

Cite as: Q. Yu *et al.*, *Science*
10.1126/science.ael0508 (2026).

A DNA Typewriter records the cell lineage history of a mouse, from zygote to late organogenesis

Qi Yu^{1,2,†}, Haedong Kim^{1,2,3,†}, Sophie Seidel^{1,2,†}, James F. Acosta-Clark^{2,4}, Beth K. Martin^{1,2}, Kyle O'Connor^{2,4}, Riza M. Daza^{1,2}, Molly Gasperini^{2,4}, Jenny F. Nathans^{1,2,5}, Maggie Lam¹, Elena Gamo^{2,4}, Shruthi Vijay Kumar^{2,4}, Lauren Kuo^{2,4}, Jean-Benoît Lallanne^{1,2,†}, Kamen P. Simeonov⁶, Marion Pepper^{2,7}, Cole Trapnell^{1,2,8,9,10}, Jesse M. Gray^{2,4}, Junhong Choi^{1,§}, Chengxiang Qiu^{11*}, Jay Shendure^{1,2,3,4,8,9,10,12*}

¹Department of Genome Sciences, University of Washington, Seattle, WA, USA. ²Seattle Hub for Synthetic Biology, Seattle, WA, USA. ³Howard Hughes Medical Institute, Seattle, WA, USA. ⁴Allen Institute, Seattle, WA, USA. ⁵Medical Scientist Training Program, University of Washington, Seattle, WA, USA. ⁶Public Health Sciences Division/Translational Research Program, Fred Hutchinson Cancer Center, Seattle, WA, USA. ⁷Department of Immunology, University of Washington, Seattle, WA, USA. ⁸Brotman Baty Institute for Precision Medicine, Seattle, WA, USA. ⁹Institute of Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA, USA. ¹⁰Allen Discovery Center for Cell Lineage Tracing, Seattle, WA, USA. ¹¹Department of Molecular and Systems Biology, Dartmouth College, Hanover, NH, USA. ¹²AI BioDesign, Seattle, WA, USA.

*Corresponding author. E-mail: shendure@uw.edu (J.S.); Chengxiang.Qiu@dartmouth.edu (C.Q.)

†These authors contributed equally to this work.

‡Present address: Dépt. de Biochimie et Médecine Moléculaire & Courtois Institute for Biomedical Innovation, Université de Montréal, Montréal, QC, Canada.

§Present address: Dev. Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

Mammalian biology unfolds over time, within tissues and organs opaque to our eyes and instruments. We applied DNA Typewriter, a sequential molecular recorder, to record the cell lineage of a mouse over nearly two weeks of development. From one embryo, we reconstruct a time-calibrated, parsimony-supported, zygote-rooted phylogeny of 1.28 million transcriptionally profiled cells. A burst of editing unequivocally marks the daughters of the first cleavage, which serve as inline replicates. We quantify clonal dominance arising during gastrulation. Tree siblings share cell type far above chance; heterotypic siblings mark terminal differentiations. Temporal sweeps of clade co-occurrence recover a dated hierarchy of cell-type couplings; imputed labels for internal nodes recapitulate known state paths. A lineage-anchored ontogeny of mammalian development, long out of reach, is coming into view.

A half-century ago, Sulston and colleagues watched *Caenorhabditis elegans* develop under the microscope, painstakingly drawing each cell division until they had mapped its complete lineage, zygote to adult (1, 2). It remains one of the singular accomplishments of modern biology, and no comparable lineage has been mapped for any other animal since. But *C. elegans* is exceptional: transparent, small, and developmentally invariant.

Nearly all other animals are opaque, and many are vastly larger, for example, an adult mouse contains billions and an adult human contains trillions of cells (3). They also develop in the opposite, regulative mode. Cell fates are shaped by extrinsic cues rather than lineage alone, and the precise sequence of divisions varies between individuals. Lineage is nonetheless far from irrelevant, and the outstanding mystery of regulative development is how a variable lineage is canalized into reproducible form and function.

To study mouse development at the whole-organism scale with genomic tools, we and others have performed molecular profiling of cells sampled from temporally staged individuals (4–9). However, sampling is destructive, severing each cell from its past and future. Although a cell type ontogeny can

be inferred in broad strokes from such time-series, lineage relationships cannot (10).

That the genome is a digital medium, faithfully copied at each cell division, makes it natural to exploit for recording cell lineage. Many studies have leveraged natural somatic variation, static barcodes or recombinase-diversified barcodes to this end (11–18), but these strategies have not delivered dense cellular phylogenies (19, 20).

Around 2016, we and others began to pursue a paradigm wherein genome editors introduce heritable marks that accrue throughout development (19–25). Read at an endpoint by single-cell profiling of endogenous transcripts and barcodes (25–28), a single destructive snapshot informs cells' present states and their relationship to other cells. Implementations group by editor class: Cas9 nucleases, whose error-prone repair leaves heritable scars (e.g., GESTALT, MEMOIR) (21–28); base editors, which install point mutations (e.g., SMALT, baseMEMOIR) (29–32); and prime editors, which write symbolic insertions (e.g., DNA Typewriter, peCHYRON) (33–35).

Despite progress, dense cellular phylogenies of large animals have remained out of reach for a confluence of technical

reasons (19), including: (i) delivery and expression: all components must reach the zygote and be expressed stably throughout development; (ii) information loss: double-strand breaks at densely packed sites trigger deletions that destroy adjacent records and media; (iii) instability of temporal resolution: as editable sites are consumed, the recording rate naturally falls; (iv) event ordering: if edits are unordered, their temporal order must be inferred; (v) scope: lineage histories alone don't reveal the historical signals and states that shaped cells' fates; (vi) cellular coverage: dense reconstructions require profiling a substantial proportion of an organism's cells; (vii) records retrieval: recovering barcodes from profiled cells/nuclei is technically demanding; (viii) tree building: algorithms developed for 10^2 - 10^4 taxa must scale to 10^5 - 10^9 cells/nuclei; and (ix) design-test cycle time: optimizing recorders (or anything) in vivo is slow.

DNA Typewriter (33) addresses, partly or wholly, four of these hurdles (**Fig. 1A**). Prime editing does not require double-strand breaks (ii). Each edit consumes a monomer and activates a new one, stabilizing the number of editable monomers (iii). Insertional edits are introduced 5'→3', so their order is preserved (iv). And specific insertions can be rendered biologically conditional, extending what can be recorded beyond lineage (v) (36). To date, however, we have demonstrated DNA Typewriter only in vitro, in cultured cells and monoclonal gastruloids, to build phylogenies of up to ~50,000 cells (33, 37).

Here, contending with the remaining obstacles and several more that arose along the way, we apply DNA Typewriter in vivo to reconstruct a 1.28 million-cell phylogeny of a single mouse embryo, from zygote to late organogenesis.

Results

Pronuclear injection of DNA Typewriter yields an extensively edited E13.5 embryo

DNA Typewriter comprises a prime editor, guide RNAs, and tape [here, PEmax, engineered prime editing guide RNA (epegRNAs), TAPE-1] (38). TAPE-1 is composed of six tandem repeats of a 14-base pair (bp) monomer preceded by a 3-bp key (here, GGA). Initially, the 5'-most portion of the tape comprises an active 20-bp protospacer ("typehead"). "Key-strokes" insert an information-bearing "symbol" (here, NNNGGA) that disrupts the typehead while also activating a new, downstream typehead (fig. S1A). Exactly one monomer per tape is editable until the array is exhausted.

We introduced DNA Typewriter by pronuclear injection (PNI) of two constructs and piggyBac transposase-encoding mRNA into zygotes. One construct encoded constitutively expressed PEmax, and the other an NNNGGA-encoding epegRNA, 12-bp barcode (tape-BC) and 6-monomer tape (Fig. 1A; fig. S1A). Differentiated tissues preclude unbiased cellular dissociation of post-gastrulation mouse embryos, which we

circumvent by profiling nuclei rather than cells (39). But mRNA is scarcer in nuclei, and targeted transcript recovery in single-nucleus RNA sequencing (RNA-seq) workflows has been a persistent challenge (37, 40). Because circular RNA reporters are recovered better than their linear equivalents (41), we reengineered TAPE-1 for expression as a Pol-III-driven circular transcript ("cirtape"; fig. S1A). Following PNI, we isolated nuclei from ten embryos harvested at E13.5 (fig. S1B). Extraembryonic tissues were discarded. By droplet digital polymerase chain reaction (ddPCR) and bulk amplicon sequencing of tape-BCs, three bore epegRNA::tape-BC::cirtape integrations, and DNA Typewriter activity at their tapes tracked PEmax copy number (Fig. 1, B and C, and fig. S1, C and D).

We focused on embryo #3 (Fig. 1D). With 11 epegRNA::tape-BC::cirtape integrations, we expected a "vocabulary" of at most 11 symbols; we instead observed 13 at $\geq 0.5\%$ frequency, including 5 abundant at monomers 1-2 and 8 at monomers 2-6 (Fig. 1E and fig. S2A). Amplicon sequencing linking tape-BCs to their *cis*-encoded symbols recovered 10 of 11, some redundant and collectively encoding exactly the 8 symbols seen at monomers 2-6. Symbols were distributed independently of whether they were encoded *in cis* (fig. S2B). A parsimonious explanation of these patterns is that the 5 "extra" symbols abundant only at monomers 1-2 were written by transiently expressed epegRNAs shortly after PNI.

With 8 symbols at 6 monomers, 8^6 ($\sim 10^5$) "tape genotypes" (an ordered set of symbols at a given tape) are possible. To assess how much of this diversity was realized, we performed rarefaction on tape genotypes (fig. S2C) (42). Each integration was represented by an estimated 180-370 genome equivalents distributed across 76-165 distinct tape genotypes, and every accumulation curve remained steeply rising. Recorded diversity thus appears limited by input material rather than by recorder capacity.

We reconstructed neighbor-joining trees from tape genotypes with a DNA Typewriter-specific distance metric (Fig. 1F and fig. S3) (43, 44). For 9 integrations, $>95\%$ of tape genotypes bore only one of two symbols at monomer-1, e.g., at integration "TCCGGAAGACCC", all 139 tape genotypes bore CAC (55%) or AAG (45%). A parsimonious explanation is that a burst of prime editing occurred at the 2-cell stage.

To ask whether these patterns are reproducible, we generated four additional DNA Typewriter embryos and characterized their tapes in bulk (supplementary text, note S1). Each was comparable or superior to embryo #3 in integration number, editing rate and recorded diversity (fig. S4). Across 85 integrations in 5 well-edited embryos, edits inferred to arise at the 1-cell and 2-cell stages were common, while 4-cell-stage edits were uncommon, with the caveat that there are limits on how readily the latter can be discerned from bulk tape genotypes. Notably, only embryo #3 lacked 1-cell-stage edits.

We conclude that PNI-delivered DNA Typewriter records heritable, ordered, diverse marks in vivo. Because each tape is edited independently, their combinatorics expands the upper bound on potential configurations from 8^6 per integration to $8^{(6 \times 11)}$ ($\sim 10^{59}$) per cell. However, realizing this potential requires determining which edited tapes co-occur within which cells.

Single-nucleus RNA-seq of a DNA Typewriter mouse embryo

We subjected ~ 10 M nuclei isolated from embryo #3 to single-nucleus combinatorial indexing (sci-RNA-seq3) (39), modified to co-amplify transcriptome and circRNA libraries (Fig. 1A). Following quality control, we obtained transcriptional profiles for 1,584,848 nuclei. To annotate cell types, we co-embedded these profiles with 1.65M E12.75-E14.25 cells from our 11.4M-cell atlas of mouse development (5), and transferred labels by k-nearest neighbors; cell type and major trajectory labels were transferred separately (Fig. 2A). Embryo #3's cell type proportions were maximally concordant with E13.5, the intended harvest stage (Fig. 2B). Stress, apoptosis, interferon and DNA-damage modules were unelevated in embryo #3 across 26 abundant cell types (fig. S5, A and B). However, of 91 genes in the vicinity of a mapped integration site (table S1), 8 were differentially expressed, most prominently the chromatin remodeler *Chd3*, which exhibited reduced expression across 20 cell types (fig. S5C).

Summed across all integrations, we obtained a median of 51 quality-passing circRNA UMIs (unique molecular identifiers) per cell. Tape genotypes were called for 54% of cell $\times$ tape combinations among genotyped cells (49% to 70% for nine integrations; 22% and 4% for the remaining two). CircRNA and transcriptome recovery were correlated, with limited evidence for cell-type-specific silencing of individual tapes (fig. S6). 95% of tape genotypes bore only expected prime editing events; the remaining 5% included nine classes of unexpected events (supplementary text, note S2).

Recording capacity was heavily utilized. Among the 4,371 cells genotyped at all 11 tapes, 55 ± 5 monomers had been written (83%), unimodally distributed and peaking at 56, with only 23 cells (0.5%) bearing 66 recovered edits (Fig. 2C). Repeating the estimate across the entire matrix by summing the per-tape means, we obtain a similar estimate of 53/66 monomers. A caveat is that although TAPE-1 was designed with six monomers, array length by E13.5 exhibits substantial variation consistent with in vivo slippage within the tandem array (5.2 ± 1.2 monomers; supplementary text, note S2), such that our estimate of utilization is a lower bound.

We define a cell's "lineage genotype" as its constellation of tape genotype calls. Among 353,681 cells genotyped at the five best recovered circRNAs, only 15% shared an identical lineage genotype with another cell, and the accumulation curve

was again linear (Fig. 2D). Resolution is thus limited by the number of cells sampled rather than by recording activity. Reassuringly, of 11,460 lineage genotypes shared by exactly two cells, 57% were of the same cell type, 11-fold above chance, an early indication that lineage genotypes capture biologically meaningful structure.

Recording dynamics shape dating and branch support of a 655,701-cell backbone phylogeny

The burst of prime editing activity at the two-cell stage allows us to resolve their descendants directly. From 401,367 cells with ≥ 8 genotyped tapes, we derived the edit state of each blastomere, hereafter "A" and "B" (A: 23 edits, B: 19 edits; table S2) and routed cells to their blastomere-of-origin. Of 1,584,848 transcriptionally profiled nuclei, 25,429 yielded no genotyped tape. Of 1,559,419 remaining cells, 889,849 (57.1%) routed to A and 653,335 (41.9%) to B, 1,543,184 (99.0%) in total; the remaining 16,235 (1.0%) lacked an unambiguous founder assignment and were set aside.

The cell-type compositions of A and B were highly correlated (Spearman's $\rho = 0.99$; Fig. 2E), suggesting contributions to the embryo proper, although unequal, are fate-neutral. However, our data do not speak to the relative or compositional contributions of A and B to extraembryonic lineages, not sampled here (45, 46).

We reconstructed phylogenies for A and B independently, and hereafter treat them as inline replicates. Because $>10^6$ taxa exceed the practical limits of direct phylogenetic reconstruction, we proceeded in two stages: (i) tree-building and dating of backbone trees from high-tape-recovery cells; then (ii) distance-based placement of moderate-tape-recovery cells onto the backbones.

Backbone trees were built by scalable neighbor-joining (47–49) on distance matrices for 410,925 (A) and 244,776 (B) cells with ≥ 7 genotyped tapes (43), rooted on the founders' states. We dated internal nodes by least-squares dating under a strict molecular clock (50, 51), fixing root (E1.5) and tips (E13.5), imposing a 2-hour minimum branch length, and capping the number of ancestral lineages for each blastomere's tree at half of the embryo's estimated cell number, where available (table S3). Joining the dated trees at a common zygote root yielded a backbone phylogeny of 655,701 tips (Fig. 2, F to H).

The floor and ceiling are principled, the former set by the fastest cell cycles of the gastrulating embryo (52) and the latter by the fact that a subsample can contain no more lineages at any given time than the embryo had cells. In practice, the ceiling binds only near gastrulation, where ceiling-free dating places 97,566 lineages at E6.5 against an epiblast of ~ 660 cells (fig. S7A), yet costs only $\sim 1.8\%$ in the least-squares objective, shifting early nodes ~ 4 days later, then decaying monotonically toward the tips (fig. S7B). Although we do not know

when the 1.36-fold excess of A vs. B cells arising by E13.5 accrues, node dates are robust to the ratio at which the ceiling is applied (fig. S7C). However, that such a correction is needed at all localizes the clock's violation to the interval between the 2-cell stage and gastrulation, where the ceiling substitutes external knowledge of embryo size for a temporal signal we could not recover for this interval, due to limitations of the recorder, the reconstruction, or some combination thereof.

In theory, DNA Typewriter's sequential design should hold the number of active recording monomers, and thus the rate at which independent marks accumulate, constant over time. To ask whether it did, we compared the observed accumulation against three idealized architectures, each tuned to reach the observed mean of 55/66 monomers written by E13.5 (Fig. 2I). Observed dynamics matched none of them, instead showing three phases (Fig. 2J): a ~20 edit burst at the 2-cell stage, marking each founding blastomere; a pronounced slowdown through subsequent cleavages and blastocyst formation; and recovery around ~E6, peaking at ~9 edits/cell/day then declining as tapes exhaust. A minimal three-phase model reproduces these dynamics (Fig. 2I). Furthermore, although we lack single-cell resolution for the four additional embryos, their bulk editing patterns are consistent with similar dynamics (supplementary text, note S1). A caveat is that, because our parsimony approach assigns an edit to the ancestral state of a node only when all of its descendants share that edit, noise in the genotype calls can bias the reconstruction toward assigning too few edits, with earlier nodes, which have far more descendants, more likely to be affected; therefore, the magnitude of this slowdown should be read as an upper bound.

How reliable is the topology of this backbone phylogeny? Node support measures are a cornerstone of phylogenetics (53), but their use in the lineage recording field is inconsistent. Resampling-based bootstraps (53), a conventional measure of node support, were computationally infeasible at this scale. We therefore devised a simple independent-edit-count support statistic for each split in the inferred phylogeny (materials and methods). Its distribution was right-skewed (median 3, tail beyond 20; fig. S7D), with 78% (A) and 80% (B) of internal branches carrying ≥ 1 , and 64% and 68% carrying ≥ 2 , supporting edits. As further checks, we recounted support using only edits on distinct tapes (fig. S8A) and confirmed that subtrees reconstructed with one tape withheld recover the held-out tape's pairwise distances well above a label-shuffled control (fig. S8B).

Recording dynamics and topological confidence are necessarily related, both because an interval in which few marks are written is one in which few divisions can be resolved, and because both are estimated from the same ancestral-state re-

construction. Correspondingly, node support varied with developmental time, mirroring editing dynamics (Fig. 2K and fig. S8A): maximal at the first division, weakest through the more sparsely recorded window that follows, then rising as editing recovers, before falling again slightly toward the tips as tape exhaustion reduces the marks available to resolve recent splits. Notably, the low-confidence early window is precisely where unconstrained dating inflates ancestral lineages beyond the embryo's cell number, and where the imposition of a ceiling does the most work (fig. S7, A and B). Support recovers by E6.5-E7.5, over which the dated tree traces an expansion from 656 to 13,181 lineages, and is maintained thereafter.

A third clock violation, addressed by neither the ceiling nor the timing correction, is that prime editing rate varies across cell types (54–57). In this embryo, cell type-stratified editing rates fall within a tight 1.6-fold range, but the slowest types are overwhelmingly hematopoietic (supplementary text, note S3; fig. S9). Our dates should therefore be read as approximations rather than exact ages, least reliable at the earliest divisions, underestimated toward the tips, and least reliable again in hematopoietic cell types.

Distance-based placement doubles the tree to 1,281,141 cells

In the second reconstruction stage, we placed 625,440 additional cells with 4 to 6 genotyped tapes onto the 655,701-cell backbone; the remaining 262,043 blastomere-routed cells were excluded, 88% for carrying < 4 genotyped tapes and the rest by the doublet and founder-contradiction filters. For placements, each query cell was attached to its single closest relative under the same distance metric used for the backbone (43), whether a backbone tip or a previously placed query; cells were processed in order of decreasing tape recovery (fig. S10A). This strategy rests on two assumptions: that the backbone samples the embryo densely enough for each query cell to have a near relative among its tips, and that 4 tapes suffice to identify that relative. Neither is guaranteed, so we assessed our placements in two orthogonal ways.

First, the edit distances themselves. Query cells share post-burst, sequentially accrued edits with their anchor and rarely contradict it (fig. S10, B and C), indicating that anchors are genuine close relatives rather than arbitrary attachment points. But most queries are genotype-identical to their anchor at every recovered monomer, so placements add little topological information and imply short divergence times (fig. S10, D and E).

Second, because placement uses edit distances alone, with no reference to transcriptional state, cell type concordance among neighboring tips serves as independent evidence that genuine lineage relationships are recovered. Because insert-

ing placements reshapes the terminal topology, adding polytomies where queries share identical genotypes (fig. S10, F and G), we compared cells on the fraction of their equally related nearest neighbors sharing their cell type. Backbone- and placement-derived cells were similarly concordant, both far above a label-permuted null, with placed cells somewhat lower in every terminal context (fig. S10H).

Placed cells therefore sit at short edit distances from their anchors and recover cell-type concordance nearly as well as backbone cells. Furthermore, they are compositionally similar to backbone cells (fig. S11). On the other hand, their accuracy is sensitive to both retained tape number and edit count (fig. S12). Placements should thus be cautiously interpreted as being in the correct local neighborhood, but not necessarily at the correct depth, a limitation that follows from the algorithm attaching queries along terminal branches rather than internal edges in order to remain computationally tractable.

The final tree of embryo #3 is ultrametric, with 1,281,141 tips (E13.5), each carrying a single-cell transcriptome and one of 138 cell type annotations (Fig. 3A). Although these cells represent only ~10% of the E13.5 embryo (4, 5, 58), coverage at depth is far higher, *e.g.* each of the 13,181 lineages dated to E7.5 is represented by a mean of 97 sampled descendants. Inspired by NextStrain (59), we built NextCell (60), an interactive explorer with which one can click any ancestral node to visualize its subclade. For example, Fig. 3B shows n1703121, an ancestral node dated to E7.4 whose descendants are endodermally fated, with subclades clonally partitioned among visceral organs.

A wave of clonal dominance arises during gastrulation

Clonal dominance, the progeny of some progenitors accruing disproportionately relative to others, is recurrent in vertebrate development, mostly studied within individual tissues (61–65). At the whole-organism scale, a few early GESTALT-marked progenitors accounted for most non-blood cells in each organ of an adult zebrafish (21), while LoxCode barcoding of the E5.5 mouse epiblast revealed clones differing widely in size and fate (66). In both cases, however, marks were established at one point and read out at another, such that dominance could be quantified but not dated.

To investigate when dominance arises, we tracked the Gini coefficient of clonal contribution across developmental time, taking lineages present at one timepoint as founders and their descendants at a later timepoint as clones. We defined founders at E7.0, by which point node support has recovered and the ceiling is no longer binding (Fig. 2K). This yields 3,849 clones (A: 1,970; B: 1,879), each contributing a mean of 333 profiled E13.5 descendants, but with marked inequality, *e.g.* the 170 largest account for 50% of tips, versus 720 expected under a neutral Yule (pure-birth) null (67),

while the observed Gini coefficient of 0.80 grossly exceeds the null's 0.50 (Fig. 4, A to C). These patterns are similar in the backbone tree (fig. S13, A to C), similar between A and B (fig. S13, D to F), and robust to the ratio at which the ceiling is applied to A vs. B during dating (fig. S14A).

If we vary the timepoint at which clonal contributions of E7.0 founders are assessed, we find that the Gini coefficient rises steeply and plateaus by ~E9.0 (Fig. 4D and fig. S14B). Because the Yule null rises similarly but asymptotes at 0.50, the informative quantity is the departure from it, which exceeds the null's 95% envelope almost immediately after founders are defined, and remains above it thereafter. Similar dynamics are observed if we define founders at timepoints ranging from E5.5 to E7.5, but not E10.0, from which inequality instead accrues gradually (fig. S14, C and D).

What might drive it? Inequality exceeding the simplistic, pure-birth Yule null requires net expansion rates that differ between lineages and persist within them, whether through division or survival. Cell cycles around gastrulation are among the shortest of the mammalian life cycle and heterogeneous, fastest in the epiblast's proliferative zone (52, 68). Over this same interval, apoptosis is elevated in the pregastrula and spatially patterned in the gastrula (69), where a substantial fraction of epiblast cells are eliminated by cell competition, which acts on relative fitness and, where the underlying differences are heritable, can drive non-neutral clonal dynamics (62, 65). Here, the 10% most proliferative E7.0-defined founders account for 39% of E8.0-dated lineages, broadly consistent with Snow's 1977 speculation that the proliferative zone's contribution could rise from 10% to 50% by E7.5 (52). Our tree records only surviving lineages, so it cannot apportion the observed inequality between differential division and differential survival, which are not mutually exclusive. Regardless, the inequality at E13.5 is much smaller than such a skew would produce if fully inherited, suggesting that the plateau marks not the end of heterogeneity but the end of its heritability.

In contrast with A and B, whose contributions to the embryo proper are fate-neutral, many E7.0-defined founders exhibit highly skewed output (Fig. 4E and fig. S14E), *e.g.*, 20 founders account for only 5% of the E13.5 embryo but 51% of its 4,178 olfactory sensory neurons. Resolved by trajectory rather than by founder, nearly all of the well-represented major trajectories exhibit elevated Gini coefficients, in a manner that is reproducible between blastomeres and between the full and backbone trees (Fig. 4F and fig. S14F). However, founders already occupy distinct positions in the gastrulating embryo by E7.0, so both trajectory-specific dynamics and founder skews plausibly reflect spatial position rather than fate bias, and we are wary of interpreting either further. This also complicates comparison with prospective clonal marking at E5.5 (66), both because we define founders 1.5 days

later (avoiding the low-confidence pre-gastrulation interval) and because gastrulation-associated clonal dominance and the later-arising fate-specific enrichments are distinct quantities that are likely decoupled.

Tree siblings share fate far in excess of chance and mark terminal differentiations

We next turned to the tips of the tree, beginning with tree siblings (cherries and terminal polytomies; fig. S10F), to ask how strongly lineage predicts fate at the finest scale available. For annotations observed in both subtrees, we asked how often tree siblings share it, relative to a tip-label-permuted null. Because our tree captures at best ~10% of the estimated 13-61M cells of an E13.5 embryo (4, 5, 58), most tree siblings are not true sisters but nearest sampled relatives, so any concordance we observe is a lower bound on true sibling concordance. Globally, 51% of tree-sibling pairs share a cell type label, a 9.4-fold enrichment over a permuted null (5.9-fold for strict two-tip cherries). Enrichment varied widely across fates but was reproducible between blastomeres (Spearman's $\rho = 0.94$; fig. S15, A and B, and tables S4 and S5). Extreme concordances corresponded to cell types understood to commit from a small, spatially restricted progenitor pool and then expand clonally within an anatomical niche, *e.g.* lung progenitor cells at 145-fold (fig. S15B).

Half of tree siblings were cell-type discordant. To ask whether heterotypic pairs might inform terminal differentiations, we extracted the most closely related subset (shared ancestor dated $\geq$ E12.5; at most one discordant tape monomer) wherein exactly one member was among 27 post-mitotic cell types (fig. S15C and table S6). This yielded 82,300 post-mitotic cells distributed across 640 progenitor-derivative pairings. Against a within-blastomere permutation null, 37 pairings were $\geq$ 3-fold enriched at 1% false discovery rate (FDR) (max 29-fold), with enrichments reproducible between subtrees (Spearman's $\rho = 0.82$; fig. S15D). Together these draw on 24 distinct progenitor states and assign at least one progenitor to 18/27 post-mitotic types, and recover canonical differentiation events at the resolution of individual dated divisions (Fig. 5A) spanning neurogenesis, terminal maturation, and hematopoiesis; the latter warrant caution, as the blood-specific slowdown may leave our edit-distance threshold less able to isolate recent divisions.

A hierarchy of fate couplings dated from gastrulation to organogenesis

Tree siblings probe only the immediate neighborhood of each tip. To reach older relationships, we partitioned each blastomere subtree into "maximal clades" (70), *i.e.* subtrees of ≥ 3 but no more than K cells whose parent exceeds K and asked which of 88 abundant cell types co-occur within them more often than expected under a tip-label-permuted null (fig. S16)

(27, 71). We ran this at $K=15$ (median ancestor age E11.2) and $K=200$ (E8.8).

Hierarchical clustering of the $K=15$ enrichment matrix revealed clear blocks corresponding to hematopoiesis, anterior and posterior neuroectoderm, neural crest, mesoderm/mesenchyme/endothelium, endoderm, and surface ectoderm epithelia (Fig. 5B). Because the tree was built without reference to cell identity, recovering this organization from clade co-occurrence alone further indicates that its topology reflects genuine lineage relationships, as does its reproducibility between the A and B sub-trees (off-diagonal Spearman's $\rho = 0.79$ at $K=15$; fig. S17A).

Block structure was much sharper at $K=15$ (recent clades; fig. S17A) than $K=200$ (ancient clades; fig. S17B). Of 3,828 cell-type pairs, 143 were significantly enriched in both blastomeres at $K=15$, but only 22 at $K=200$, and 18 at both (fig. S17C). Whereas pairs enriched only at $K=15$ overwhelmingly reflect post-gastrulation relationships, those enriched at both remain coupled within clades whose common ancestor predates E9, implying couplings dating to gastrulation or earlier. That some couplings persist across a 13-fold change in the clade-size cap while most do not, suggested to us that the depth at which an enrichment disappears is itself informative, and that sweeping developmental time directly might allow the coupling of cell-type pairs to be dated systematically.

To this end, for each time t from E7.0 to E13.25 in 0.25-day steps, we defined the clade of every lineage present at t as all of its sampled descendants, and computed the same pairwise co-occurrence enrichment as above against a tip-label permutation null. For each qualifying pair we recorded its coupling depth, *i.e.* the earliest t at which the two remained significantly co-enriched in both blastomere subtrees. This yielded coupling depths for 129 cell-type pairs, touching 75/88 cell types with dates spanning E8.25-E13.25 (Fig. 5C and table S7). ; Significance formed a single contiguous interval in time for 122 of these 129 pairs (95%). Average-linkage clustering of these depths yielded a hierarchy in which most branch points carry a developmental date; clusters generally corresponded to major germ layers or organs, and the most deeply coupled groups "closed" first (Fig. 5D). The order and dating of the cell type couplings were biologically sensible and reproducible between blastomeres (fig. S17D).

Because our placement heuristic could in principle manufacture the terminal structure these analyses rest on, and because global tip-label permutation is a permissive null, we repeated the sibling, heterotypic-pairing and clade co-occurrence analyses on the backbone tree alone, and against a within-clone null (supplementary text, note S4; tables S5 to S7). Enrichments fell in magnitude but persisted, and the rank order of coupling depths was preserved throughout, while their absolute dates shifted by a median of 0.75 days,

deeper on the backbone and shallower under the within-clone null. We conclude that sibling enrichments and the ordering of cell type couplings are robust to both tree and null, but also that absolute coupling dates should not be overinterpreted.

Ancestral state imputation recovers known developmental transitions

Each of the tree's 1,281,141 tips carries a transcriptional state, but its 1,109,894 internal nodes are unobserved. We sought to develop an exploratory, heuristic imputation strategy that leverages our single-cell RNA-seq (scRNA-seq) atlas of mouse development (5). Within individual cell types, tree siblings are on average closer in transcriptional state than non-siblings of the same type, and transcriptional similarity tracks lineal relatedness (Fig. 6, A and B). However, these trends are subtle, such that we favored approaches that aggregate information across many descendants of each internal node.

Traversing from tips to root, we assigned each internal node the mean transcriptional profile of its children, then linked ancestral nodes to atlas cells by mutual nearest neighbors in successive 6-hour windows from E13.5 back to E8.5, assigning each matched node an updated profile and a cell type by majority vote (Fig. 6, C and D). Critically, the lineage relationships here are given by the tree, with the atlas serving only to identify the contemporaneous state most consistent with each ancestral node's descendants.

This heuristic yielded cell type labels for 926,852 of the 1,054,879 internal nodes within the atlas interval (88%), with coverage lowest early (46% at E8.75) and best late (96-98% between E11.5 and E13.0). Imputed labels are robust to random masking of 20% of tips, albeit less so for rare annotations (fig. S18), and are almost entirely lost when cell identities are permuted across tips (0.06% label agreement). But they are also noisy, implying 132,536 distinct "state paths" that are overwhelmingly rare and biologically implausible.

We therefore defined a state path as "qualifying" if it was heterotypic, non-recursive, carried by at least 5 instances and $\geq 1\%$ of its terminal cell type's traced cells, and composed of constituent steps that were individually well-supported (≥ 50 overall transitions) and asymmetric ($>60/40$ preference for observed direction) in the overall tree. This yielded 275 qualifying state paths, 98% of which are recovered independently in both blastomeres, covering 82 of 136 cell types and 117,704 tips (table S8). These criteria use only the phylogeny and imputed labels, and no external knowledge of developmental biology.

Although qualifying state paths are short (193 with two states, 77 with three, 5 with four), imputed rather than recorded, and lack validation, they are clearly enriched for well-established differentiations (*e.g.* lateral plate and intermedi-

ate mesoderm $\rightarrow$ posterior intermediate mesoderm $\rightarrow$ metanephric mesenchyme $\rightarrow$ nephron progenitors). To evaluate this systematically, and with the caveat that this is not fully independent validation, we scored imputed paths against our previously reported, manually curated cell-type transition graph (5). Of 360 testable steps, 160 (44%) connect states that are directly adjacent in the 474-edge transition graph and 207 (58%) connect states within two edges of one, 15.7-fold and 7.4-fold enrichment over chance. Imputed annotations and qualifying paths can also be explored in NextCell (60), including as temporal flow-allocation diagrams (9) (Fig. 6, E to G).

Discussion

Molecular technologies, like algorithms, have a theoretical scaling that is clean and a practical scaling that is not. At the outset of this field around 2016, we and others imagined that the path from proof-of-concept to a "Sulston mouse" would be relatively straightforward. But the devil, of course, is in the details.

Of the five obstacles described in the introduction that are not addressed by DNA Typewriter itself, we make progress on all but fully resolve none. Delivery by PNI and piggyBac established a stable, constitutively expressed system through E13.5, but was mosaic and uncontrolled in integration number and location (i); PNI offers a faster design-test loop than cell culture or organoids, directly in vivo, but does not result in a distributable "recorder mouse" line (ix). sci-RNA-seq3 recovered $>10^6$ cells, $\sim 10\%$ of this embryo but orders of magnitude short of what is required for dense sampling of an adult mouse (vi). CircTape exhibited subjectively improved recovery relative to our prior attempts at targeted transcript recovery from single nuclei (37, 40), yet only 54% of the cell $\times$ tape matrix was filled, possibly because circular RNAs are exported to the cytoplasm (72) (vii). Finally, although we succeeded in building a time-calibrated, parsimony-supported phylogeny of 1.28 million cells, its dating and topology are limited by clock violations and the placement heuristic (viii).

Four further obstacles, not on our list, loomed large in practice. First, cell type annotation of this phylogeny's tips was tractable only because of a parallel, multi-year effort to generate a highly resolved scRNA-seq atlas of mouse development (4–6). Second, a barren phylogeny does not directly inform the identities of ancestral nodes; our co-occurrence and state-imputation heuristics are first approximations, and the toolkit for "developmental statistics" remains sparse (73). Third, with a single embryo, reproducibility rests on comparisons of blastomere A vs. B subtrees, a useful but limited form of replication, *e.g.* we cannot assess how integration-mediated disruption of genes such as *Chd3* (74) impacted the phylogeny. As such, our results here and any follow-on analyses of the phylogeny should be regarded as preliminary, pending confirmation in independent animals.

The fourth new obstacle was the frustrating-yet-fascinating dynamics of PNI-delivered DNA Typewriter. A massive editing burst, helpful for rooting, spanned the 1-cell (minor ZGA) and 2-cell (major ZGA) stages (supplementary text, note S1), which may be permissive for genome editing (75, 76). Less helpfully, the editing rate then sharply fell, recovering by gastrulation, albeit not to the level of the 1-cell/2-cell burst. We speculate that post-ZGA epigenetic silencing of integrated cassettes and/or a lower intrinsic editing rate in epiblast precursors, which dominate the embryo proper at this stage, reduce prime editing during this window. Single-cell profiling of extraembryonic tissues from DNA Typewriter embryos may shed further light.

From its perigastrulation recovery through E13.5, DNA Typewriter recorded robustly, although rates varied modestly between cell types. The marked blood-specific slowdown could be experiment-specific (*e.g.* silencing of specific PEmax integration sites), setup-specific (*e.g.* silencing of the CAG promoter driving PEmax) or biological (*e.g.* dNTP limitation in hematopoietic cells); further experiments are required to distinguish these. Regardless, cell type-specific rates violate the strict molecular clock. Looking forward, we favor a transition to maximum likelihood methods that more explicitly accommodate molecular biology, missing data, and complexity in both phylodynamics and recorder dynamics, as well as divide-and-conquer algorithms that may scale to 10^8 cells and beyond (43, 77–80).

Other avenues for improvement are arithmetic and scale. Recording capacity is the product of tape number and length. Many of the 11×6 -monomer tapes were saturated by E13.5, eroding the rate stability that DNA Typewriter's sequential recording affords. By contrast, 50×20 -monomer tapes would afford 1,000 positions, enough to stably record to adulthood. Sampling DNA Typewriter embryos at more stages, in conjunction with chemical, genetic or environmental perturbations, and/or with integration of signal-, state- and spatially-resolved recording, would each add dimensions that our single tree lacks (36, 81, 82).

Large, dense cellular phylogenies of regulatively developing animals are a genuinely new kind of data object, enabling investigation of the statistical regularities that recur within an individual, or hold across individuals, even when no single lineage does. As we have argued elsewhere, such phylogenies are poised to anchor the “consensus ontogeny” that we envision for organisms such as the mouse, a lineage-anchored reference onto which multimodal measurements can be layered (83). The 1.28 million tree reported here, one embryo at one stage comprising ~10% of its cells, is at best a single viewpoint on such an ontogeny. Yet it recovers, from clonal co-occurrence alone and without reference to transcriptional state, a dated hierarchy of cell-type couplings that agrees with canonical relationships of mouse development accrued over

decades of investigation. That so much can be read from a single, incomplete lineage is itself encouraging. We have only scratched the surface of what might be learned from this one dataset, and we welcome its further exploration, and its integration with other cellular phylogenies of the developing mouse (84).

Following years of work across many groups (19), we hope that this and contemporaneous studies (84) mark an inflection point for the molecular recording field. As more sophisticated in vivo recording devices are brought to bear, the phenotypic equivalent of a mammalian reference genome, from zygote to birth to adulthood to death, comes into view (83).

The materials and methods are available in the supplementary materials.

REFERENCES AND NOTES

1. J. E. Sulston, H. R. Horvitz, Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* **56**, 110–156 (1977). [doi:10.1016/0012-1606\(77\)90158-0](https://doi.org/10.1016/0012-1606(77)90158-0) [Medline](#)
2. J. E. Sulston, E. Schierenberg, J. G. White, J. N. Thomson, The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983). [doi:10.1016/0012-1606\(83\)90201-4](https://doi.org/10.1016/0012-1606(83)90201-4) [Medline](#)
3. I. A. Hatton, E. D. Galbraith, N. S. C. Merleau, T. P. Miettinen, B. M. Smith, J. A. Shander, The human cell count and size distribution. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2303077120 (2023). [doi:10.1073/pnas.2303077120](https://doi.org/10.1073/pnas.2303077120) [Medline](#)
4. J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F. J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019). [doi:10.1038/s41586-019-0969-x](https://doi.org/10.1038/s41586-019-0969-x) [Medline](#)
5. C. Qiu, B. K. Martin, I. C. Welsh, R. M. Daza, T.-M. Le, X. Huang, E. K. Nichols, M. L. Taylor, O. Fulton, D. R. O'Day, A. R. Gomes, S. Ilcisin, S. Srivatsan, X. Deng, C. M. Disteche, W. S. Noble, N. Hamazaki, C. B. Moens, D. Kimelman, J. Cao, A. F. Schier, M. Spielmann, S. A. Murray, C. Trapnell, J. Shendure, A single-cell time-lapse of mouse prenatal development from gastrula to birth. *Nature* **626**, 1084–1093 (2024). [doi:10.1038/s41586-024-07069-w](https://doi.org/10.1038/s41586-024-07069-w) [Medline](#)
6. C. Qiu, J. Cao, B. K. Martin, T. Li, I. C. Welsh, S. Srivatsan, X. Huang, D. Calderon, W. S. Noble, C. M. Disteche, S. A. Murray, M. Spielmann, C. B. Moens, C. Trapnell, J. Shendure, Systematic reconstruction of cellular trajectories across mouse embryogenesis. *Nat. Genet.* **54**, 328–341 (2022). [doi:10.1038/s41588-022-01018-x](https://doi.org/10.1038/s41588-022-01018-x) [Medline](#)
7. B. Pijuan-Sala, J. A. Griffiths, C. Guibentif, T. W. Hiscock, W. Jawaid, F. J. Calero-Nieto, C. Mulas, X. Ibarra-Soria, R. C. V. Tyser, D. L. L. Ho, W. Reik, S. Srinivas, B. D. Simons, J. Nichols, J. C. Marioni, B. Göttgens, A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**, 490–495 (2019). [doi:10.1038/s41586-019-0933-9](https://doi.org/10.1038/s41586-019-0933-9) [Medline](#)
8. C. Qiu, R. M. Daza, I. C. Welsh, R. P. Patwardhan, B. K. Martin, T. Li, S. Yang, C. C. A. Mannens, S. De Winter, N. Kempynck, M. L. Taylor, O. Fulton, T.-M. Le, D. R. O'Day, J.-B. Lallanne, S. Domcke, S. A. Murray, S. Aerts, C. Trapnell, J. Shendure, Evolutionary transfer learning enables organism-wide inference of mammalian enhancer landscapes. *bioRxiv* (2026)p. 2026.04.07.717039.
9. M. Mittnenzweig, Y. Mayshar, S. Cheng, R. Ben-Yair, R. Hadas, Y. Rais, E. Chomsky, N. Reines, A. Uzonyi, L. Lumerman, A. Lifshitz, Z. Mukamel, A.-H. Orenbuch, A. Tanay, Y. Stelzer, A single-embryo, single-cell time-resolved model for mouse gastrulation. *Cell* **184**, 2825–2842.e22 (2021). [doi:10.1016/j.cell.2021.04.004](https://doi.org/10.1016/j.cell.2021.04.004) [Medline](#)
10. D. E. Wagner, A. M. Klein, Lineage tracing meets single-cell omics: Opportunities and challenges. *Nat. Rev. Genet.* **21**, 410–427 (2020). [doi:10.1038/s41576-020-0223-2](https://doi.org/10.1038/s41576-020-0223-2) [Medline](#)
11. K. K. Chow, M. W. Budde, A. A. Granados, M. Cabrera, S. Yoon, S. Cho, T.-H. Huang, N. Koulana, K. L. Frieda, L. Cai, C. Lois, M. B. Elowitz, Imaging cell lineage with a synthetic digital recording system. *Science* **372**, eabb3099 (2021). [doi:10.1126/science.abb3099](https://doi.org/10.1126/science.abb3099) [Medline](#)

12. F. Farzadfard, T. K. Lu, Synthetic biology. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. *Science* **346**, 1256272 (2014). [doi:10.1126/science.1256272](https://doi.org/10.1126/science.1256272) [Medline](#)
13. J. Livet, T. A. Weissman, H. Kang, R. W. Draft, J. Lu, R. A. Bennis, J. R. Sanes, J. W. Lichtman, Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56–62 (2007). [doi:10.1038/nature06293](https://doi.org/10.1038/nature06293) [Medline](#)
14. W. Pei, T. B. Feyerabend, J. Rössler, X. Wang, D. Postrach, K. Busch, I. Rode, K. Klapproth, N. Dietlein, C. Quedenau, W. Chen, S. Sauer, S. Wolf, T. Höfer, H.-R. Rodewald, Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature* **548**, 456–460 (2017). [doi:10.1038/nature23653](https://doi.org/10.1038/nature23653) [Medline](#)
15. C. Walsh, C. L. Cepko, Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. *Science* **255**, 434–440 (1992). [doi:10.1126/science.1734520](https://doi.org/10.1126/science.1734520) [Medline](#)
16. S. J. Salipante, M. S. Horwitz, Phylogenetic fate mapping. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5448–5453 (2006). [doi:10.1073/pnas.0601265103](https://doi.org/10.1073/pnas.0601265103) [Medline](#)
17. K.-M. Kim, D. Shibata, Tracing ancestry with methylation patterns: Most crypts appear distantly related in normal adult human colon. *BMC Gastroenterol.* **4**, 8 (2004). [doi:10.1186/1471-230X-4-8](https://doi.org/10.1186/1471-230X-4-8) [Medline](#)
18. D. Frumkin, A. Wasserstrom, S. Kaplan, U. Feige, E. Shapiro, Genomic variability within an organism exposes its cell lineage tree. *PLOS Comput. Biol.* **1**, e50 (2005). [doi:10.1371/journal.pcbi.0010050](https://doi.org/10.1371/journal.pcbi.0010050) [Medline](#)
19. A. Askary, W. Chen, J. Choi, L. Y. Du, M. B. Elowitz, J. A. Gagnon, A. F. Schier, S. Seidel, J. Shendure, T. Stadler, M. Tran, The lives of cells, recorded. *Nat. Rev. Genet.* **26**, 203–222 (2025). [doi:10.1038/s41576-024-00788-w](https://doi.org/10.1038/s41576-024-00788-w) [Medline](#)
20. T. Stadler, O. G. Pybus, M. P. H. Stumpf, Phylodynamics for cell biologists. *Science* **371**, eaah6266 (2021). [doi:10.1126/science.aah6266](https://doi.org/10.1126/science.aah6266) [Medline](#)
21. A. McKenna, G. M. Findlay, J. A. Gagnon, M. S. Horwitz, A. F. Schier, J. Shendure, Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* **353**, aaf7907 (2016). [doi:10.1126/science.aaf7907](https://doi.org/10.1126/science.aaf7907) [Medline](#)
22. K. L. Frieda, J. M. Linton, S. Hormoz, J. Choi, K. K. Chow, Z. S. Singer, M. W. Budde, M. B. Elowitz, L. Cai, Synthetic recording and in situ readout of lineage information in single cells. *Nature* **541**, 107–111 (2017). [doi:10.1038/nature20777](https://doi.org/10.1038/nature20777) [Medline](#)
23. R. Kalhor, K. Kalhor, L. Mejia, K. Leeper, A. Graveline, P. Mali, G. M. Church, Developmental barcoding of whole mouse via homing CRISPR. *Science* **361**, eaat9804 (2018). [doi:10.1126/science.aat9804](https://doi.org/10.1126/science.aat9804) [Medline](#)
24. M. M. Chan, Z. D. Smith, S. Grosswendt, H. Kretzmer, T. M. Norman, B. Adamson, M. Jost, J. J. Quinn, D. Yang, M. G. Jones, A. Khodaverdian, N. Yosef, A. Meissner, J. S. Weissman, Molecular recording of mammalian embryogenesis. *Nature* **570**, 77–82 (2019). [doi:10.1038/s41586-019-1184-5](https://doi.org/10.1038/s41586-019-1184-5) [Medline](#)
25. S. Bowling, D. Sritharan, F. G. Osorio, M. Nguyen, P. Cheung, A. Rodriguez-Fraticelli, S. Patel, W.-C. Yuan, Y. Fujiwara, B. E. Li, S. H. Orkin, S. Hormoz, F. D. Camargo, An Engineered CRISPR-Cas9 Mouse Line for Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. *Cell* **181**, 1693–1694 (2020). [doi:10.1016/j.cell.2020.06.018](https://doi.org/10.1016/j.cell.2020.06.018) [Medline](#)
26. B. Raj, D. E. Wagner, A. McKenna, S. Pandey, A. M. Klein, J. Shendure, J. A. Gagnon, A. F. Schier, Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nat. Biotechnol.* **36**, 442–450 (2018). [doi:10.1038/nbt.4103](https://doi.org/10.1038/nbt.4103) [Medline](#)
27. B. Spanjaard, B. Hu, N. Mitic, P. Olivares-Chauvet, S. Janjuha, N. Ninov, J. P. Junker, Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat. Biotechnol.* **36**, 469–473 (2018). [doi:10.1038/nbt.4124](https://doi.org/10.1038/nbt.4124) [Medline](#)
28. A. Alemany, M. Florescu, C. S. Baron, J. Peterson-Maduro, A. van Oudenaarden, Whole-organism clone tracing using single-cell sequencing. *Nature* **556**, 108–112 (2018). [doi:10.1038/nature25969](https://doi.org/10.1038/nature25969) [Medline](#)
29. K. Liu, S. Deng, C. Ye, Z. Yao, J. Wang, H. Gong, L. Liu, X. He, Mapping single-cell-resolution cell phylogeny reveals cell population dynamics during organ development. *Nat. Methods* **18**, 1506–1514 (2021). [doi:10.1038/s41592-021-01325-x](https://doi.org/10.1038/s41592-021-01325-x) [Medline](#)
30. D. M. Chadly, K. L. Frieda, C. Gui, L. Klock, M. Tran, M. Y. Sui, Y. Takei, R. Bouckaert, C. Lois, L. Cai, M. B. Elowitz, Reconstructing cell histories in space with image-readable base editor recording. *bioRxiv* 2024.01.03.573434 [Preprint] (2024); <https://doi.org/10.1101/2024.01.03.573434>.
31. D. Chadly, R. Hadas, L. Klock, J. Yue, F. Horns, A. Askary, A. Granados, R. Bouckaert, C. Lois, L. Cai, M. B. Elowitz, Regenerative base editing enables deep lineage recording. *bioRxiv* 2026.02.06.703857 [Preprint] (2026); <https://doi.org/10.64898/2026.02.06.703857>.
32. E. Winter, F. Emiliani, A. Cook, A. Abderrahim, A. McKenna, BASELINE: A CRISPR Base Editing Platform for Mammalian-Scale Single-Cell Lineage Tracing. *bioRxiv* 2025.03.19.644238 [Preprint] (2025); <https://doi.org/10.1101/2025.03.19.644238>.
33. J. Choi, W. Chen, A. Minkina, F. M. Chardon, C. C. Suiter, S. G. Regalado, S. Domcke, N. Hamazaki, C. Lee, B. Martin, R. M. Daza, J. Shendure, A time-resolved, multi-symbol molecular recorder via sequential genome editing. *Nature* **608**, 98–107 (2022). [doi:10.1038/s41586-022-04922-8](https://doi.org/10.1038/s41586-022-04922-8) [Medline](#)
34. T. B. Loveless, C. K. Carlson, C. A. Dentzel-Helmy, V. J. Hu, S. K. Ross, M. C. Demelo, A. Murtaza, G. Liang, M. Ficht, A. Singhai, M. J. Pajoh-Casco, C. C. Liu, Open-ended molecular recording of sequential cellular events into DNA. *Nat. Chem. Biol.* **21**, 512–521 (2025). [doi:10.1038/s41589-024-01764-5](https://doi.org/10.1038/s41589-024-01764-5) [Medline](#)
35. L. W. Koblan, K. E. Yost, P. Zheng, W. N. Colgan, M. G. Jones, D. Yang, A. Kumar, J. Sandhu, A. Schnell, D. Sun, C. Ergen, R. A. Saunders, X. Zhuang, W. E. Allen, N. Yosef, J. S. Weissman, High-resolution spatial mapping of cell state and lineage dynamics in vivo with PETracer. *Science* **390**, eadx3800 (2025). [doi:10.1126/science.adx3800](https://doi.org/10.1126/science.adx3800) [Medline](#)
36. W. Chen, J. Choi, X. Li, J. F. Nathans, B. Martin, W. Yang, N. Hamazaki, C. Qiu, J.-B. Lalanne, S. Regalado, H. Kim, V. Agarwal, E. Nichols, A. Leith, C. Lee, J. Shendure, Symbolic recording of signalling and cis-regulatory element activity to DNA. *Nature* **632**, 1073–1081 (2024). [doi:10.1038/s41586-024-07706-4](https://doi.org/10.1038/s41586-024-07706-4) [Medline](#)
37. S. G. Regalado, C. Qiu, S. Kottapalli, C. Huynh, R. M. Daza, A. Keith, J. Parks, B. K. Martin, W. Chen, H. Liao, H. Kim, X. Li, J.-B. Lalanne, N. Hamazaki, S. Domcke, J. Choi, J. Shendure, A fluctuation test for gastruloid heterogeneity. *Cell* **189**, 2026.09.057 (2026). [doi:10.1016/j.cell.2026.09.057](https://doi.org/10.1016/j.cell.2026.09.057)
38. H. Liao, J. Choi, J. Shendure, Molecular recording using DNA Typewriter. *Nat. Protoc.* **19**, 2833–2862 (2024). [doi:10.1038/s41596-024-01003-0](https://doi.org/10.1038/s41596-024-01003-0) [Medline](#)
39. B. K. Martin, C. Qiu, E. Nichols, M. Phung, R. Green-Gladden, S. Srivatsan, R. Blecher-Gonen, B. J. Beliveau, C. Trapnell, J. Cao, J. Shendure, Optimized single-nucleus transcriptional profiling by combinatorial indexing. *Nat. Protoc.* **18**, 188–207 (2023). [doi:10.1038/s41596-022-00752-0](https://doi.org/10.1038/s41596-022-00752-0) [Medline](#)
40. X. Li, W. Chen, B. K. Martin, D. Calderon, C. Lee, J. Choi, F. M. Chardon, T. A. McDiarmid, R. M. Daza, H. Kim, J.-B. Lalanne, J. F. Nathans, D. S. Lee, J. Shendure, Chromatin context-dependent regulation and epigenetic manipulation of prime editing. *Cell* **187**, 2411–2427.e25 (2024). [doi:10.1016/j.cell.2024.03.020](https://doi.org/10.1016/j.cell.2024.03.020) [Medline](#)
41. J.-B. Lalanne, S. G. Regalado, S. Domcke, D. Calderon, B. K. Martin, X. Li, T. Li, C. C. Suiter, C. Lee, C. Trapnell, J. Shendure, Multiplex profiling of developmental cis-regulatory elements with quantitative single-cell expression reporters. *Nat. Methods* **21**, 983–993 (2024). [doi:10.1038/s41592-024-02260-3](https://doi.org/10.1038/s41592-024-02260-3) [Medline](#)
42. S. H. Hurlbert, The Nonconcept of Species Diversity: A Critique and Alternative Parameters. *Ecology* **52**, 577–586 (1971). [doi:10.2307/1934145](https://doi.org/10.2307/1934145) [Medline](#)
43. S. Seidel, A. Zwaans, S. Regalado, J. Choi, J. Shendure, T. Stadler, SciPhy: A Bayesian phylogenetic framework using sequential genetic lineage tracing data. *Nat. Commun.* **17**, 7398 (2026). [doi:10.1038/s41467-026-73377-6](https://doi.org/10.1038/s41467-026-73377-6) [Medline](#)
44. N. Saitou, M. Nei, The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987). [Medline](#)
45. K. Piotrowska-Nitsche, A. Perea-Gomez, S. Haraguchi, M. Zernicka-Goetz, Four-cell stage mouse blastomeres have different developmental properties. *Development* **132**, 479–490 (2005). [doi:10.1242/dev.01602](https://doi.org/10.1242/dev.01602) [Medline](#)
46. B. Plusa, A.-K. Hadjantonakis, D. Gray, K. Piotrowska-Nitsche, A. Jedrusik, V. E. Papaioannou, D. M. Glover, M. Zernicka-Goetz, The first cleavage of the mouse zygote predicts the blastocyst axis. *Nature* **434**, 391–395 (2005). [doi:10.1038/nature03388](https://doi.org/10.1038/nature03388) [Medline](#)
47. P. Cardoso, F. Rigal, J. C. Carvalho, BAT – Biodiversity Assessment Tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. *Methods Ecol. Evol.* **6**, 232–236 (2015). [doi:10.1111/2041-210X.12310](https://doi.org/10.1111/2041-210X.12310)
48. M. K. Kuhner, J. Felsenstein, A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Mol. Biol. Evol.* **11**, 459–468 (1994). [Medline](#)
49. W. Wang, J. Barbetti, T. Wong, B. Thornlow, R. Corbett-Detig, Y. Turakhia, R. Lanfear, B. Q. Minh, DecentTree: Scalable Neighbour-Joining for the genomic era.

- Bioinformatics* **39**, btad536 (2023). [doi:10.1093/bioinformatics/btad536](https://doi.org/10.1093/bioinformatics/btad536) [Medline](#)
50. L. Bromham, D. Penny, The modern molecular clock. *Nat. Rev. Genet.* **4**, 216–224 (2003). [doi:10.1038/nrg1020](https://doi.org/10.1038/nrg1020) [Medline](#)
 51. T.-H. To, M. Jung, S. Lycett, O. Gascuel, Fast Dating Using Least-Squares Criteria and Algorithms. *Syst. Biol.* **65**, 82–97 (2016). [doi:10.1093/sysbio/syv068](https://doi.org/10.1093/sysbio/syv068) [Medline](#)
 52. M. H. L. Snow, Gastrulation in the mouse: Growth and regionalization of the epiblast. *Development* **42**, 293–303 (1977). [doi:10.1242/dev.42.1.293](https://doi.org/10.1242/dev.42.1.293)
 53. J. Felsenstein, CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evolution* **39**, 783–791 (1985). [doi:10.1111/j.1558-5646.1985.tb00420.x](https://doi.org/10.1111/j.1558-5646.1985.tb00420.x) [Medline](#)
 54. K. Ponnienselvan, P. Liu, T. Nyalile, S. Oikemus, A. T. Joynt, K. Kelly, D. Guo, Z. Chen, J. M. Lee, C. A. Schiffer, C. P. Emerson, N. D. Lawson, J. K. Watts, E. J. Sontheimer, J. Luban, S. A. Wolfe, Addressing the dNTP bottleneck restricting prime editing activity. *bioRxiv* [Preprint] 2023.10.21.563443 (2023); <https://doi.org/10.1101/2023.10.21.563443>
 55. P. J. Chen, J. A. Hussmann, J. Yan, F. Knipping, P. Ravisankar, P.-F. Chen, C. Chen, J. W. Nelson, G. A. Newby, M. Sahin, M. J. Osborn, J. S. Weissman, B. Adamson, D. R. Liu, Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* **184**, 5635–5652.e29 (2021). [doi:10.1016/j.cell.2021.09.018](https://doi.org/10.1016/j.cell.2021.09.018) [Medline](#)
 56. J. Ferreira da Silva, G. P. Oliveira, E. A. Arasa-Verge, C. Kagiou, A. Moretton, G. Timelthaler, J. Jiricny, J. I. Loizou, Prime editing efficiency and fidelity are enhanced in the absence of mismatch repair. *Nat. Commun.* **13**, 760 (2022). [doi:10.1038/s41467-022-28442-1](https://doi.org/10.1038/s41467-022-28442-1) [Medline](#)
 57. P. Liu, K. Ponnienselvan, T. Nyalile, S. Oikemus, A. T. Joynt, S. Iyer, K. Kelly, D. Guo, P. P. Kyawe, E. Vanderleeden, S. D. Redick, L. Huang, Z. Chen, J. M. Lee, C. A. Schiffer, D. M. Harlan, J. P. Wang, C. P. Emerson Jr., N. D. Lawson, J. K. Watts, E. J. Sontheimer, J. Luban, S. A. Wolfe, Increasing intracellular dNTP levels improves prime editing efficiency. *Nat. Biotechnol.* **43**, 539–544 (2025). [doi:10.1038/s41587-024-02405-x](https://doi.org/10.1038/s41587-024-02405-x) [Medline](#)
 58. J. Mu, J. C. Slevin, D. Qu, S. McCormick, S. L. Adamson, In vivo quantification of embryonic and placental growth during gestation in mice using micro-ultrasound. *Reprod. Biol. Endocrinol.* **6**, 34 (2008). [doi:10.1186/1477-7827-6-34](https://doi.org/10.1186/1477-7827-6-34) [Medline](#)
 59. K. R. Andrews, J. Chang, C. Roemer, J. Hadfield, V. Lin, A. F. Brito, R. O. Daodu, I. A. Joia, K. Kistler, A. Li, L. H. Moncla, M. I. Paredes, D. Kühnert, L. M. Torres, L. Voith, I. Aksamentov, E. B. Hodcroft, J. Huddleston, J. T. McCrone, J. S. J. Anderson, T. R. Sibley, J. Lee, R. A. Neher, T. Bedford, Nextstrain automates real-time phylogenetic analysis of open data for endemic and emerging pathogens. *bioRxiv* [Preprint] 2026.03.23.713807 (2026); <https://doi.org/10.64898/2026.03.23.713807>
 60. NextCell — the cell lineage of a mouse embryo, zygote to E13.5; <https://seattlehub.org/nextcell>
 61. K. Kamimoto, K. Kaneko, C. Y.-Y. Kok, H. Okada, A. Miyajima, T. Itoh, Heterogeneity and stochastic growth regulation of biliary epithelial cells dictate dynamic epithelial tissue remodeling. *eLife* **5**, e15034 (2016). [doi:10.7554/eLife.15034](https://doi.org/10.7554/eLife.15034) [Medline](#)
 62. N. Shakiba, A. Fahmy, G. Jayakumaran, S. McGibbon, L. David, D. Trcka, J. Elbaz, M. C. Puri, A. Nagy, D. van der Kooy, S. Goyal, J. L. Wrana, P. W. Zandstra, Cell competition during reprogramming gives rise to dominant clones. *Science* **364**, eaan0925 (2019). [doi:10.1126/science.aan0925](https://doi.org/10.1126/science.aan0925) [Medline](#)
 63. V. Gupta, K. D. Poss, Clonally dominant cardiomyocytes direct heart morphogenesis. *Nature* **484**, 479–484 (2012). [doi:10.1038/nature11045](https://doi.org/10.1038/nature11045) [Medline](#)
 64. C. Lopez-Garcia, A. M. Klein, B. D. Simons, D. J. Winton, Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330**, 822–825 (2010). [doi:10.1126/science.1196236](https://doi.org/10.1126/science.1196236) [Medline](#)
 65. J. Nichols, A. Lima, T. A. Rodríguez, Cell competition and the regulative nature of early mammalian development. *Cell Stem Cell* **29**, 1018–1030 (2022). [doi:10.1016/j.stem.2022.06.003](https://doi.org/10.1016/j.stem.2022.06.003) [Medline](#)
 66. T. S. Weber, C. Biben, D. C. Miles, S. P. Glaser, S. Tomei, C.-Y. Lin, A. Kueh, M. Pal, S. Zhang, P. P. L. Tam, S. Taoudi, S. H. Naik, LoxCode in vivo barcoding reveals epiblast clonal fate bias to fetal organs. *Cell* **188**, 3882–3896.e19 (2025). [doi:10.1016/j.cell.2025.04.026](https://doi.org/10.1016/j.cell.2025.04.026) [Medline](#)
 67. G. Udny Yule, A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F.R.S. *Philos. Trans. R. Soc. Lond. B* **213**, 21–87 (1925); <https://www.jstor.org/stable/92117>
 68. A. Mac Auley, Z. Werb, P. E. Mirkes, Characterization of the unusually rapid cell cycles during rat gastrulation. *Development* **117**, 873–883 (1993). [doi:10.1242/dev.117.3.873](https://doi.org/10.1242/dev.117.3.873) [Medline](#)
 69. K. Manova, C. Tomihara-Newberger, S. Wang, A. Godelman, S. Kalantry, K. Witty-Blease, V. De Leon, W. S. Chen, E. Lacy, R. F. Bachvarova, Apoptosis in mouse embryos: Elevated levels in pregastrulae and in the distal anterior region of gastrulae of normal and mutant mice. *Dev. Dyn.* **213**, 293–308 (1998). [doi:10.1002/\(SICI\)1097-0177\(199811\)213:3<293::AID-AJA6>3.0.CO;2-D](https://doi.org/10.1002/(SICI)1097-0177(199811)213:3<293::AID-AJA6>3.0.CO;2-D) [Medline](#)
 70. M. Balaban, N. Moshiri, U. Mai, X. Jia, S. Mirarab, TreeCluster: Clustering biological sequences using phylogenetic trees. *PLOS ONE* **14**, e0221068 (2019). [doi:10.1371/journal.pone.0221068](https://doi.org/10.1371/journal.pone.0221068) [Medline](#)
 71. C. Weinreb, A. Rodriguez-Fraticelli, F. D. Camargo, A. M. Klein, Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science* **367**, eaaw3381 (2020). [doi:10.1126/science.aaw3381](https://doi.org/10.1126/science.aaw3381) [Medline](#)
 72. L. H. Ngo, A. G. Bert, B. K. Dredge, T. Williams, V. Murphy, W. Li, W. B. Hamilton, K. T. Carey, J. Toubia, K. A. Pillman, D. Liu, J. Desogus, J. A. Chao, A. J. Deans, G. J. Goodall, V. O. Wickramasinghe, Nuclear export of circular RNA. *Nature* **627**, 212–220 (2024). [doi:10.1038/s41586-024-07060-5](https://doi.org/10.1038/s41586-024-07060-5) [Medline](#)
 73. J. A. Farrell, Y. Wang, S. J. Riesenfeld, K. Shekhar, A. Regev, A. F. Schier, Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, eaar3131 (2018). [doi:10.1126/science.aar3131](https://doi.org/10.1126/science.aar3131) [Medline](#)
 74. L. Snijders Blok, J. Rousseau, J. Twist, S. Ehresmann, M. Takaku, H. Venselaar, L. H. Rodan, C. B. Nowak, J. Douglas, K. J. Swoboda, M. A. Steeves, I. Sahai, C. T. R. M. Stumpel, A. P. A. Stegmann, P. Wheeler, M. Willing, E. Fiala, A. Kochhar, W. T. Gibson, A. S. A. Cohen, R. Agbahovbe, A. M. Innes, P. Y. B. Au, J. Rankin, I. J. Anderson, S. A. Skinner, R. J. Louie, H. E. Warren, A. Afenjar, B. Keren, C. Nava, J. Buratti, A. Isapof, D. Rodriguez, R. Lewandowski, J. Propst, T. van Essen, M. Choi, S. Lee, J. H. Chae, S. Price, R. E. Schnur, G. Douglas, I. M. Wentzensen, C. Zweier, A. Reis, M. G. Bialer, C. Moore, M. Koopmans, E. H. Brilstra, G. R. Monroe, K. L. I. van Gassen, E. van Binsbergen, R. Newbury-Ecob, L. Bownass, I. Bader, J. A. Mayr, S. B. Wortmann, K. J. Jakielski, E. A. Strand, K. Kloth, T. Bierhals, J. D. Roberts, R. M. Petrovich, S. Machida, H. Kurumizaka, S. Lelieveld, R. Pfundt, S. Jansen, P. Deriziotis, L. Faivre, J. Thevenon, M. Assoum, L. Shriberg, T. Kleefstra, H. G. Brunner, P. A. Wade, S. E. Fisher, P. M. Campeau; The DDD study, CHD3 helicase domain mutations cause a neurodevelopmental syndrome with macrocephaly and impaired speech and language. *Nat. Commun.* **9**, 4619 (2018). [doi:10.1038/s41467-018-06014-6](https://doi.org/10.1038/s41467-018-06014-6) [Medline](#)
 75. R. P. Kim-Yip, R. McNulty, B. Joyce, A. Mollica, P. J. Chen, P. Ravisankar, B. K. Law, D. R. Liu, J. E. Toettcher, E. A. Ivakine, E. Posfai, B. Adamson, Efficient prime editing in two-cell mouse embryos using PEEmbryo. *Nat. Biotechnol.* **42**, 1822–1830 (2024). [doi:10.1038/s41587-023-02106-x](https://doi.org/10.1038/s41587-023-02106-x) [Medline](#)
 76. B. Gu, E. Posfai, J. Rossant, Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat. Biotechnol.* **36**, 632–637 (2018). [doi:10.1038/nbt.4166](https://doi.org/10.1038/nbt.4166) [Medline](#)
 77. J. Pilarski, T. Stadler, S. Seidel, Assessing the inference of single-cell phylogenies and population dynamics from CRISPR lineage recordings. *PLOS Comput. Biol.* **22**, e1014370 (2026). [doi:10.1371/journal.pcbi.1014370](https://doi.org/10.1371/journal.pcbi.1014370) [Medline](#)
 78. P. Zaharias, T. Warnow, Recent progress on methods for estimating and updating large phylogenies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **377**, 20210244 (2022). [doi:10.1098/rstb.2021.0244](https://doi.org/10.1098/rstb.2021.0244) [Medline](#)
 79. N. Konno, Y. Kijima, K. Watano, S. Ishiguro, K. Ono, M. Tanaka, H. Mori, N. Masuyama, D. Pratt, T. Ideker, W. Iwasaki, N. Yachie, Deep distributed computing to reconstruct extremely large lineage trees. *Nat. Biotechnol.* **40**, 566–575 (2022). [doi:10.1038/s41587-021-01111-2](https://doi.org/10.1038/s41587-021-01111-2) [Medline](#)
 80. W. Fang, Y. Yang, H. Ji, R. Kalhor, Reconstructing Progenitor State Hierarchy and Dynamics Using Lineage Barcoding Data. *Methods Mol. Biol.* **2886**, 177–199 (2025). [doi:10.1007/978-1-0716-4310-5_9](https://doi.org/10.1007/978-1-0716-4310-5_9) [Medline](#)
 81. X. Huang, J. Henck, C. Qiu, V. K. A. Sreenivasan, S. Balachandran, O. V. Amarie, M. Hrabě de Angelis, R. Y. Behncke, W.-L. Chan, A. Despang, D. E. Dickel, M. Duran, A. Feuchtinger, H. Fuchs, V. Gailus-Durner, N. Haag, R. Hägerling, N. Hansmeier, F. Hennig, C. Marshall, S. Rajderkar, A. Ringel, M. Robson, L. M. Saunders, P. da Silva-Buttkus, N. Spielmann, S. R. Srivatsan, S. Ulberts, L. Wittler, Y. Zhu, V. M.

- Kalscheuer, D. M. Ibrahim, I. Kurth, U. Kornak, A. Visel, L. A. Pennacchio, D. R. Beier, C. Trapnell, J. Cao, J. Shendure, M. Spielmann, Single-cell, whole-embryo phenotyping of mammalian developmental disorders. *Nature* **623**, 772–781 (2023). [doi:10.1038/s41586-023-06548-w](https://doi.org/10.1038/s41586-023-06548-w) [Medline](#)
82. E. Barkan, M. Duran, N. Lammers, A. Tresenrider, D. Jackson, H. Lee, B. Haagen, L. Saunders, P. Abitua, D. Kimelman, C. Trapnell, Embryo-scale single-cell chemical transcriptomics reveals dependencies between cell types and signaling pathways. *bioRxiv* 2025.04.03.646423 [Preprint] (2025). <https://doi.org/10.1101/2025.04.03.646423>
83. S. Domcke, J. Shendure, A reference cell tree will serve science better than a reference cell atlas. *Cell* **186**, 1103–1114 (2023). [doi:10.1016/j.cell.2023.02.016](https://doi.org/10.1016/j.cell.2023.02.016) [Medline](#)
84. W. N. Colgan, L. W. Koblan, J. Villagrana, T.-C. J. Hou, M. Wang, G. Gowri, W. Chandler, L. A. Sepúlveda, D. Ciftci, K. Smolyar, K. E. Yost, A. Young, L. Wittler, S. Markoulaki, K. M. Loh, X. Zhuang, N. Yosef, Z. D. Smith, J. S. Weissman, Comprehensive Lineage Tracing Maps the Landscape of Cell Fate Decisions in Mouse Embryogenesis. *Cell* **189**, 2026.09.050 (2026). [doi:10.1016/j.cell.2026.09.050](https://doi.org/10.1016/j.cell.2026.09.050)
85. S. Seidel, C. Qui, J. Shendure, H. Kim, shendurelab/dtt-mouse: v1.1.0, Zenodo (2026); <https://doi.org/10.5281/zenodo.22801270>
86. H. Kim, Figure 1A, BioRender (2026); <https://BioRender.com/r2oheqm>
87. J. L. Litke, S. R. Jaffrey, Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. *Nat. Biotechnol.* **37**, 667–675 (2019). [doi:10.1038/s41587-019-0090-6](https://doi.org/10.1038/s41587-019-0090-6) [Medline](#)
88. W. Wang, A. Bradley, Y. Huang, A piggyBac transposon-based genome-wide library of insertionally mutated BLM-deficient murine ES cells. *Genome Res.* **19**, 667–673 (2009). [doi:10.1101/gr.085621.108](https://doi.org/10.1101/gr.085621.108) [Medline](#)
89. B. Martin, sci-RNA-seq3_pipeline, Github (2026); https://github.com/bethmartin/sci-RNA-seq3_pipeline/tree/MEGAsci
90. G. Renaud, U. Stenzel, T. Maricic, V. Wiebe, J. Kelso, deML: Robust demultiplexing of Illumina sequences using a likelihood-based approach. *Bioinformatics* **31**, 770–772 (2015). [doi:10.1093/bioinformatics/btu719](https://doi.org/10.1093/bioinformatics/btu719) [Medline](#)
91. J. Cao, J. S. Packer, V. Ramani, D. A. Cusanovich, C. Huynh, R. Daza, X. Qiu, C. Lee, S. N. Furlan, F. J. Steemers, A. Adey, R. H. Waterston, C. Trapnell, J. Shendure, Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017). [doi:10.1126/science.aam8940](https://doi.org/10.1126/science.aam8940) [Medline](#)
92. F. Krueger, TrimGalore, Github (2019); <https://github.com/FelixKrueger/TrimGalore>
93. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). [doi:10.1093/bioinformatics/bts635](https://doi.org/10.1093/bioinformatics/bts635) [Medline](#)
94. S. Anders, P. T. Pyl, W. Huber, HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015). [doi:10.1093/bioinformatics/btu638](https://doi.org/10.1093/bioinformatics/btu638) [Medline](#)
95. S. L. Wolock, R. Lopez, A. M. Klein, Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* **8**, 281–291.e9 (2019). [doi:10.1016/j.cels.2018.11.005](https://doi.org/10.1016/j.cels.2018.11.005) [Medline](#)
96. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018). [doi:10.1186/s13059-017-1382-0](https://doi.org/10.1186/s13059-017-1382-0) [Medline](#)
97. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018). [doi:10.1093/bioinformatics/bty560](https://doi.org/10.1093/bioinformatics/bty560) [Medline](#)
98. Y. Turakhia, B. Thornlow, A. S. Hinrichs, N. De Maio, L. Gozashti, R. Lanfear, D. Haussler, R. Corbett-Detig, Ultrafast Sample placement on Existing tRees (USHER) enables real-time phylogenetics for the SARS-CoV-2 pandemic. *Nat. Genet.* **53**, 809–816 (2021). [doi:10.1038/s41588-021-00862-7](https://doi.org/10.1038/s41588-021-00862-7) [Medline](#)
99. A. Zeileis, Measuring Inequality, Concentration, and Poverty [R package ineq version 0.2-13] (2014).
100. C. Gini, Measurement of Inequality of Incomes. *The Economic Journal* **31**, 124–126 (1921). [doi:10.2307/2223319](https://doi.org/10.2307/2223319)
101. P. Paul, S. Rajput, P. Joshi, M. Naithani, N. Chowdhury, S. Rao, M. O. Pai, Comparison of Fluorometric and UV Spectrophotometric Findings for DNA Isolated From Formalin-Fixed Paraffin-Embedded Blocks, Fine Needle Aspiration Cytology Smears, and Blood. *Cureus* **13**, e19583 (2021). [doi:10.7759/cureus.19583](https://doi.org/10.7759/cureus.19583) [Medline](#)
102. A. M. Zorn, J. M. Wells, Vertebrate endoderm development and organ formation. *Annu. Rev. Cell Dev. Biol.* **25**, 221–251 (2009). [doi:10.1146/annurev.cellbio.042308.113344](https://doi.org/10.1146/annurev.cellbio.042308.113344) [Medline](#)
103. W. V. Cardoso, J. Lü, Regulation of early lung morphogenesis: Questions, facts and controversies. *Development* **133**, 1611–1624 (2006). [doi:10.1242/dev.02310](https://doi.org/10.1242/dev.02310) [Medline](#)
104. A. Medvinsky, E. Dzierzak, Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell* **86**, 897–906 (1996). [doi:10.1016/S0092-8674\(00\)80165-8](https://doi.org/10.1016/S0092-8674(00)80165-8) [Medline](#)
105. M. Ganuza, T. Hall, D. Finkelstein, A. Chabot, G. Kang, S. McKinney-Freeman, Lifelong haematopoiesis is established by hundreds of precursors throughout mammalian ontogeny. *Nat. Cell Biol.* **19**, 1153–1163 (2017). [doi:10.1038/nch3607](https://doi.org/10.1038/nch3607) [Medline](#)
106. M. E. Bronner, N. M. LeDouarin, Development and evolution of the neural crest: An overview. *Dev. Biol.* **366**, 2–9 (2012). [doi:10.1016/j.ydbio.2011.12.042](https://doi.org/10.1016/j.ydbio.2011.12.042) [Medline](#)
107. A. Streit, The preplacodal region: An ectodermal domain with multipotential progenitors that contribute to sense organs and cranial sensory ganglia. *Int. J. Dev. Biol.* **51**, 447–461 (2007). [doi:10.1387/ijdb.072327as](https://doi.org/10.1387/ijdb.072327as) [Medline](#)
108. Y. Kojima, O. H. Tam, P. P. L. Tam, Timing of developmental events in the early mouse embryo. *Semin. Cell Dev. Biol.* **34**, 65–75 (2014). [doi:10.1016/j.semcdb.2014.06.010](https://doi.org/10.1016/j.semcdb.2014.06.010) [Medline](#)

ACKNOWLEDGMENTS

We thank the members of the Shendure lab, the Seattle Hub for Synthetic Biology, Allen Discovery Center for Cell Lineage, and UW Genome Sciences communities for helpful discussions over the course of many years (particularly the Shendure/SeaHub Recording subgroup, A. Schier, M. Elowitz, R. Waterston, and W. Colgan). We also thank the Transgenic Colony Management, Neurosurgery & Behavior, Lab Animal Services, and the Veterinary Services teams at the Allen Institute for their contributions to animal work. Q.Y., H.K., and S.S. contributed equally to this work as co-first authors. When referencing this paper in their curriculum vitae or other professional materials, each co-first author may list their own name first. Several authors used AI-based tools during this work. The authors built the NextCell interactive browser with Claude Code (Anthropic). Data exploration, data analysis, data visualization, and code development were additionally supported by Claude Opus 5, Claude Opus 4.7, and Claude Sonnet 5 (Anthropic), which the authors also used to improve the readability and language of the manuscript. The authors did not use AI-assisted tools to generate or alter primary data, draw scientific conclusions, or perform other key authoring tasks. All AI-assisted code, analyses, and text were reviewed, tested, and edited by the authors, who take full responsibility for the data, code, analyses, and conclusions presented here. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health (NIH). This research was supported in part by the Brotman Baty Institute for Precision Medicine as well as the Allen Institute, founded by J. Allen, chair and cofounder of Allen Family Philanthropies, and the late P. G. Allen, investor, philanthropist, and cofounder of Microsoft. We gratefully acknowledge their vision and generosity, which made this work possible. J.S. is an investigator of the Howard Hughes Medical Institute (HHMI). **Funding:** Allen Institute (J.S.); Brotman Baty Institute for Precision Medicine (J.S.); Cancer Research Society of Canada grant 1155581 (J.-B.L.); Chan Zuckerberg Initiative Foundation, Seattle Hub for Synthetic Biology award CZIF2023-008738 (J.S., C.T., M.P.); Charles H. Hood Foundation, Inc., MA (C.Q.); Damon Runyon Cancer Research Foundation grant DFS-64-24 (J.C.); Damon Runyon Cancer Research Foundation grant DRG-2435-21 (J.-B.L.); HHMI (J.S.); NIH grant R00HG012973 (J.C.); NIH grant P30CA008748 (J.C.); NIH grant P20GM130454 through Dartmouth's Center for Quantitative Biology (C.Q.); Paul G. Allen Frontiers Group, Allen Discovery Center for Cell Lineage Tracing (J.S., C.T.); Searle Scholars Award SSP-2025-101 (J.C.); Swiss National Science Foundation Postdoc. Mobility fellowship grant 239394 (S.S.); Washington Research Foundation Postdoctoral Fellowship (H.K.). **Author contributions:** Project initiation: Q.Y., J.C., J.S.; Circatope development: H.K. with assistance from J.F.N., Q.Y., J.-B.L., K.P.S.; Design of in vivo experiment: Q.Y., J.F.A.-C., M.G.; Design and execution of ddPCR: E.G., L.K., S.V.K.; Pronuclear zygotic injection platform development: J.F.A.-C., K.O.; Pronuclear zygotic

injections and embryo manipulations: K.O., Q.Y.; Bulk amplicon sequencing and data analysis: Q.Y., H.K., R.M.D., J.S.; Sci-RNA-seq3 experiments: B.K.M., Q.Y.; Sci-RNA-seq3 data analysis: C.Q. with assistance from J.S., S.S., C.T.; Phylogenetics: S.S. with assistance from J.S., C.Q., M.L.; Analysis of timed, tip-annotated trees: S.S., C.Q., J.S.; Software: S.S., C.Q., J.S.; Funding acquisition: J.S., C.T., M.P.; Project administration: J.S., J.M.G., J.F.A.-C.; Visualization: C.Q., S.S., H.K., Q.Y., J.S.; Manuscript writing: S.S., C.Q., H.K., Q.Y., J.S. with assistance from J.F.A.-C., B.K.M., R.M.D.; Project oversight: J.S. **Competing interests:** The University of Washington has filed a patent application related to DNA Typewriter (US20240355418A1) on which J.C. and J.S. are listed as inventors. J.S. is on the scientific advisory board, is a consultant, and/or is a cofounder of 10x Genomics, Cellular Intelligence, Guardant Health, Pacific Biosciences, and Phase Genomics. C.T. is a paid consultant for 10x Genomics and holds equity in the company. K.P.S. is a member of the Translational Advisory Council at SeaBridge, Allen Institute. All other authors declare that they have no competing interests. **Data, code, and materials availability:** Bulk and scRNA-seq sequencing data are available at Gene Expression Omnibus with identifier GSE341627. Intermediate data files (e.g., full and backbone trees in Newick format) are available at NextCell (60) or at our GitHub site for this project (85). All other data are available in the main text or the supplementary materials. Code is available at GitHub and Zenodo for this project (85). PEmax and epegRNA::Tape-BC::cirtape constructs are available from author J.S. and Addgene. **License information:** Copyright © 2026 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the Author Accepted Manuscript (AAM) of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/cgi/content/full/science.ael0508/DC1

Materials and Methods

Supplementary Text

Figs. S1 to S18

Tables S1 to S8

References (87–108)

MDAR Reproducibility Checklist

Data S1 and S2

Submitted 30 July 2026; accepted 22 September 2026

Published online 08 October 2026

10.1126/science.ael0508

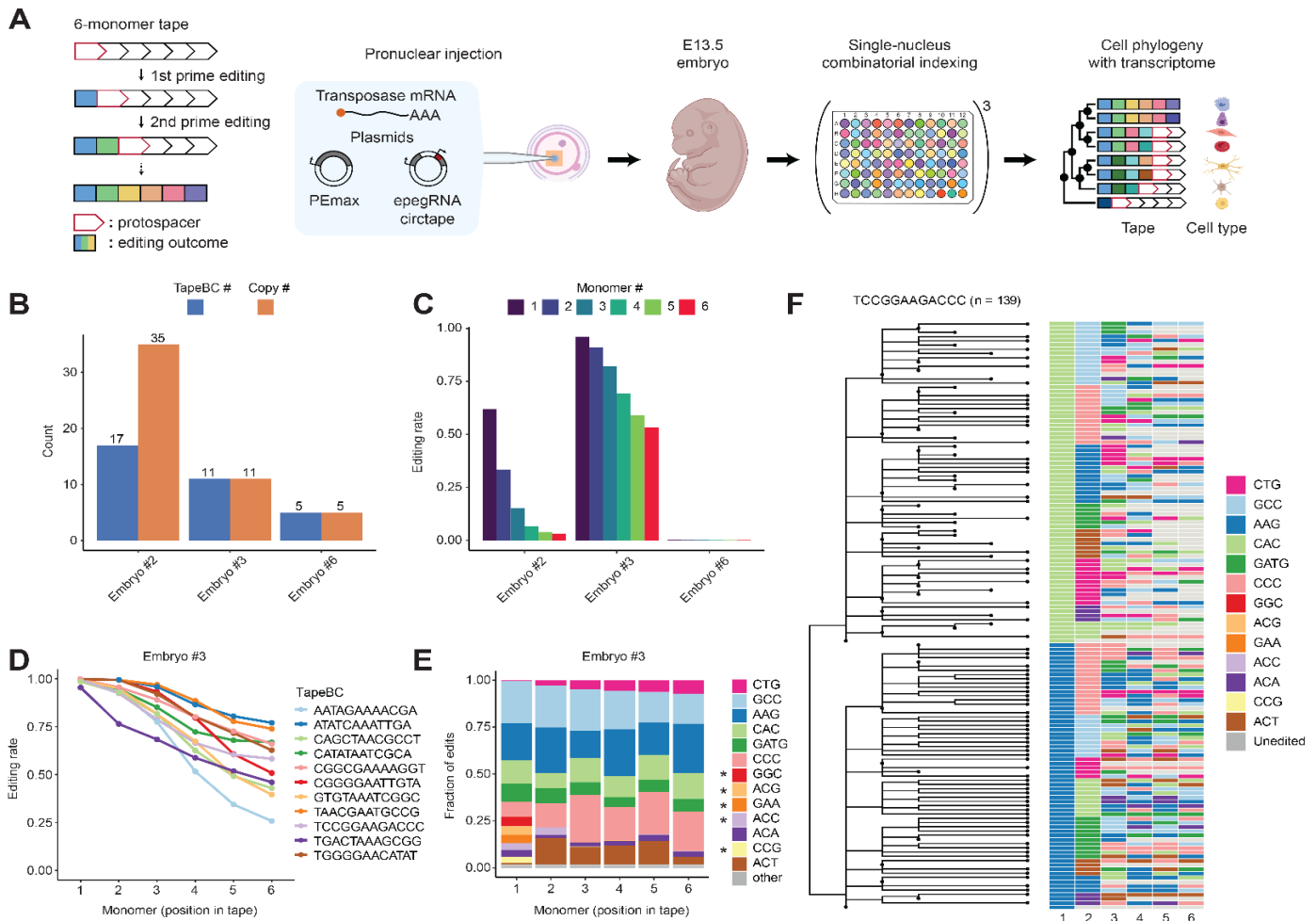


Fig. 1. PNI of DNA Typewriter components identifies an E13.5 embryo with robust levels of sequential editing. (A) Schematic. (Left) A 6-monomer tape is edited sequentially by prime editing, with each event installing an insertional symbol (colored) at the active monomer while regenerating the protospacer (red outline) at the next. (Middle) Constructs encoding prime editor (PEmax), guides (epegRNAs), and barcoded tape (tape-BC, circatpe), were delivered to the zygote by PNI, together with piggyBac transposase mRNA. The tape is transcribed by Pol-III and circularized via the Tornado system, and additionally bears a 30-nt poly(A) region and a custom capture sequence, enabling oligo-dT-based and (optionally) targeted reverse transcription, respectively. (Right) Embryos were harvested at E13.5 and transcriptionally profiled by three-level single-nucleus combinatorial indexing (sci-RNA-seq3) (39), yielding paired circatpe genotypes and transcriptomes that together facilitate the reconstruction of a cell phylogeny annotated by cell type. Some elements were created in BioRender (86). (B) Number of distinct tape-BC sequences (unique integration barcodes; blue) and the estimated total number of genomic integrations after accounting for multi-copy barcodes (orange). (C) Per-monomer editing rates for each embryo, expressed as the fraction of resolved reads edited at each of the six monomers (colored 1-6), collapsed with respect to independent integrations. (D) Per-monomer editing rate in embryo #3, shown separately for each of 11 single-copy tape-BC integrations. Editing declines progressively from monomer-1 to monomer-6, consistent with sequential editing. (E) Frequency of each insertional symbol at each of the six monomers in embryo #3, aggregated across integrations. The 13 symbols present at $\geq 0.5\%$ frequency at one or more monomer positions are shown individually; the remainder are grouped as “other”. The 5 symbols appearing at $\geq 0.5\%$ frequency only at monomer-1 and/or monomer-2 are marked with an asterisk. (F) Neighbor-joining lineage tree reconstructed from tape genotypes observed at a single representative integration (tape-BC “TCCGGAAGACCC”; n = 139 genotypes), with the corresponding edit patterns shown at right (colored by symbol; gray, unedited).

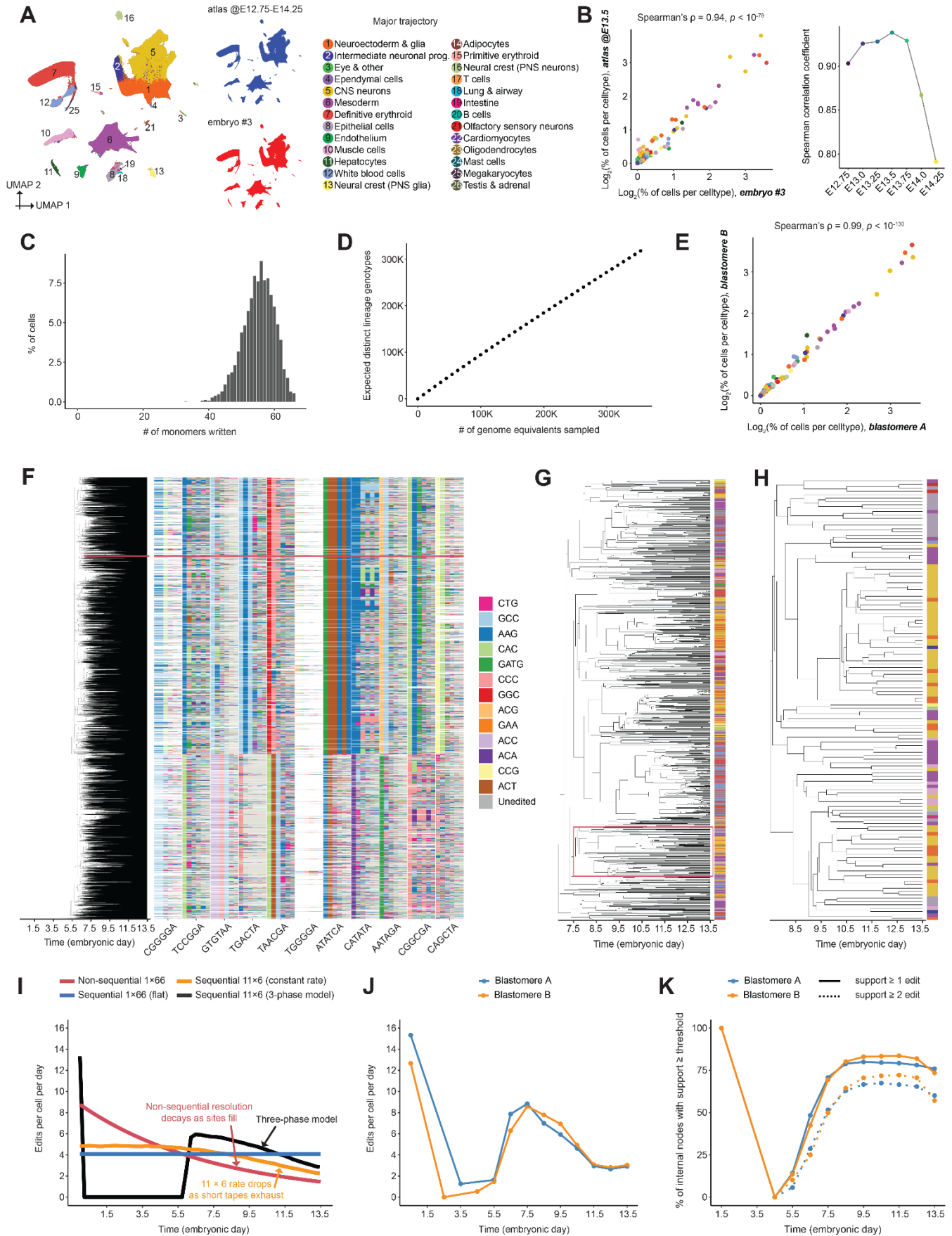


Fig. 2. Single-nucleus RNA-seq of a DNA Typewriter embryo enables reconstruction of a time-calibrated, zygote-rooted, tip-annotated phylogeny. (A) Two-dimensional (2D) uniform manifold approximation and projection (UMAP) co-embedding of 1.58M scRNA-seq profiles from embryo #3 (this study; median UMI: 906, genes: 654) and 1.65M from 7 timepoints of our high-temporal-resolution single-cell atlas of mouse development (E12.75 - E14.25 in 6-hours increments) (5). (Left) All cells, colored by the 26 major trajectories (numbered key); major trajectories accounting for more than 0.1% of cells are labeled. (Right) The same embedding split by dataset of origin. (B) (Left) Log₂-scaled percentage of cells per cell type, for each of the 169 cell types detected among cells shown in panel (A). Each point is one cell type, colored by its major trajectory. Right: Spearman's ρ between cell type proportions in embryo #3 vs. each wildtype atlas timepoint from E12.75 to E14.25. (C) Histogram of the number of edited monomers per cell in embryo #3, for the 4,371 cells with tape genotypes at all 11 tape integrations. Although cells genotyped at all 11 tapes are enriched for cell types with higher tape recovery, reweighting this distribution to the embryo's major-trajectory composition shifts it only from 55.2 ± 4.8 to 54.9 ± 5.1 monomers written (mean $\pm$ s.d.). (D) Cell-level rarefaction analysis of lineage-genotype diversity (Hurlbert interpolation), based on the 353,681 cells with a lineage genotype call for all 5 of the most highly recovered integrations. The curve remains near-linear. (E) Log₂-scaled percentages of cells per cell type in blastomere A versus blastomere B, for the 141 cell types detected in the union of A and B. Each point is one cell type, colored by its major trajectory. (F) Timed backbone phylogeny (root fixed E0; A and B subtrees fixed at E1.5, tips fixed at E13.5). The tree (left) is shown beside the per-cell tapes, colored by the observed edit outcome (right). (G and H) Zoom-in on clades outlined by red rectangles in panels (F) (1,167 cells; clade root at E6.8) and (G) (130 cells; clade root at E7.5), respectively. Major trajectory annotations of tips are shown at right [color legend in panel (A)]. (I) Mean edits accumulated per cell versus developmental time, for four models, each tuned to reach 55 of 66 edits by E13.5. Non-sequential editing to a 1×66 -monomer tape under a constant per-site, per-day rate (red) yields ~6-fold variation in temporal resolution, from ~8.8 edits/cell/day at the outset to ~1.5 by E13.5, as sites are consumed. Sequential editing to a 1×66 -monomer tape under a constant per-site, per-day rate (blue) holds flat at ~4.1 edits/cell/day. Sequential editing to 11×6 -monomer tapes under a constant per-site, per-day rate (orange) behaves similarly, except that the aggregate rate falls late as individual tapes reach their limit. Sequential editing to 11×6 -monomer tapes under a three-phase model (black) comprises a 20-edit burst, a slowdown interval approximated as quiescence until E6, and then a constant per-site, per-day rate. (J) Observed tape editing rate across developmental time for blastomere A and blastomere B. On the backbone tree, ancestral tape states were inferred as the longest edit prefix shared by all descendant genotypes of each node (a conservative rule at nodes with many descendants), and branches assigned to 1-day bins from E0 to E14 by their temporal midpoint; the rate for each bin is the sum of edits across integrations divided by the total branch time falling in that bin. The peak in the E0-E1.0 bin corresponds to the founder burst, placed on each blastomere's stem lineage. (K) Internal node support across developmental time for blastomere A and B, binned by node date over the same 1-day windows. Support is the number of independent edits inferred to occur along the branch entering a node, from the same ancestral-state reconstruction as panel (J); panels show the percentage of nodes with support ≥ 1 or ≥ 2 . In panels (J) and (K), points are plotted at bin midpoints, and bins containing no branch midpoint (J) or dated node (K) are omitted rather than plotted as zero. (J) E1.5 in both blastomeres, E2.5, E4.5 in A, E3.5 in B; (K) E2.5, 3.5 in both blastomeres; panel (K) additionally begins at E1.5, the age of the blastomere subtree roots. These gaps, and the sparse sampling of neighboring early bins, reflect the small number of resolvable divisions during the slowdown interval.

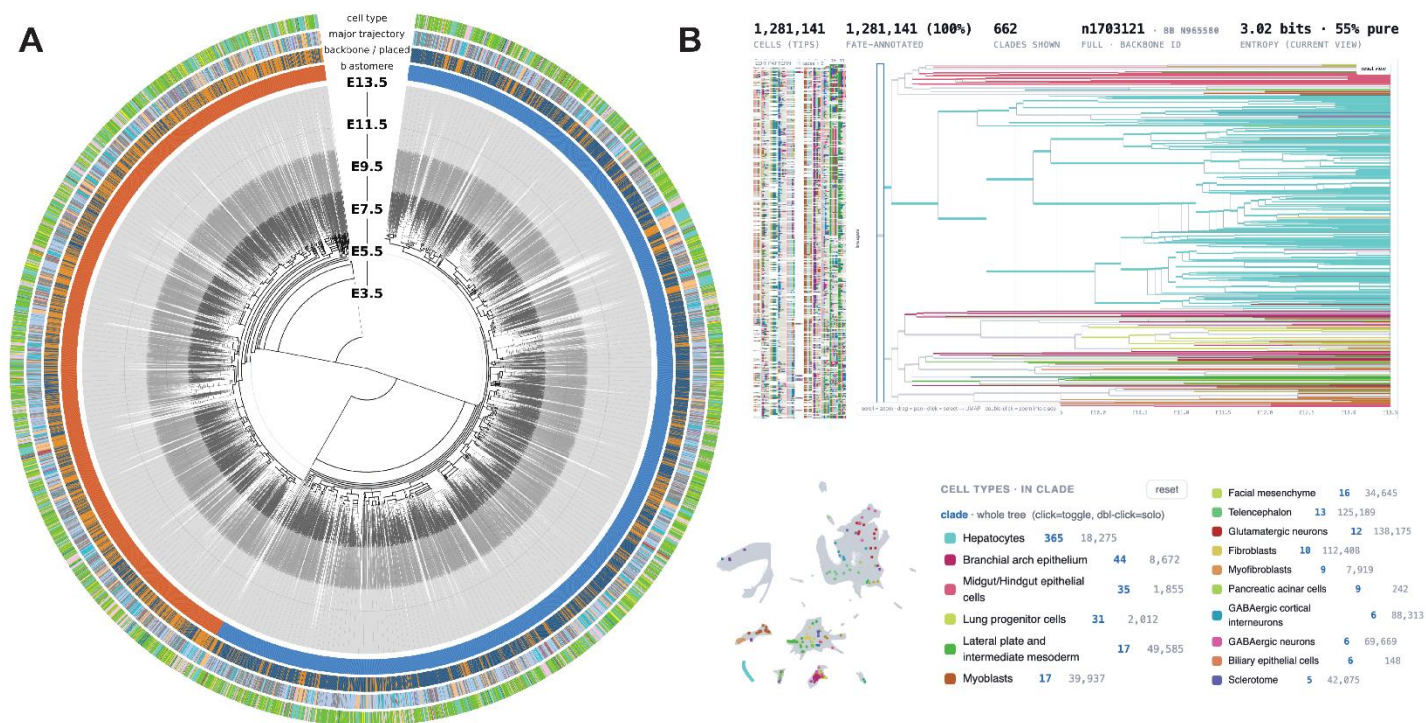


Fig. 3. A $\sim 10^6$ cell phylogeny of a developing mouse from zygote to late organogenesis and an interactive browser to explore it. (A) 1,281,141 cells shown as a circular dendrogram in which radial position gives inferred node age, from zygote (center) to E13.5 (perimeter). Branch widths are scaled in 2-day steps to render the sparse early backbone legible. The two blastomere sub-lineages, A and B, occupy the right and left halves. Concentric rings encode, from inner to outer: (i) blastomere identity (A, blue; B, orange); (ii) reconstruction origin (backbone, dark blue; placement, orange); (iii) major trajectory (legend in Fig. 2A); and (iv) cell type. The red asterisk and wedge mark the clade shown in (B); the wedge's angular width is exaggerated ~ 5 -fold for visibility. **(B)** Interactive visualization in NextCell (60), focused on ancestral node n1703121 in the full tree, an endodermally fated progenitor (E7.4) whose 662 sampled descendants are clonally partitioned among cell types associated with visceral organs, including hepatocytes ($n = 365$), branchial arch epithelium ($n = 44$), midgut/hindgut epithelium ($n = 35$), and lung progenitors ($n = 31$). Panels show the edit states of 11 genotyped tapes (top left); a dendrogram of the subclade, with branches colored where $\geq 60\%$ of descendant cells share a cell type (top right); a UMAP of all cells in the final tree, highlighting subclade members (bottom left); and cell types observed ≥ 5 times, with counts in the subclade and among all tips (bottom right).

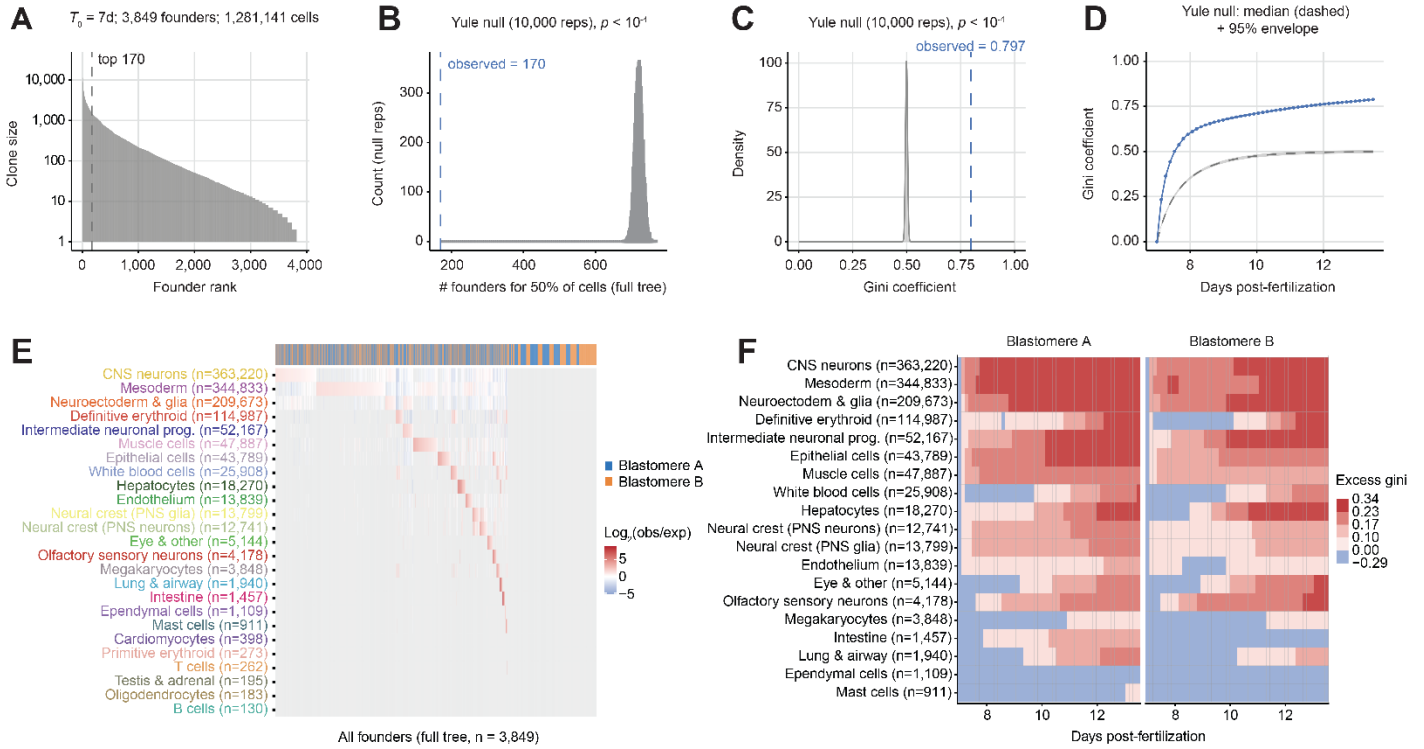


Fig. 4. Quantitation of clonal dominance arising during gastrulation. (A) Rank-abundance of founder clone sizes ($\log_{10}$ scale). Founder clones ($n = 3,849$) are defined as the lineages present at $T_0 = 7.0$ days post-fertilization in the dated tree; each founder's clone size is its number of E13.5 descendants. The dashed line marks the top 170 founders. (B) Number of founders required to generate 50% of all E13.5 cells, in observed data (blue dashed line, $n = 170$) vs. the Yule null (gray, 10,000 replicates). $p = P(\text{null} \leq \text{observed})$. (C) Gini coefficient of founder clone sizes on the full tree ($0 = \text{every founder contributes equally}$, $1 = \text{all cells derive from a single founder}$); the observed value (dashed line) is compared with its distribution under the Yule null (10,000 replicates). $p = P(\text{null Gini} \geq \text{observed})$. (D) Emergence of clonal dominance on the backbone tree. The Gini coefficient computed across the same 3,848 founders' descendant-lineage counts at each time T , tracking the fixed founder set forward from T_0 to E13.5 (blue), and comparing against the Yule null (gray) and its 95% envelope (gray). (E) Founder clone contribution to each major trajectory relative to its overall contribution to the embryo, full placement tree ($n = 3,849$ founders, both blastomeres combined). Color shows $\log_2(\text{observed/expected})$ cells of each trajectory (rows, ordered by sampled cells) from each founder (columns, ordered by most-enriched trajectory), where expected assumes a founder contributes to every trajectory in proportion to its overall clone size. Founders are ordered by the trajectory in which they are most enriched, then by enrichment strength within it. Top strip, blastomere of origin (blue, blastomere A; orange, blastomere B); gray, fewer than 10 sampled descendants of that founder in that trajectory. (F) Excess Gini, the observed Gini index minus the 97.5% upper bound of the blastomere-specific Gini null, on the backbone tree for major cell-type trajectories with ≥ 100 cells in both blastomeres ($n = 19$), over developmental time, split by blastomere of origin. Red, higher inequality than the blastomere's overall neutral expectation; blue, less. Smaller trajectories may show less signal due to estimator noise, not necessarily true absence of Gini excess.

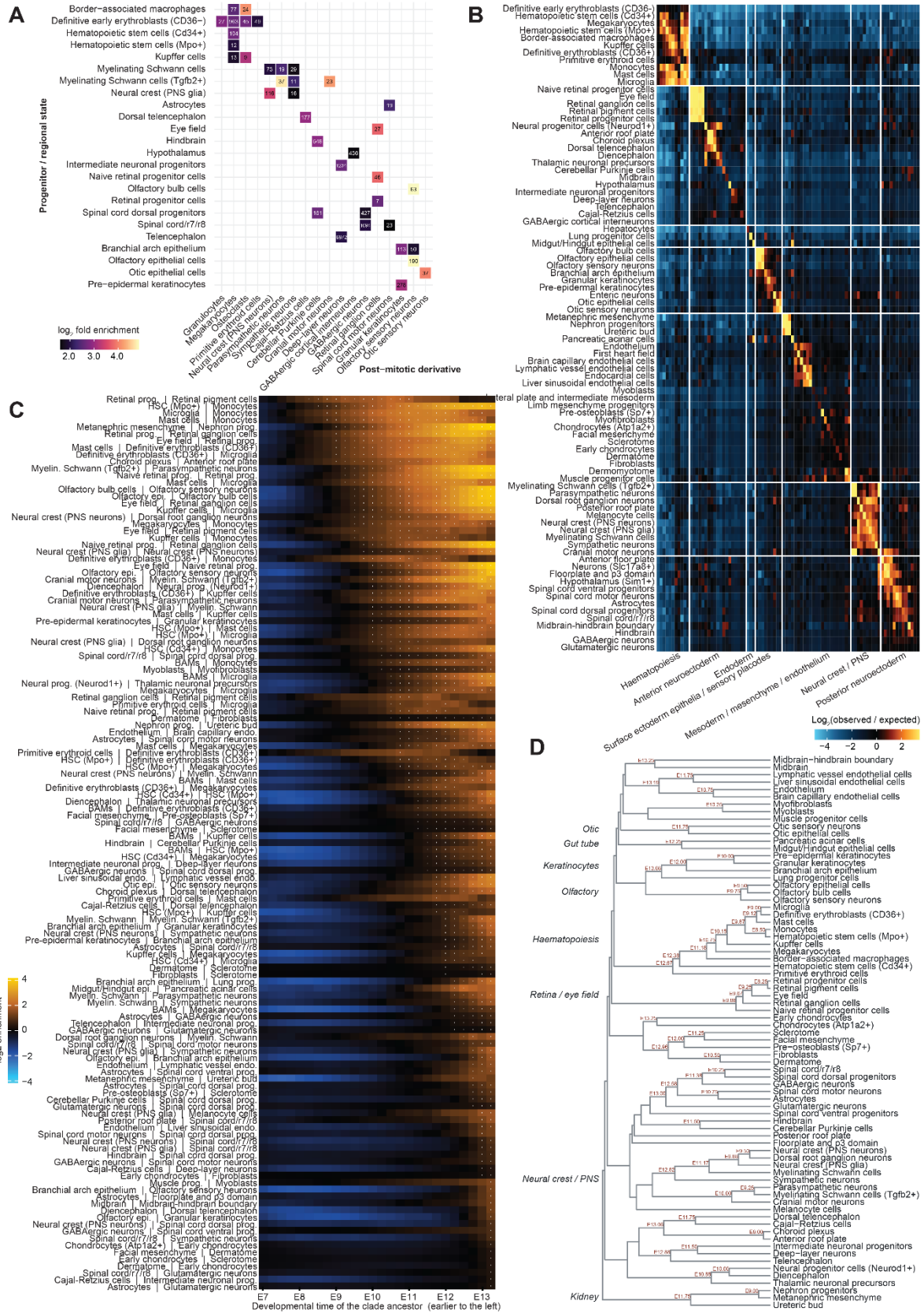


Fig. 5. Lineage predicts fate at every scale, from sibling pairs to a dated hierarchy of cell-type couplings. (A) Heatmap of the 37 progenitor → post-mitotic heterotypic pairings that significantly and reliably recur among late-splitting tree siblings above chance. Rows, 24 progenitor or regional states; columns, 18 post-mitotic cell types; both ordered by germ layer. Color, $\log_2$ -fold enrichment of the pairing over a tip-label-permuted null. Numbers within cells, post-mitotic cells supporting that pairing (range 7-6,972; 13,549 in total). White, pairing not captured. All pairings shown are supported by at least 3 post-mitotic cells, are at least 3-fold enriched, and clear a Benjamini-Hochberg FDR of 1% across the 640 pairings tested. (B) Briefly, we partitioned the tree into “maximal clades” (fig. S16), and for each pair of abundant cell types, counted the number of clades containing both (including self-pairs) and compared this to the expected count under random cell-type label shuffling (100 permutations). $\log_2$ -enrichment (observed / expected), clipped to $[-5, 3.5]$ is shown at $K=15$, capturing local family groups. Cell types are ordered by hierarchical clustering (Ward's linkage on Euclidean distance) of the $K=15$ matrix, with 4 of 88 cell types manually repositioned to make the annotated groupings contiguous (see materials and methods for details of manual repositionings); the same order is used across rows and columns. White lines were manually added and demarcate clusters of cell types corresponding to major groupings shown at bottom. (C) Timed fate couplings between cell types. Each row is one of the 129 cell-type pairs with a defined coupling depth. Each column is one of the 26 developmental time slices from E7.0 to E13.25 in 0.25-day steps, with earlier ancestors to the left. Color, $\log_2$ -enrichment (observed / expected under tip-label permutation null) at that slice, averaged over timings calculated on each blastomere subtree independently. White dots mark slices at which the pair is co-enriched in both blastomere sub-trees independently. The leftmost white dot in each row therefore corresponds to a given pair's coupling depth. Rows are ordered by coupling depth, deepest first, then by peak enrichment. Within each pair label, the cell type appearing earlier in the Qiu *et al.* developmental graph (5) is listed first (a presentational guidepost, not a claim of directionality, which this analysis does not compute). Source values are provided in table S7. (D) A draft hierarchy of timed fate couplings. Average-linkage hierarchy over the matrix of pairwise coupling depths, built on 75 cell types that couple with at least one partner. The tree is drawn unscaled and cannot be read as a phylogeny; horizontal distance is merge rank, not time. Relatedly, the hierarchy appears inverted with respect to a lineage splitting through time, because branches here are joined, rather than split, at their deepest coupling. Of its 74 branch points, the 51 that close at or before E13.25, the last slice, are labeled with that coupling depth; the remaining 23 merge only beyond the last slice and so carry no coupling depth. As coupling depth reports how far back two fates remain confined to a shared pool of lineages, not when they diverged, groups closing earliest stay co-restricted deepest. Labels at the right of cell types name manually curated coherent groups. 13 of 88 cell types (≥ 100 cells in each blastomere subtree) couple with no partner at any slice by this method and are omitted.

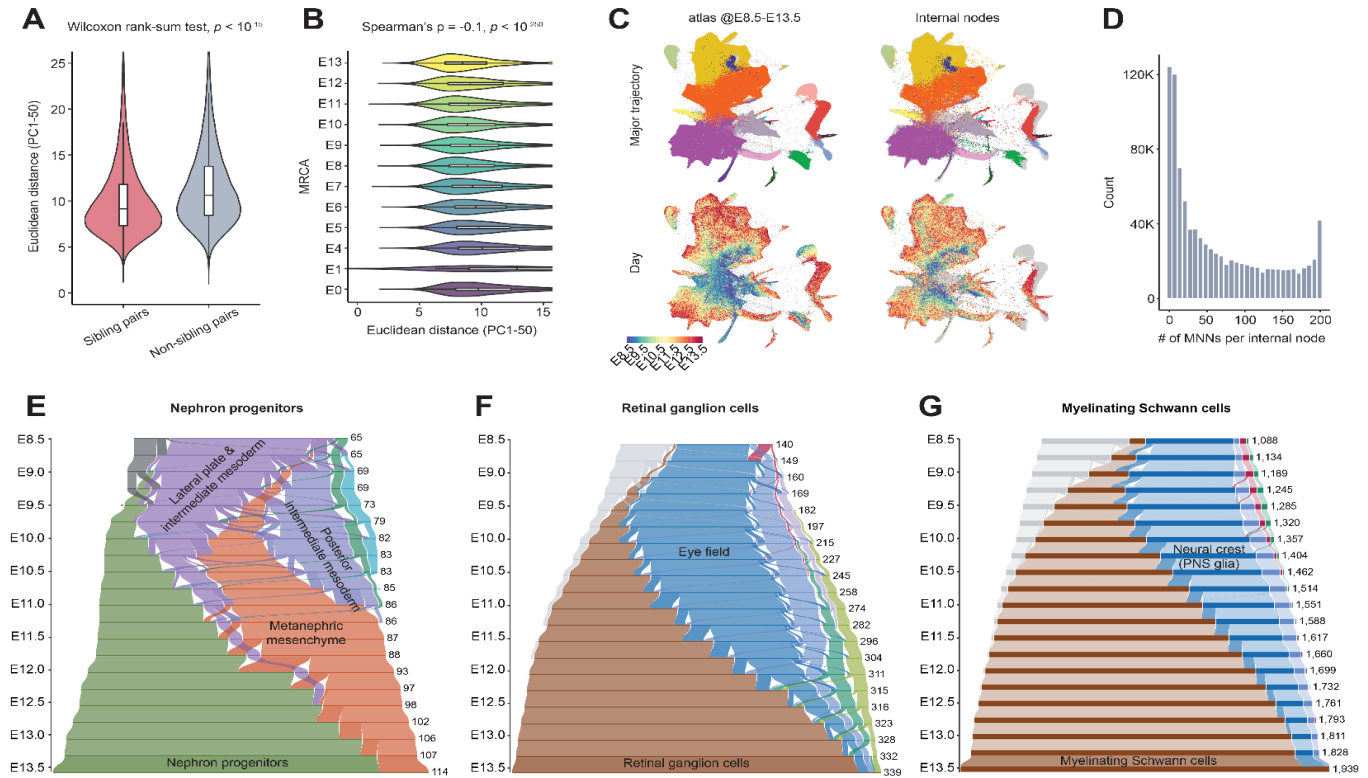


Fig. 6. Heuristic inference of the transcriptional states and cell type annotations of ancestral nodes. (A) For each cell type, up to 10,000 sibling pairs were randomly sampled, along with a matched number of non-sibling control pairs of the same cell type (excluding known sibling and self-pairs). Transcriptional distance was defined as the Euclidean distance in the first 50 principal components (PCs). Inner box plots indicate the median and interquartile range. (B) Same-cell type pairs, binned by most recent common ancestor (MRCA) depth on the lineage tree, were sampled at up to 10,000 pairs per unit-depth bin. Transcriptional distance was defined as the Euclidean distance in the first 50 PCs. Inner box plots indicate the median and interquartile range. Spearman's ρ between transcriptional distance and MRCA unit-depth bin are reported. (C) 2D UMAP co-embedding of 1,973,105 single-cell profiles, comprising 926,852 internal nodes with imputed transcriptomes from embryo #3 and ~1 million cells from 21 timepoints of our high-temporal-resolution single-cell atlas of mouse development (E8.5–E13.5, 6-hour increments; each timepoint downsampled to 50,000 cells). UMAP was first computed on the atlas cells, and internal nodes were then projected onto this embedding. Both internal nodes and atlas cells were represented in a joint 50-dimensional PCA space, computed by integrating embryo #3 (all E13.5) with ~5.2 million single-cell profiles from the atlas (E8–E13.5); each internal node's transcriptome was defined as the mean of the PCA coordinates of its assigned mutual nearest neighbors (MNNs). (D) Histogram of the number of MNNs assigned to each of 926,852 internal nodes. (E to G) Examples illustrating ancestral state views in NextCell (60). (E) Developmental origins of a cell type at E13.5, illustrated for nephron progenitors. Cells of the indicated type were traced back through the lineage tree, and at each 0.25-day slice from E13.25 to E8.5, every lineage was assigned the inferred state of its nearest annotated ancestor; ancestral state calls extend back only to E8.5. Cells whose lineage carries no called ancestor, and those on rare paths (carrying <1% of the cell type's cells), are set aside rather than plotted, leaving 114 of 607 nephron progenitors. Vertical position indicates developmental time. Bar width is proportional to the number of ancestral lineages, given at right, and segments are colored by inferred state (up to the 8 most abundant shown individually, the remainder grouped as "Other"; gray, the portion of a lineage earlier than its own earliest called ancestor). Within each gap, one ribbon per state links lineages between adjacent slices (faint for same-state, solid for a state change) so ribbon crossings mark inferred state transitions. Single-slice A→B→A excursions (0.25 d) were removed. Going backward in time, lineages coalesce into fewer ancestors; the dashed bottom row is the E13.5 cells themselves, one lineage per cell and therefore a single state by construction. (F) As in panel (E), for retinal ganglion cells (339 of 1,645 plotted). (G) As in panel (E), for myelinating Schwann cells (1,939 of 7,663 plotted).



A DNA Typewriter records the cell lineage history of a mouse, from zygote to late organogenesis

Qi Yu, Haedong Kim, Sophie Seidel, James F. Acosta-Clark, Beth K. Martin, Kyle O'Connor, Riza M. Daza, Molly Gasperini, Jenny F. Nathans, Maggie Lam, Elena Gamo, Shruthi Vijay Kumar, Lauren Kuo, Jean-Benoît Lallanne, Kamen P. Simeonov, Marion Pepper, Cole Trapnell, Jesse M. Gray, Junhong Choi, Chengxiang Qiu, and Jay Shendure

Science Ahead of Print DOI: 10.1126/science.ael0508

View the article online

<https://www.science.org/doi/10.1126/science.ael0508>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science (ISSN 1095-9203) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2026 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works