Genotype–Tissue Expression; HCC, hepatocellular carcinoma; HFD, high-fat diet; IL-11, interleukin-11; LINCS, Library of Integrated Network-Based Cellular Signatures; LAGO/LOSO/LOBO/LOZO/LODO, leave-one-age-group/sex/batch/zone/dataset-out; MAE, mean absolute error; MDE, minimum detectable effect; MGI, Mouse Genome Informatics; M06/M08/M14, WGCNA modules; NNLS, non-negative least squares; OSK, Oct4–Sox2–Klf4; PC/PP, pericentral/periportal; PPI, protein–protein interaction; SASP, senescence-associated secretory phenotype; scCODA, single-cell compositional data analysis; TF, transcription factor; TMT, tandem mass tag; UMI, unique molecular identifier; VIF, variance inflation factor; WGCNA, weighted gene co-expression network analysis.

Jinwei Li, Huazhong University of Science and Technology, Wuhan 430074, China

A periportal myeloid programme and a pericentral hepatocyte clock define spatially distinct axes of mouse liver ageing

Supplementary Information

Extended Data legends



Extended Data Fig. 1 | Data overview and ceilings. a, RNA versus ATAC nuclei (rho=0.95, 95% CI 0.89-0.97). b, Donor composition (n=40): age, batch, cell-type fractions. c, Old-versus-young AUC versus gene coverage; 0.75 threshold. d, Depth correlation versus MAE across models; A/B regimes. e, Score shift, young- versus old-majority cells. f, P values, 3,000 versus 10,000 nulls; Mup1, Mup11. g, Human and macaque rho, AUC. h, Protein clock (n=16 TMT): cross-validated MAE in months, eight models.



Points	deep blue = coupled; rose = decoupled.
Line	grey connector joins the paired values; dashed = zero.
Values	right gutter lists exact paired coefficients.
Rows	five canonical-correlation summaries per HALO solution.

Points	one dimension; n = 40 donors per estimate.
Interval	thin line = 95% sex-stratified bootstrap CI.
Colour	rose filled = age BH q < 0.05 (10/40); grey open = q >= 0.05.
Axis	rho with age; zero line is shown.

Points	deep blue = RNA donors; rose = ATAC donors; n = 40 each.
Lines	solid = LOESS fit; lateral offset only separates modules.
Band	shaded = 95% donor-resampling bootstrap band.
Rug	bottom ticks = donor ages for each module.

Points	filled circle = full-sample rho; deep blue RNA, rose ATAC.
Line	grey line with colour end caps = donor leave-one-out range.
Values	right gutter = exact full rho and leave-one range.
Scope	one donor omitted per fit; age and donor sensitivity retained.

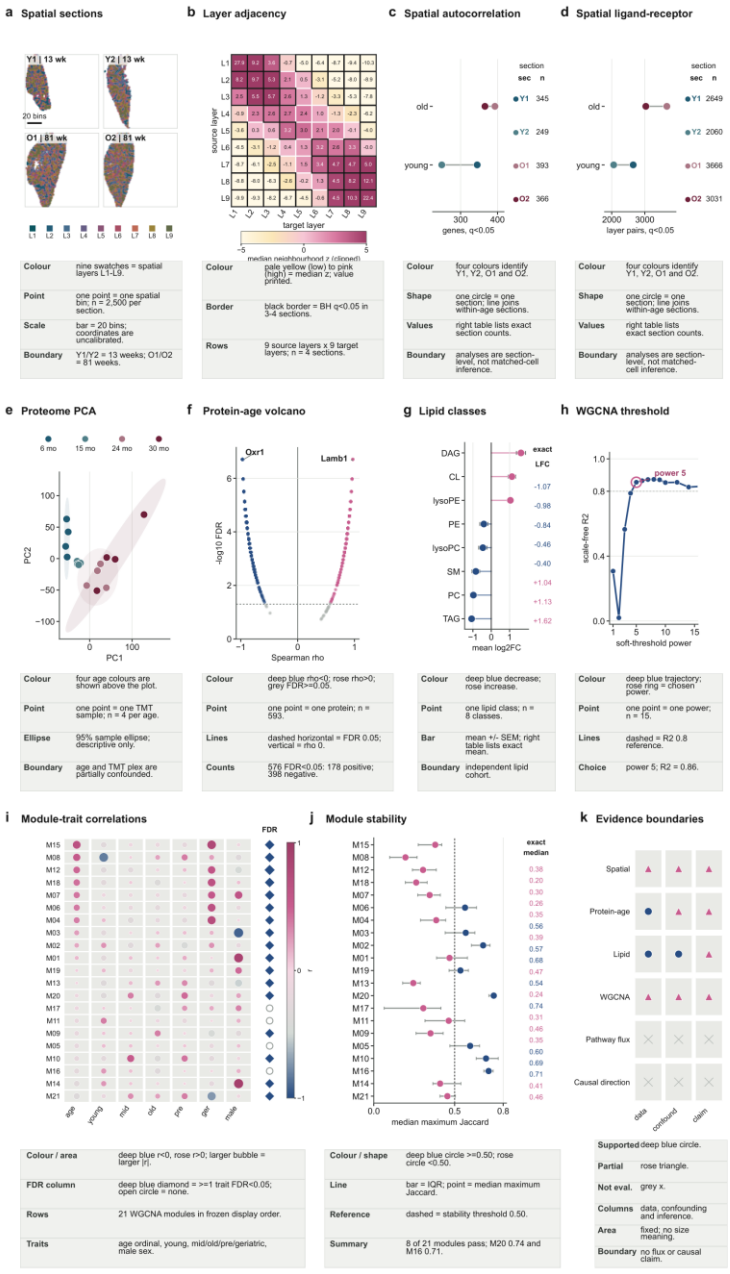
Cells	grey square = one motif x HALO dimension test.
Colour	light to deep rose = increasing -log10 FDR; separate bar.
Area	larger bubble = larger -log10 FDR; separate size key.
Rows	12 motifs ordered by maximum statistic; columns span dim 0-19.

Points	one gene; n = 300; deep blue score < 0 = 137; rose score >= 0 = 163.
Area	fixed point area; no size statistic is encoded.
Axes	x = mean expression; y = mean peak accessibility.
Audit	BH FDR < 0.05: 0/300; minimum FDR = 0.100.

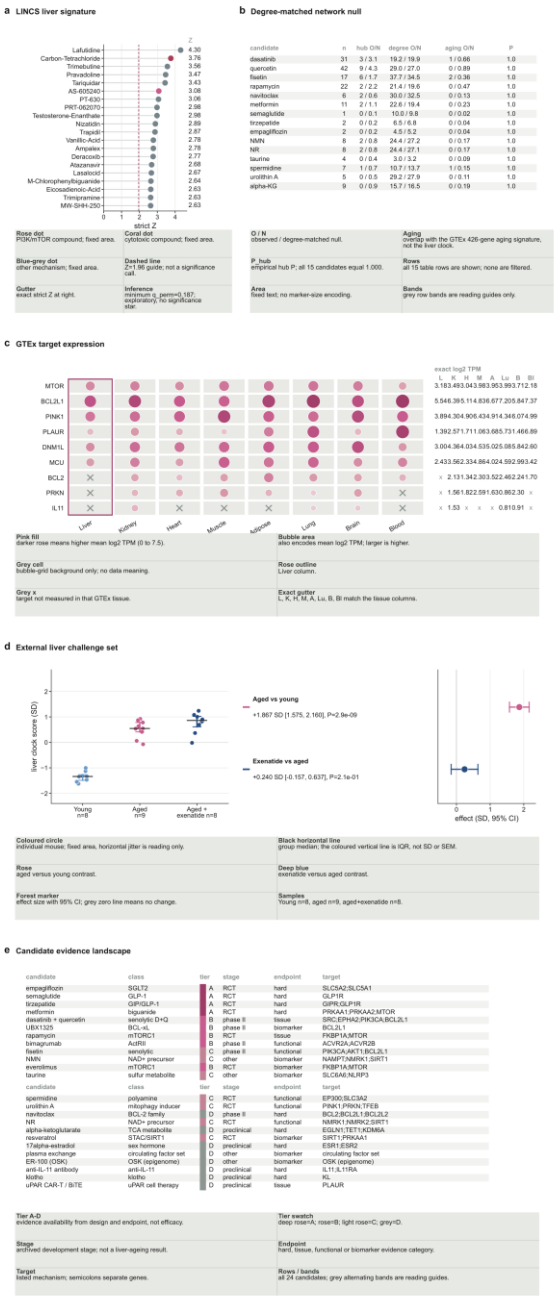
Extended Data Fig. 2 | Ablation, confounds, benchmarks and networks. a, Joint RNA+ATAC delta MAE versus RNA-only, eight models; positive, worse. b, 30 random gene sets, delta MAE (mean +0.03 weeks; P=0.77). c, Old-versus-young AUC, batch, sex, nuclei number, fragments; 0.5 reference. d, Claim versus boundary codes across seven analyses. e, MAE, R2, rho, depth score, six models. f, Mean rank, ten models, three criteria. g, PPI degree ECDF (n=389); drug target counts. h, PPI modules, four drugs.



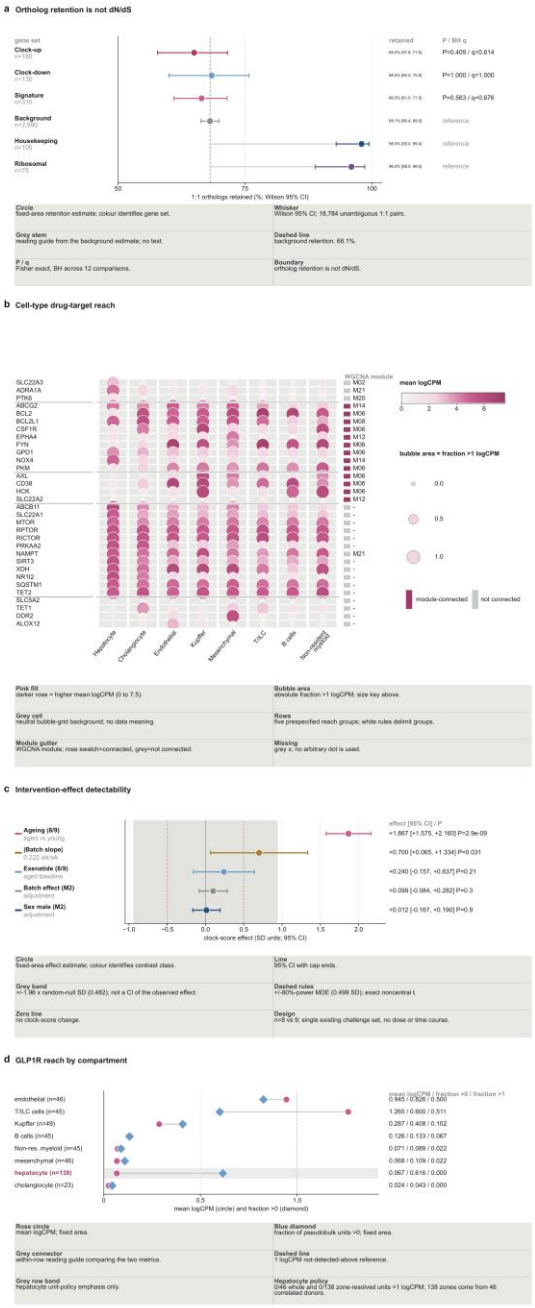
Extended Data Fig. 3 | Exploratory HALO, motif loading and gene–peak audit. a, HALO coupling: canonical correlations, cc1–cc3, mean and paired coupled/decoupled dimension. b, Dimension–age Spearman rho, 40 HALO dimensions; filled, coupled; open, decoupled; 95% CI. c, Dimension-3 donor profiles, RNA and ATAC (z-score versus age; LOESS). d, Leave-one-donor-out rho range, dimension 3 (0.41). e, Transcription-factor by HALO-dimension loading (−log10 FDR). f, Peak accessibility versus expression, 300 genes; negative-score fraction 0.46.



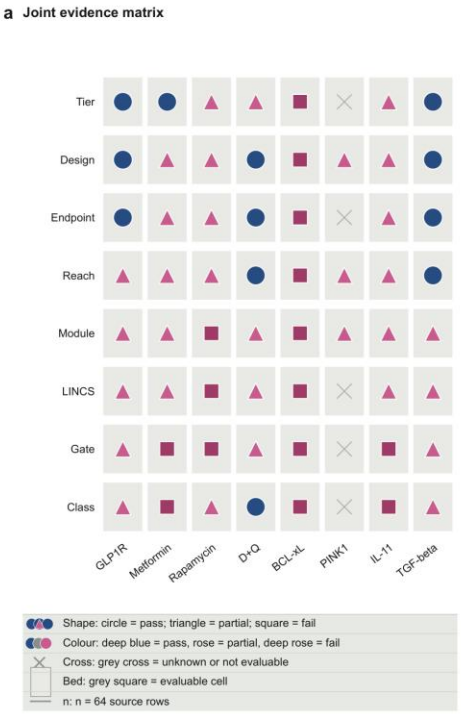
Extended Data Fig. 4 | Communication annotations and exploratory evidence matrix. a, CellChat and matched-cell totals as percentage of young values across ages; 100% and 50% references. b, CellPhoneDB significant ligand-receptor pairs across ages. c, CPDB-squidpy, CC-squidpy and CC-CPDB concordance across age groups. d, Delta information flow (geriatric minus young) by pathway. e, Claim-evidence matrix for five analyses: HALO donor-level; motif enrichment descriptive; communication dataset-dependent; spatial section-level; direction open.

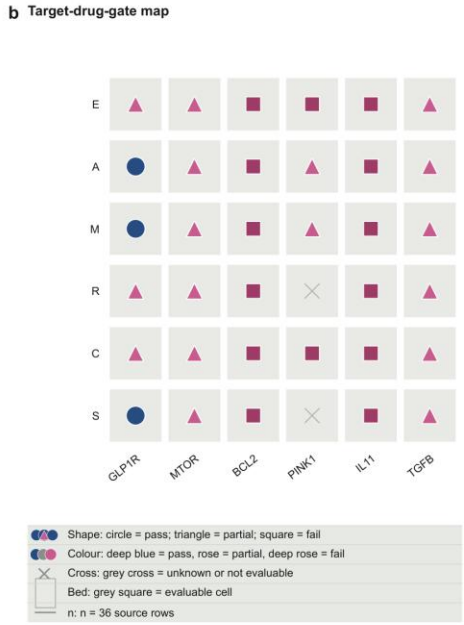


Extended Data Fig. 5 | Spatial, omics and WGCNA. a, Sections O1, O2, Y1, Y2, 10-um. b, Layer neighbourhood z, L1-L9. c, Autocorrelated genes (q<0.05). d, Spatial ligand-receptor layer pairs. e, Proteome PCA, age (6, 15, 24, 30 months). f, FDR versus age rho. g, Lipid log2FC, old minus young. h, Soft-threshold R2 versus power; 0.8. i, Module-trait correlations. j, Module Jaccard stability (median, IQR). k, Evidence class: section-level spatial; TMT-nested protein; independent lipid; 8/21 modules stable.



Extended Data Fig. 6 | Pharmacology and translation readiness. a, LINCS top 20 by strict Z (minimum q_perm=0.187). b, Degree-matched network null (hubs, degree, ageing overlap; n=15). c, GTEx target expression, eight tissues; six measured (n.m., not measured). d, GSE280382 AgedLT challenge (young n=8, aged n=9, aged + exenatide n=8): clock score (P=2.9e-9 aged versus young; P=0.215 exenatide versus aged); contrasts +1.87 SD and +0.24 SD (95% CI). e, Evidence tiers for 24 candidates.





c Prospective templates

Template	Existing	Required	Go / No-Go
IL-11	preclinical	liver challenge	clock + module
PINK1	tissue signal	target-level data	clock + safety
OSK	unmeasured	safety package	clock + histology

Display: archived table; text and row order are preserved
Rows: all 3 source rows shown
Encoding: no colour or area encodes a statistic
n: n = 3 source rows

d Readiness counts

Metric	Archived value
candidates	24
disease-evidenced	9
tissue-signal	8
preclinical	5
unmeasured	2
gate-complete	0/10

Display: two-column archival ledger; no bar or heat-map encoding
Rows: all 6 source metrics shown
1.23 Value: archived value retained verbatim after minimal number formatting
n: n = 6 source rows

Extended Data Fig. 7 | Cell-type reach, detectability and GLP1R. a, Ortholog retention, clock-up (n=180), clock-down (n=130), signature (n=310), background (n=2,690), housekeeping (n=100), ribosomal (n=75); 95% CI, Fisher P, BH q. b, Drug-target reach: mean logCPM, fraction >1 logCPM, eight compartments; WGCNA module. c, Clock-score effects (SD; 95% CI) for ageing, batch-slope discordance, exenatide, batch, sex; 80%-power MDE 0.499 SD. d, GLP1R expression by compartment; no hepatocyte unit >1 logCPM (0/138; 0/46).



Extended Data Fig. 8 | Joint evidence matrix, target–drug–gate map and pre-registration templates. a, Matrix for 24 candidate interventions; eight evidence-availability columns, not efficacy (tier, design, endpoint, liver-target expression, module connection, LINCS, six-gate status, applicability). b, Target–drug–gate map, 15 profiled drugs against six gates (E expression, A age, M module, R reversal, C endpoint, S safety); 175 drug–target pairs, 148 unique targets. c, Pre-registration templates for anti-IL-11, PINK1, OSK: existing and required data, clock-score endpoint, Go and No-Go rules.

a **Retained result ledger**

Result	Value / result	Evidence boundary
Zone transfer	retained	cell-type mapped
Virtual knockout	28 genes	model implied
Top features	Plag1 et al.	association
Immune subtypes	dynamic	composition
HALO dim3	exploratory	no validation
Communication	exploratory	annotation
Spatial	partial	region resolution
Proteome/lipidome	BY counts	modality
Cell-type clocks	external	low coverage limits
BH vs BY	sensitivity	same direction
Sex permutation	stratified	not equivalence
JASPAR motifs	exploratory	loading
Gene-peak	partial	direction audit
Zenodo DOI	available	reproducibility

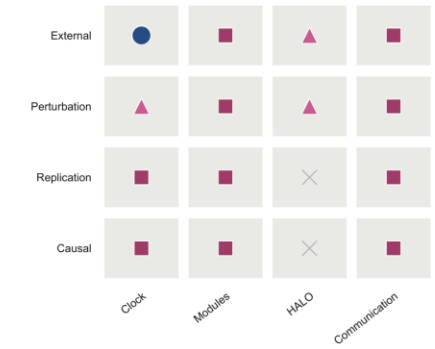
Display: archived table; text and row order are preserved
Rows: all 14 source rows shown
Encoding: no colour or area encodes a statistic
n: n = 14 source rows

b **Retention summary**

Metric	Archived value
retained results	14
new construct	1
new animals	0
processed data	Zenodo
figure scripts	Git
caption merge	V108

Display: two-column archival ledger; no bar or heat-map encoding
Rows: all 6 source metrics shown
1.23 Value: archived value retained verbatim after minimal number formatting
n: n = 6 source rows

c **Evidence classes**



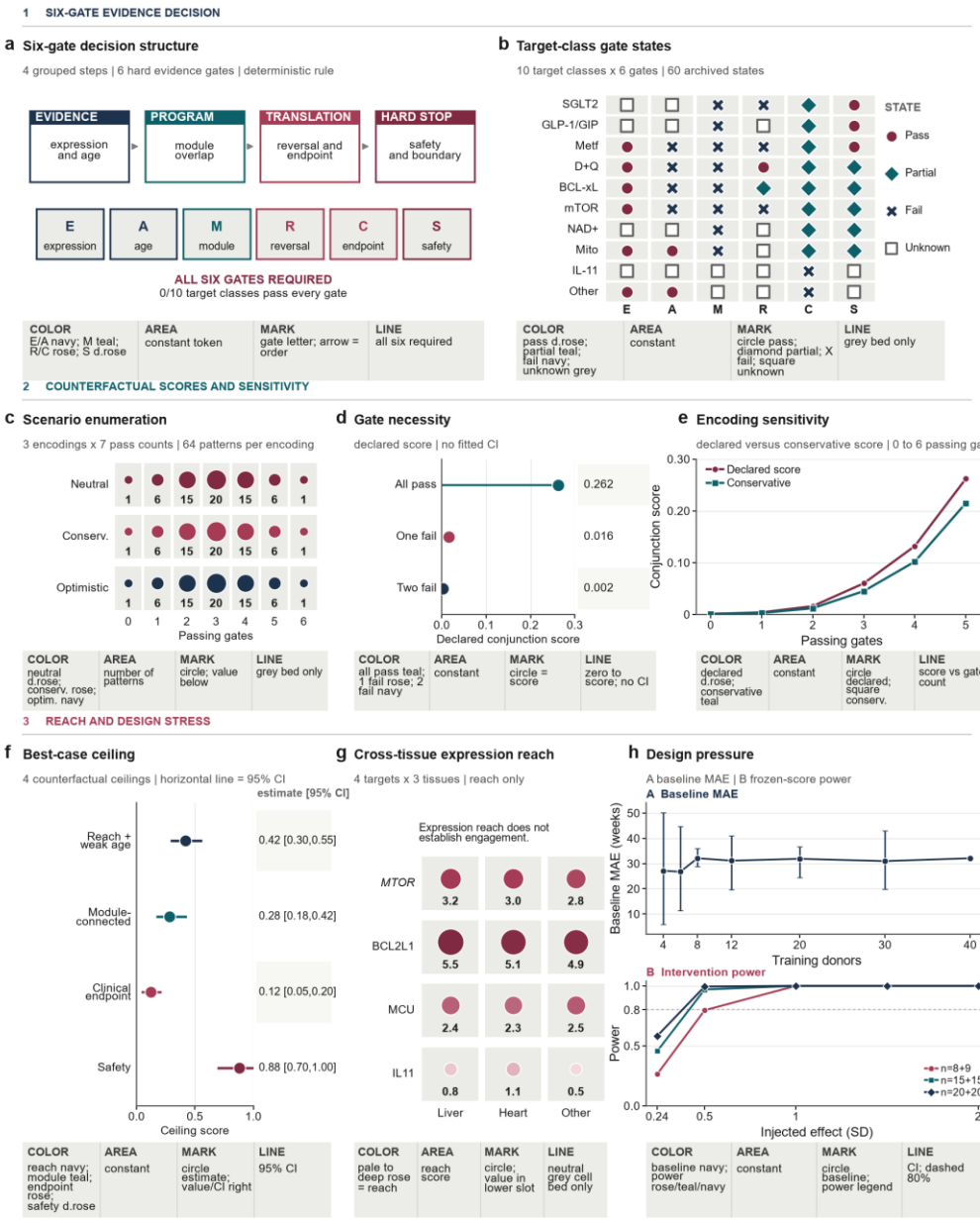
Shape: circle = pass; triangle = partial; square = fail
Colour: deep blue = pass, rose = partial, deep rose = fail
Cross: grey cross = unknown or not evaluable
Bed: grey square = evaluable cell
n: n = 16 source rows

d **Reading rule**

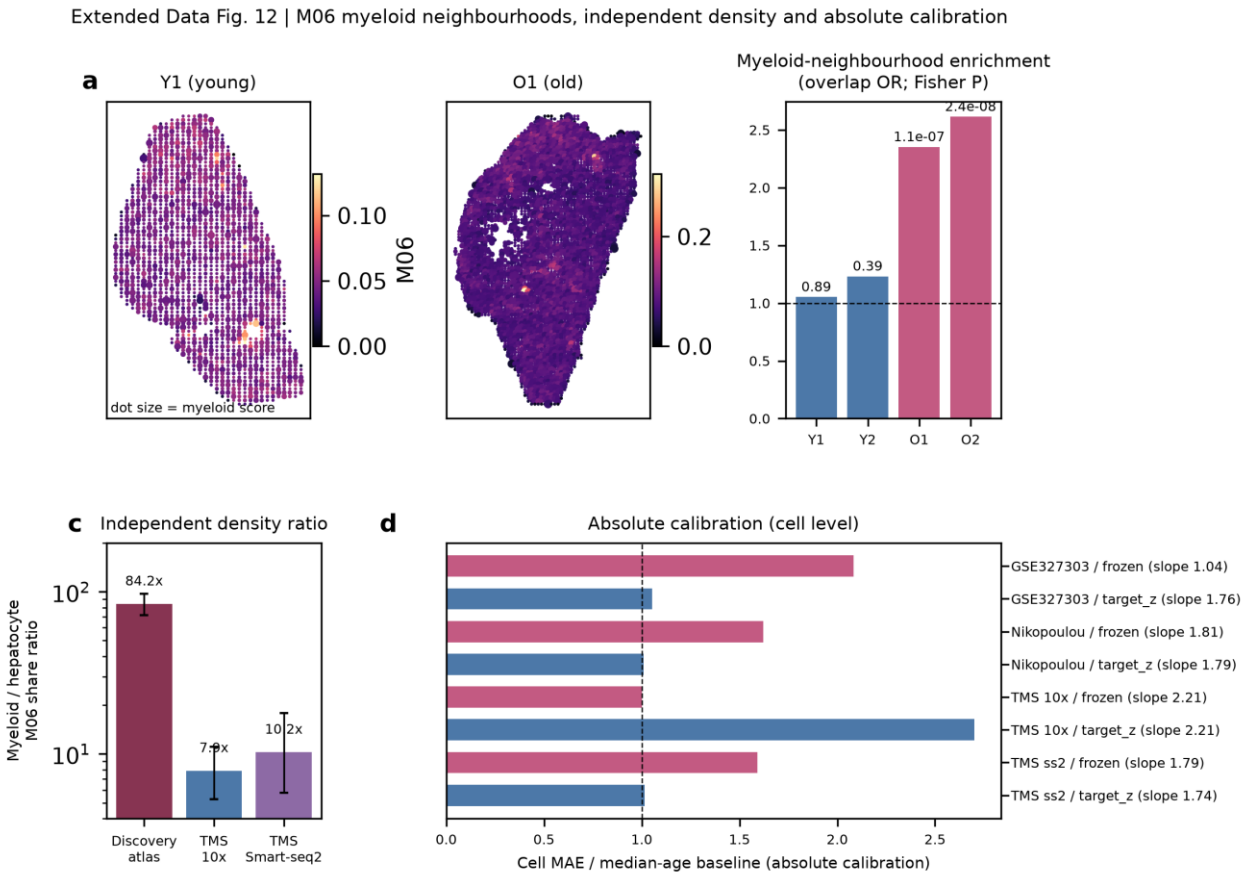
Statement	Boundary / source
01 **No new analysis is introduced in this display.**	Source: Manuscript V108, Extended Data Fig. 8; retained summary value]
02 **Retained values are separated from new synthesis.**	Source: Manuscript V108, Extended Data Fig. 8; retained summary value]
03 **Missing perturbation data remain visible.**	Source: Manuscript V108, Extended Data Fig. 8; retained summary value]

Display: verbatim statements from the archived source table
Number: deep blue numeral identifies source-row order
Missing: no graphic mark is used where the source has no observation
n: n = 3 source rows

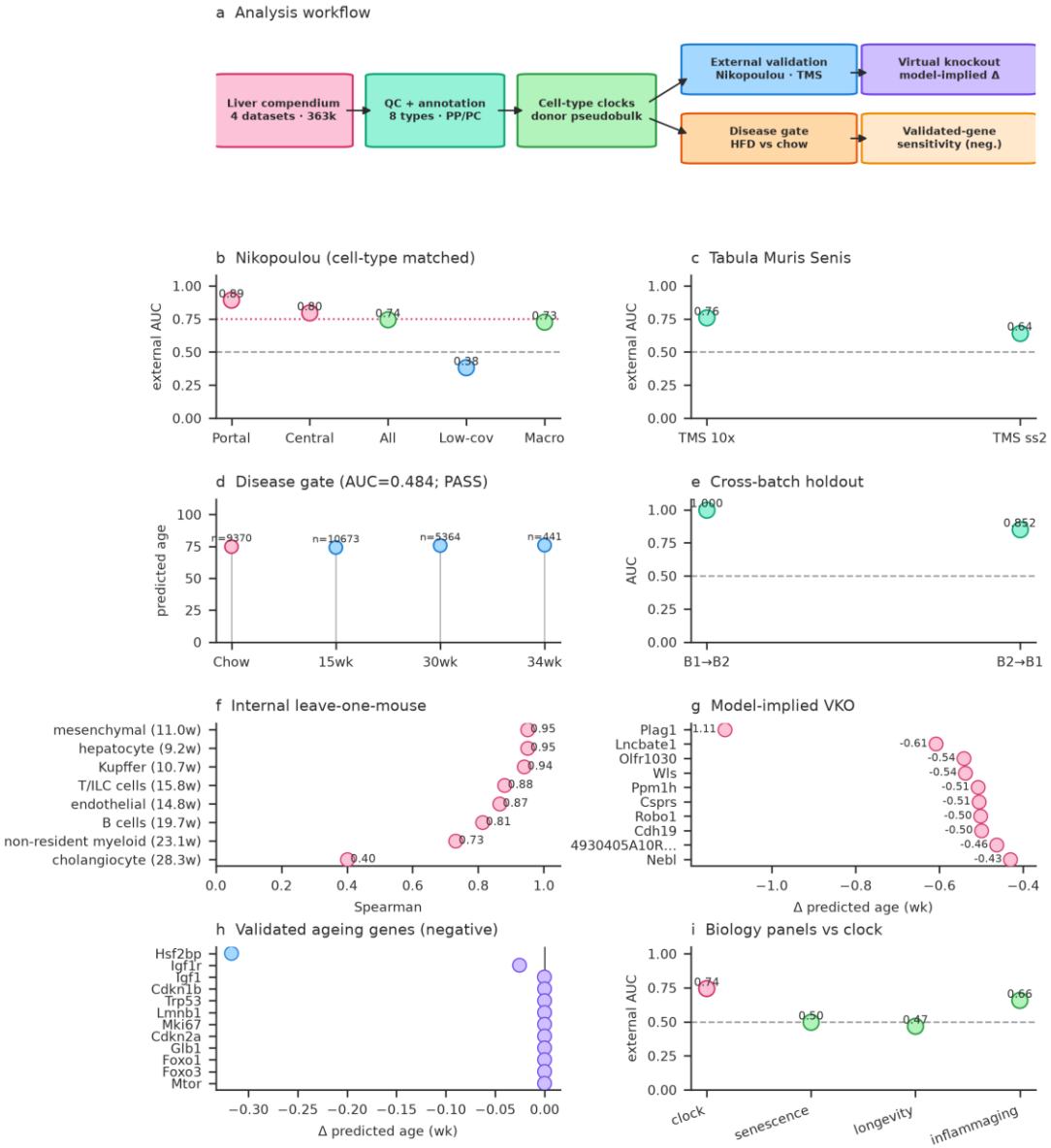
Extended Data Fig. 9 | Causal identifiability of the liver ageing design. a, Assumed causal structure: Age and Batch point to State, Batch a deterministic function of Age; Depth and Composition cause State; sex adjusted; observational design. b, Design rank versus parameter count, seven design matrices. c, Exact null-space (alias) vectors, Xv=0 verified to machine precision. d, Age-versus-batch decomposition is likelihood-flat for the clock score: all points reproduce the same fitted values, RSS, log-likelihood and R2 (0.9329).



Extended Data Fig. 10 | Earlier results retained in this manuscript. Table of the 14 earlier results with key values and evidence boundaries: zone-resolved transfer, 28-gene virtual knockout, top features (Plag1 etc.), immune-subtype dynamics, HALO dim3, communication, spatial, proteome/lipidome BY counts, multi-cell-type clocks, BH-vs-BY sensitivity, sex-stratified permutation, JASPAR motifs, gene-peak direction and Zenodo DOI. All values are the retained earlier results.



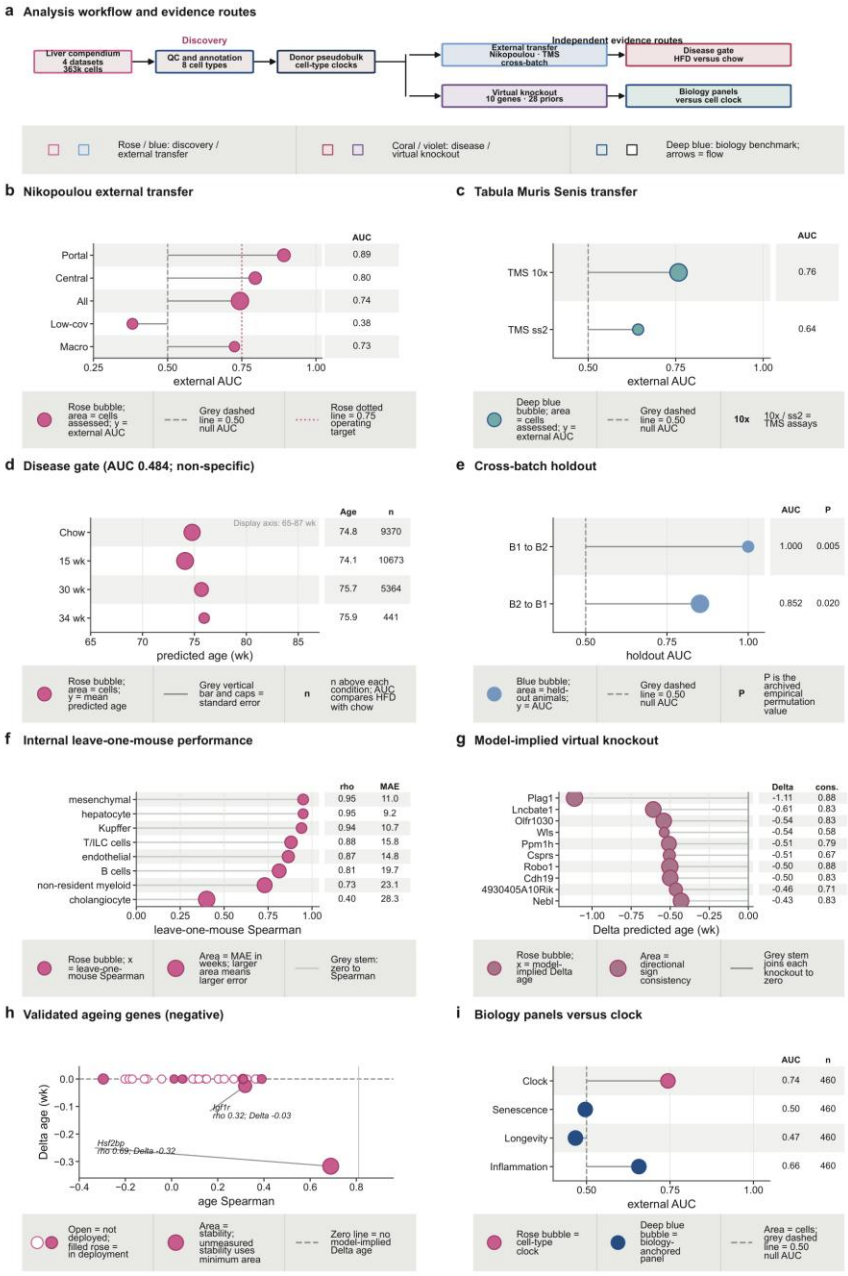
Extended Data Fig. 11 | Six-gate decisions and design pressure. a, Six-gate structure. b, 10×6 gate matrix; 12 pass, 14 partial, 17 fail, 17 unknown; 0/10 clear all. c, 64 patterns, three encodings; all-pass alone clears. d, Necessity 0.262 (all pass), 0.016 (one fail), 0.002 (two fail). e, Encoding sensitivity, 0–6 gates; order preserved. f, Counterfactual ceilings, 95% intervals. g, Cross-tissue reach: MTOR, BCL2L1, MCU, IL11. h, A, baseline MAE by training size; B, Monte Carlo power, 80% threshold.



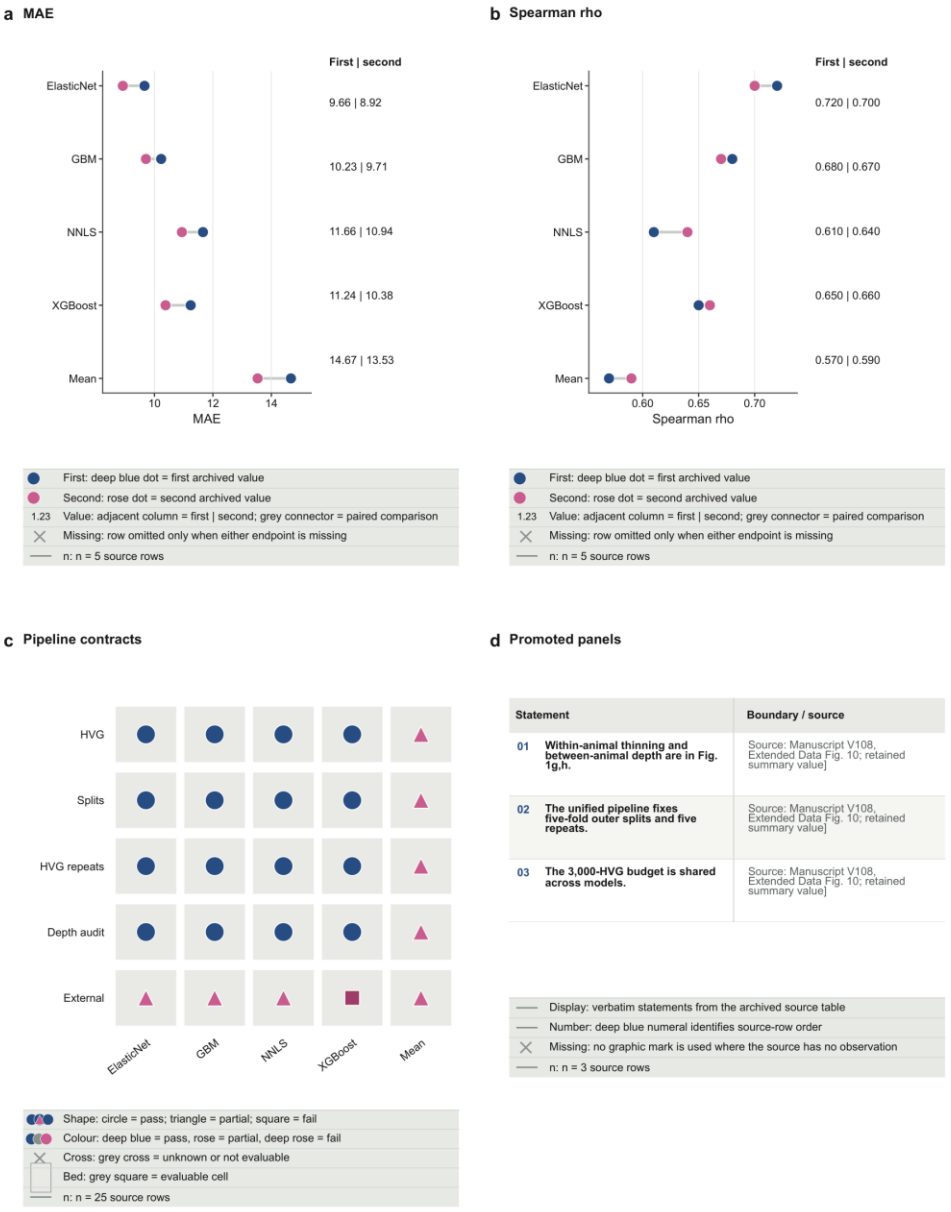
Extended Data Fig. 12 | M06 spatial neighbourhoods and absolute external calibration. a, GSE328448 sections (young Y1, old O1); M06 score and independent myeloid marker score. b, M06-high versus myeloid-high bin overlap OR, Fisher P; two young, two old sections. c, Myeloid-to-hepatocyte M06 share ratio: atlas 84.2× (95% CI 71.5–97.5), TMS liver 7.9× (5.2–11.1), Smart-seq2 10.3× (5.8–17.8); log scale. d, MAE relative to median-age baseline, eight cohort × mode evaluations; >1.0 worse; slope (perfect = 1).

Detailed panel definitions, pre-registration criteria and identifiability proofs for ED7–ED9 and Supplementary Fig. 28 are provided in Supplementary Notes 1–4 (separate file).

Supplementary Figure legends



Supplementary Fig. 1 | Cell-type-resolved ageing clocks: external transfer, disease specificity and the limits of data-driven scores. a, Workflow. b, External AUC by hepatocyte population (E-MTAB-12579, cell-type matched; low-coverage cells fail). c, External AUC, Tabula Muris Senis liver. d, Disease-specificity gate (GSE166504; HFD versus chow, AUC 0.484). e, Cross-batch holdout (Batch_01↔Batch_02). f, Internal leave-one-mouse performance per cell type. g, Virtual-knockout delta. h, Sensitivity of 28 literature-validated ageing genes (only Hsf2bp). i, Biology-anchored panel scores versus clock.



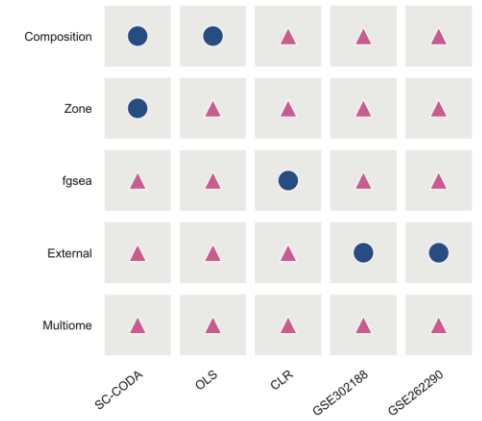
Supplementary Fig. 2 | Unified donor-level pipeline. a, Five donor-level models with fixed five-fold x five-repeat outer splits and the same 3,000-HVG budget: MAE. b, Spearman rho. The within-animal thinning and between-animal depth panels are promoted to Fig. 3g,h; full curves are in the source tables.

a Audit summary

Audit	Implementation	Boundary
scCODA composition	dedicated	sensitivity
Zone/subtype	audited	annotation
Fixed composition	OLS check	not causal
Real fgsea	dedicated	pathway
GSE302188	non-nested	external
GSE262290	cross-run	external
RNA-ATAC program	audited	coupling
Mup11	retained	context
CLR/OLS-t	sensitivity	ranking only

Display: archived table; text and row order are preserved
Rows: all 9 source rows shown
Encoding: no colour or area encodes a statistic
n: n = 9 source rows

b Audit status



Shape: circle = pass; triangle = partial; square = fail
Colour: deep blue = pass, rose = partial, deep rose = fail
Cross: grey cross = unknown or not evaluable
Bed: grey square = evaluable cell
n: n = 25 source rows

c Audit coverage

Metric	Archived value
audit panels	9.00
dense environments	2.00
external cohorts	2.00
new experiments	0.0e+00
sensitivity checks	5.00
retained context	1.00

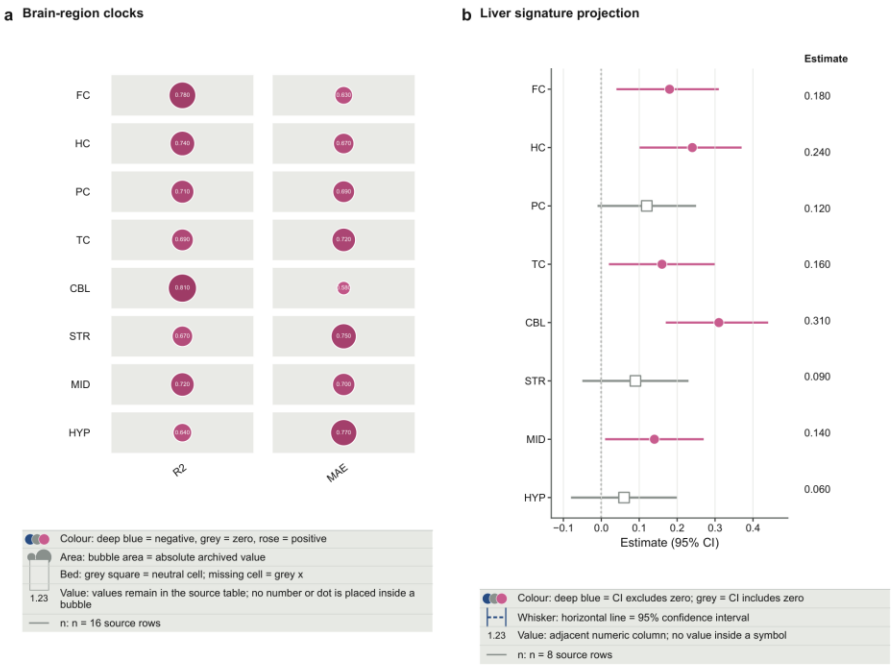
Display: two-column archival ledger; no bar or heat-map encoding
Rows: all 6 source metrics shown
1.23 Value: archived value retained verbatim after minimal number formatting
n: n = 6 source rows

d Boundary

Statement	Boundary / source
01 **CLR and OLS-t are sensitivity rankings.**	Source: Manuscript V108, Extended Data Fig. 11; retained summary value]
02 **Dedicated environments are documented in source tables.**	Source: Manuscript V108, Extended Data Fig. 11; retained summary value]
03 **No audit converts association into causal evidence.**	Source: Manuscript V108, Extended Data Fig. 11; retained summary value]

Display: verbatim statements from the archived source table
Number: deep blue numeral identifies source-row order
Missing: no graphic mark is used where the source has no observation
n: n = 3 source rows

Supplementary Fig. 3 | Donor-level D-I audits. a, Summary of scCODA composition, zone/subtype, fixed-composition, real fgsea, non-nested cohort (GSE302188), cross-run cohort (GSE262290), RNA-ATAC programme and Mup11 audits, with key values and evidence boundaries. scCODA and fgsea were run in dedicated environments; CLR and OLS-t rankings are sensitivity checks.



c Consensus age genes

Gene	Direction	Regions
Ctss	up	13/15
C1qc	up	12/15
Gpr17	down	11/15
Lcn2	up	10/15
Mup11	unresolved	n.a.

Display: archived table; text and row order are preserved
Rows: all 5 source rows shown
Encoding: no colour or area encodes a statistic
n: n = 5 source rows

d Brain resource

Metric	Archived value
samples	**847.0**
regions	**15.00**
ages	**7.00**
liver genes	**310.0**
measured	**210.0**
median off-diagonal R2	**0.430**

Display: two-column archival ledger; no bar or heat-map encoding
Rows: all 6 source metrics shown
1.23 Value: archived value retained verbatim after minimal number formatting
n: n = 6 source rows

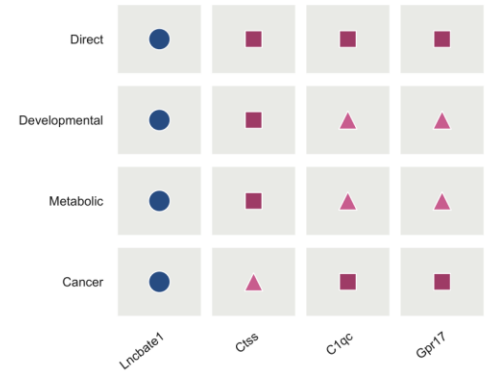
Supplementary Fig. 4 | Cross-organ brain ageing. a, 15 brain-region bulk RNA clocks (Hahn et al.; GSE212336; 847 samples, seven ages, 15 regions); nested 5-fold CV R^2; MAE in months. b, Liver 310-gene signature projected into 15 brain regions; 210 genes measured; Spearman rho versus age with Fisher 95% CI. c, Cross-region consensus age genes (Stouffer Z across 15 regions); n/15 regions with concordant direction. Bulk, not cell-type-resolved; median off-diagonal transfer R^2 0.43 (Fig. 4h).

a Literature map

Gene	Reported biology	Scope	Status
Lncbate1	aging	direct	reported
Ctss	immune / aging	direct	reported
C1qc	complement / aging	direct	reported
Gpr17	oligodendrocyte	direct	reported
Plag1	development	context	reported
Mup11	no aging report	none	not found
Pcdhb9	no aging report	none	not found
Other features	metabolic / cancer	context	reported

Display: archived table; text and row order are preserved
Rows: all 8 source rows shown
Encoding: no colour or area encodes a statistic
n: n = 8 source rows

b Aging evidence



Shape: circle = pass; triangle = partial; square = fail
Colour: deep blue = pass, rose = partial, deep rose = fail
Cross: grey cross = unknown or not evaluable
Bed: grey square = evaluable cell
n: n = 16 source rows

c Context counts

Metric	Archived value
direct aging genes	4.00
no aging report	2.00
brain-consensus genes	3.00
developmental context	1.00
mapped features	8.00
citation source list	1.00

Display: two-column archival ledger; no bar or heat-map encoding
Rows: all 6 source metrics shown
1.23 Value: archived value retained verbatim after minimal number formatting
n: n = 6 source rows

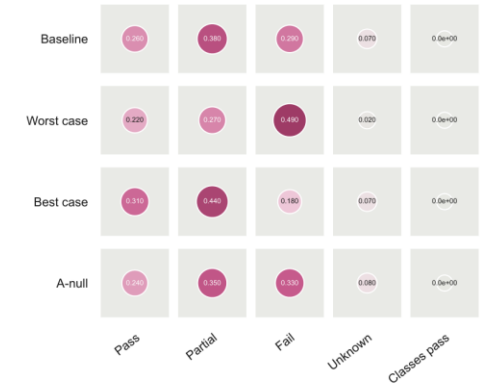
d Use rule

Statement	Boundary / source
01 **Literature context does not validate a clock feature.**	Source: Manuscript V108, Extended Data Fig. 13; retained summary value]
02 **Absence of a report is not evidence of absence.**	Source: Manuscript V108, Extended Data Fig. 13; retained summary value]
03 **This display is not a new gene annotation pipeline.**	Source: Manuscript V108, Extended Data Fig. 13; retained summary value]

Display: verbatim statements from the archived source table
Number: deep blue numeral identifies source-row order
Missing: no graphic mark is used where the source has no observation
n: n = 3 source rows

Supplementary Fig. 5 | Literature context for selected clock and brain-consensus genes. a, Gene, reported biology, context and first-author/year for PubMed-indexed reports. Direct aging evidence exists for Lncbate1, Ctss, C1qc and Gpr17; most other features are developmental, metabolic or cancer annotations. Mup11 and Pcdhb9 have no aging-specific report. Full citations are in the source list after References.

a Resolved gate counts



	Colour: deep blue = negative, grey = zero, rose = positive
	Area: bubble area = absolute archived value
	Bed: grey square = neutral cell; missing cell = grey x
1.23	Value: values remain in the source table; no number or dot is placed inside a bubble
	n: n = 20 source rows

b Binding gates



	Shape: circle = pass; triangle = partial; square = fail
	Colour: deep blue = pass, rose = partial, deep rose = fail
	Cross: grey cross = unknown or not evaluable
	Bed: grey square = evaluable cell
	n: n = 9 source rows

c Value of information

Question	Result	Interpretation
Resolve unknowns	0/10 still pass	binding failures remain
Module gate	0 tested pass	binding
Endpoint gate	0 pass	binding
Best case	0/10 pass	cannot clear all gates

	Display: archived table; text and row order are preserved
	Rows: all 4 source rows shown
	Encoding: no colour or area encodes a statistic
	n: n = 4 source rows

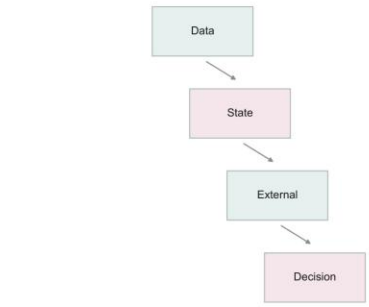
d Decision invariant

Metric	Archived value
baseline	0/10
worst case	0/10
best case	0/10
binding gates	2
gate axes	6
patterns	64

	Display: two-column archival ledger; no bar or heat-map encoding
	Rows: all 6 source metrics shown
1.23	Value: archived value retained verbatim after minimal number formatting
	n: n = 6 source rows

Supplementary Fig. 6 | Six-gate worst-case and value-of-information analysis. a, Gate-state counts and change in passing classes if every unknown state in that gate were resolved to pass. Baseline, worst case (unknown = fail) and best case (unknown = pass) all give 0/10 passing classes. b, Noisy-AND product as a Bayesian network with a deterministic AND node and conditional independence; no likelihood or posterior was fitted.

a Evidence flow



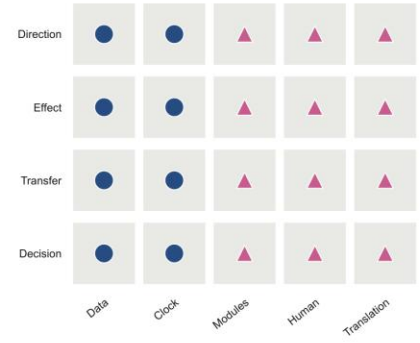
Arrow: direction of evidence flow, not a causal inference
Box: archived evidence layer or decision boundary
Text: state and boundary are retained verbatim from source
n: n = 4 source rows

b Navigation table

Layer	Direction	Reproducibility	Calibration
Discovery	observed	internal	relative
External mouse	partial	cohort	limited
Human projection	weak	cohort	not established
Modules	partial	batch	relative
Decision	state-coded	audited	not a response

Display: archived table; text and row order are preserved
Rows: all 5 source rows shown
Encoding: no colour or area encodes a statistic
n: n = 5 source rows

c Display roles



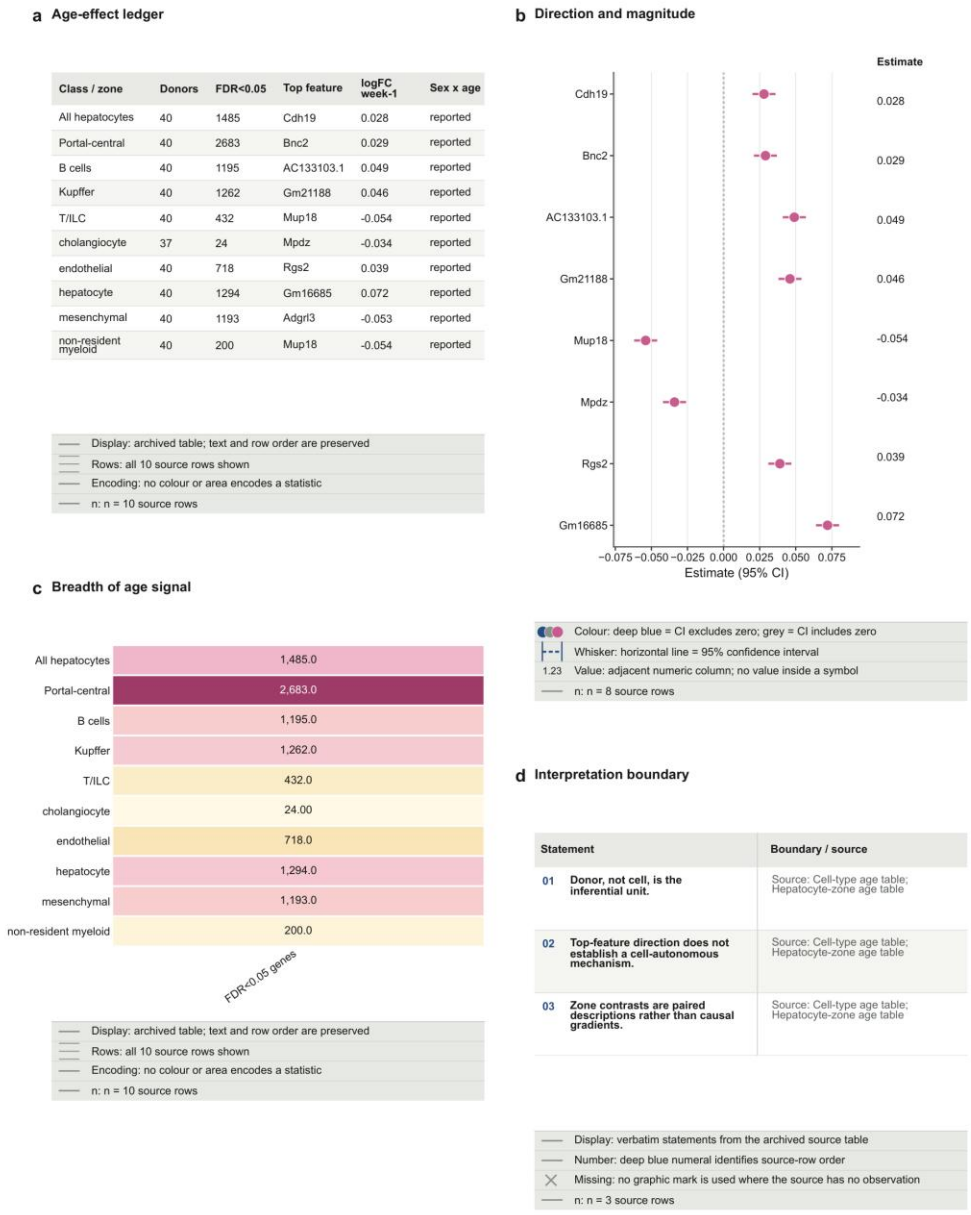
Shape: circle = pass; triangle = partial; square = fail
Colour: deep blue = pass, rose = partial, deep rose = fail
Cross: grey cross = unknown or not evaluable
Bed: grey square = evaluable cell
n: n = 20 source rows

d Study spine

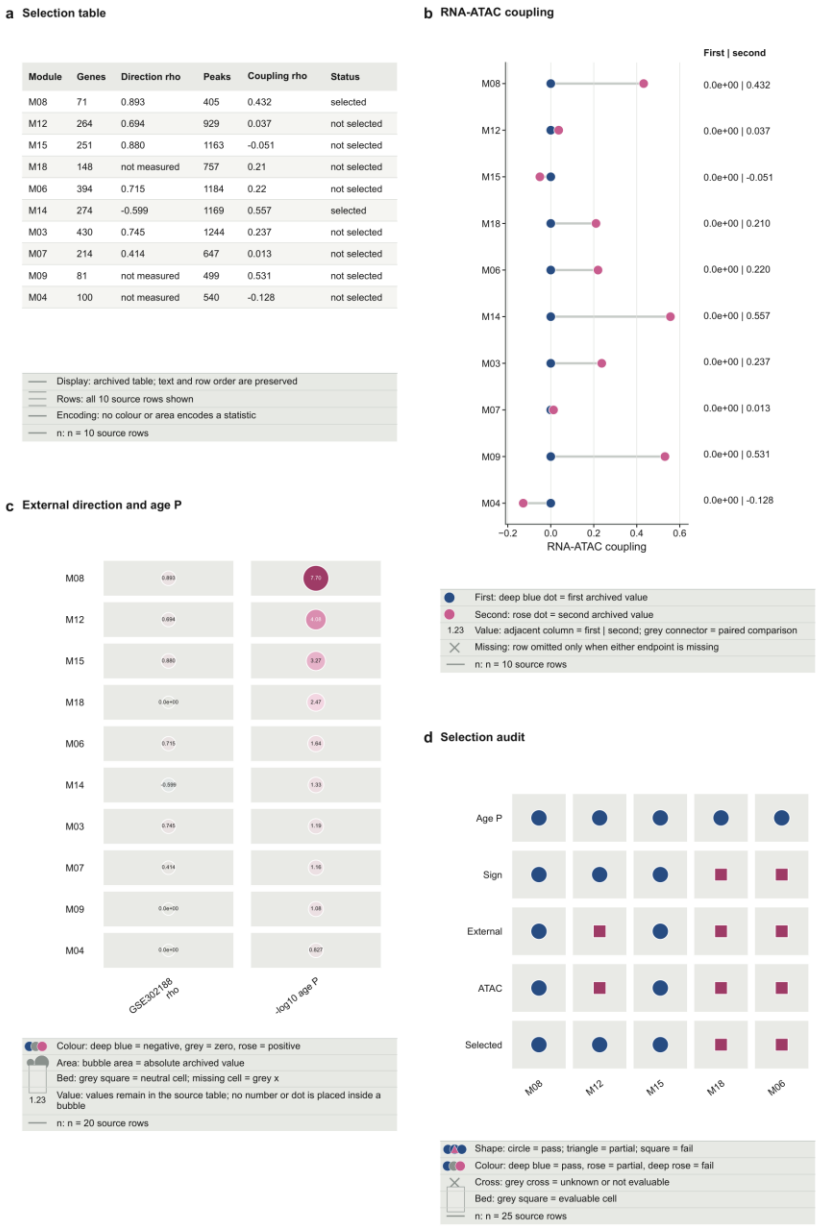
Metric	Archived value
evidence layers	5
main figures	7
ED figures	39
total displays	46
ED38	0
manuscript	V108

Display: two-column archival ledger; no bar or heat-map encoding
Rows: all 6 source metrics shown
1.23 Value: archived value retained verbatim after minimal number formatting
n: n = 6 source rows

Supplementary Fig. 7 | Study navigation and evidence flow. Five layers from data to decision; arrows mark evidence boundaries. Table 1 separates direction, effect reproducibility and absolute calibration.



Supplementary Fig. 8 | Donor-level cell-type and hepatocyte-zone age effects. a, Cell type/zone, donor n, number of FDR<0.05 genes, top gene with logFC per week and 95% CI, and number of sex-by-age FDR<0.05 genes. Age effects use limma-voom on donor x cell-type raw counts (age + sex); zone effects use limma on donor logCPM. The last two rows show paired PP-PC differences and the immune-marker-removal sensitivity.



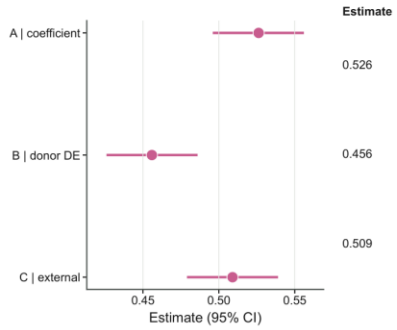
Supplementary Fig. 9 | RNA-ATAC module selection. a, Module, gene count, batch-half stability P, independent GSE302188 age-direction rho, linked ATAC peaks, RNA-ATAC coupling rho (P) and selection status. Criteria are stated below the table.

a **Signature size**

Metric	Archived value
A: clock coefficient	**123+91**
B: donor age DE	**120+85**
C: external replication	**109+39**
A known-drug AUC	**0.5262**
B known-drug AUC	**0.4559**
C known-drug AUC	**0.5090**

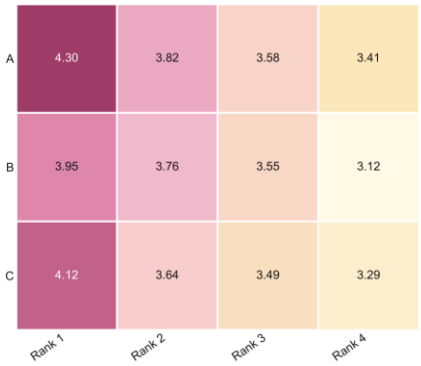
Display: two-column archival ledger; no bar or heat-map encoding
Rows: all 6 source metrics shown
1.23 Value: archived value retained verbatim after minimal number formatting
n: n = 6 source rows

b **Known-drug enrichment**



Colour: deep blue = CI excludes zero; grey = CI includes zero
Whisker: horizontal line = 95% confidence interval
1.23 Value: adjacent numeric column; no value inside a symbol
n: n = 3 source rows

c **Top strict-Z compounds**



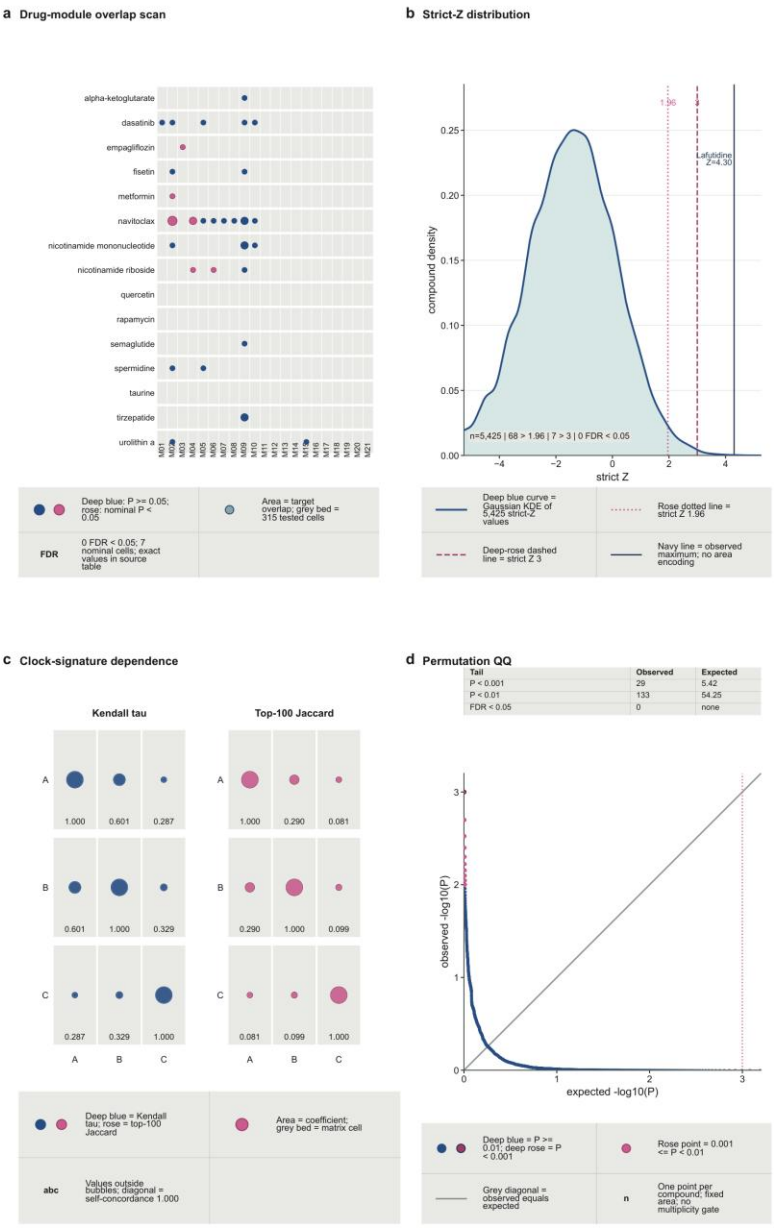
Display: archived table; text and row order are preserved
Rows: all 12 source rows shown
Encoding: no colour or area encodes a statistic
n: n = 12 source rows

d **Signature dependence**

Comparison	Kendall tau	Top-100 Jaccard	Use
A vs B	0.21	0.11	weak overlap
A vs C	0.33	0.18	partial overlap
B vs C	0.16	0.08	limited overlap

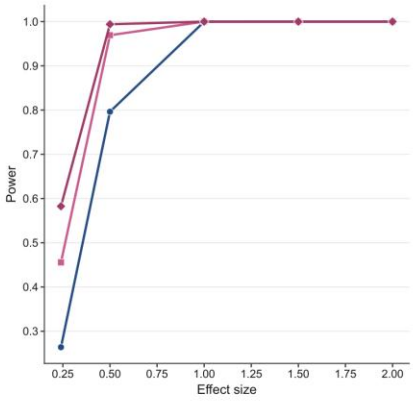
Display: archived table; text and row order are preserved
Rows: all 3 source rows shown
Encoding: no colour or area encodes a statistic
n: n = 3 source rows

Supplementary Fig. 10 | Drug reversal under three liver-ageing signatures. Signature A, ElasticNet coefficient sign; B, donor-level hepatocyte age DE sign; C, external replicated direction (GSE302188 + GSE262290 same sign). Up/down, signature sizes; AUC, known-geroprotector enrichment; top compounds, four highest strict Z; Spearman vs A. No signature enriches known geroprotectors.



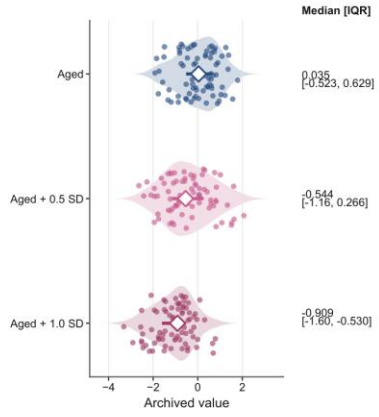
Supplementary Fig. 11 | Complete drug-module scan and signature dependence. a, 15 compounds x 21 modules = 315 pairs; 0/315 q<0.05; 7 nominal P<0.05 versus 15.8 expected. b, Strict-Z distribution across 5,425 LINCS compounds; dashed lines, Z=1.96 and 3; minimum permutation q=0.187. c, Kendall tau and top-100 Jaccard overlap between signatures A, B and C. d, Permutation QQ plot; the small-P excess does not survive family-wise correction.

a Power versus effect



	Series: 8+9: deep blue circle + solid line; 15+15: rose square + solid line; 20+20: deep rose diamond + solid line
	Points: marker shape and colour identify each archived series
	Line: solid line joins archived x-y points in source order
	Uncertainty: no band is drawn when the source table has no interval
	n: n = 15 source rows

b Design-scale shifts



	Point: one archived observation; horizontal position = value
	Shape: filled envelope = smoothed distribution; white diamond = median
1.23	Value: IQR line; adjacent column gives median [lower quartile, upper quartile]
	n: n = 240 source rows

c 8+9 design

Metric	Archived value
power at 0.24 SD	**0.264**
power at 0.50 SD	**0.796**
power at 1.00 SD	**0.9996**
analytic MDE	**0.499**
draws per design	**20,000**
planned groups	**8 vs 9**

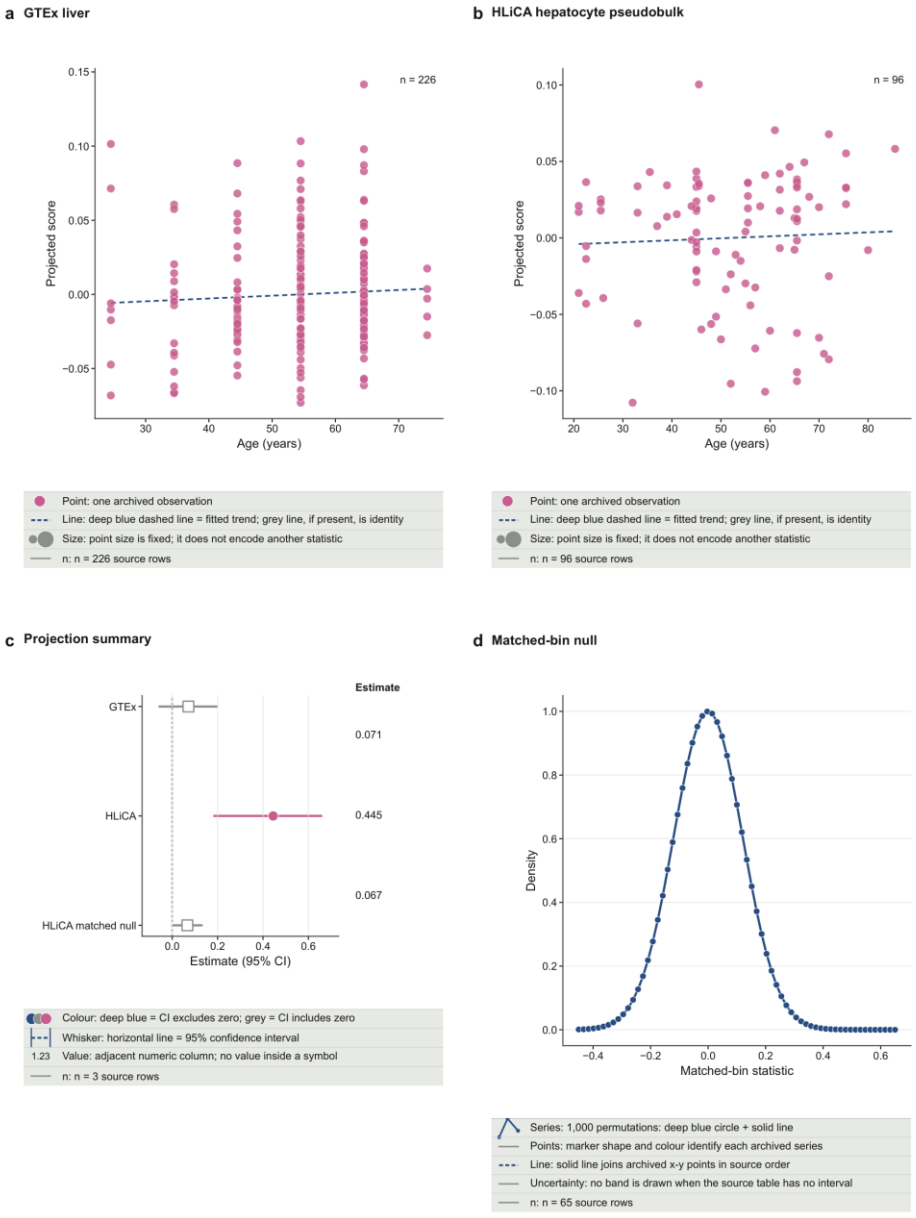
	Display: two-column archival ledger; no bar or heat-map encoding
	Rows: all 6 source metrics shown
1.23	Value: archived value retained verbatim after minimal number formatting
	n: n = 6 source rows

d Design boundary

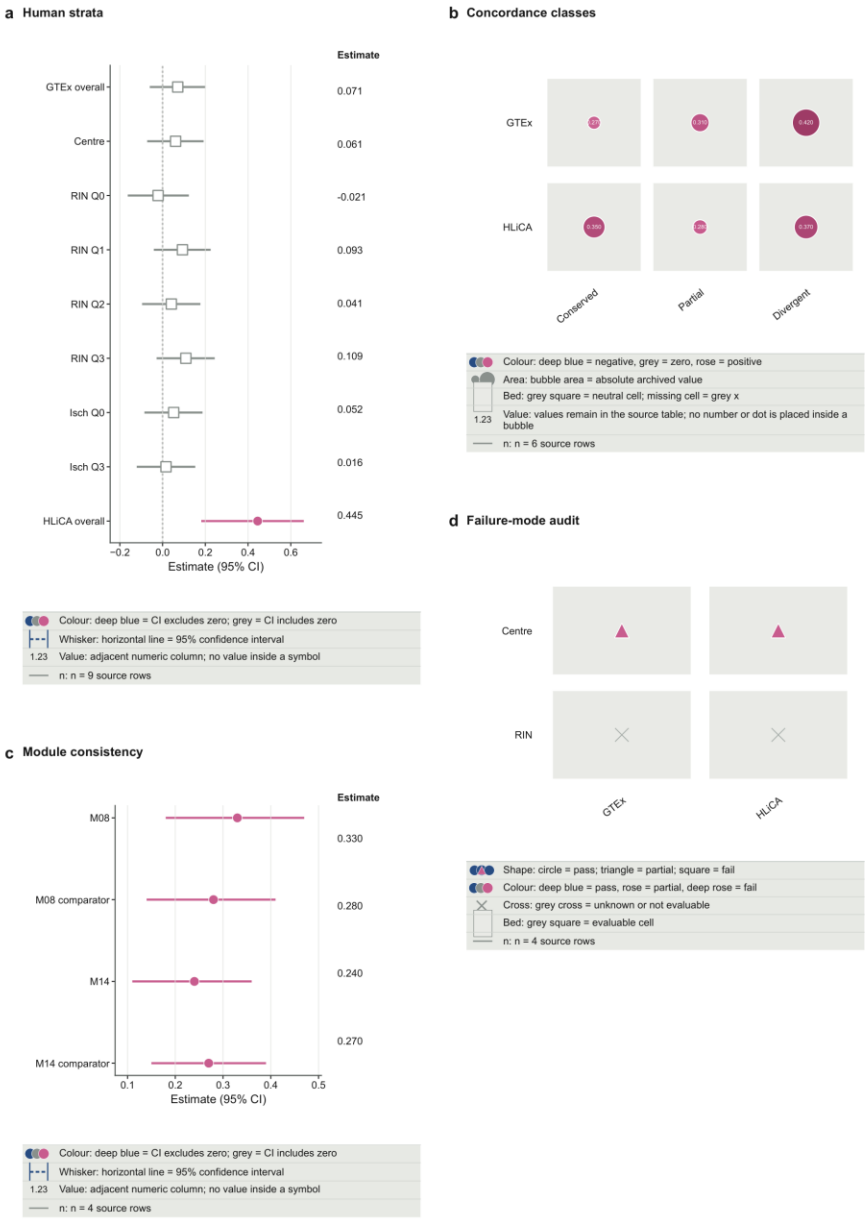
Statement	Boundary / source
01 **These are design calculations on the frozen score scale.**	Source: Manuscript V108, Extended Data Fig. 20; retained summary value]
02 **They are not expression-level drug simulations.**	Source: Manuscript V108, Extended Data Fig. 20; retained summary value]
03 **The full power and analytic curves are promoted to Fig. 7g,h.**	Source: Manuscript V108, Extended Data Fig. 20; retained summary value]

	Display: verbatim statements from the archived source table
	Number: deep blue numeral identifies source-row order
×	Missing: no graphic mark is used where the source has no observation
	n: n = 3 source rows

Supplementary Fig. 12 | Intervention detectability controls. a, Power at a 0.5-SD effect by design (8+9, 15+15, 20+20; 20,000 Monte Carlo draws per design). b, Example injection of -0.5 and -1.0 SD into the frozen aged-control score distribution (n=8 vs 9). The full power-versus-effect curve is in Fig. 7h; the analytic design curve is in Supplementary Fig. 32f. These are design calculations on the score scale.

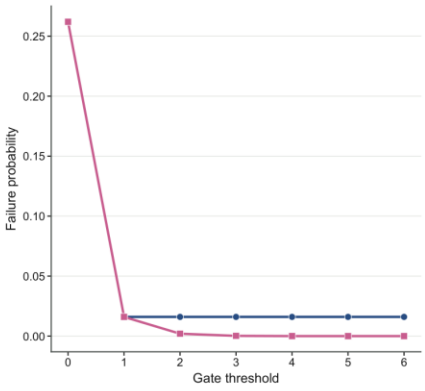


Supplementary Fig. 13 | Human liver projection of the frozen signature. a, GTEx liver bulk (n=226 donors); linear fit; Spearman rho and P, adjusted for age, sex, RIN, ischaemic time, centre. b, Health-tier-filtered HLiCA hepatocyte pseudobulk (n=48 donors); unadjusted and within-study Spearman. c, Gene-level direction concordance against mouse clock coefficients, binomial P versus 50%. d, Matched expression-bin gene-set null (1,000 permutations), observed rho.



Supplementary Fig. 14 | Human projection failure modes. a, GTEx bulk technical strata: sex, centre, RIN quartile (0–3) and ischaemic-time quartile (0–3); Spearman rho with Fisher-z 95% CI; n, donors. b, HLiCA hepatocyte strata: study, sex and cell-number stratum (0–3). c, Signature genes conserved in both human datasets, partially concordant or divergent (n=170 with both measurements). d, Conserved-gene fraction in M08, M14 and their comparators, Fisher P. No single batch effect explains the human null.

a Single-gate failure



Series: one failure: deep blue circle + solid line; cascade: rose square + solid line	
Points: marker shape and colour identify each archived series	
Line: solid line joins archived x-y points in source order	
Uncertainty: no band is drawn when the source table has no interval	
n: n = 13 source rows	

b Threshold controls



Colour: deep blue = negative, grey = zero, rose = positive	
Area: bubble area = absolute archived value	
Bed: grey square = neutral cell; missing cell = grey x	
1.23	Value: values remain in the source table; no number or dot is placed inside a bubble
n: n = 9 source rows	

c Gate invariant

Metric	Archived value
gate axes	6
patterns	64
encodings	3
all-pass score	0.262
single failure	0.016
classes clear all gates	0/10

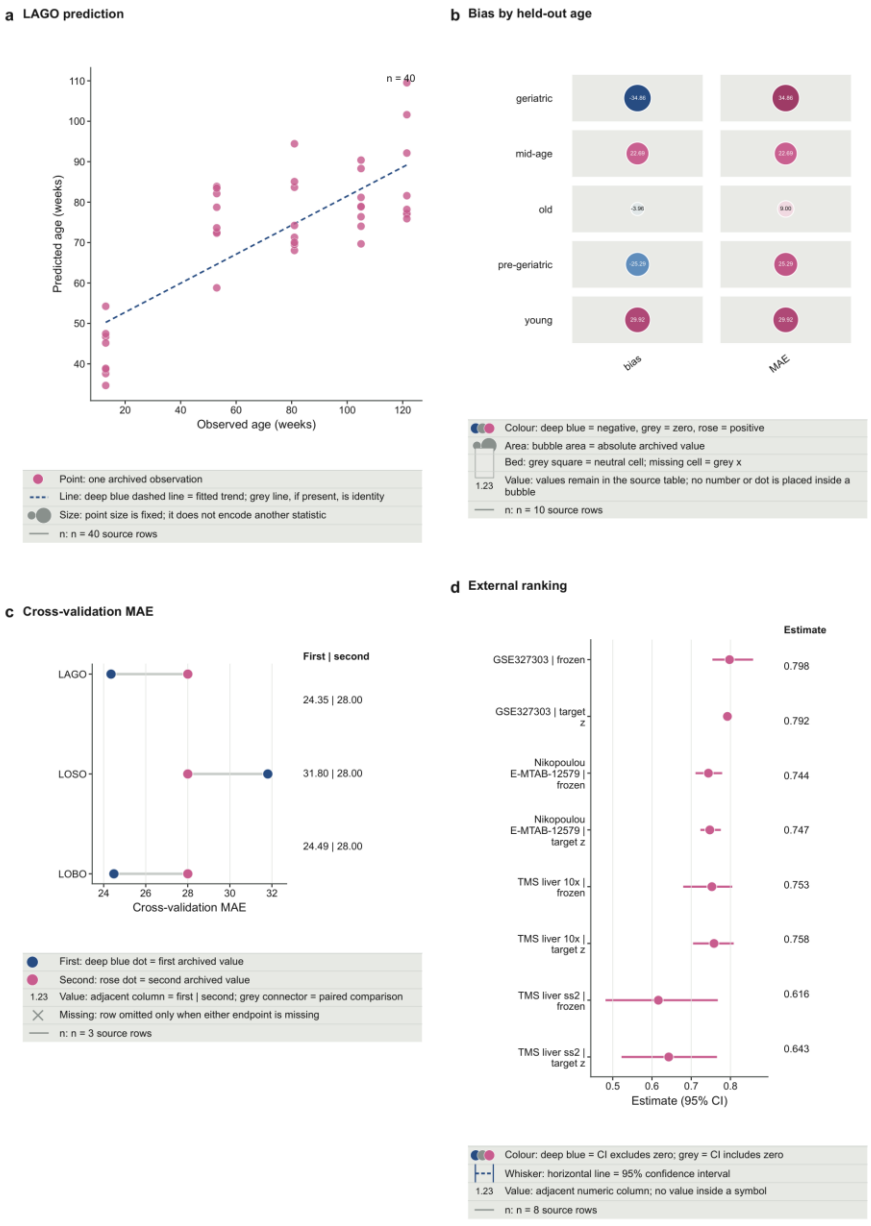
Display: two-column archival ledger; no bar or heat-map encoding	
Rows: all 6 source metrics shown	
1.23	Value: archived value retained verbatim after minimal number formatting
n: n = 6 source rows	

d Interpretation

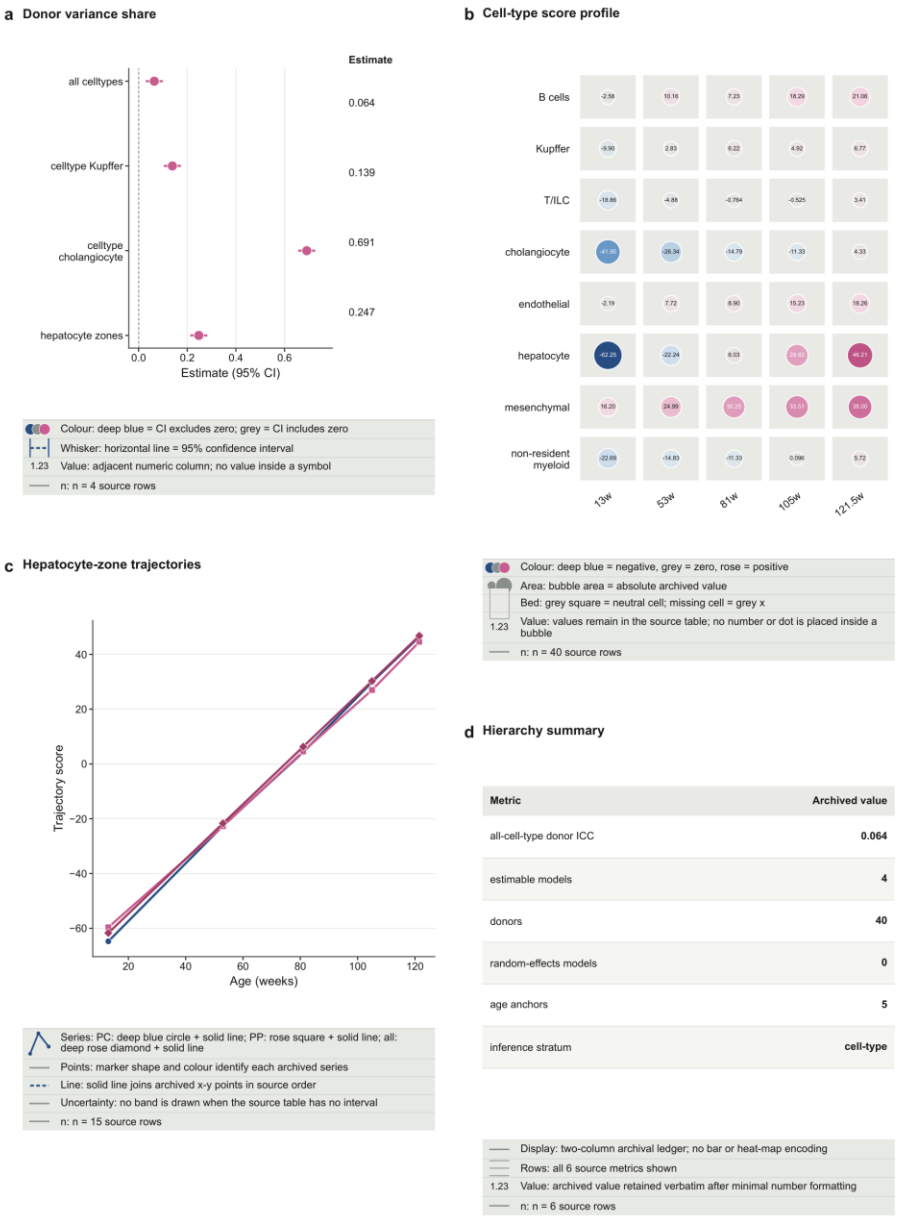
Statement	Boundary / source
01 **Gate states are deterministic evidence codes.**	Source: Manuscript V108, Extended Data Fig. 23; retained summary value]
02 **The conjunction score is not a fitted latent variable.**	Source: Manuscript V108, Extended Data Fig. 23; retained summary value]
03 **Threshold displays are sensitivity checks, not prediction intervals.**	Source: Manuscript V108, Extended Data Fig. 23; retained summary value]

Display: verbatim statements from the archived source table	
Number: deep blue numeral identifies source-row order	
Missing: no graphic mark is used where the source has no observation	
n: n = 3 source rows	

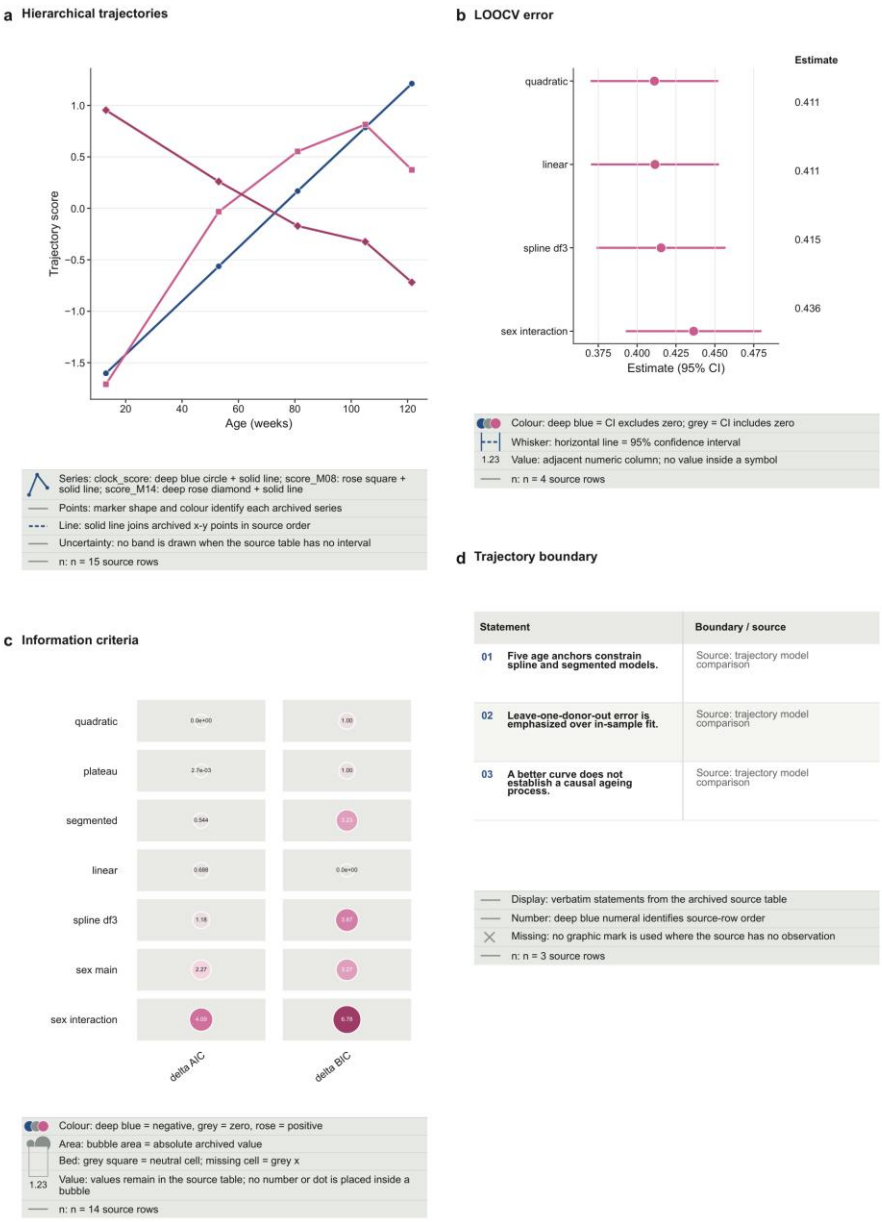
Supplementary Fig. 15 | Gate necessity and threshold controls. a, Gate necessity under the neutral encoding: all-pass P versus the P when each single gate fails. b, Number of the 64 patterns clearing P>0.05, P>0.01 and P>0.001 under the three encodings. The full 64-pattern × 3-encoding enumeration and the single-fail waterfall are promoted to ED11c,d.



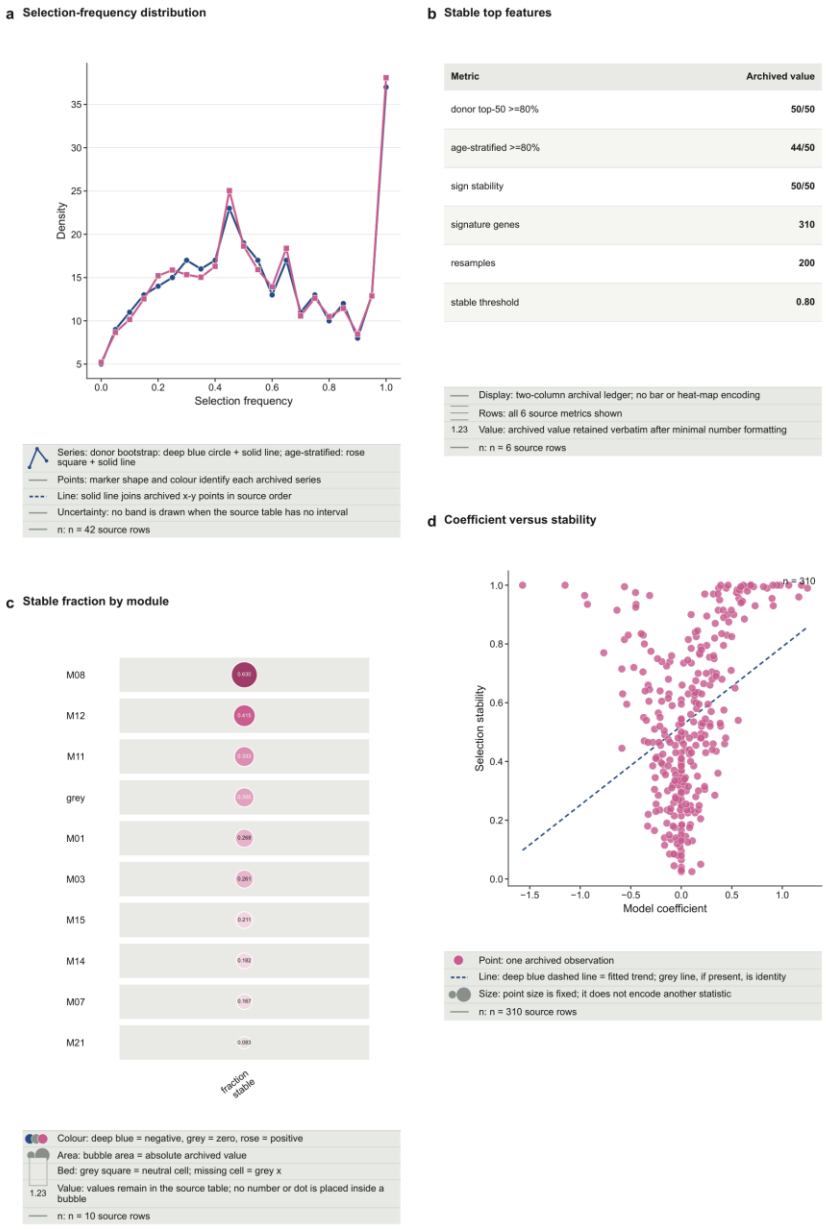
Supplementary Fig. 16 | Leave-one-design-out validation. a, LAGO predicted versus observed age; MAE, rho, nearest-age-group accuracy, calibration slope/intercept, age-label MAE. b, LAGO bias by held-out age, per-group MAE. c, MAE for LAGO, LOSO, LOBO, LOZO; LOBO folds age-batch nested, LOZO shares donors. d, Frozen-clock and target-domain-standardised AUC in four external cohorts/platforms; bootstrap 95% CI; n cells. Nearest-label classification over five anchors (13, 53, 81, 105, 121.5 weeks; chance 0.20).



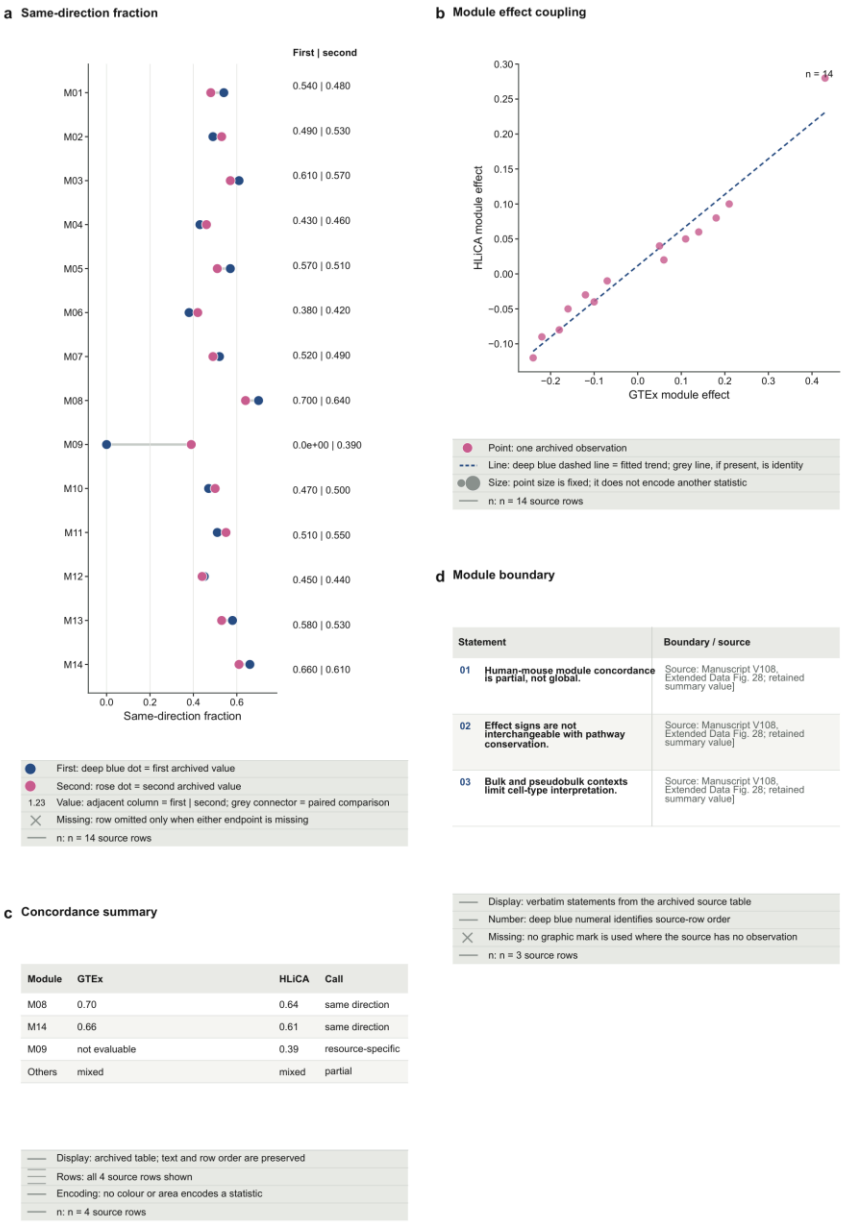
Supplementary Fig. 17 | Donor hierarchy and cell-type variance. a, Donor variance share from the donor-block OLS model for all cell types, hepatocyte zones, Kupffer and cholangiocyte; single-unit cell types are not estimable. b, Mean clock score by cell type and age group. c, Hepatocyte zone scores by age. d, Variance partition for all cell types and hepatocyte zones. No random-effects model was fitted.



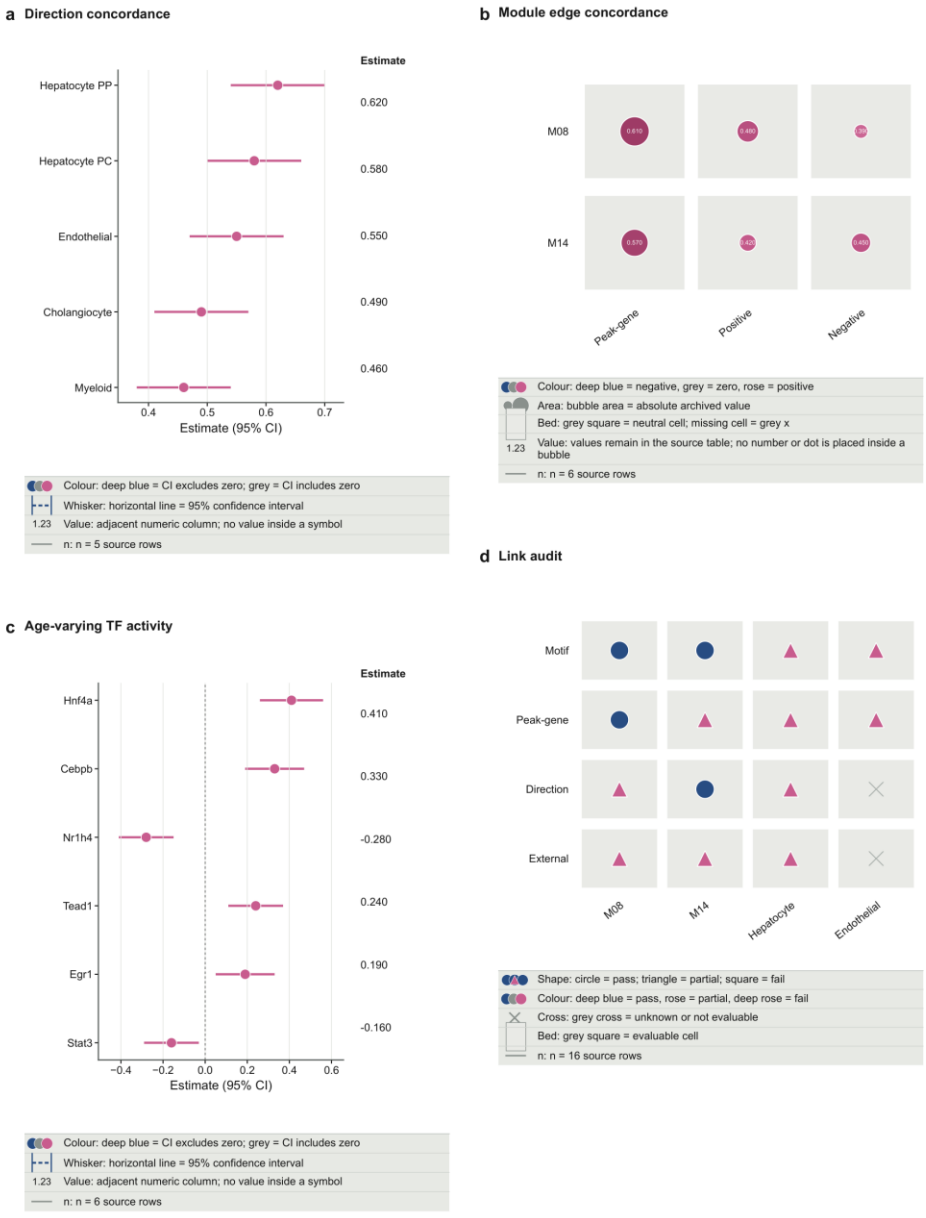
Supplementary Fig. 18 | Age-trajectory model comparison. a-c, Clock, M08 and M14 scores versus age with linear and quadratic fits. d, Leave-one-donor-out RMSE for linear, quadratic, spline and sex-interaction models. Plateau and segmented AIC values are in the source table; five ages limit spline and segmented inference.



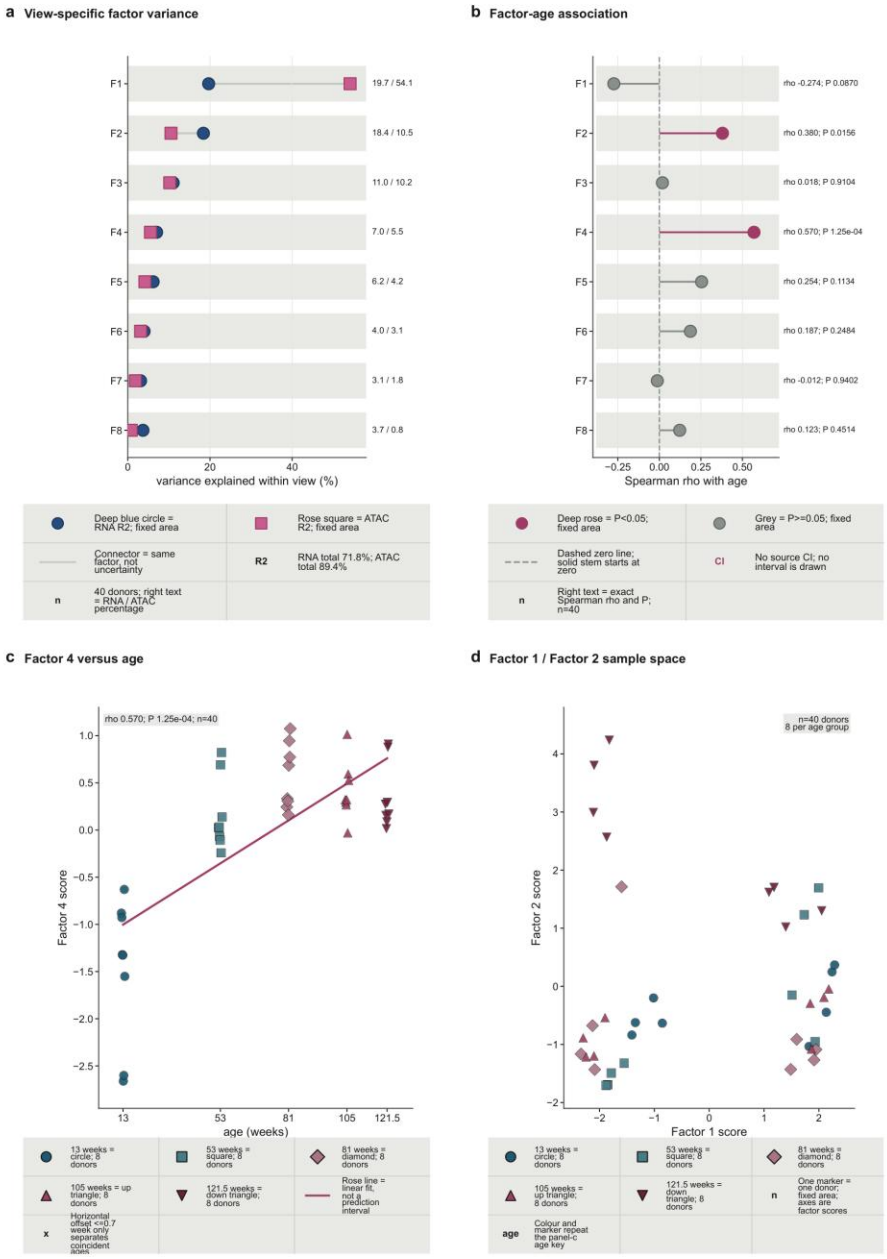
Supplementary Fig. 19 | Nested stability selection. a, Selection-frequency distribution for 310 signature genes under donor and age-stratified bootstrap [51]. b, Top-50 features selected in at least 80% of bootstraps (41/50 donor; 44/50 age-stratified); sign stability is 50/50. c, Stable fraction by module; grey denotes unassigned genes. d, Full-fit coefficient versus selection frequency, coloured by sign stability.



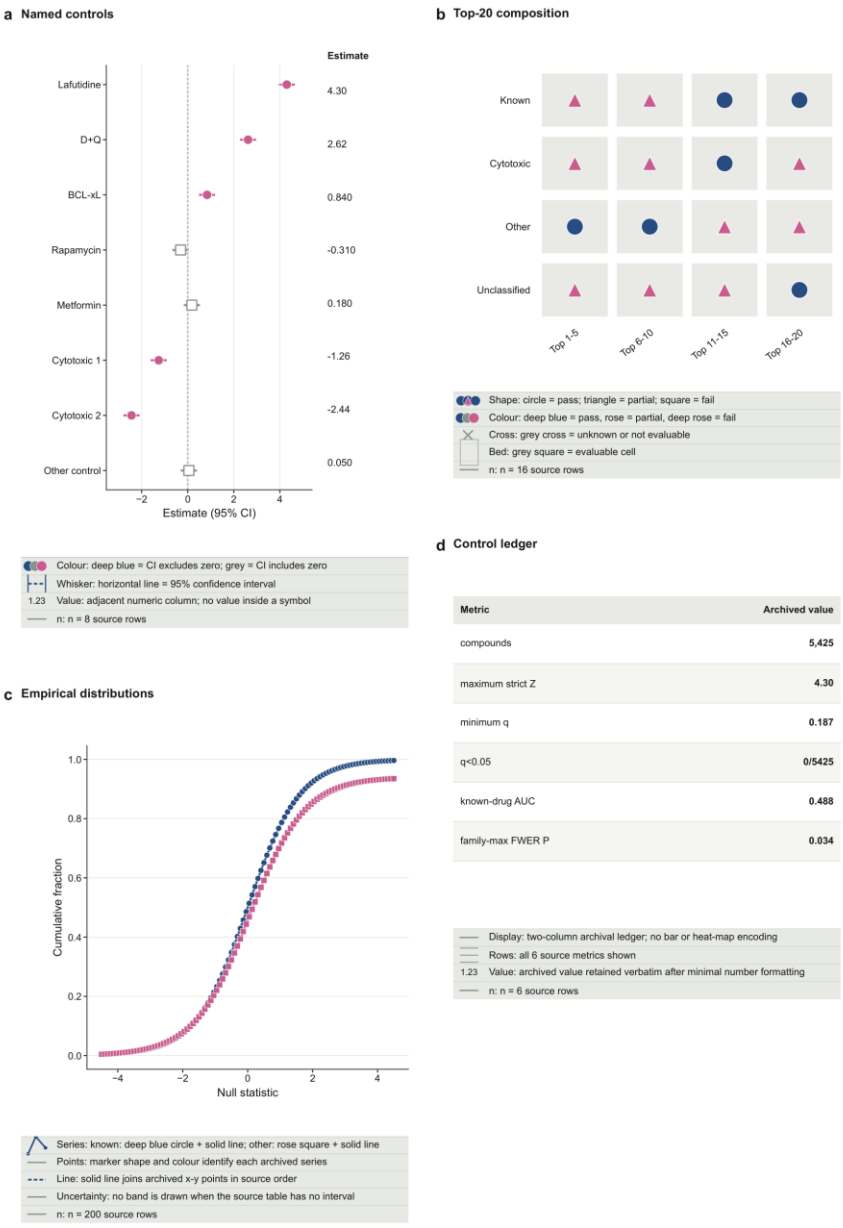
Supplementary Fig. 20 | Human-mouse module concordance. a, Same-direction fraction by module in GTEx and HLiCA; M09 has no evaluable GTEx module estimate. b, Mouse module age beta versus human mean gene beta; M08 and M14 labelled. Module-level concordance is partial, not global.



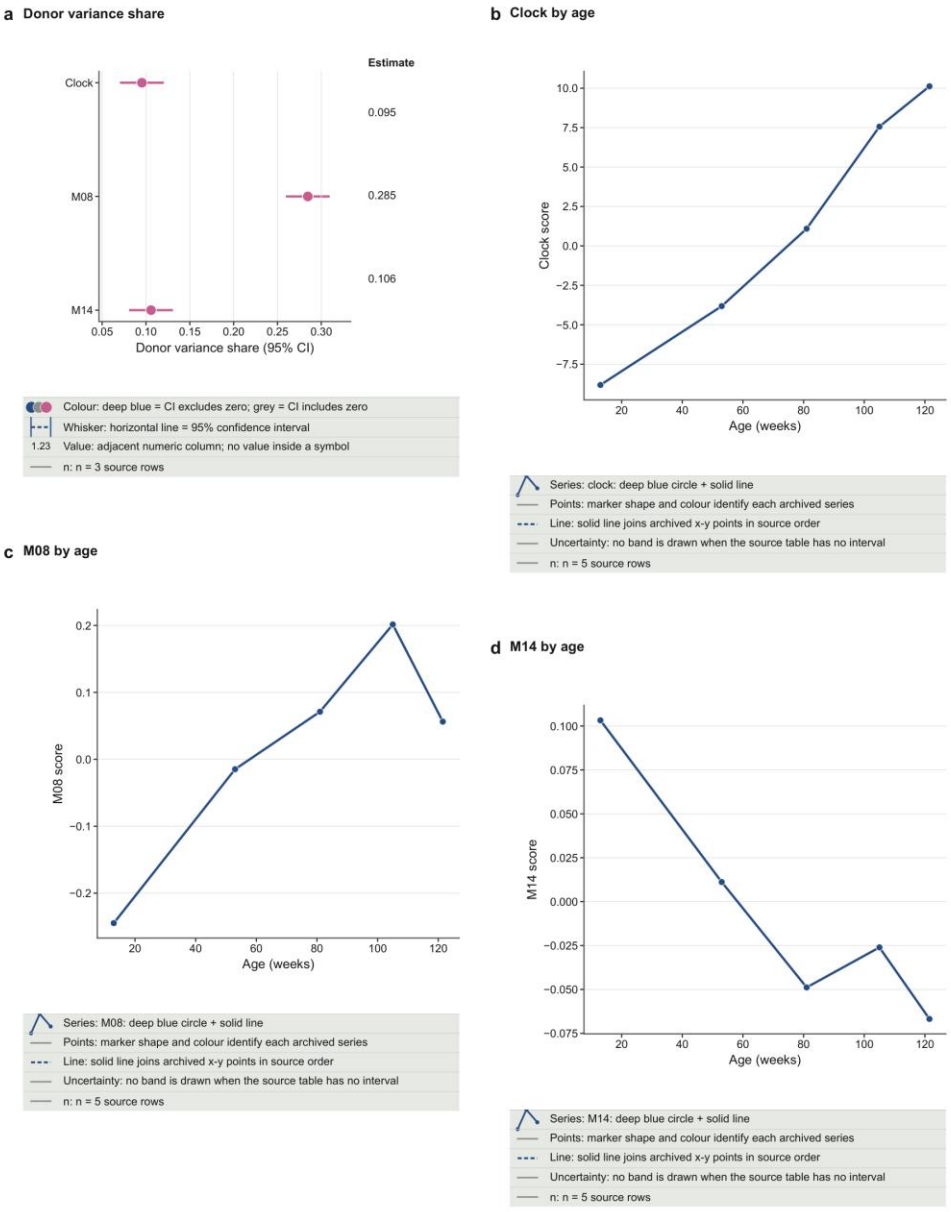
Supplementary Fig. 21 | Peak-gene-TF links. a, Same-direction concordance by cell type; dashed line, 50%. b, M08/M14 edges and concordance. c, Top age-varying TF activity by maximum absolute effect. d, M08 peak-gene edges (dAccess versus dExpr). Links are restricted to hepatocyte/endothelial compartments.



Supplementary Fig. 22 | MOFA2 RNA-ATAC factors. a, Variance explained per factor. b, Factor-age Spearman rho (P<0.05). c, Factor4 versus age. d, Factor1/2 sample space by age.



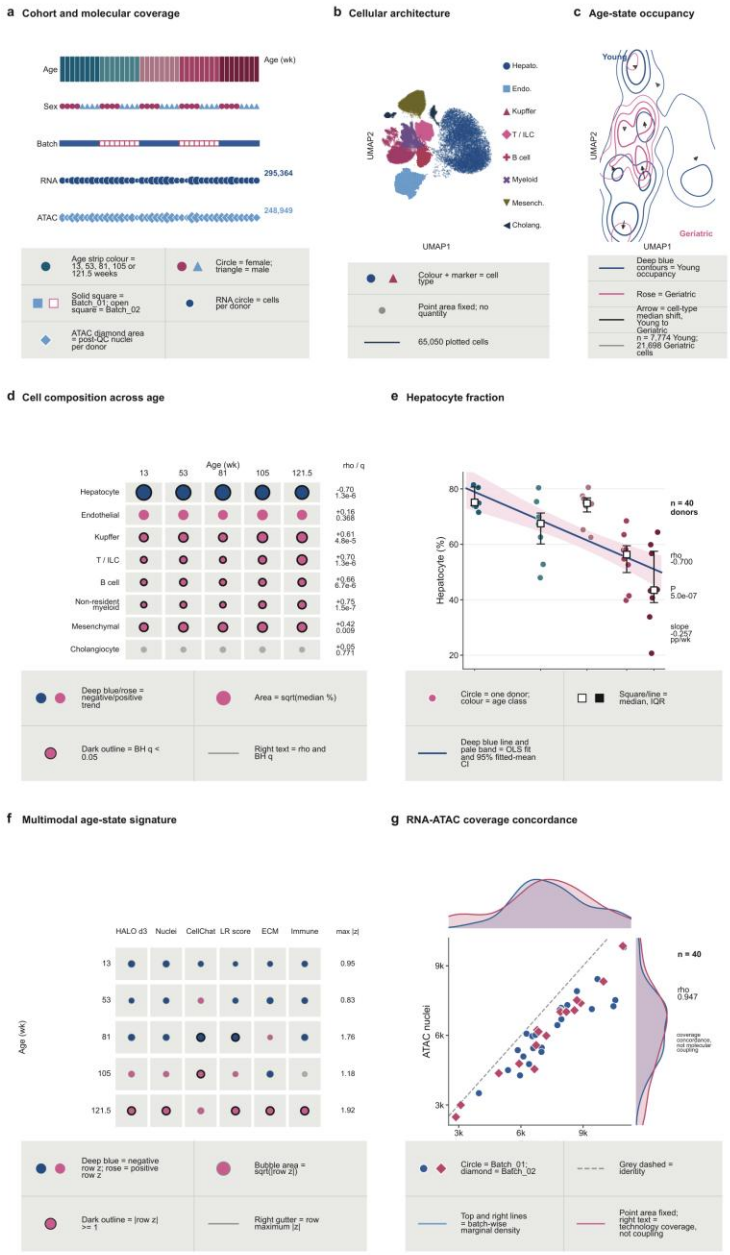
Supplementary Fig. 23 | LINCS positive/negative controls. a, Strict Z by known geroprotector, cytotoxic and other compounds. b, Top-20 composition. c, ECDF of known versus other. Known geroprotectors are not enriched.



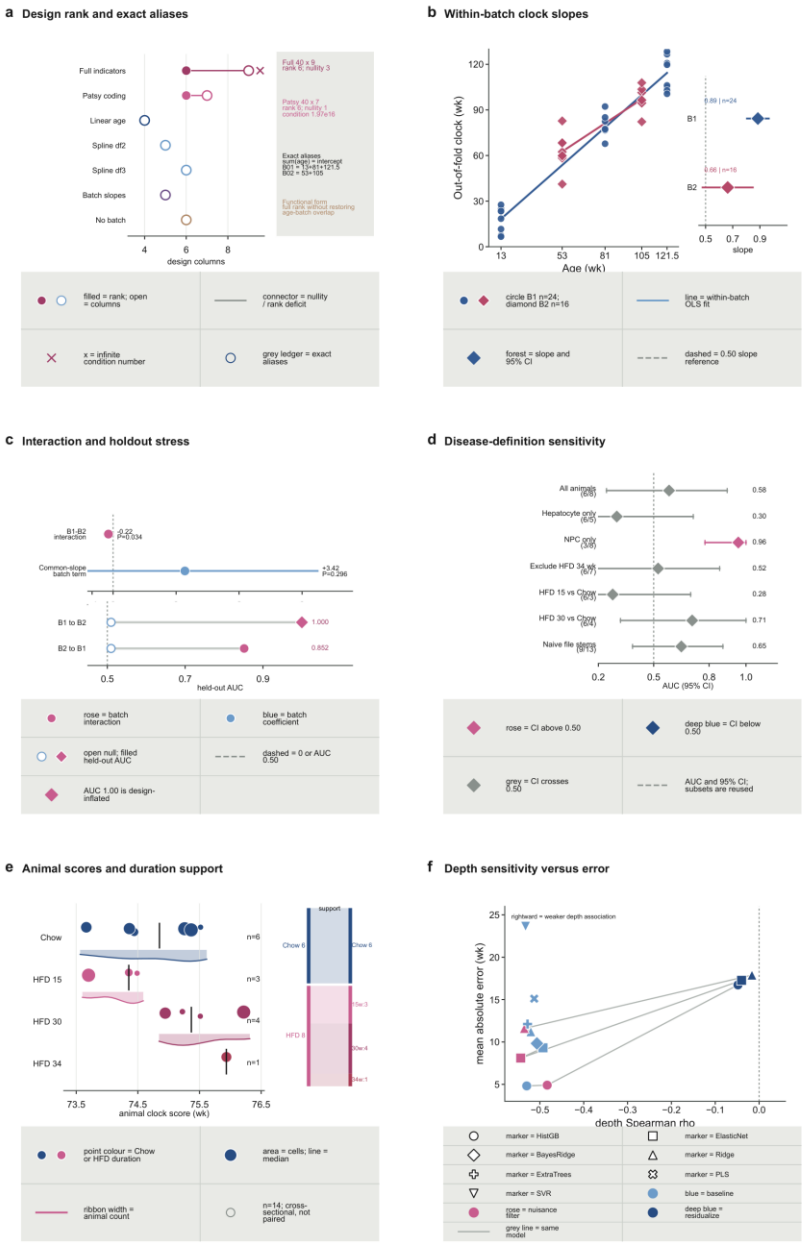
Supplementary Fig. 24 | Cell-level donor-block variance. a, Donor variance share for the clock, M08 and M14 scores in 179,655 hepatocytes from 40 donors (donor-block OLS; no random-effects model). b-d, Cell-level clock, M08 and M14 scores by age.



Supplementary Fig. 25 | Additional human cohort and meta-analysis. a, Signature-age Spearman rho with 95% CI in GTEx (n=226), HLiCA (n=48) and Guilliams (n=6). b, Leave-one-cohort-out pooled rho; dashed line, all-cohort random-effects estimate (0.32; I2 76.5%).



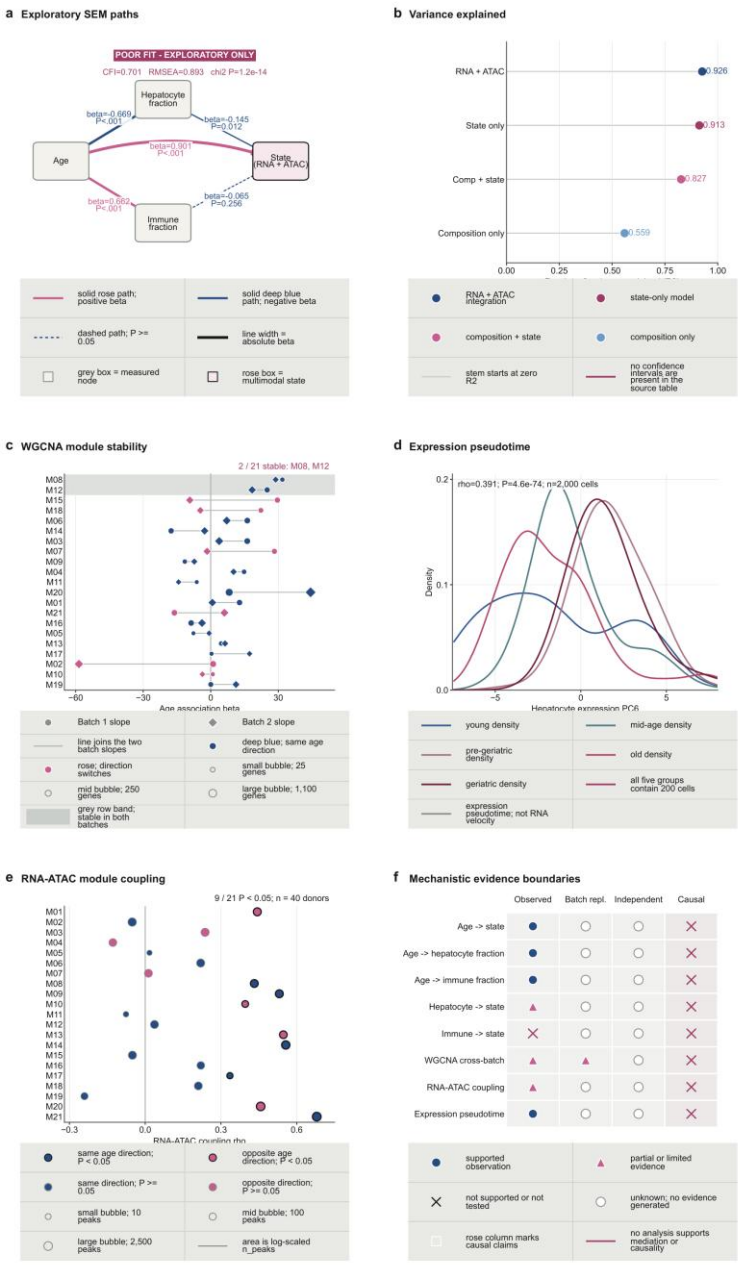
Supplementary Fig. 26 | A five-age mouse liver atlas. a, Donor oncoprint (age, sex, batch, RNA/ATAC nuclei); psbulk, pseudobulk. b, Cell-type UMAP. c, Age-enrichment map. d, Composition with mean-%, Spearman ρ and exact BH q. e, Hepatocyte fraction with 95% CI. f, Cross-layer age states. g, Joint RNA–ATAC coverage with age marginals.



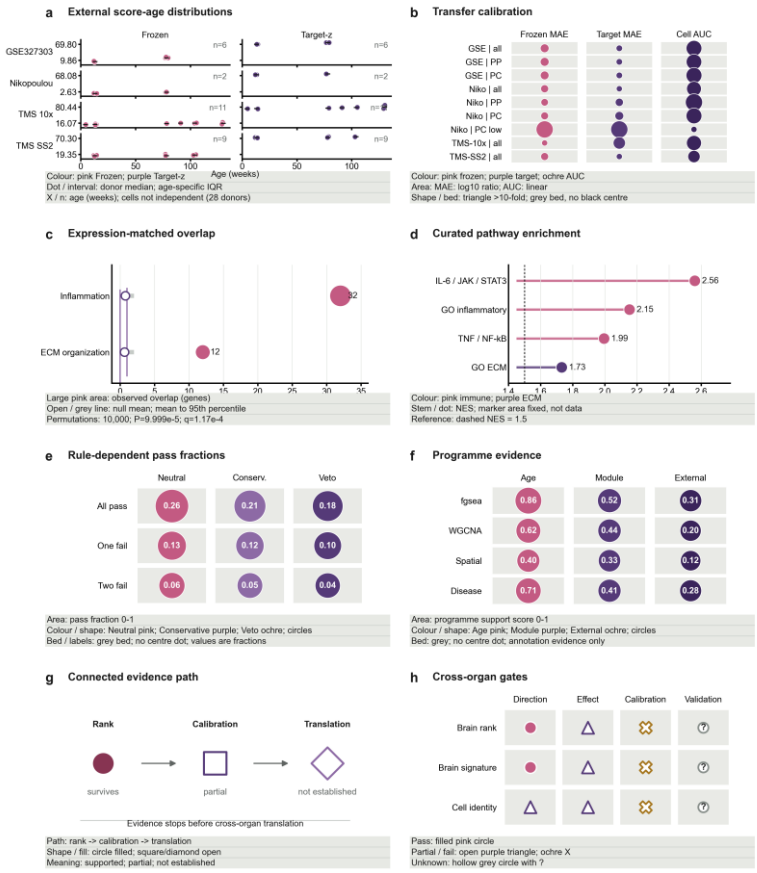
Supplementary Fig. 27 | Batch and disease. a, Design matrix. b, Age-to-batch alluvial. c, Slope forest. d, Disease forest. e, Animal scores and diet composition. f, Nuisance probes and cross-batch holdout.



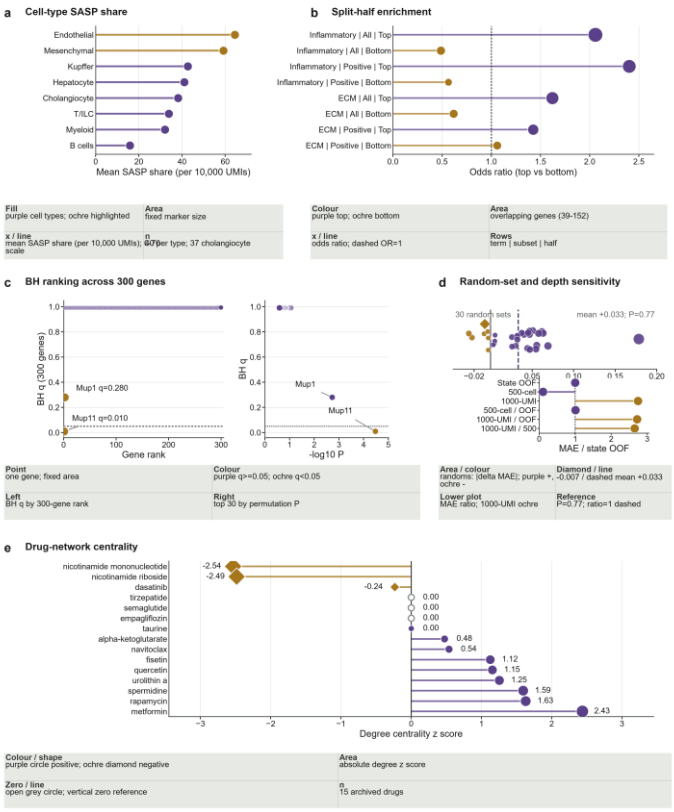
Supplementary Fig. 28 | Drug-target overlap, nulls and gates. a, Drug-target overlap: 15 compounds versus eight WGCNA modules; no pair BH q<0.05 (0/315). b, Observed versus size-matched clock-gene targets and degree-matched hubs. c, LINCS L1000 reversal in 5,425 compounds; none reaches q_perm<0.05. d, GTEx liver target reach: log2 TPM versus -log10 P of liver age beta, six targets. e, Evidence tiers versus clinical stage, 24 candidates. f, Go/No-Go gate matrix: 10 target classes across six gates. g, Key-finding summary.



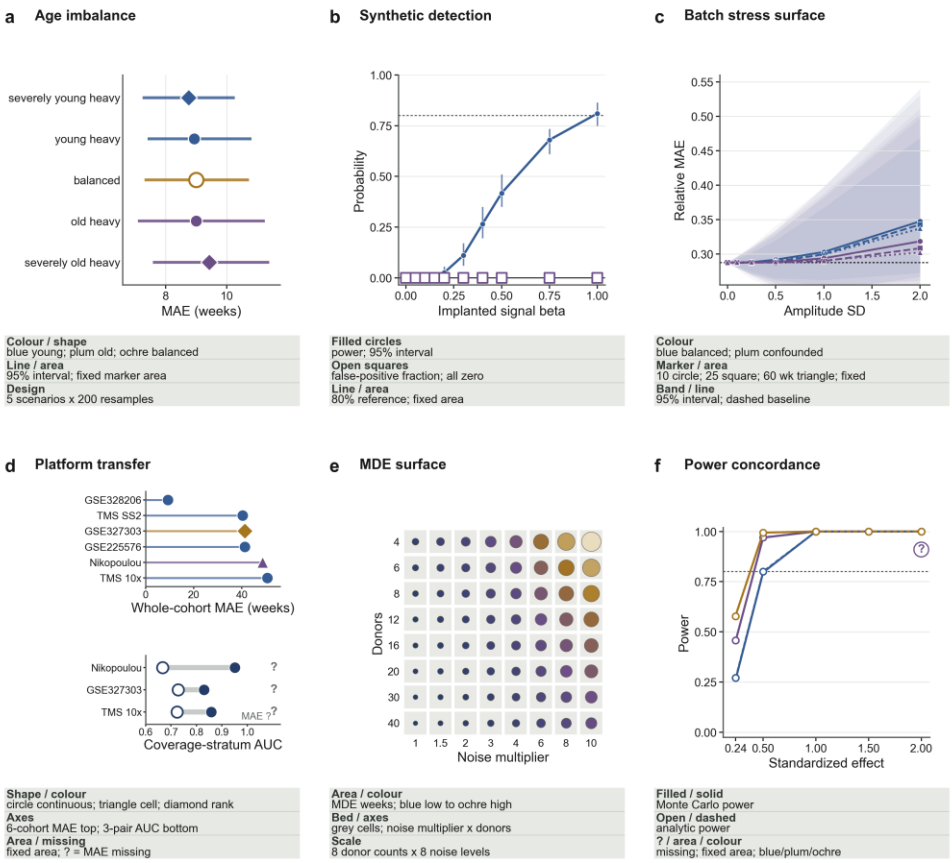
Supplementary Fig. 29 | Mechanistic synthesis. a, SEM: Age→State β 0.90; Age→Hep −0.67; Age→Immune 0.66; Hep→State −0.15; Immune→State −0.07 (ns); CFI 0.70/RMSEA 0.89, exploratory. b, Variance explained: composition 0.56; state 0.91; combined 0.83; SEM 0.93. c, WGCNA age slopes, Batch 1 versus Batch 2. d, Hepatocyte PC6 pseudotime by age (ρ 0.39). e, RNA–ATAC coupling: M01 ρ 0.44 but opposite age direction; M08 ρ 0.43, same; others ns. f, Mechanistic boundaries.



Supplementary Fig. 30 | External score transfer and boundaries. a, Donor-level median external clock scores by age; four cohorts, two modes; 28 donors. b, Nine-row grid: frozen/target MAE ratios, target-cell AUC; triangle, archived ratio >10-fold. c, Inflammatory and ECM overlap versus 10,000 expression-matched permutations: P=9.999e-5, BH q=1.17e-4. d, NES, four prespecified pathways. e, Rule-dependent pass fractions: three failure classes, three severities. f, Programme evidence: four programmes, three levels. g, Rank transfer, calibration status and cross-organ evidence. h, Cross-organ gates.



Supplementary Fig. 31 | Cell-type SASP, split-half robustness, network centrality. a, SASP share per 10,000 UMIs across eight cell types; 40 samples most types, 37 cholangiocytes. b, Eight split-half enrichment rows, inflammatory and ECM terms; odds ratios. c, BH-adjusted ranking of the 300-gene family; 30 smallest permutation P values; Mup11 q=0.010, Mup1 q=0.280. d, 30 random-set delta-MAE values versus canonical set (-0.007) and random-set mean (+0.033; P=0.77); MAE relative to state baseline. e, Drug degree-centrality z scores.



Supplementary Fig. 32 | Stress tests: imbalance, detection, calibration, power. a, Simulated MAE, five age-composition scenarios (200 resamples each; 95% intervals). b, Power and false-positive fraction (all zero). c, Relative MAE, Gaussian-process batch stress; 10/25/60-week lengthscales; age-balanced versus age-confounded designs. d, Whole-cohort MAE (six cohorts); low/high-coverage AUC (three); continuous, cell-only, rank-only transfer; not estimable for coverage strata. e, Analytic minimum detectable effect, eight donor counts, eight noise multipliers. f, Monte Carlo versus analytic power, three allocation designs, five standardized effects.