

Fluctuating DNA methylation tracks cancer evolution at clinical scale

<https://doi.org/10.1038/s41586-025-09374-4>

Received: 17 October 2023

Accepted: 9 July 2025

Published online: 10 September 2025

Open access

 Check for updates

Calum Gabbutt^{1,2,3,16}✉, Martí Duran-Ferrer^{4,5,16}, Heather E. Grant¹, Diego Mallo⁶, Ferran Nadeu^{4,5}, Jacob Househam^{1,3}, Neus Villamor^{4,5,7}, Madlen Müller⁸, Simon Heath⁸, Emanuele Raineri⁸, Olga Krali⁹, Jessica Nordlund⁹, Thorsten Zenz^{10,11}, Ivo G. Gut^{8,12}, Elias Campo^{4,5,7,12}, Armando Lopez-Guillermo^{4,5,7}, Jude Fitzgibbon³, Chris P. Barnes¹³, Darryl Shibata¹⁴, José I. Martin-Subero^{4,5,12,15,17}✉ & Trevor A. Graham^{1,3,17}✉

Cancer development and response to treatment are evolutionary processes^{1,2}, but characterizing evolutionary dynamics at a clinically meaningful scale has remained challenging³. Here we develop a new methodology called EVOFLUX, based on natural DNA methylation barcodes fluctuating over time⁴, that quantitatively infers evolutionary dynamics using only a bulk tumour methylation profile as input. We apply EVOFLUX to 1,976 well-characterized lymphoid cancer samples spanning a broad spectrum of diseases and show that initial tumour growth rate, malignancy age and epimutation rates vary by orders of magnitude across disease types. We measure that subclonal selection occurs only infrequently within bulk samples and detect occasional examples of multiple independent primary tumours. Clinically, we observe faster initial tumour growth in more aggressive disease subtypes, and that evolutionary histories are strong independent prognostic factors in two series of chronic lymphocytic leukaemia. Using EVOFLUX for phylogenetic analyses of aggressive Richter-transformed chronic lymphocytic leukaemia samples detected that the seed of the transformed clone existed decades before presentation. Orthogonal verification of EVOFLUX inferences is provided using additional genetic data, including long-read nanopore sequencing, and clinical variables. Collectively, we show how widely available, low-cost bulk DNA methylation data precisely measure cancer evolutionary dynamics, and provides new insights into cancer biology and clinical behaviour.

Cancer development is an evolutionary process^{1,2}; consequently, the evolutionary history of a cancer may set its future trajectory and allow inference of the clinical path of a patient⁵. However, testing this hypothesis directly is challenging because longitudinal patient samples are required to document evolutionary history. Consequently, evolutionary histories are typically inferred from single timepoint data; for example, somatic (epi)mutations are patterned in distinctive ways by differing evolutionary dynamics³. In the haematological system, genome sequencing of single cells or single-cell colonies have been used to infer the phylogenetic relationships among cells^{6–8}. The expense of this approach has restricted analyses to small numbers of cases, limiting suitability for clinical translation.

DNA methylation can serve as a lineage marker, recording the clonal architecture of cell populations^{9–13} or the proliferative history^{14,15}.

We have recently identified DNA methylation at some CpG sites, which stochastically fluctuates over time at a timescale measured in years⁴. These fluctuating CpGs (fCpGs) function as a ‘methylation barcode’, providing a low-cost strategy to provide high temporal resolution lineage tracing in patient samples⁴. In this study, we constructed a quantitative modelling framework called evolutionary inference using fluctuating methylation (EVOFLUX). This framework enables precise quantitative inference of the evolutionary history of cancer cells from input fCpG data derived from clinical specimens, at scale (Fig. 1a).

EVOFLUX works by considering the heterogeneity of fCpG methylation values within a sample. At a diploid locus, each fCpG can take one of three states: neither allele methylated, one allele methylated or both alleles methylated (0%, 50% or 100% methylated, respectively), so n fCpG sites can take 3^n possible methylation patterns. fCpGs fluctuate

¹Centre for Evolution and Cancer, Institute of Cancer Research, London, UK. ²X Centre for AI in Science, Imperial College London, London, UK. ³Centre for Genomics and Computational Biology, Barts Cancer Institute, Queen Mary University of London, London, UK. ⁴Fundació de Recerca Clínic Barcelona-Institut d'Investigacions Biomèdiques August Pi i Sunyer (FRCB-IDIBAPS), Barcelona, Spain. ⁵Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain. ⁶Arizona Cancer Evolution Center, Biodesign Institute and School of Life Sciences, Arizona State University, Tempe, AZ, USA. ⁷Hospital Clínic de Barcelona, Barcelona, Spain. ⁸Centro Nacional de Análisis Genómico (CNAG), Barcelona, Spain. ⁹Department of Medical Sciences, Molecular Precision Medicine and Science for Life Laboratory, Uppsala University, Uppsala, Sweden. ¹⁰Department of Medical Oncology and Hematology, University Hospital and University of Zürich, Zurich, Switzerland. ¹¹The LOOP Zurich, Medical Research Center, Zurich, Switzerland. ¹²Universitat de Barcelona, Barcelona, Spain. ¹³Department of Cell and Developmental Biology, University College London, London, UK. ¹⁴Department of Pathology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. ¹⁵Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. ¹⁶These authors contributed equally: Calum Gabbutt, Martí Duran-Ferrer. ¹⁷These authors jointly supervised this work: José I. Martin-Subero, Trevor A. Graham. ✉e-mail: c.gabbutt@imperial.ac.uk; imartins@recerca.clinic.cat; trevor.graham@icr.ac.uk

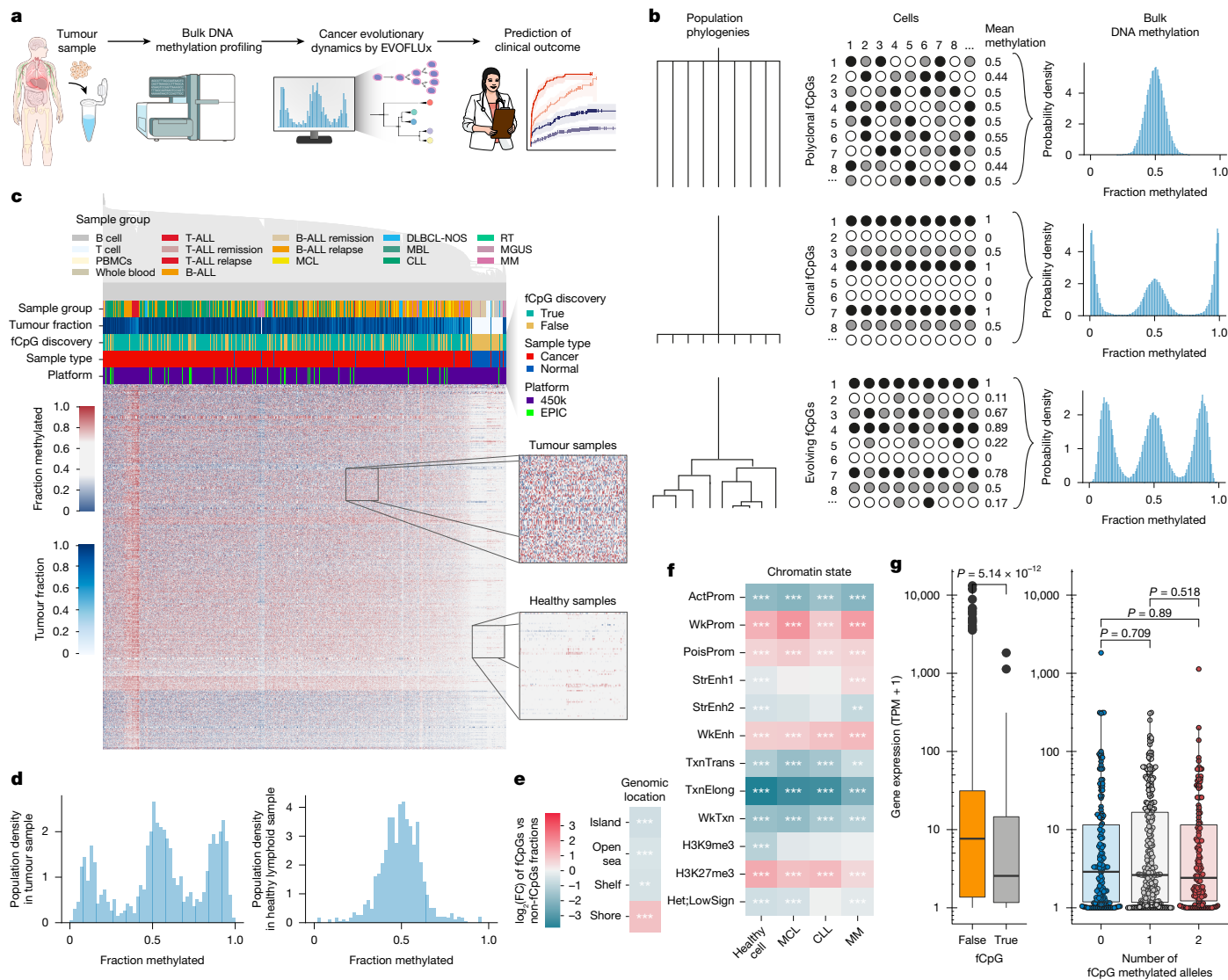


Fig. 1 | Selection and characterization of fCpG loci. **a**, Schematic representing the study design. Bulk cancer tissue samples were collected, DNA extracted and methylation arrays were performed. Evolutionary dynamics were inferred and these were correlated with clinical variables and outcomes. The illustrations of the body, tumour, eppendorf tube and machine were reproduced courtesy of NIAID. The illustration of the physician was adapted from Science Figures, under an Open Design Licence 1.0. The illustration of the screen was created by Simon Dürr under a CC0 1.0 licence. **b**, fCpG methylation patterns reflect population evolutionary dynamics. Three scenarios are depicted: in a polyclonal population with a distant MRCA, diverse fCpG methylation (unmethylated in white, heterozygous in grey and methylated in black) results in a unimodal distribution of average methylation near the steady state (top). Following a recent clonal expansion (cell number 3 of the top panel), identical fCpG methylation across cells yields a characteristic ‘W-shape’ in the bulk methylation distribution (middle). Post-bottleneck, ongoing fluctuations generate diverse fCpG methylation patterns, changing the distribution of bulk methylation values (bottom). **c**, A hierarchically clustered heatmap of the

methylation status independently, meaning that they function as an ‘evolving barcode’ to track clonal evolution: two somatic cells with close ancestry will share a near-identical pattern of fCpG methylation, whereas distantly related cells will have divergent fCpG methylation patterns (Fig. 1b). In bulk populations of clonal somatic cells, the dominant fCpG pattern represents the fCpG state of the founder cell of the population. Therefore, the precise distribution of fCpG methylation is determined by the evolutionary history of the population, meaning

978 fCpGs identified in lymphoid cells ($n = 2,204$ samples). Magnified regions represent the homogeneous intermediate methylation pattern in normal lymphoid cells and the speckled pattern in cancer samples. PBMC, peripheral blood mononuclear cell; MGUS, monoclonal gammopathy of undetermined significance; MM, multiple myeloma; NOS, not otherwise specified; RT, Richter transformation. **d**, Example histograms of fCpG methylation distributions from healthy and neoplastic lymphoid cells (CLLs). **e, f**, Heatmaps showing the log₂ fold change (FC) in genomic locations (**e**) and chromatin status (**f**) of fCpGs compared with non-fCpGs. **g**, Comparison between fCpG-associated and non-fCpG-associated genes in a single CLL sample (left; $P = 5.14 \times 10^{-12}$, Wilcoxon test, sample SCLL-328), and expression of fCpG-associated genes separated by discretized allele methylation status (right; P values were determined by two-sided Wilcoxon test, no multiple correction; $n = 505$ for fCpG genes and $n = 15,736$ for non-fCpG genes). The boxplot centre shows the median, the box shows the quartiles and whiskers represent $\pm 1.5 \times$ interquartile range (IQR). TPM, transcripts per million.

that mathematical modelling can be used to recover the evolutionary history of a sample from input fCpG data.

Here we focused on lymphoid neoplasms, which cover a broad spectrum of diseases and subtypes with highly variable clinicobiological features, from highly proliferative acute disease to indolent chronic leukaemia, arising in infants to older adults, with tumour samples across disease stages^{16,17}. These tumours have been extensively profiled by DNA methylation arrays, which have provided insights into their

cellular origin, pathogenesis and clinical behaviour¹⁸. Although their temporal clonal dynamics has been partially analysed in few patients^{19,20}, their precise evolutionary histories remain poorly characterized. Applying EVOFLUX to 1,976 well-characterized lymphoid malignancies, we precisely measured individual tumour evolutionary history and show that these histories are associated with disease outcome.

Characterization of fCpGs in lymphoid cancers

We assembled bulk Illumina methylation array data of normal and neoplastic lymphoid cells from 2,430 samples^{14,21–30} (Methods; Supplementary Tables 1–3). Following quality control, we retained 2,204 samples from 2,054 patients and 389,180 CpGs. As fCpG loci are tissue specific⁴, we constructed a pipeline to identify lymphoid-specific fCpGs (Methods; Extended Data Fig. 1a) using 1,471 samples from multiple lymphoid tumour entities (Supplementary Tables 1 and 2). We identified 978 pan-lymphoid cancer fCpGs (Supplementary Table 4). Methylation at fCpGs shows a characteristic ‘speckled’ pattern across cancers (Fig. 1c) because (de)methylation occurs independently in each tumour, in stark contrast to the orderly patterns observed for traditional methylation clocks or a random subset of CpGs (Extended Data Fig. 1b–d). fCpGs did not cluster the samples by disease (Fig. 1c), and there was also no clustering based on disease subtype or array platform (Extended Data Fig. 1e–i), except for some B and T cell acute lymphoblastic leukaemias (B/T-ALLs) and multiple myeloma cases that exhibited global hypermethylation or hypomethylation, respectively, as previously reported¹⁴. This is consistent with fCpGs behaving as a stochastic ‘barcode’ encoding lineage information. By comparison, methylation at CpGs excluded by our selection filters either did cluster by disease, had very low heterogeneity across samples or had unequal methylation and demethylation rates (Supplementary Figs. 1 and 2).

We then examined the methylation value distribution of these 978 fCpG in individual samples (Supplementary Table 5). In each cancer sample, the fCpGs followed a characteristic ‘W-shaped’ distribution that depicts the fCpG methylation pattern of the founder cell of the cancer sample (Fig. 1d). By contrast, the healthy B cell subpopulations, which were not included in the discovery set, had unimodal distributions with intermediate methylation levels consistent with these being polyclonal populations (that is, average of the three methylation states; Fig. 1d).

Methylation values across fCpGs were uncorrelated, except for a small number of fCpGs located within 1 kb of another fCpG (74 of 978; Extended Data Fig. 2a). In whole-genome bisulfite sequencing (WGBS) data of sorted bulk B and T cell populations³¹, methylation at fCpG loci in these polyclonal normal samples was largely intermediate. Over a small window of 100 bp, as the distance from the fCpGs increased, an increasing fraction of the neighbouring CpGs were either hypermethylated or hypomethylated (Extended Data Fig. 2b). Together, these analyses suggest that the local 3D genome structure influences (de) methylation processes.

We sought to verify fCpGs as ‘evolving barcodes’. Analysis of fCpG methylation fluctuation over time confirmed that inter-patient fCpG heterogeneity was not caused by common single-nucleotide polymorphisms (SNPs; Methods; Extended Data Fig. 2c and Supplementary Fig. 3). We generated long-read nanopore sequencing³² on normal B cells ($n = 6$) and matched chronic lymphocytic leukaemia (CLL)-Richter transformation samples ($n = 2$ pairs) to simultaneously detect genetic mutations and DNA methylation, and confirmed that fCpG methylation variation is not a consequence of underlying somatic mutation (Extended Data Fig. 2d and Supplementary Fig. 4). In matched data, fCpG methylation levels measured by bead array or long-read sequencing were highly concordant (Extended Data Fig. 2e), and similar excellent concordance was observed in additional WGBS data (Extended Data Fig. 2f). We constructed fCpG methylation haplotypes using long-read sequencing (Extended Data Fig. 2g) and additional single-cell reduced representation bisulfite sequencing⁶, and detected lower

intra-haplotype heterogeneity within CLL samples than normal B cell samples (Extended Data Fig. 2h and Supplementary Fig. 5), consistent with the leukaemia being a clonal expansion, whereas normal B cells are polyclonal.

We utilized somatic copy number alterations in 492 CLL and 85 mantle cell lymphoma (MCL) samples^{26,33} (Supplementary Tables 6 and 7) to distinguish between alleles and show the (de)methylation at fCpGs occurred independently on each allele (Extended Data Fig. 3a), despite copy number alterations being rare in our cohorts (Supplementary Fig. 6). Thus, fCpGs show independent ongoing allele-specific changes to methylation, uniquely labelling cell lineages.

As the DNA methylome is influenced by age³⁴, we tested whether fCpGs showed evidence of age-dependent epigenetic modulation. In normal blood samples, mean fCpG methylation was not correlated with age, suggesting that fluctuations continue throughout life, whereas fCpG methylation variance increased with age (Extended Data Fig. 3b and Supplementary Fig. 7). Variance is higher in samples where there has been a recent clonal expansion (that is, homozygous methylated or unmethylated alleles become more prominent; Extended Data Fig. 3c), suggesting that fCpGs were detecting age-related clonal expansions of cells of the haematopoietic system^{4,35,36}.

We analysed the genomic features of fCpG sites. fCpGs were enriched on the shores of CpG islands (Fig. 1e), underrepresented in gene-associated regions (Extended Data Fig. 3d) and, notably, were distinct from CpGs used in other epigenetic clocks (Extended Data Fig. 3e and Supplementary Tables 8 and 9). At the chromatin level, fCpGs were enriched in normal and neoplastic B cell weak promoters and enhancers as well as H3K27me3-marked regions, and significantly underrepresented in active promoters and H3K36me3-marked regions (Fig. 1f). RNA-sequencing analysis of CLL samples demonstrated that genes associated with fCpGs have significantly lower expression levels (Fig. 1g and Extended Data Fig. 3f), with no association between fCpG methylation status and associated gene expression in matched cases (Fig. 1g). No correlation was observed between fCpG methylation and the expression of key DNA methylation modifier genes (Extended Data Fig. 3g). Pathway enrichment analysis revealed that fCpG-associated genes were underrepresented in pathways ubiquitously expressed across multiple tissue types but enriched in developmental pathways (Supplementary Tables 10 and 11). Although these results do not provide a detailed molecular understanding of the mechanisms underpinning fCpG fluctuation; together, they indicate that fCpGs tend to be located in silent regions of the genome and do not regulate transcription, so are likely to be neutral lineage markers.

EVOFLUX measures clonal evolution

We developed EVOFLUX, a stochastic mathematical modelling and Bayesian inference framework, to simulate how clonal evolution quantitatively determines fCpG methylation values and enable inference of evolutionary history of individual tumour samples (Fig. 2a).

EVOFLUX simulates the ongoing gain and loss of methylation at fCpGs within a lineage from the birth of a patient until the beginning of a cancer-associated clonal expansion at some specified time, and then continues to simulate methylation fluctuations within the growing population of cancer cells until the cancer sample was collected at time T (Methods; Fig. 2a and Supplementary Information). The key parameters in the model are:

- Cancer growth rate per year (θ), assuming an exponentially growing population.
- Cancer age, measured in terms of the age of the patient in years at the time the cancer started growing (τ).
- fCpG switching rates per allele per year. Four parameters corresponding to the four possible transitions between homozygous unmethylated, heterozygous methylated and homozygous methylated (μ , ν , γ and ζ).

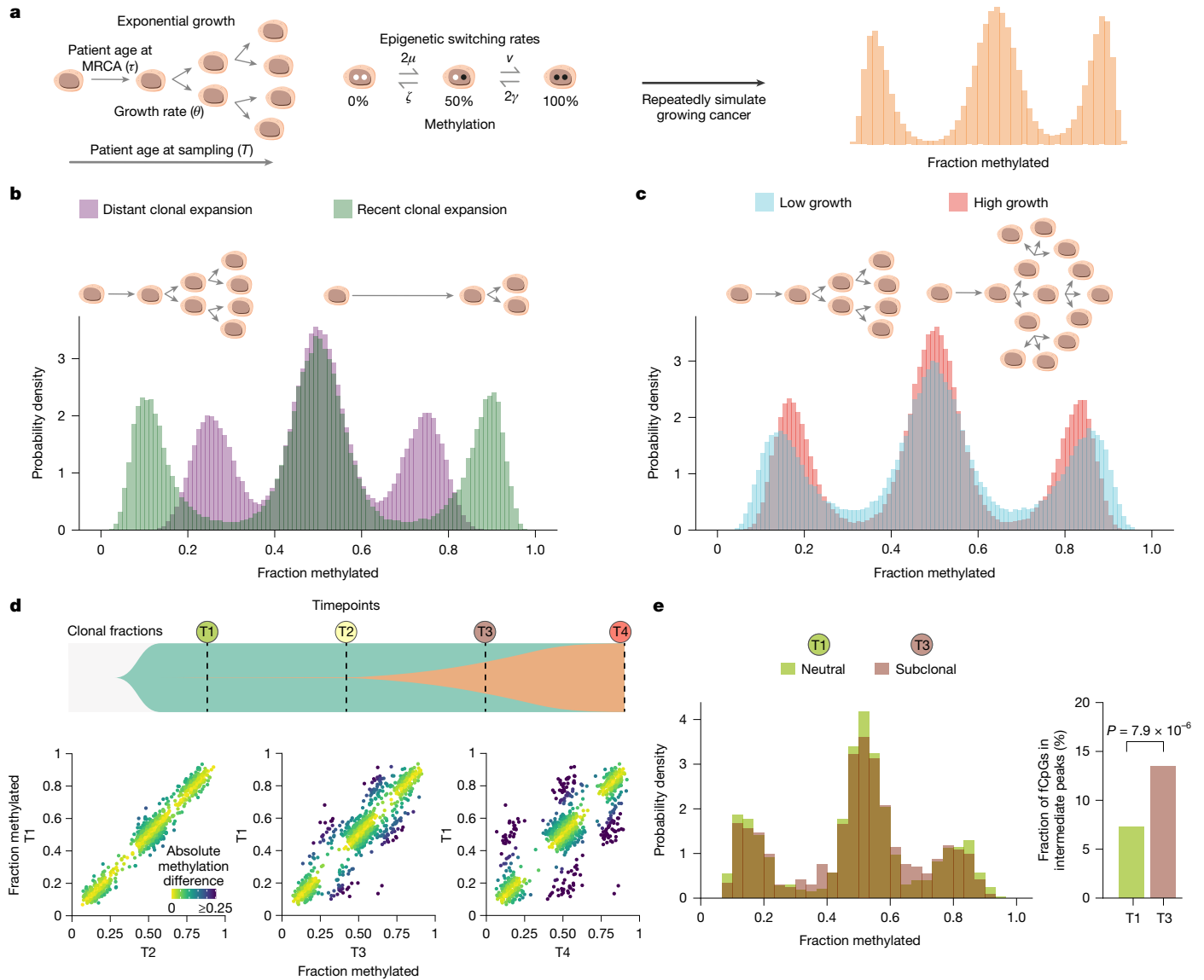


Fig. 2 | The EVOFLUX model accurately captures fCpG data patterns.

a, Schematic of the EVOFLUX method that simulates the evolutionary dynamics of a growing cancer. EVOFLUX relies on a mathematical model describing how the fCpG distributions vary with the evolutionary parameters of a cancer. The model is split into two phases: before the MRCA (τ) in which methylation changes occur in the single-cell lineage that is subsequently transformed, and following the MRCA in which the population grows exponentially (θ). At each time step, epigenetic switching is allowed to occur between the three possible states (rates μ , ν , γ and ζ). **b,c**, Simulated fCpG methylation distributions for cancer with a distant versus recent MRCA (**b**), and rapid versus slow growth

(**c**). **d**, Fisher plot showing the change in the subclonal composition within a simulated set of longitudinal samples (top). Scatter plots showing the marginal fCpG methylation distribution between simulated pairwise longitudinal samples with the subclonal proportions depicted in each corresponding sample timepoint ($n=1,000$ simulated fCpGs; bottom). Points in the scatter plot are coloured according to the difference in methylation between the first and the subsequent timepoints. **e**, Histogram overlaying the fCpG methylation distributions of data simulated at T1 (neutral) and T3 (subclonal; left), and comparison of the fraction of fCpGs at intermediate values between T1 and T3 (right; the P value was determined by chi-squared test).

By combining the cancer growth rate and age, the cancer effective population size (N_e) – the number of long-lived lineages in the cancer – is calculated as $N_e = e^{\theta(T-\tau)}$.

Computational simulations of the model recapitulated the observed W-shaped fCpG methylation distribution. Altering model parameters caused notable shifts in the distribution: increasing cancer age caused the flanking peaks of the W-shaped distribution to move towards the central peak, whereas slower growth broadened peak width (Fig. 2b,c). Hence, the distribution of fCpGs encoded the evolutionary history of a tumour.

We added simulations of a single advantageous subclone within the cancer (Methods). Sampling longitudinally from model simulations and comparing fCpG methylation between timepoints showed that

subclone outgrowth was marked by the small number of fCpGs with distinct methylation status in the subclone, becoming detectable only when the subclone was sufficiently large (Fig. 2d,e).

EVOFLUX contained an extensive Bayesian inference method to learn model parameters from input fCpG methylation distribution data, accounting for tumour purity and the technical noise introduced by the methylation array (Methods; Fig. 3a). We generated simulated fCpG distributions with prespecified (that is, known) model parameters and validated the ability of EVOFLUX to accurately recover the ‘ground truth’ in these simulated data (Extended Data Fig. 4a,b and Supplementary Information), even when the assumptions underlying the method were weakened (Extended Data Fig. 4c–h and Supplementary Information).

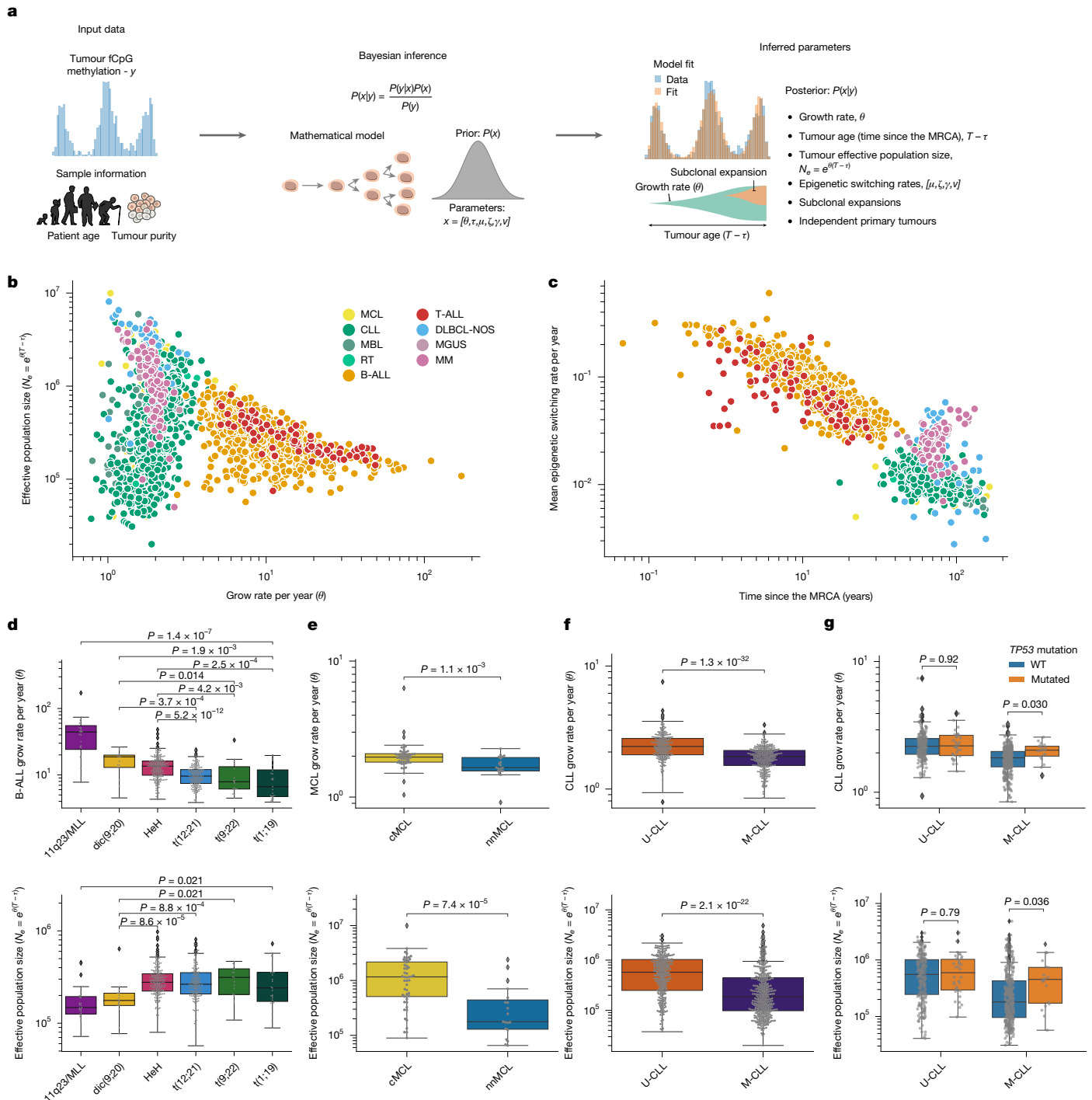


Fig. 3 | EVOFLux reveals the evolutionary dynamics of lymphoid cancers.

a, Schematic showing the EVOFLux Bayesian inference pipeline used to infer patient-specific tumour evolutionary history from input bulk methylation data, patient age and tumour purity. The illustration of human ageing was adapted from Pixabay under a CC0.1.0 licence. The illustration of the tumour was adapted courtesy of NIAID. **b**, Scatter plot of the inferred growth rate (θ) versus effective population size ($N_e = e^{\theta(T-\tau)}$) per sample from patients with cancer, using the posterior median as a descriptive summary statistic of the probability distribution inferred for each evolutionary variable ($n = 1,885$ samples; coloured according to cancer type). **c**, Scatter plot of the inferred time since the MRCA (τ) versus the mean epigenetic switching rate (averaged across different rate parameters) per sample ($n = 1,885$ samples). **d–f**, Inferred growth rate (top) and effective population size (N_e ; bottom) of individual cancer

samples separated by molecular subtype in B-ALL (**d**; two-sided MWU test, $n = 767$ samples), MCL (**e**; two-sided MWU test, $n = 126$ samples) and CLL (**f**; two-sided MWU test, $n = 702$ samples). Differences between subtypes were tested using MWU tests, with Holm–Bonferroni corrections applied. In panel **d**, B-ALL cases with MLL rearranged had significantly higher growth rate than each of the other B-ALL subtypes (all $P < 0.005$); however, for ease of presentation, only the comparison with t(1;19) is shown here. cMCL, conventional MCL; nnMCL, non-nodal MCL. **g**, Inferred growth rate (top) and effective population size (bottom) for CLL cases with a TP53 driver mutation in each IGHV CLL subtype (two-sided MWU, $n = 646$ samples, multiple corrected using false discovery rate together with analysis of other driver mutations; Extended Data Fig. 7). WT, wild type. The boxplot centre shows the median, the box denotes the quartiles and whiskers represent $\pm 1.5 \times$ IQR.

Evolution of lymphoid malignancies

We applied EVOFLUX to 1,976 samples of lymphoid cancers (including T-ALL, B-ALL, CLL, MCL, diffuse large B cell lymphoma (DLBCL) and multiple myeloma) and premalignant conditions (that is, monoclonal B cell lymphocytosis (MBL) and monoclonal gammopathy of undetermined significance) for which we had age and tumour cell purity information, to infer the individual growth rate of each cancer, time since the most recent common ancestor (MRCA) and epigenetic switching rates (Supplementary Table 12). Posterior distributions were well formed (Extended Data Fig. 5a) and posterior predictive distributions recapitulated the input data well (Extended Data Fig. 5b), emphasizing the excellent fit of the model to data. Inferred parameters were not significantly affected by tumour cell content of samples (Supplementary Fig. 8) or exclusion of copy number alteration-altered regions (Extended Data Fig. 5c–e and Supplementary Fig. 9). Most parameters were also insensitive to the number of fCpGs excluded (Extended Data Fig. 5f,g), except the effective population size (Extended Data Fig. 5h).

Paediatric ALL and adult lymphoid neoplasms exhibited markedly different evolutionary histories (Fig. 3b). ALLs demonstrated much higher growth rates (θ ; Extended Data Fig. 6a; $P = 9.3 \times 10^{-306}$, Mann–Whitney U (MWU) test, Holm–Sidak correction), smaller effective population sizes (N_e ; Extended Data Fig. 6b; $P = 8.1 \times 10^{-25}$) and shorter times since the MRCA (Extended Data Fig. 6c; $P = 6.0 \times 10^{-306}$) than other lymphoid malignancies. T-ALL grew faster than B-ALL ($P = 0.0017$, Holm–Sidak correction) and showed more homogenous growth rates ($P = 0.00044$, Levene test). In adult cancers, MBL (a precursor to CLL) displayed lower growth rates and longer time since the MRCA than CLL (Extended Data Fig. 6a,c; $P = 9.7 \times 10^{-10}$ and $P = 9.9 \times 10^{-13}$, respectively). DLBCL notably had the largest N_e despite comparable growth rates.

fCpG switching rates varied significantly across diseases, with paediatric ALLs showing much faster switching than adult malignancies (Fig. 3c and Extended Data Fig. 6d). Paediatric ALLs also demonstrated a strong positive correlation between fCpG switching rate and growth rate (Extended Data Fig. 6e; $P = 2.4 \times 10^{-98}$ and $R^2 = 0.44$ in B-ALL and $P = 5.9 \times 10^{-6}$ and $R^2 = 0.22$ in T-ALL) and a strong negative correlation with patient age (Extended Data Fig. 6f–k; $P = 3.3 \times 10^{-137}$ and $R^2 = 0.56$ in B-ALL and $P = 3.6 \times 10^{-18}$ and $R^2 = 0.6$ in T-ALL). These findings suggest that fCpG (de)methylation rates are decreased in adult lymphoid cancers.

In CLL, we estimated the ‘contemporary’ growth rate derived from multiple longitudinal clinical measurements of the lymphocyte count preceding treatment. The EVOFLUX inferred growth rate, which represents the rate of the initial clonal expansion of the disease, was moderately correlated with the contemporary growth rate ($P = 2 \times 10^{-5}$, $R = 0.27$; Extended Data Fig. 6l and Supplementary Fig. 10).

Evolution varies by cancer subtype

We examined how cancer evolutionary history related to molecular subtypes. In B-ALL, MLL-rearranged cases had a significantly higher growth rate (Fig. 3d; $P = 1.3 \times 10^{-13}$, MWU, 44.3 ± 6.1 versus 11.7 ± 0.2 per year (mean \pm standard error)), but lower effective population size N_e ($P = 3.7 \times 10^{-7}$, $1.8 \times 10^5 \pm 0.2 \times 10^5$ versus $3.0 \times 10^5 \pm 0.06 \times 10^5$ cells) than the other subtypes, consistent with their distinct clinical behaviour³⁷. In MCL, the generally more indolent leukaemic non-nodal MCL²⁶ had a lower growth rate (Fig. 3e; $P = 1.1 \times 10^{-3}$, 1.7 ± 0.1 versus 2.1 ± 0.1 per year) and N_e ($P = 7.4 \times 10^{-5}$, $4.7 \times 10^5 \pm 1.4 \times 10^5$ versus $1.5 \times 10^6 \pm 0.2 \times 10^6$ cells) than the more aggressive conventional MCL. In DLBCL transcriptomic subtypes³⁸, there was no significant differences, probably due to the smaller number of cases and the lower sample purity (Supplementary Fig. 11).

In CLL, two major molecular subtypes are defined based on the extent of somatic hypermutation in the heavy-chain variable region of the IG gene (IGHV): unmutated CLL (U-CLL) and mutated CLL (M-CLL). The

more aggressive U-CLL subtype^{39,40} showed significantly higher growth rates (Fig. 3e; $P = 1.3 \times 10^{-32}$, 2.3 ± 0.04 versus 1.8 ± 0.02 per year) and larger N_e ($P = 2.1 \times 10^{-22}$, $7.2 \times 10^5 \pm 0.3 \times 10^5$ versus $4.1 \times 10^5 \pm 0.3 \times 10^5$ cells) than M-CLL, independent of tumour purity (Supplementary Fig. 12). Similar results were obtained when analysing its precursor condition MBL (Supplementary Fig. 12).

Patients with mutations in specific driver genes, such as *TP53*, are well known to have a worse prognosis⁴¹. We compared the inferred growth rates and effective population sizes accounting for IGHV status for the most prevalent driver genetic alterations in CLL: *TP53*, *SF3B1*, *NOTCH1*, *ATM*, *POT1* and *IGLV3-21*^{R110}, *del(11)(q22.3)*, *del(13)(q14.3)*, *del(17)(p13.1)* and trisomy 12 (Fig. 3g and Extended Data Fig. 7). Patients with M-CLL with *TP53* mutations had a higher growth rate and effective population size ($P = 0.030$ and $P = 0.036$, respectively, MWU test, false discovery rate corrected; Fig. 3g). Trisomy 12 was associated with increased effective population size in both U-CLL and M-CLL ($P = 0.036$ and $P = 0.036$, respectively), but no difference in growth rate.

Most lymphoid cancers grow effectively-neutrally

As new advantageous subclones can arise during cancer evolution, we also examined subclonal architecture in our cohort. In CLL, a small fraction of cases presents two or more clones with independent origins⁴². In genetic data, subclonal, independent and monoclonal (Extended Data Fig. 8a) architectures are distinguishable by characteristic patterning of mutation allele frequencies^{43,44}. Similarly, in fCpG data, clonal architectures are depicted by additional intermediate peaks in the methylation distribution (Fig. 2d,e). Simulations showed that subclone inference by EVOFLUX was limited to detect only strongly selected subclones arising at an intermediate timepoint in the history of the tumour (Supplementary Fig. 13 and Supplementary Information), for reasons analogous to limitations of subclone detection in genetic sequencing data⁴⁵. We describe the evolution in tumours without detectable subclones as effectively-neutral.

Applying EVOFLUX in our cohort revealed that most cancers (1,610 of 1,976) showed no evidence of either subclonal or independent clones (Extended Data Fig. 8b and Supplementary Table 12). The frequency of subclone detection varied considerably between cancer types, ranging from over 30% in CLL (232 of 718) to less than 5% in DLBCL (1 of 57).

We verified EVOFLUX inferences with matched whole-exome sequencing (WES) data from 425 CLL cases (Supplementary Table 13). Using the MOBSTER subclonal deconvolution tool⁴³, subclones were detected in 78 of 425 cancers (Supplementary Table 14), and these cancers had significantly higher EVOFLUX subclonality weights ($P = 2.0 \times 10^{-4}$, MWU test; Extended Data Fig. 8c). MOBSTER was more likely to detect subclones in cancers with more mutations (Extended Data Fig. 8d), suggesting limited power to detect subclones in WES. We therefore obtained matched whole-genome sequencing (WGS) data for 127 CLL samples (Supplementary Table 15) and observed better agreement between EVOFLUX and MOBSTER subclone calls ($P = 3.9 \times 10^{-4}$; Extended Data Fig. 8e and Supplementary Table 16), in which MOBSTER subclone calls were then independent of single-nucleotide variant count ($P > 0.05$). A classifier to predict WGS subclone architecture using EVOFLUX outperformed a WES-based classifier (area under the curve (AUC) = 0.73 versus AUC = 0.62) and performed equivalently to a classifier using both EVOFLUX and WES (AUC = 0.74; Extended Data Fig. 8f). Hence, EVOFLUX was more effective at detecting ongoing subclonal selection than MOBSTER applied to WES data.

CLLs with two independent clonal origins were detected in 22 of 718 cases. Validation through comparing IG gene rearrangements from WES or WGS and RNA sequencing⁴⁶ (Supplementary Table 17) showed that patients with multiple IG gene rearrangements had elevated independent origin model weightings (Extended Data Fig. 8g; $P = 0.028$).

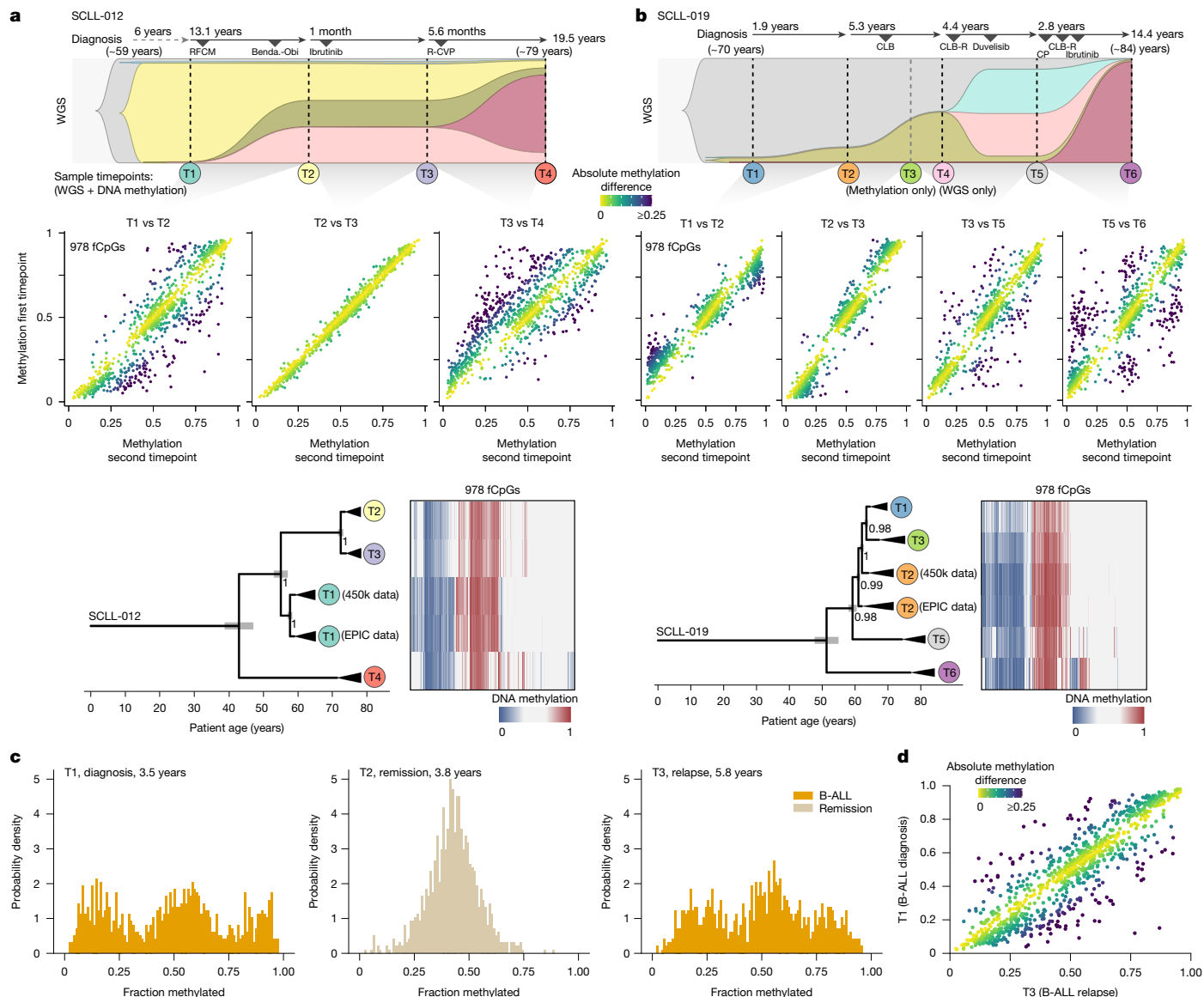


Fig. 4 | fCpGs allow for phylogenetic reconstruction of longitudinal lymphoid cancer samples. **a, b**, Timelines and Fisher plots derived from WGS of two patients with CLL with longitudinal samples, annotated with treatment received (top). The Richter-transformed clone is shown in puce. Scatter plots showing the marginal fCpG methylation distribution between pairwise samples from above (middle). The points are coloured according to the difference in methylation between the first and the subsequent timepoints. Reconstructed phylogenies of the relationship between samples, annotated with the sample timepoints (bottom). The black triangles represent the time that occurred since the MRCA ($T - \tau$), taken as the posterior median from the

single-sample EVOFLUX inferences. The methylation fraction of the 978 fCpG loci are presented as heatmaps (0% blue and 100% red). Benda-Obi, bendamustine, obinutuzumab; CLB, chlorambucil; CLB-R, chlorambucil and rituximab; CP, cyclophosphamide and prednisone; R-CVP, rituximab, cyclophosphamide, vincristine and prednisone; RFCM, rituximab, fludarabine, cyclophosphamide and mitoxantrone. **c**, An example of longitudinal samples (SW-BCP-ALL-375) showing the development of the fCpG distribution from diagnostic B-ALL, through remission and relapse. **d**, Scatter plot showing the marginal fCpG methylation distribution between diagnosis and relapse from panel **c**.

fCpGs record clonal dynamics over time

Some patients with CLL undergo Richter transformation, the emergence of an aggressive phenotype with dismal prognosis. We assembled longitudinal matched WGS and methylation data²⁷ for two patients with CLL developing Richter transformation followed for 19.5 and 14.5 years, respectively. WGS data provided ground-truth measurement of clonal evolution during the decades of longitudinal follow-up, which we contrasted with clonal inference from methylation data (Fig. 4a,b and Supplementary Table 18).

In CLL case 12, between T1 and T2 (13.1 years), WGS data showed that a subclonal expansion occurred, and this was mirrored in fCpGs (Fig. 4a). In the short (1 month) period between T2 and T3, the patient received

ibrutinib treatment, but there was no clonal expansion detected in WGS nor methylation data. The patient then received rituximab, cyclophosphamide, vincristine and prednisone (R-CVP) combination treatment, but by T4 (5.6 months after T3) presented a clinical manifestation of Richter transformation, and WGS showed that the Richter transformation clone had expanded to form 77% of the tumour. This very large clonal expansion was clearly evident in fCpG methylation data.

In CLL case 19, the initial samples at T1–T4 spanned a period of 7.1 years. WGS showed gradual expansion of a subclone that was mirrored in fCpG methylation data (Fig. 4b). At T5 (5.5 years later), there had been a large nested subclonal expansion detected by WGS that was also evident in fCpG data. At T6 (2.8 years later), the patient was diagnosed with Richter transformation. WGS showed near-fixation of the

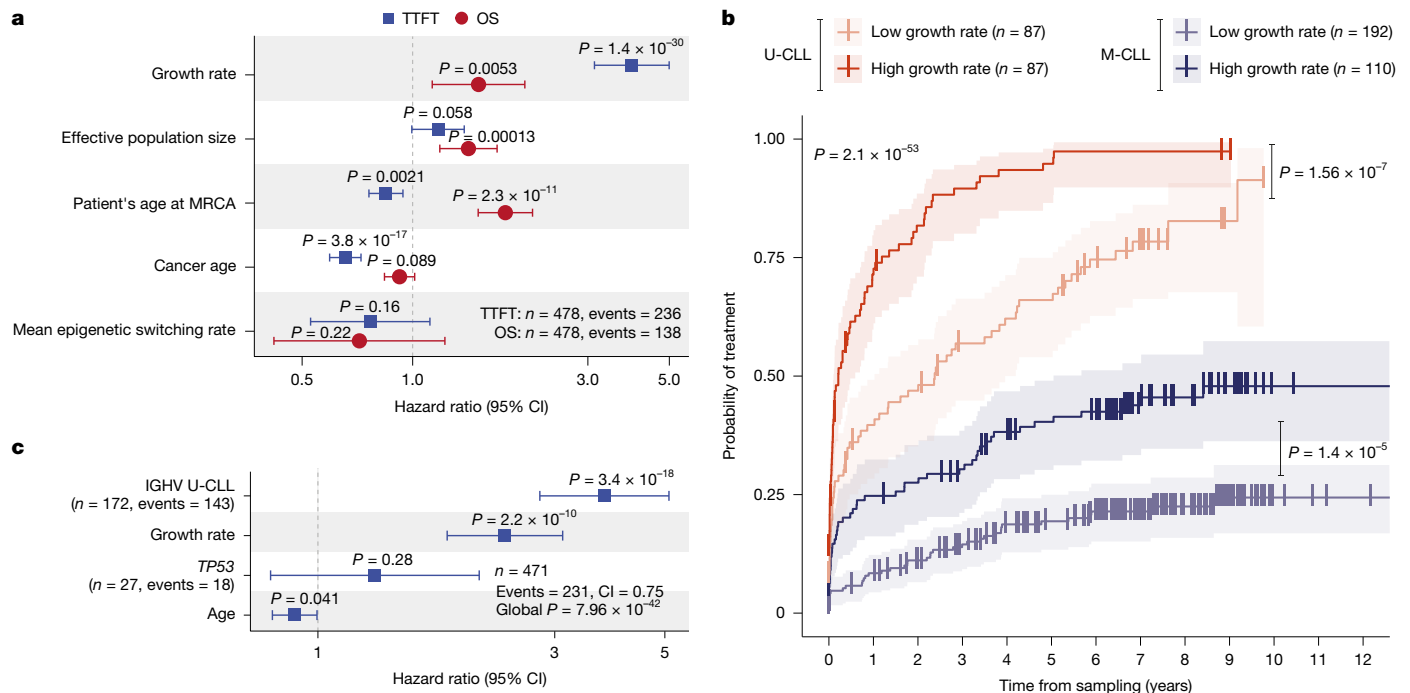


Fig. 5 | The evolutionary history of a tumour is prognostic of clinical outcome.

a, Univariate survival analysis of the TTFT (blue) and overall survival (OS; red) in the discovery CLL cohort for evolutionary variables inferred via EVOFLUX. **b**, Kaplan–Meier curves comparing the TTFT between patients with high versus low inferred cancer growth rates, separated by IGHV mutational status. **c**, Multivariate Cox regression model of the TTFT shows that the cancer

growth rate is significant when controlling for IGHV status, *TP53* alterations and age at sampling. The error bars represent 95% confidence intervals. A log-rank test was used in the Kaplan–Meier curves and Wald tests for Cox models. A Schoenfeld residuals test was used to test proportional hazard assumptions. No multiple comparison adjustments were done.

Richter transformation clone and there was a correspondingly stark signal in fCpG data.

We used EVOFLUX on these longitudinal fCpG methylation data to construct phylogenetic relationships between samples (Methods). In both cases, the phylogenies (Fig. 4a,b) showed that the Richter transformation clone diverged exceptionally early, roughly a decade before the MRCA of the samples containing non-transformed CLL cells (9 and 12 years for cases 12 and 19, respectively). This was consistent with our previous analysis of single-cell RNA sequencing and DNA sequencing that detected Richter transformation cells at low frequencies within the diagnostic CLL sample²⁷, but suggests, remarkably, that the initial Richter transformation divergence occurred well before diagnosis, over 30 years before the clinical presentation of Richter transformation.

We validated the fCpG phylogenetic inferences by comparing them with phylogenetic trees from matched WGS data²⁷ and comparing it with methylation data (Methods; Extended Data Fig. 9a,b). fCpGs exactly recapitulated the WGS tree topology and had highly similar branch lengths unlike other CpG sets (Supplementary Figs. 14–16). Hence, fCpGs are a high-resolution phylogenetic character.

We also performed phylogenetic analysis on B-ALL cases from diagnosis to relapse (Extended Data Fig. 9c–e). All relapse samples formed a separate clade from the initial diagnostic sample, suggesting a major treatment-induced evolutionary bottleneck. In patients with B-ALL with matched cancer–remission samples, we consistently observed that the W-shape (indicating a clonal expansion) was replaced by a unimodal distribution (indicating no clonal expansion) following successful treatment, with similar variance as normal blood (Extended Data Fig. 9f,g). In two patients with matched diagnosis, remission and relapse samples, we found that the unimodal fCpG distribution during remission was replaced with a W-shaped distribution at relapse similar in shape to the diagnostic sample, due to the clonal expansion driving recurrence (Fig. 4c and Extended Data Fig. 9h). Comparing the fCpG distributions between diagnostic versus relapse samples

revealed subclonal evolution through treatment (Fig. 4d and Extended Data Fig. 9i).

Evolutionary history and clinical outcome

To investigate the relationship between the evolutionary history and future clinical trajectory of a cancer, we leveraged a series of 478 CLL cases with well-annotated follow-up data. Using univariate Cox models, we tested the effect of evolutionary parameters on time to first treatment (TTFT), which reflects the natural cancer biology, and then on overall survival, which is a more complex end point as it convolves disease biology with treatment responses.

In univariate analysis, faster growing CLLs had markedly shorter TTFT (Fig. 5a; $P = 1.4 \times 10^{-30}$, hazard ratio (HR) = 3.95) and worse overall survival ($P = 0.0053$, HR = 1.51). The N_e of a cancer did not have a strong effect on TTFT ($P = 0.058$, HR = 1.17), but was associated with shorter overall survival ($P = 1.3 \times 10^{-4}$, HR = 1.41). The patient age at the time of the MRCA of the CLL population was highly correlated with the age of the patient (Supplementary Fig. 17), so unsurprisingly, older patients had worse overall survival ($P = 2.3 \times 10^{-11}$, HR = 1.79). The decrease in risk of progression with cancer age ($P = 3.8 \times 10^{-17}$, HR = 0.65), measured by the time since the MRCA, was probably due to confounding with the growth rate, as these parameters were negatively correlated (Supplementary Fig. 17). The epigenetic switching rate parameters were largely uninformative of prognosis.

As the growth rate was different between U-CLL and M-CLL (Fig. 5b), we analysed its prognostic impact within each group separately. Higher growth rates consistently correlated with shorter TTFT in both subtypes (Fig. 5b; $P = 1.4 \times 10^{-5}$ for M-CLL, $P = 1.56 \times 10^{-7}$ for U-CLL and overall $P = 2.1 \times 10^{-53}$). In a multivariate Cox regression model, growth rate maintained a strong independent prognostic impact as quantitative variable ($P = 2.2 \times 10^{-10}$, HR = 2.28) even when controlling for the IGHV mutational status and *TP53* aberrations, as well as age (Fig. 5c).

Of note, the effect of *TP53* mutations on TTFT appeared mediated by increased growth rate. The cancer N_e was more significantly correlated with overall survival than the growth rate, and this effect was preserved in the U-CLL subtype (Extended Data Fig. 10a; $P = 0.55$ for M-CLL, $P = 9.90 \times 10^{-7}$ for U-CLL and overall $P = 4.72 \times 10^{-9}$) and in the multivariate setting with UGHV status, *TP53* aberrations and age at sampling (Extended Data Fig. 10b; $P = 0.025$, HR = 1.33).

Although the inference of the evolutionary parameters on our initial cohort was wholly blinded to the clinical outcomes, we also validated our findings using a second independent cohort of 209 patients with CLL (135 untreated at sampling)^{28,29} (Supplementary Table 2). These results verified tumour initial growth rate as a predictor of TTFT (Extended Data Fig. 10c–e). Furthermore, the EVOFLUX-derived growth rate was prognostic even when controlling for the contemporary rate of change of lymphocyte counts ($P = 0.018$; Extended Data Fig. 10f).

These results demonstrate that the evolutionary parameters inferred from a cost-effective methylation array could have a direct clinical application by contributing to predict the clinical behaviour of patients with CLL independently from well-established prognostic variables.

Discussion

Our study establishes a computational framework called EVOFLUX, which enables quantitative measurement of the evolutionary history of human malignancies at massive scale using only widely available and low-cost bulk methylation data as input. Evolutionary histories are fundamentally distinct from characterizations of the contemporary phenotype of cancer cells, such as the fraction of proliferating cells. EVOFLUX methodology should also work identically for sequencing-based methods to measure methylation such as bifulfite-based^{47,48} and bisulfite-free approaches^{32,49} (for example, long-read nanopore), which show an excellent correlation with our array data. In theory, these methods allow for assessment of many more fCpGs, increasing inference accuracy. EVOFLUX should also be applicable to tumour-derived cell-free DNA extracted from blood.

Evolutionary histories are strongly associated with disease phenotype and clinical outcomes across lymphoid disease types. We consider this strong evidence that clonal evolution, the fundamental cellular process of disease development, underlies the clinical course of the disease. Consequently, we expect these results to generalize across all cancer types. We note that genome-wide DNA methylation analyses also measure other important biological features of a cancer (for example, molecular subtype¹⁴) that could be combined with EVOFLUX-based inference of evolutionary history to further improve the prognostic value of DNA methylation data.

In summary, we present a cost-effective high-throughput platform for measuring cancer evolutionary dynamics at the population scale in patient samples. These fundamental measurements of the disease biology hold substantial prognostic value and represent an innovative asset in the field of precision oncology.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-025-09374-4>.

- Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
- Merlo, L. M. F., Pepper, J. W., Reid, B. J. & Maley, C. C. Cancer as an evolutionary and ecological process. *Nat. Rev. Cancer* **6**, 924–935 (2006).
- Turajlic, S., Sottoriva, A., Graham, T. & Swanton, C. Resolving genetic heterogeneity in cancer. *Nat. Rev. Genet.* **20**, 404–416 (2019).
- Gabbutt, C. et al. Fluctuating methylation clocks for cell lineage tracing at high temporal resolution in human tissues. *Nat. Biotechnol.* **40**, 720–730 (2022).

- Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).
- Gaiti, F. et al. Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. *Nature* **569**, 576–580 (2019).
- Lee-Six, H. et al. Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473–478 (2018).
- Williams, N. et al. Life histories of myeloproliferative neoplasms inferred from phylogenies. *Nature* **602**, 162–168 (2022).
- Yatabe, Y., Tavaré, S. & Shibata, D. Investigating stem cells in human colon by using methylation patterns. *Proc. Natl Acad. Sci. USA* **98**, 10839–10844 (2001).
- Hong, Y. J., Marjoram, P., Shibata, D. & Siegmund, K. D. Using DNA methylation patterns to infer tumor ancestry. *PLoS ONE* **5**, e12002 (2010).
- Brocks, D. et al. Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. *Cell Rep.* **8**, 798–806 (2014).
- Hao, J. J. et al. Spatial intratumoral heterogeneity and temporal clonal evolution in esophageal squamous cell carcinoma. *Nat. Genet.* **48**, 1500–1507 (2016).
- Siegmund, K. D., Marjoram, P., Woo, Y. J., Tavaré, S. & Shibata, D. Inferring clonal expansion and cancer stem cell dynamics from DNA methylation patterns in colorectal cancers. *Proc. Natl Acad. Sci. USA* **106**, 4828–4833 (2009).
- Duran-Ferrer, M. et al. The proliferative history shapes the DNA methylome of B-cell tumors and predicts clinical outcome. *Nat. Cancer* **1**, 1066–1081 (2020).
- Endicott, J. L., Nolte, P. A., Shen, H. & Laird, P. W. Cell division drives DNA methylation loss in late-replicating domains in primary human cells. *Nat. Commun.* **13**, 6659 (2022).
- de Leval, L. et al. Genomic profiling for clinical decision making in lymphoid neoplasms. *Blood* **140**, 2193–2227 (2022).
- Arber, D. A. et al. International consensus classification of myeloid neoplasms and acute leukemias: integrating morphologic, clinical, and genomic data. *Blood* **140**, 1200–1228 (2022).
- Duran-Ferrer, M. & Martín-Subero, J. I. Epigenomic characterization of lymphoid neoplasms. *Annu. Rev. Pathol.* **19**, 371–396 (2024).
- Gruber, M. et al. Growth dynamics in naturally progressing chronic lymphocytic leukaemia. *Nature* **570**, 474–479 (2019).
- Gutierrez, C. et al. Multifunctional barcoding with ClonMapper enables high-resolution study of clonal dynamics during tumor evolution and treatment. *Nat. Cancer* **2**, 758–772 (2021).
- Kulis, M. et al. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat. Genet.* **47**, 746–756 (2015).
- Nordlund, J. et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. *Genome Biol.* **14**, r105 (2013).
- Reinius, L. E. et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS ONE* **7**, e41361 (2012).
- Lee, S. T. et al. Epigenetic remodeling in B-cell acute lymphoblastic leukemia occurs in two tracks and employs embryonic stem cell-like signatures. *Nucleic Acids Res.* **43**, 2590–2602 (2015).
- Queirós, A. C. et al. Decoding the DNA methylome of mantle cell lymphoma in the light of the entire B cell lineage. *Cancer Cell* **30**, 806–821 (2016).
- Nadeu, F. et al. Genomic and epigenomic insights into the origin, pathogenesis, and clinical behavior of mantle cell lymphoma subtypes. *Blood* **136**, 1419–1432 (2020).
- Nadeu, F. et al. Detection of early seeding of Richter transformation in chronic lymphocytic leukemia. *Nat. Med.* **28**, 1662–1671 (2022).
- Oakes, C. C. et al. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat. Genet.* **48**, 253–264 (2016).
- Dietrich, S. et al. Drug-perturbation-based stratification of blood cancer. *J. Clin. Invest.* **128**, 427–445 (2018).
- Agirre, X. et al. Whole-epigenome analysis in multiple myeloma reveals DNA hypermethylation of B cell-specific enhancers. *Genome Res.* **25**, 478–487 (2015).
- Loyfer, N. et al. A DNA methylation atlas of normal human cell types. *Nature* **613**, 355–364 (2023).
- Clarke, J. et al. Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* **4**, 265–270 (2009).
- Puente, X. S. et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* **526**, 519–524 (2015).
- Seale, K., Horvath, S., Teschendorff, A., Eynon, N. & Voisin, S. Making sense of the ageing methylome. *Nat. Rev. Genet.* **23**, 585–605 (2022).
- Boddicker, N. J. et al. Relationship among three common hematological premalignant conditions. *Leukemia* **37**, 1719–1722 (2023).
- Jaiswal, S. & Ebert, B. L. Clonal hematopoiesis in human aging and disease. *Science* **366**, eaan4673 (2019).
- Rice, S. & Roy, A. MLL-rearranged infant leukaemia: a ‘thorn in the side’ of a remarkable success story. *Biochim. Biophys. Acta Gene Regul. Mech.* **1863**, 194564 (2020).
- Alizadeh, A. A. et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511 (2000).
- Hamblin, T. J., Davis, Z., Gardiner, A., Oscier, D. G. & Stevenson, F. K. Unmutated Ig VH genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* **94**, 1848–1854 (1999).
- Damle, R. N. et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* **94**, 1840–1847 (1999).
- Zenz, T. et al. TP53 mutation and survival in chronic lymphocytic leukemia. *J. Clin. Oncol.* **28**, 4473–4479 (2010).
- Plevova, K. et al. Multiple productive immunoglobulin heavy chain gene rearrangements in chronic lymphocytic leukemia are mostly derived from independent clones. *Haematologica* **99**, 329–338 (2014).
- Caravagna, G. et al. Subclonal reconstruction of tumors by using machine learning and population genetics. *Nat. Genet.* **52**, 898–907 (2020).

44. Williams, M. J. et al. Quantification of subclonal selection in cancer from bulk sequencing data. *Nat. Genet.* **50**, 895–903 (2018).
45. Heide, T. et al. Reply to ‘Neutral tumor evolution?’. *Nat. Genet.* **50**, 1633–1637 (2018).
46. Knisbacher, B. A. et al. Molecular map of chronic lymphocytic leukemia and its impact on outcome. *Nat. Genet.* **54**, 1664–1674 (2022).
47. Leitão, E. et al. Locus-specific DNA methylation analysis by targeted deep bisulfite sequencing. *Methods Mol. Biol.* **1767**, 351–366 (2018).
48. Meissner, A. et al. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877 (2005).
49. Füllgrabe, J. et al. Simultaneous sequencing of genetic and epigenetic bases in DNA. *Nat. Biotechnol.* **41**, 1457–1464 (2023).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

Methods

Assembly and quality control of DNA methylation data

We assembled and processed with a harmonized pipeline¹⁴ (v4.1; see Code availability section) 2,430 bulk sample Illumina methylation array data of normal and neoplastic lymphoid cells from previous publications^{14,21–30}. As healthy control samples, this dataset contained sorted CD19⁺ B cells ($n = 40$), CD3⁺ T cells ($n = 35$), peripheral blood mononuclear cells ($n = 6$) and whole-blood samples ($n = 6$). As tumour samples, we included precursor 797 B-ALLs and 90 T-ALLs at diagnosis, 28 B-ALLs and 2 T-ALLs at relapse, as well as 74 B-ALLs and 12 T-ALLs at complete remission (that is, normal blood); 149 MCLs; 722 CLLs, 55 of its precursor condition MBL and 6 samples from patients with CLL undergoing a DLBCL transformation called Richter transformation; 62 primary DLBCL, not otherwise specified; and 104 multiple myeloma and 16 of its precursor condition monoclonal gammopathy of undetermined significance. In brief, raw idat files were loaded and processed with R (v4.3.1) using the minfi package^{50,51} (v1.46.0) in batches as specified in the column 'SSNOB_NORMALIZATION_BATCH' of Supplementary Table 2. In brief, the data were processed for each batch as follows. First, idats files were loaded into a RGChannelSet object, and minfi quality metrics using the qcReport function were performed, removing samples with unexpected distributions of methylation values (that is, distributions markedly distinct from a bimodal centred around 0 and 1 β -values and/or from the remaining samples) and low signal intensities of internal control probes for each sample, including bisulfite conversions I and II, extension hybridization, hybridization, non-polymorphic, specificities I and II, and target removal probes.

Next, further quality metrics were derived using the function minfiQC on the unnormalized RGChannelSet object. Those samples with median signal intensities of unmethylated and methylated channels of at least 10.5 in log₂ scale were considered as having good signal intensities. Subsequently, detection P values were calculated across all CpGs and samples using the detectionP function for the unnormalized RGChannelSet object. Samples were considered as good if having a mean detection P value across all CpGs of $P \leq 0.01$. On a CpG level, we retained CpGs with a detection $P \leq 1 \times 10^{-16}$ in 90% or more of the samples, which has been shown to improve the quality of downstream analyses^{52,53}. The RGChannelSet object was normalized with the single-sample batch-independent preprocessNoob function with dye bias correction. We next retained only CpGs (excluding CH probes) that did not contain any SNP neither in the interrogated CpGs nor in the probe extension using the dropMethylationLoci and dropLociWithSnps functions with default options (minor allele frequency (MAF) = 0). Further analyses using long-read nanopore data, Illumina array control probes, annotation packages and a data-driven approach were used to ensure the lack of any genetic confounding in the methylation values of the resulting fCpGs (see the next sections).

Furthermore, CpGs with any previous evidence of potential cross-hybridization were excluded⁵⁴ and only CpGs mapping to autosomal chromosomes were subsequently retained for downstream analyses. Finally, to further confirm the accuracy of the filtering criteria, we checked the distribution of normalized methylation values and performed principal component analyses separately for samples passing all quality checks as well as those considered as bad samples. The final DNA methylation matrix contained 2,204 samples and 389,180 CpGs passing all the aforementioned quality controls, and included 2,054 patients (22 technical replicates, 3 synchronic and 125 longitudinal samples from the same patients)⁵⁵ (Supplementary Table 2).

To determine the purity of samples, we used our previously deconvolution strategy to infer tumour cell content by DNA methylation¹⁴, which was used as a consensus purity in all the tumour samples except for DLBCL and multiple myeloma. In these two tumour entities, we have previously identified a DNA methylation signature loss causing

inaccurate tumour purity predictions using DNA methylation data, and therefore we used available genetic or flow cytometry data for DLBCL and multiple myeloma, respectively.

Pipeline to select fluctuating CpGs

We constructed a pipeline to identify fCpGs in lymphoid tumours, based on the following criteria:

- (1) Heterogeneous across different participants with the same disease (by accepting CpG loci with the top 5% of standard deviation of methylation value within a cancer type).
- (2) Equally likely to be methylated or unmethylated (by selecting CpGs with average methylation of approximately 0.5 within a cancer type).
- (3) Unlikely to be associated with specific cell or cancer types. We used an unsupervised Laplacian score feature selection metric⁵⁶ to rank CpG loci by their tendency to preserve the nearest-neighbour graph, and accepted the 5% least-informative CpGs.

Exclusion of genetic confounding on fCpGs

We performed a series of analyses to exclude the potential genetic confounding (germline SNPs and somatic SNVs) on our fCpGs. We first excluded the possibility that common germline SNPs caused methylation heterogeneity at fCpG sites between individuals. We observed very distinct methylation dynamics of array control probes containing SNPs (which had been removed during the initial array processing) versus fCpGs. SNP probes showed the same distribution in all samples (Extended Data Fig. 2c), including longitudinally followed cases (Supplementary Fig. 3), whereas fCpGs only showed a W distribution in cancer samples with ongoing fluctuations over time. Thus, although SNPs reflect the stable genetic identity of the individual, fCpGs reflect the identity of a single cell and its evolving lineage. In addition, we used the packages SNPlocs.Hsapiens.dbSNP155.GRCh38 (v0.99.24) and MafH5.gnomAD.v4.0.GRCh38 (v3.19) to check for any known significant germline or somatic genetic confounding on the resulting 978 fCpGs. We found approximately 60% of fCpGs reported in the gnomAD v4 database (with the array background having approximately 65%), but with a very low MAF (median of 1×10^{-5} and mean of 1×10^{-3}). To exclude the possibility of unknown or very rare genetic confounding, we used the data-driven gaphunting algorithm⁵⁷ available in the minfi R package, which further discarded a possible cancer-specific single-nucleotide variation (SNV) that could confound the methylation values at the 978 identified fCpGs. Finally, Oxford Nanopore long read of a subset of normal and neoplastic samples further validated that fCpGs represent de/methylated cytosines (Extended Data Fig. 2d,e; see next section).

Generation and analyses of long-read nanopore data

For long-read methylation sequencing in CLL and Richter transformation samples, concentration was assessed using the Qubit assay and DNA integrity was analysed either with the Femto Pulse System (Agilent) or the Fragment Analyzer (Agilent). When more than 6 μg of material with good integrity was available, DNA was additionally treated with the Short Fragment Eliminator Kit XS (PacBio) and eluted in EB buffer. Approximately 4 μg of DNA was used for library preparation according to the standard LSK114 kit and protocol from Oxford Nanopore. The time for DNA repair and end-prep was increased up to 30 min at 20 °C and 30 min at 65 °C. Adapter ligation was performed for 1 h at room temperature. All elutions were performed at 37 °C for 1.5 h, and 550–600 ng of DNA was loaded onto a FLO-PRO114M (CLL cells) flow cells. Flow cells were washed (EXP-WSH004) after 1–2 days, if pore count decreased to less than 30%. A total of 1–4 washes were performed for each flow cell. Flow cells were run for 100 (CLL cells) hours in total with the Fast model (MinKNOW 23.11.7, Dorado 7.2.13). The raw data were rebase-called using dorado duplex (v0.5.3) and applying the SUP and modified call to detect 5mC and 5hmC, (model dna_r10.4.1_e8.2_400bps_sup@v4.3.0_5mCG_5hmCG@v1).

In normal B cell samples, 1–3 µg of DNA was used for WGS. Libraries were prepared with the DNA ligation kit LSK110 with no modifications. Libraries were loaded onto a flow cell version FLO-PRO002 (R9.4) and were run for 90–110 h. The basecalling was performed on live mode with the Guppy basecaller (v6.2.7), included in the MinKNOW (v22.08.6), using the SUP model for base modification detection of 5mC and 5hmC (dna_r9.4.1_450bps_modbases_5hmc_5mc_cg_sup.cfg).

In all samples, the generated unmapped BAM files after the basecalling were converted to FASTQ files using the SAMtools fastq -T Mm, MI command. The FASTQ files were then mapped to BAM files using the command minimap2 -ax map-ont -y./GCA_000001405.15_GRCh38_no_alt_analysis_set.fna.mmi. The methylation values were extracted from the BAMs into bedMethyl files using the in-house tool bam2bedmethyl (v0.3.2) and compressed/indexed using bgzip/tabix. Reads from each strand were combined to generate DNA matrices for each CpG and were used for obtaining the methylation values of all fCpGs.

In addition, mini BAM files containing all reads from the 976 fCpGs were generated (in hg38 genome assembly). The reads showed excellent mappability, with a mean of perfect nucleotide matches (NM tag; Levenshtein distance) for all fCpGs across samples of 96.41% (range of 73.31–97.90), and mean mapping quality (MAPQ) of all the reads covering all fCpGs across samples of 59.510 (range of 2–60). Subsequently, long reads were phased using variants called using Clair 3 (v1.0.9, model r941_prom_hac_g360 + g422)⁵⁸ with the Longphase package (v1.7)⁵⁹. The methylation status of each CpG was called using the modcall function within the Longphase package. At fCpGs, only 2.7% of the reads were non-canonical bases (Extended Data Fig. 2d). The variant allele frequency (VAF) of these mutations tended to be low and was negatively correlated with the coverage at that site (Supplementary Fig. 4a). Hence, the majority of these non-canonical base pairs are probably due to errors in nucleotide assignment. There is also no association between the methylation status of different reads and the variants present within a 50-bp window of each fCpG locus (Supplementary Fig. 4b). Hence, assessment of fCpG methylation via bead array was not majorly confounded by miscalled variants. The fCpG methylation patterns seen in the bead array data were replicated in the long-read data (Extended Data Fig. 2e) and the correlation between the fraction methylated measured via bead array and long-read sequencing at fCpGs was excellent (Extended Data Fig. 2e). The same correspondence was observed in WGBS data (Extended Data Fig. 2f).

To assess the intra-sample long-read diversity for each sample, the pairwise Hamming distances were calculated between every read on both haplotypes. The two lists of Hamming distances were concatenated, and the mean calculated as a summary statistic of the read diversity for each sample. One normal B cell sample contained only two reads from one haplotype, and zero from the other, and so was excluded from further analysis.

Analysis of scRRBS data

Previously published single-cell reduced representation bisulfite sequencing (scRRBS) data were obtained⁶ and the fCpG methylation values extracted. Methylation values for normal B cells from 6 donors and CLL cells from 12 patients. There was a high dropout rate, so to extract meaningful patterns we plotted a subset of 40 cells and 20 fCpGs with a high density and overlap of fCpGs across single cells as examples (Supplementary Fig. 5a,b).

To compare the full set of data accounting for the high degree of missing data, we used a metric of heterogeneity at a given fCpG that weights by the number of non-missing fCpGs according to:

$$d_i = \sqrt{\frac{n_i(n_i - 1)}{2}} \sigma(\beta_i)$$

Where n_i is the number of non-NaN values for the i th fCpG, $\frac{n_i(n_i - 1)}{2}$ is the total possible pairwise comparisons between a set of n objects and

$\sigma(\beta_i)$ is the standard deviation across the methylation values of the i th fCpG (Supplementary Fig. 5c).

Characterization and annotation of fCpGs

To characterize the genomic and regulatory context of fCpGs, we used a series of statistical analyses and database annotations. We annotated fCpGs using Illumina manifest and other genomic annotation packages available at Bioconductor including IlluminaHumanMethylation450kanno.ilmn12.hg19 (v0.6.1) and IlluminaHumanMethylationEPICanno.ilm10b2.hg19 (v0.6.0). We additionally used the packages SNPlocs.Hsapiens.dbSNP155.GRCh38 (v0.99.24) and MafH5.gnomAD.v4.0.GRCh38 (v3.19) to check any possible germline or somatic genetic confounding on the resulting 978 fCpGs. We found approximately 60% of fCpGs reported in the gnomAD v4 database (with the array background having approximately 65%), but with a very low MAF (median of 1×10^{-5} and mean of 1×10^{-3}). In addition, we used the Illumina 450k and EPIC array internal SNP probes and showed a dramatically distinct methylation dynamics compared with fCpGs in single-timepoint (Extended Data Fig. 2c) and longitudinal (Supplementary Fig. 3) samples. Finally, the data-driven gaphunting algorithm available in the minfi R package was applied with all the previously published thresholds and cut-offs⁵⁷, which further discarded possible cancer-specific SNV that could confound the methylation values at the 978 identified fCpGs.

We used Chi-squared tests to assess the enrichment of fCpGs in distinct genomic regions or elements. We performed gene-set enrichment analysis on the fCpG-associated genes using gProfiler⁶⁰, specifically focusing on the Gene Ontology biological processes⁶¹ and the Human Protein Atlas⁶². The statistical domain space was limited to genes targeted by at least one CpG in the 389,180 candidate CpG set and significance was determined using the g:SCS algorithm⁶³. Previous chromatin segmentation of normal and neoplastic B cells was used to assess the chromatin-state enrichment of fCpG^{14,64}.

fCpGs were checked for their overlap with previous ‘epigenetic clocks’, including mitotic^{14,65–68}, chronological age^{69–78}, gestational age^{79–83}, biological age and mortality^{84–86} and trait predictors^{87,88}. The package methylCIPHER (<https://github.com/MorganLevineLab/methylCIPHER>) was used to obtain the CpGs for most of the epigenetic clocks. The package methylclock (v1.10.0) was used to calculate all epigenetic clocks but epiCMIT, which was derived as previously described¹⁴.

CLL RNA sequencing data

Previously available RNA sequencing data for 294 patients with CLL were obtained³³ and processed as previously described²⁶. Matched RNA sequencing data and DNA methylation data for the same patients at the same timepoint were available for 224 patients with CLL. Transcript per million counts were used to represent differential gene expression values across genes and samples. We used the gene annotation provided in the R Bioconductor package IlluminaHumanMethylationEPICanno.ilm10b2.hg19 to classify genes associated with fCpGs. Genes targeted by any fCpG were considered as ‘fCpG genes’.

In each methylation sample, the 978 fCpGs were discretized as homozygous demethylated, heterozygous methylated or homozygous methylated (coded as [0,1,2], respectively). This was done by separately fitting a β -mixture model with three components to each sample using Stan⁸⁹ and extracting the component mixture probability. The gene expression value for genes classified as having and fCpG with 0, 1 or 2 alleles methylated were plotted as previously described.

DNA methylation data from normal blood samples

External DNA methylation data were downloaded from the Gene Expression Omnibus database using the GEOquery R package (v2.72.0). For sorted immune cells, these include GSE137594 and GSE184269. For whole-blood samples, these include GSE72773, GSE55763, GSE40279

Article

and GSE36054. Data were analysed with the normalization procedure used in each study together with the metadata provided. Mean and standard deviation for fCpGs were calculated with fCpGs present in the provided normalized matrices.

A stochastic model of fCpGs in a growing population

We built a generative computational model of how the patterns of fCpGs vary over time (t) according to the evolutionary history of a cancer. Initially, our model focused on neutral evolution, before expanding to non-neutral modes of tumour evolution below. For the full explanation of the model, see the Supplementary Information.

Our model was parameterized in terms of the age of the patient at which the MRCA emerged (τ), the exponential growth rate of the cancer (θ) and the epigenetic switching rates of the fCpGs (μ , ν , γ and ζ). The model was partitioned into two phases: before and after the emergence of the MRCA. At time $t = 0$, the fCpGs were assumed to be equally likely to be homozygously methylated or demethylated. The fCpG status of the MRCA at time $t = \tau$ was calculated by applying matrix exponentiation.

The second phase of the model consisted of a discrete time Markov process. The effective population size of the growing cancer was modelled as growing according to a deterministic exponential growth equation, $N_e = e^{\theta(T-\tau)}$. Each fCpG was considered independently; at each time step, $t \rightarrow t + \delta t$, the number of homozygous-methylated (m), heterozygous-methylated (k) and homozygous-demethylated cells (w) at a specific fCpG was updated according to the epigenetic switching rates.

At the time of sample, T , the fraction methylation of each simulated fCpG was calculated by summing the number of methylated alleles and normalizing by the total number of alleles in the population:

$$\beta_c = \frac{k + 2m}{2N_e}$$

We further accounted for contaminating normal cells and the technical noise introduced by the methylation bead array. The methylation of the contaminated samples was assumed to be an average of the cancer methylation, $\beta_c(t)$, weighted by the tumour purity ρ , and the average of the normal population, β_n , weighted by $1 - \rho$. Following our previous work, the bead array was assumed to saturate at extreme methylation values, shifting the minimum and maximum methylation by δ and ε , respectively⁴. The noise of the bead array was assumed to be β -distributed, with precision parameter κ .

Non-neutral models of tumour evolution

Alongside our model of neutral exponentially growing cancer populations, we devised two alternative models of cancer growth:

- (1) A subclonal selection model in which a single cell within the cancer develops a selective advantage and begins to grow at an increased growth rate.
- (2) An independent clonal origins model, in which a patient has developed two distinct cancers concurrently.

For the subclonal selection model, we replaced the growth rate (θ) and the time of the MRCA (τ) with the growth rates and time of the MRCA of the initial, slower-growing population (θ_1 and τ_1 , respectively), and that of the more recently emerging, faster-growing population (θ_2 and τ_2), constraining $\tau_1 < \tau_2$ and $\theta_1 < \theta_2$ (Extended Data Fig. 8a). We assumed that the initial cancer population began exponentially growing at τ_1 as above, but at time $t = \tau_2$, we selected a single cell with a set of fCpG states drawn according to the cancer population and allowed this second population to grow concurrently with a growth rate θ_2 .

The independent-cancer model followed the same scheme as the nested subclonal selection model, except the methylation status of the emerging cancer was that of an independent cell that experienced random fluctuations between $t = 0$ and $t = \tau_2$.

If we let the number of cells in the less fit subclone in each methylation state be $\{m_1, k_1, w_1\}$ and in the fitter subclone be $\{m_2, k_2, w_2\}$, following the convention above, then in both cases the measured methylation patterns at the time of sample are:

$$\beta_c(T) = \frac{k_1(T) + 2m_1(T) + k_2(T) + 2m_2(T)}{2N_e(T)}$$

Where $N_e(T) = e^{\theta_1(T-\tau_1)} + e^{\theta_2(T-\tau_2)}$.

Adaption of simulations to a longitudinal setting

We modified the simulations of how the fCpG methylation distribution changes over time to allow for multiple sequential sample collections. These simulations allow for neutral, independent clones, a single subclonal expansion or two subclonal expansions, which can either be nested or emerge from the clonal trunk in parallel. This required pre-specification of sampling times, along with the emergence times of any subclones or independent clones, which we collected to form a set of 'landmark times'. The discrete time steps of the simulation were split into phases between the landmark times, which evolved according to the discrete time Markov process outlined above. At each sampling time, the fCpG methylation fraction was calculated as above and stored as a column in the output matrix.

Prior functions

For each methylation array blood sample, we had matched age (T) and purity (ρ) information. Hence, the parameters to be inferred are the growth rate (θ), the age of the patient when the MRCA emerged (τ), the epigenetic switching rates (μ , ν , γ , ζ), the average fraction methylated of contaminating normal cells (β_n), the β -offsets from 0 and 1 due to the background noise on the methylation array (δ and ε , respectively) and the precision of the β -distributed noise (κ).

These parameters are constrained either to be positive ($\theta, \mu, \nu, \gamma, \zeta, \kappa > 0$) or to lie within a specified range ($0 < \tau/T, \delta, \varepsilon < 1$), which we achieved using appropriate prior distributions. To better allow for priors to be set on a biologically meaningful scale, the priors for the log-normal distribution were set in terms of the real scale mean and standard deviation, rather than the standard log-scale. To reduce correlations in the posterior and make sampling more efficient, the variables ν and ζ were normalized by μ and γ , respectively.

The priors are as follows:

θ - lognormal(3,2)

$\frac{\tau}{T}$ - beta(2,2)

μ - halfnormal(0, 0.05)

γ - halfnormal(0, 0.05)

$\frac{\nu}{\mu}$ - lognormal(1, 0.7)

$\frac{\zeta}{\gamma}$ - lognormal(1, 0.7)

β_n - beta(2,2)

δ - beta(5, 95)

ε - beta(95, 5)

κ - halfnormal(100, 30)

When fitting non-neutral models of tumour growth, the inference was parameterized in terms of the relative growth of the fitter subclone, $\tilde{\theta}_2 = \frac{\theta_2}{\theta_1}$, and the fraction of the population consisting of the fitter subclone, $f = \frac{e^{\theta_2(t-\tau_2)}}{e^{\theta_1(t-\tau_1)} + e^{\theta_2(t-\tau_2)}}$. The age at which the second clone emerges is then:

$$\tau_2 = T - \frac{(T - \tau_1)\theta_1}{\theta_2} - \frac{\text{logit}(f)}{\theta_2}$$

This parameterization induces less correlation in the resulting posterior, which greatly improves the sampling efficiency. The priors on these additional parameters are:

$$\frac{\tau_1}{T} \sim \text{beta}(2, 2)$$

$$\tilde{\theta}_2 \sim \text{lognormal}(1, 0.7)$$

$$f \sim \text{beta}(2, 2)$$

All the other priors were the same as in the neutral case.

Bayesian inference

We developed a stochastic estimator of the log-likelihood function at a given set of parameters by simulating the fCpG methylation distribution a large number of times, correcting for the bias inherent with using a finite number of simulations and penalizing the log-likelihood for extreme values of the N_e (see Supplementary Information for details).

The standard Bayesian algorithms developed to infer the posterior for a given set of data (for example, Markov chain Monte Carlo (MCMC), nested sampling) are typically used when the log-likelihood is analytically tractable and can be calculated exactly. It has been shown that, as long as the stochastic approximation of the log-likelihood is unbiased, MCMC methods can obtain an exact Bayesian inference of the true posterior, as in pseudo-marginal Metropolis–Hastings⁹⁰.

Here we used a nested sampling approach using the dynesty package^{91–93}. Unlike pseudo-marginal Metropolis–Hastings, nested sampling is able to efficiently explore multimodal posterior landscapes (which can occur under the subclonal and independent cancer models).

Model selection for the mode of tumour evolution

We used an expected log pointwise predictive density⁹⁴ approach to compare our competing models of evolution for each sample using the arviz Python package⁹⁵, which uses PSIS-LOO-CV to compare the out-of-sample prediction accuracy between models while naturally penalizing more complex models. This required the log-likelihood per data point and the posterior predictive for every point in the posterior. The weights of the respective models were calculated using pseudo-Bayesian model averaging using Akaike-type weighting, stabilized using the Bayesian bootstrap⁹⁶.

CLL and Richter transformation genomic analyses

Previous mutated annotation files from WES⁴⁶ and WGS²⁷ data were used to further validate our distinct EVOFLUX evolutionary modes (that is, neutral, subclonal and independent) and Richter transformation phylogenies.

Subclonal deconvolution of WES and WGS data

To detect subclones in bulk WES and WGS data, we used MOBSTER⁴³, which fits the VAF spectrum with a mixture model containing a Pareto distribution to account for the neutral tail⁹⁷ and a variable number of β -distributions to account for the clonal and subclonal peaks.

We ran MOBSTER using the default parameters, except using a minimum 5% VAF threshold and lowering the minimum number of mutations to compose a cluster to five in WES samples due to the low number of

mutations. We then manually quality controlled all 377 WES samples and 10 WGS, tuning the fitting parameters to better represent the data (for instance, when the clonal peak had been called at a low frequency despite the median tumour purity being 95%).

Phylogenetic inference of longitudinal methylation data

A novel Bayesian phylogenetic method was used to reconstruct the evolutionary relationships and the time to MRCA of longitudinal samples from the same patients. This was carried out in the BEAST (v1.8.4) framework^{98,99} using custom models implemented in PISCA¹⁰⁰ (v1.1; available from <https://github.com/adamallo/PISCA>).

EVOFLUX provided an estimate of the age of the patient when the MRCA of each bulk sample emerged. To estimate the methylation status of each fCpG at the MRCA of the sample in each of our longitudinal samples, we discretized the fCpGs as described above (see the section ‘CLL RNA sequencing data’).

We implemented a four-parameter biallelic binary substitution model analogous to the pre-growth EVOFLUX model in PISCA. This plugin contains all the required statistical machinery to use this model for somatic phylogenetic estimation. The biallelic binary substitution model has three relative rate parameters: (1) heterozygous methylation $\tilde{\nu}$, (2) homozygous demethylation $\tilde{\gamma}$, and (3) heterozygous demethylation $\tilde{\zeta}$, where homozygous methylation $\tilde{\mu}$ was normalized to 1. For all relative transition rate parameters, a log-normal prior with mean of 1 and standard deviation of 0.6 was used, with a half-normal prior with mean of 0 and standard deviation of 0.13 for the molecular clock rate, using a strict clock model for the rate of evolution across the tree. Two demographic tree models, constant population size¹⁰¹ and exponential growth¹⁰², were compared by marginal likelihood estimation using path-sampling¹⁰³ and a constant population model was deemed more appropriate.

MCMC chains were run for 100 million generations sampled every 100,000 generations and convergence was assessed using Tracer (v1.7)¹⁰⁴, ensuring effective sample sizes (ESS) greater than 500 for all parameters. Maximum clade credibility trees were then made using 10% burn-in and medium node heights. The resulting trees were plotted using ggtree¹⁰⁵.

Phylogenetic inference of SNVs from WGS data

Each bulk sample is represented by a set of clonal mutations found during the deconvolution of WGS data (see above). Where a mutation was deemed absent in the clonal peak, the reference nucleotide was used. Mutational signature assignment¹⁰⁶ was used to select mutations in the clock-like SBS1 channel¹⁰⁷. BEAST (v1.10)¹⁰⁸ was then used with the simple binary substitution model (as SBS1 effectively represents just C-to-T substitutions), a strict clock model, a constant population size prior¹⁰¹ and a flat prior on the age of MRCA (from zero to earliest patient sample), with ancestral state estimation at the root. Chains were run and ESS values assessed as described above. The distances between the ancestral state of the root at each MCMC state and the clock rate were used to calculate the expected evolution distance between the root and the known germline. This was used to inform the length of the branch between germline (at birth) and the MRCA of the samples.

Survival analysis

Clinical analyses were performed in CLL for TTFT and overall survival from the time of sampling. Tumour growth rate (θ), effective population size (N_e) and epigenetic switching rates were analysed as continuous variables in univariate Cox regression models for both TTFT and overall survival. The effect size of HRs for each evolutionary variable were analysed considering different scaling factors. In particular, the growth rate was analysed assuming exponential growth (that is, for $\theta = 1$, the population is $e = 2.71$ times bigger per year), the N_e was considered per million cells, and the cancer age or time from the MRCA was analysed for each 10 years. Individual switching rate

Article

parameters (μ , v , γ and ζ) were largely uninformative of prognosis and were summarized into a mean epigenetic switching rate, which was scaled by a factor of 100. In addition, growth rate and effective population were analysed as continuous variables in multivariate Cox regression models together with *TP53* aberrations (considering mutations and deletions together), IGHV gene mutational status and the age of patients at sampling. Kaplan–Meier curves were generated for low and high growth rates and effective population size within IGHV subtypes using maximally selected log-rank statistic using the *maxstats* package (v0.7-25). *P* values from Kaplan–Meier curves were derived using the log-rank statistic. *Survival* (v3.5-7), *survminer* (v0.4.9) and *ggsurvfit* (v0.3.1) packages were used under R (v4.3.1). Plots were generated using *ggplot2* (v3.5.2).

Estimating the rate of change in lymphocyte counts

Historical records of the absolute number of lymphocytes in blood obtained via haemocytometer were collected for patients with CLL over the whole disease course (that is, an approximate of the number of malignant CLL cells in blood). In 231 patients with CLL, we could obtain at least 10 sample timepoints (that is, at least 10 medical appointments, median $n = 27$ and mean $n = 34$) before the first treatment, allowing us to track the natural history of the disease before treatment intervention for the tumour (Supplementary Fig. 10). We fitted a linear model to all 231 cases and obtained the slope of the observed log number of lymphocytes (that is, the coefficient of the univariate linear model) and compared it with growth rate estimates derived from EVOFLUX.

Statistical analysis

Statistical tests performed throughout the study were performed as two-sided. Appropriate multiple test correction, such as the Holm–Sidak correction, is noted when applied.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

No new methylation bead array data were generated in the course of this study. The harmonized and filtered methylation matrix has been deposited to Zenodo⁵⁵ (<https://doi.org/10.5281/zenodo.15479736>). Previously published DNA methylation data reanalysed in this study can be found under the accession codes: EGAS00001001196 for B cells; GSE56602, GSE49032, GSE76585 and GSE69229 for ALL; EGAS00001001637 and EGAS00001004165 for MCL; EGAD00010000871, EGAD00010000948 and EGAD00010001975 for CLL; EGAS00001000841 for multiple myeloma; and EGAD00010001974 for DLBCL. External DNA methylation data for sorted immune cells can be found under the accession codes GSE137594 and GSE184269. For whole-blood samples, the accession codes are GSE72773, GSE55763, GSE40279 and GSE36054. CLL gene expression data are available under the accession codes EGAS00001000374 and EGAS00001001306. Chromatin immunoprecipitation followed by sequencing datasets are available from Blueprint (<https://www.blueprint-epigenome.eu/>) under the accession code EGAS00001000326. Matched WES and WGS data are available under the accession codes EGAS00000000092 and EGAD00001008954, respectively, under controlled access. Pathway analysis was run using the Gene Ontology: Biological Processes release 2023-03-06 and the Human Protein Atlas (v10.0) databases. Matched Oxford Nanopore long-read data were generated for six normal B cells and samples and two CLL–Richter transformation sample pairs. Long-read data are available at the European Genome–Phenome Archive repository under the accession code EGAS500000001192. Source data are provided with this paper.

Code availability

The codes for EVOFLUX used to infer the evolutionary history of cancer samples from methylation array data (<https://github.com/CalumGabbutt/evoflux>), to curate DNA methylation data and to perform clinical and additional bioinformatic analyses (<https://github.com/Duran-FerrerM/evoflux>) and for the phylogenetic method (<https://github.com/adamallo/PISCA>), are available on GitHub.

50. Aryee, M. J. et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
51. Fortin, J.-P., Triche, T. J. Jr & Hansen, K. D. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* **33**, 558–560 (2017).
52. Lehne, B. et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol.* **16**, 37 (2015).
53. Zhou, W., Triche, T. J., Laird, P. W. & Shen, H. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res.* **46**, e123 (2018).
54. Chen, Y. A. et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* **8**, 203–209 (2013).
55. Duran-Ferrer, M., Gabbutt, C., Martin-Subero, J. I. & Graham, T. Harmonised methylation array matrix related to the article Fluctuating DNA methylation tracks cancer evolution at clinical scale. *Zenodo* <https://doi.org/10.5281/ZENODO.15479737> (2025).
56. He, X., Cai, D. & Niyogi, P. Laplacian score for feature selection. In *Advances in Neural Information Processing Systems 18* (eds. Weiss, Y., Schölkopf, B. & Platt, J.) (MIT Press, 2005).
57. Andrews, S. V., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D. & Fallin, M. D. “Gap hunting” to characterize clustered probe signals in Illumina methylation array data. *Epigenetics Chromatin* **9**, 56 (2016).
58. Zheng, Z. et al. Symphonizing pileup and full-alignment for deep learning-based long-read variant calling. *Nat. Comput. Sci.* **2**, 797–803 (2022).
59. Lin, J. H., Chen, L. C., Yu, S. C. & Huang, Y. T. LongPhase: an ultra-fast chromosome-scale phasing algorithm for small and large variants. *Bioinformatics* **38**, 1816–1822 (2022).
60. Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).
61. Ashburner, M. et al. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
62. Uhlén, M. et al. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
63. Reimand, J., Kull, M., Peterson, H., Hansen, J. & Vilo, J. g:Profiler — a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.* **35**, W193–W200 (2007).
64. Beekman, R. et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat. Med.* **24**, 868–880 (2018).
65. Yang, Z. et al. Correlation of an epigenetic mitotic clock with cancer risk. *Genome Biol.* **17**, 205 (2016).
66. Teschendorff, A. E. A comparison of epigenetic mitotic-like clocks for cancer risk prediction. *Genome Med.* **12**, 56 (2020).
67. Youn, A. & Wang, S. The MiAge Calculator: a DNA methylation-based mitotic age calculator of human tissue types. *Epigenetics* **13**, 192–206 (2018).
68. Zhou, W. et al. DNA methylation loss in late-replicating domains is linked to mitotic cell division. *Nat. Genet.* **50**, 591–602 (2018).
69. Bocklandt, S. et al. Epigenetic predictor of age. *PLoS ONE* **6**, e14821 (2011).
70. Garagnani, P. et al. Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell* **11**, 1132–1134 (2012).
71. Hannum, G. et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367 (2013).
72. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
73. Lin, Q. et al. DNA methylation levels at individual age-associated CpG sites can be indicative for life expectancy. *Aging* **8**, 394–401 (2016).
74. Vidal-Bralo, L., Lopez-Golan, Y. & Gonzalez, A. Simplified assay for epigenetic age estimation in whole blood of adults. *Front. Genet.* **7**, 209192 (2016).
75. Weidner, C. I. et al. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol.* **15**, R24 (2014).
76. Zhang, Q. et al. Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. *Genome Med.* **11**, 54 (2019).
77. Horvath, S. et al. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford progeria syndrome and ex vivo studies. *Aging* **10**, 1758–1775 (2018).
78. Shireby, G. L. et al. Recalibrating the epigenetic clock: implications for assessing biological age in the human cortex. *Brain* **143**, 3763–3775 (2020).
79. Bohlin, J. et al. Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biol.* **17**, 207 (2016).
80. Knight, A. K. et al. An epigenetic clock for gestational age at birth based on blood methylation data. *Genome Biol.* **17**, 206 (2016).
81. Lee, Y. et al. Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels. *Aging* **11**, 4238–4253 (2019).
82. Mayne, B. T. et al. Accelerated placental aging in early onset preeclampsia pregnancies identified by DNA methylation. *Epigenomics* **9**, 279–289 (2017).
83. McEwen, L. M. et al. The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proc. Natl Acad. Sci. USA* **117**, 23329–23335 (2020).
84. Belsky, D. W. et al. DunedinPACE, a DNA methylation biomarker of the pace of aging. *eLife* **11**, e73420 (2022).

85. Levine, M. E. et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
86. Lu, A. T. et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**, 303–327 (2019).
87. McCartney, D. L. et al. Epigenetic prediction of complex traits and death. *Genome Biol.* **19**, 136 (2018).
88. Liang, X. et al. DNA methylation signature on phosphatidylethanol, not on self-reported alcohol consumption, predicts hazardous alcohol consumption in two distinct populations. *Mol. Psychiatry* **26**, 2238–2253 (2020).
89. Carpenter, B. et al. Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1 (2017).
90. Andrieu, C. & Roberts, G. O. The pseudo-marginal approach for efficient Monte Carlo computations. *Ann. Statist.* <https://doi.org/10.1214/07-AOS574> (2009).
91. Skilling, J. Nested sampling. In *AIP Conference Proceedings* 735 (eds Fischer, R. et al.) 395–405 (AIP Publishing, 2004).
92. Skilling, J. Nested sampling for general Bayesian computation. *Bayesian Anal.* **1**, 833–860 (2006).
93. Speagle, J. S. dynesty: A dynamic nested sampling package for estimating Bayesian posteriors and evidences. *Mon. Not. R. Astron. Soc.* **493**, 3132–3158 (2019).
94. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27**, 1413–1432 (2015).
95. Kumar, R., Carroll, C., Hartikainen, A. & Martin, O. ArviZ a unified library for exploratory analysis of Bayesian models in Python. *J. Open Source Softw.* **4**, 1143 (2019).
96. Yao, Y., Vehtari, A., Simpson, D. & Gelman, A. Using stacking to average Bayesian predictive distributions. *Bayesian Anal.* **13**, 917–1007 (2017).
97. Williams, M. J., Werner, B., Barnes, C. P., Graham, T. A. & Sottoriva, A. Identification of neutral tumor evolution across cancer types. *Nat. Genet.* **48**, 238–244 (2016).
98. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).
99. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).
100. Martínez, P. et al. Evolution of Barrett's esophagus through space and time at single-crypt and whole-biopsy levels. *Nat. Commun.* **9**, 794 (2018).
101. Kingman, J. F. C. The coalescent. *Stoch. Process Appl.* **13**, 235–248 (1982).
102. Griffiths, R. C. & Tavaré, S. Sampling theory for neutral alleles in a varying environment. *Phil. Trans. R. Soc. Lond. B* **344**, 403–410 (1994).
103. Baele, G. et al. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Biol. Evol.* **29**, 2157–2167 (2012).
104. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).
105. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
106. Diaz-Gay, M. et al. Assigning mutational signatures to individual samples and individual somatic mutations with SigProfilerAssignment. *Bioinformatics* **39**, 12 (2023).
107. Tate, J. G. et al. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).
108. Suchard, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018).

Acknowledgements This research was primarily funded by an Accelerator award Cancer Research UK/AIRC/AECC joint funder-partnership (to J.F., J.I.M.-S. and T.A.G.) and the US National Institutes of Health National Cancer Institute (U54 CA217376 to D.S. and T.A.G., R01 CA140657 to D.M.). Additional funding was provided by Cancer Research UK via direct grants

(A19771 and DRCNPG-May21_100001 to T.A.G.), through the CRUK Convergence Science Centre (CTRQQR-2021\100009 to T.A.G.) and via an early detection primer award (EDDPMA-May23/100059 to T.A.G. and C.G.). Further funding was provided by the European Research Council under the European Union's Horizon 2020 Research and Innovation Program (810287, BCLLatlas, to I.G.G., E.C. and J.I.M.-S.), La Caixa Foundation (CLLEvolution LCF/PR/HR17/52150017 (HR17-00221LCF) and CLLSYSTEMS LCF/PR/HR22/52420015 (HR22-00172) Health Research 2017 and 2022 Programs to E.C.), Generalitat de Catalunya Suport Grups de Recerca AGAUR (2021-SGR-01343 to J.I.M.-S. and 2021-SGR-01172 to E.C.). C.G. was further supported by the BBSRC London Interdisciplinary Doctoral Programme (1902605), Thorton Foundation funding to the Institute of Cancer Research and the Eric and Wendy Schmidt Foundation. M.D.-F. was supported by a postdoctoral fellowship of the AECC Scientific Foundation. J.N. and O.K. were supported by the Swedish Research Council (2019-01976), the Swedish Childhood Cancer Fund (PR2019-0046/PR2022-0082) and the Swedish Cancer Society (CAN2022-2395). This research utilized the Cancer Research UK City of London High Performance Computing facility, and was partially developed at the Centro Esther Koplowitz. We acknowledge the SciLifeLab National Genomics Infrastructure, SNP&SEQ Technology Platform, funded by the Swedish Research Council and the Knut and Alice Wallenberg Foundation, for assistance with DNA methylation analyses. We thank D. Landau and F. Gaiti for supplying the single-cell reduced representation bisulfite sequencing data for analysis.

Author contributions T.A.G. and J.I.M.-S. conceived, funded and supervised the study. C.G. and T.A.G. conceived and designed the modelling scheme and the Bayesian inference framework, with support from C.P.B.; C.G. implemented and ran this. M.D.-F. processed and curated DNA methylation and patient metadata, and performed the survival analysis. M.D.-F. and C.G. conceived the phylogenetic analysis; C.G., H.E.G. and D.M. designed and ran them. C.G., M.D.-F., H.E.G., J.H. and F.N. performed the bioinformatic analysis. M.M., S.H., E.R. and I.G.G. performed the experiments and primary analyses of the long-read nanopore sequencing data. M.D.-F., F.N., N.V., O.K., J.N., T.Z., E.C. and A.L.-G. provided data and contributed to sample biological and/or clinical annotation. M.D.-F., N.V., E.C., A.L.-G., J.F., D.S. and J.I.M.-S. provided clinical insight. This study builds on the earlier conception of the fluctuating methylation phenomenon by D.S. C.G., M.D.-F., J.I.M.-S. and T.A.G. wrote the manuscript. All authors approved the final version of the manuscript.

Competing interests T.A.G., J.I.M.-S., M.D.-F. and C.G., the Institute of Cancer Research, Fundació de Recerca Clínic Barcelona-Institut d'Investigacions August Pi i Sunyer and Institutació Catalana de Recerca i Estudis Avançats have filed for a patent on a method to measure the evolutionary dynamics in cancers using DNA methylation (GB2317139.0). T.A.G. is named as a co-inventor on adjacent patent applications that describe a method for TCR sequencing (GB2305655.9), and a method to infer drug resistance mechanisms from barcoding data (GB2501439.0). T.A.G. has received an honoraria from Genentech and consultancy fees from DAIa therapeutics. F.N. received honoraria from AbbVie, AstraZeneca, Janssen and SOPHiA GENETICS for speaking at educational activities, research funding from Gilead, and licensed the use of the protected IgCaller algorithm for SOPHiA GENETICS and Diagnóstica Longwood. All other authors declare no competing interests.

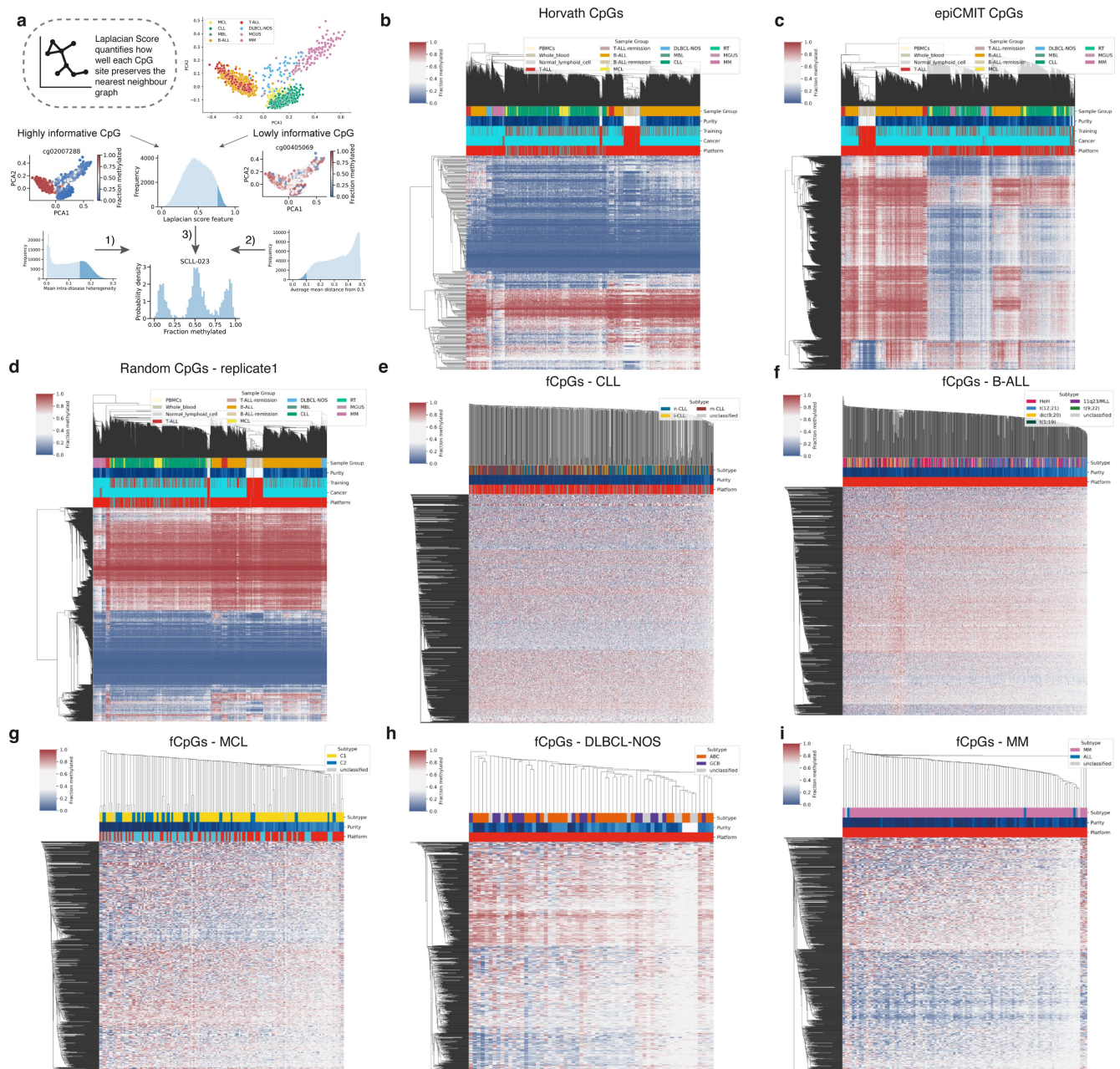
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-025-09374-4>.

Correspondence and requests for materials should be addressed to Calum Gabbutt, José I. Martin-Subero or Trevor A. Graham.

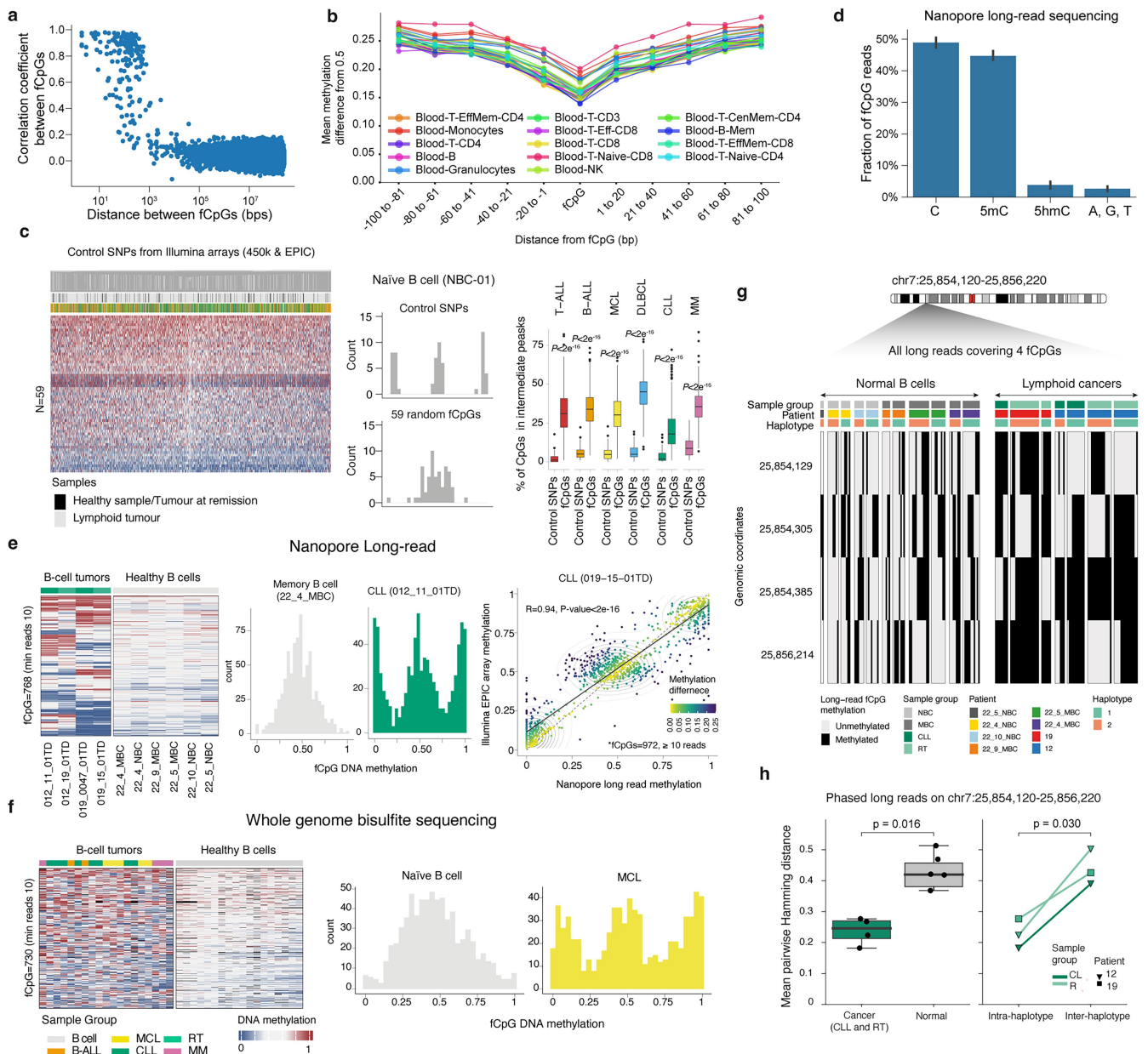
Peer review information *Nature* thanks Pavlo Lutsik, George Vassiliou and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Comparison of the methylation patterns of fCpGs vs epigenetic clocks. a: fCpGs were selected by combining three filters: (1) CpGs with low intra-disease heterogeneity were removed as sites that were likely not fluctuating, (2) CpGs with a mean methylation far from 0.5 were removed as likely belonging to CpG sites with skewed methylation/demethylation rates, and (3i) CpGs which preserved the nearest-neighbour graph (quantified using the Laplacian Score) were removed as sites that were likely under selection/strict regulation. **b-d:** Heatmaps of **b:** Horvath's CpGs⁷², **c:** epiCMT CpGs¹⁴,

and **d:** random subset of 1,000 CpGs from the candidate CpGs on the bead array. **e-i:** Heatmaps of our set of 978 pan-lymphoid fCpGs with disease-specific subtypes annotated in: **e** chronic lymphocytic leukaemia (CLL), **f** B-cell acute lymphoblastic leukaemia (B-ALL), **g** mantle cell lymphoma (MCL), **h** diffuse large B-cell lymphoma - not otherwise specified (DLBCL-NOS), and **i** multiple myeloma (MM). Hierarchical clustering with average linkage and a Euclidean metric was used in heatmaps.



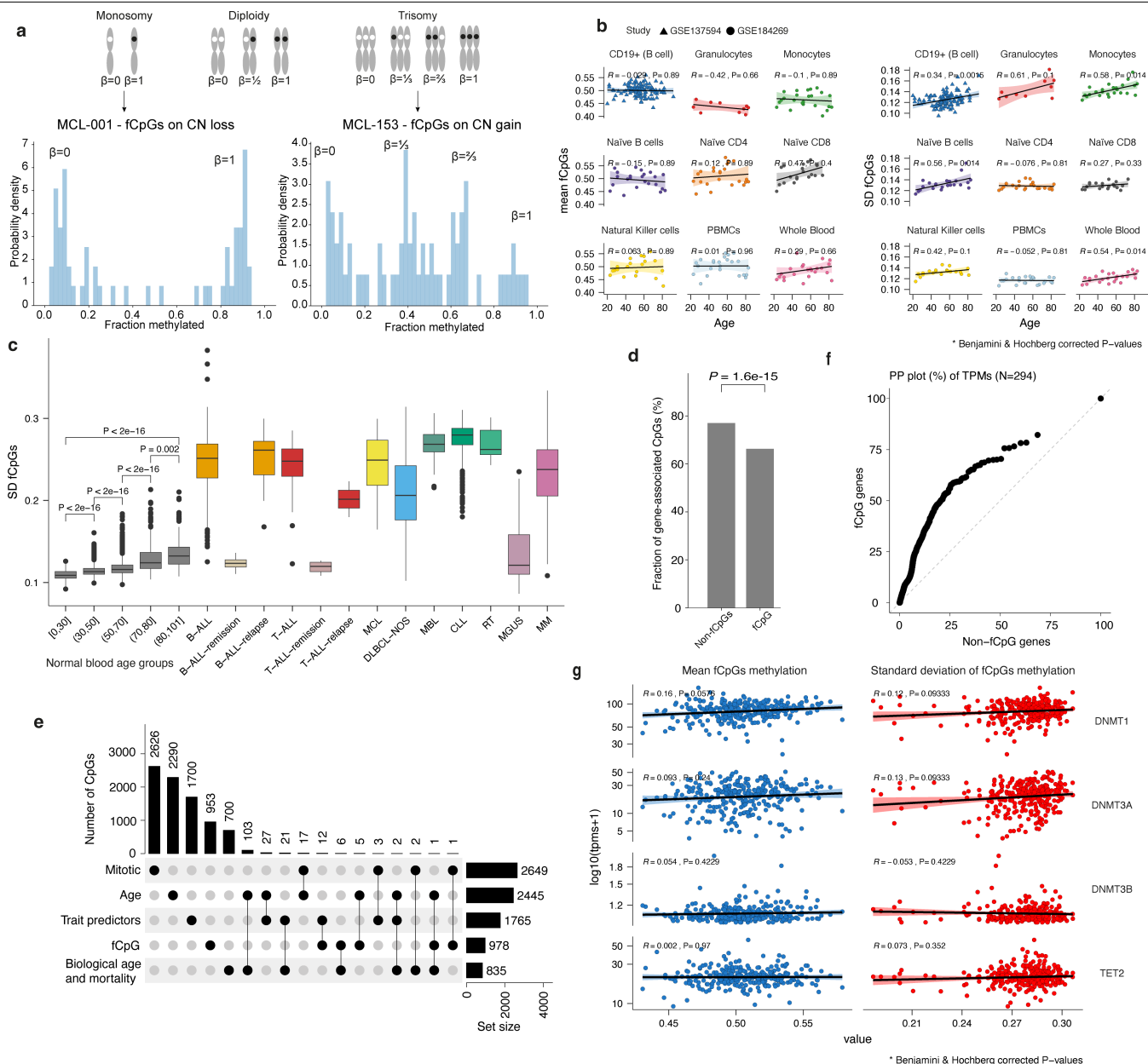
Extended Data Fig. 2 | Validation of fCpGs as an evolving barcode.

a: Scatterplot showing the pairwise correlation coefficient in the fCpG methylation values of lymphoid cancer samples with other fCpG methylation values on the same chromosome, and the genomic distance between them.

b: The average absolute difference from 0.5 methylation value of fCpGs with their local neighbourhood in whole genome bisulphite sequencing (WGBS) data in different B- and T-cell populations³¹. The fraction of hypo- and hypermethylated CpG loci increases as a function of distance from the reference fCpG locus.

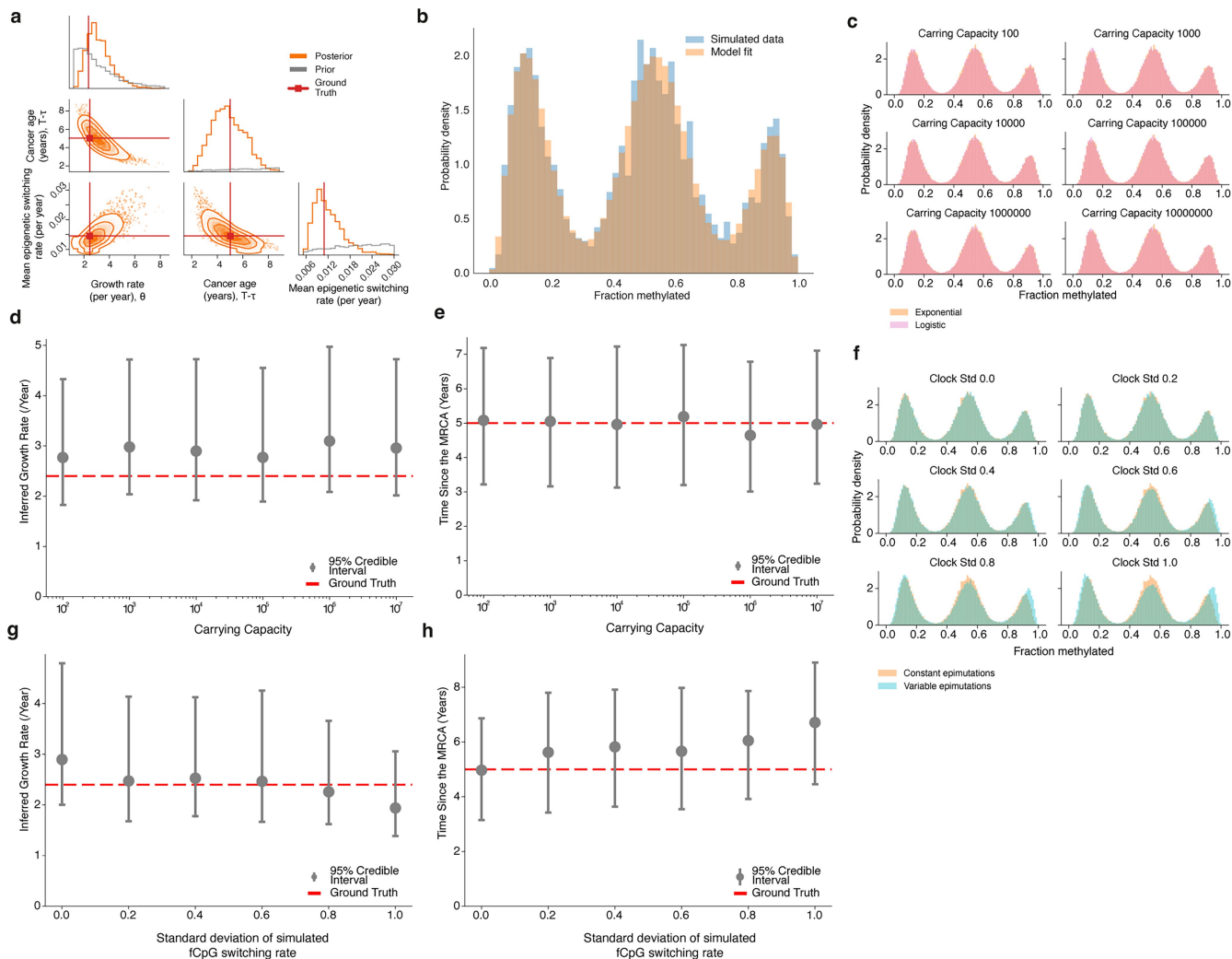
c: (left) Heatmap of control single nucleotide polymorphism (SNP) probes from Illumina arrays, showing distinct methylation dynamics compared to fCpGs, with normal and remission samples intermingled with tumours. (middle) A distinct methylation distribution of control SNP probes and the same number of random fCpGs is shown for one healthy naïve B cell sample. (right) The percentage of CpGs in intermediate peaks (i.e., ≥ 0.2 and ≤ 0.4 and ≥ 0.6 and ≤ 0.8) is notably higher in fCpGs compared to control

SNP probes. **d**: A bar plot showing the fraction of fCpGs that were called as unmodified cytosine (C), 5-Methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC) or a non-canonical base (A, G or T). **e-f**: Heatmap, histograms and scatterplots showing that the methylation status of fCpGs is validated in Oxford nanopore long-read (e) and WGBS data (f). **g**: fCpG methylation of phased Oxford nanopore long-reads of 2 CLL and 2 matched RT samples, 3 sorted memory and 3 naïve B cells samples. Long reads which cover the 4 fCpGs within the region chr7:25,854,120-25,856,220 are shown. **h**: (left) A box plot comparing the mean intra-haplotype Hamming distance of phased reads from cancer and normal samples ($p = 0.016$, MW-U test). One normal B-cell sample contained only 2 reads and was therefore removed. (right) Paired comparison of the mean intra- and inter-haplotype Hamming distances in 1 CLL and 2 RT samples ($p = 0.030$, paired T-test). One of the CLL samples only had reads from one haplotype, and thus was not informative for this plot.



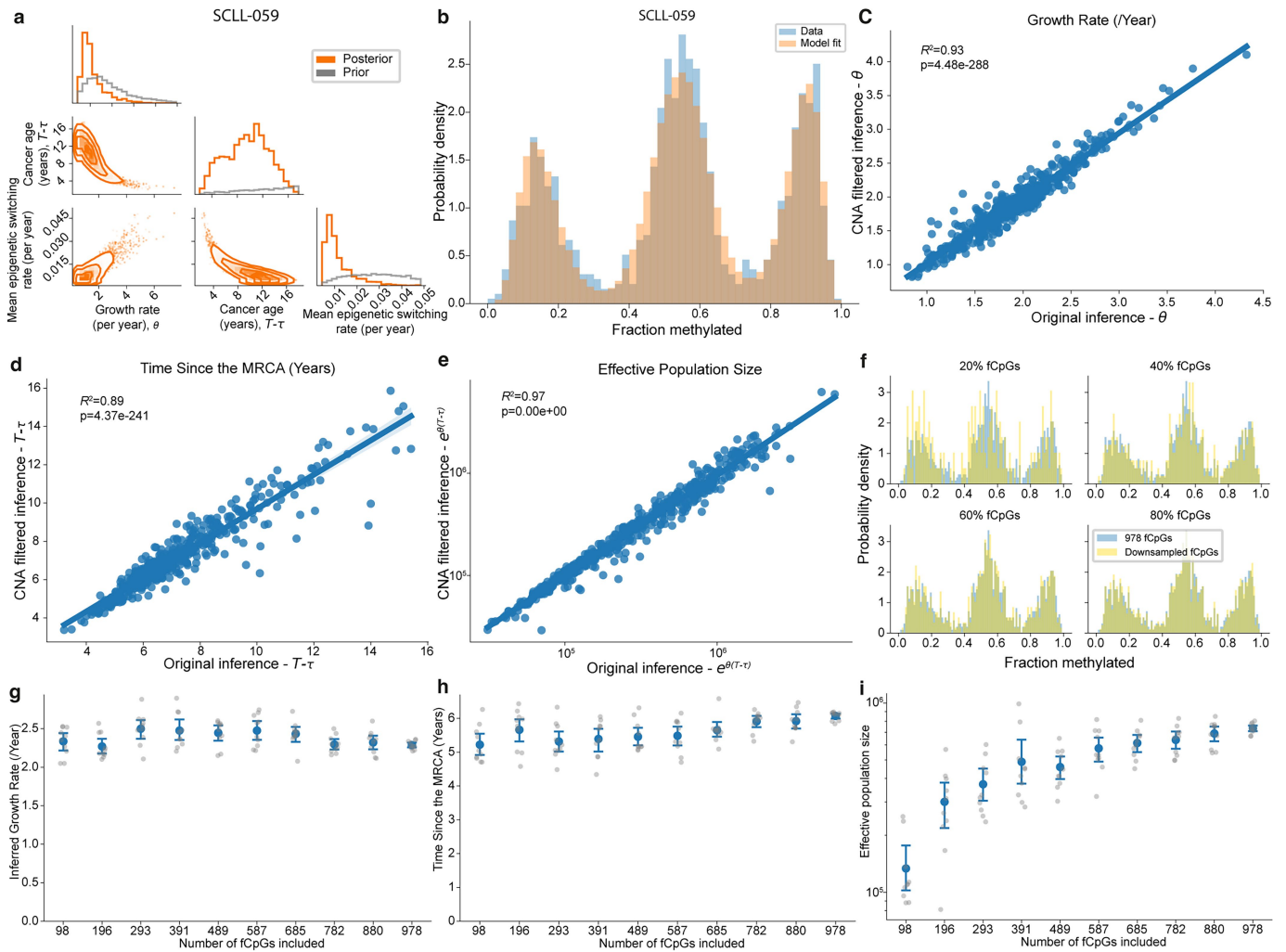
Extended Data Fig. 3 | fCpGs change methylation in an allele specific and are selectively neutral manner. **a:** (top) Illustration of the possible combinations of fCpG states in monosomic and trisomic regions, along with the corresponding fraction methylation value. (bottom) Example histograms showing the fCpG methylation distribution of fCpGs located on just monosomic) or trisomic regions within 2 patient samples. **b:** Mean and standard deviation of fCpGs as a function of age in 2 independent cohorts of normal lymphoid cells, showing consistent methylation fluctuations but increased variance during aging. **c:** Boxplots of standard deviation of fCpG methylation values of whole-blood samples divided by age groups together with lymphoid tumour samples. P -values were derived using two-tailed t tests.

Boxplots whiskers represent ± 1.5 IQR. **d:** fCpG loci are significantly less likely to be associated with a gene (annotations provided by UCSC, $p = 1.6e-15$, chi-sq test). **e:** A comparison of the mutual overlap between different epigenetic clocks: mitotic^{14,65-68}, chronological age⁶⁹⁻⁷⁸, gestational age⁷⁹⁻⁸³, biological age and mortality⁸⁴⁻⁸⁶, and trait predictors^{87,88} and fCpGs. **f:** P-P plot demonstrating fCpG associated genes have a lower expression (in transcripts per million) than non-fCpG associated genes. **g:** Linear regressions of the RNA expression of key methylation maintenance genes (DNMT1, DMT3a, DNMT3B and TET2) vs and the mean (left) and standard deviation (right) of the fCpG distribution.



Extended Data Fig. 4 | Synthetic data analysis shows EVOFLUX is robust to varying modelling assumptions. **a:** Posterior distributions after fitting the simulated data in Fig. 2b with the Bayesian inference method. The posterior (orange) displays tightening around the ground truth parameter values (red) compared to the prior (grey). **b:** Histogram of the simulated fCpG methylation distribution (blue) with the posterior predictive of the model fit overlaid (orange). **c:** Histograms showing the methylation distributions of 10,000 synthetic fCpGs simulated under the logistic growth model with varying carrying capacity (Supplementary Information). **d-e:** EVOFLUX posterior median and 95% credible interval were run on a subset of 2,000 of the simulated

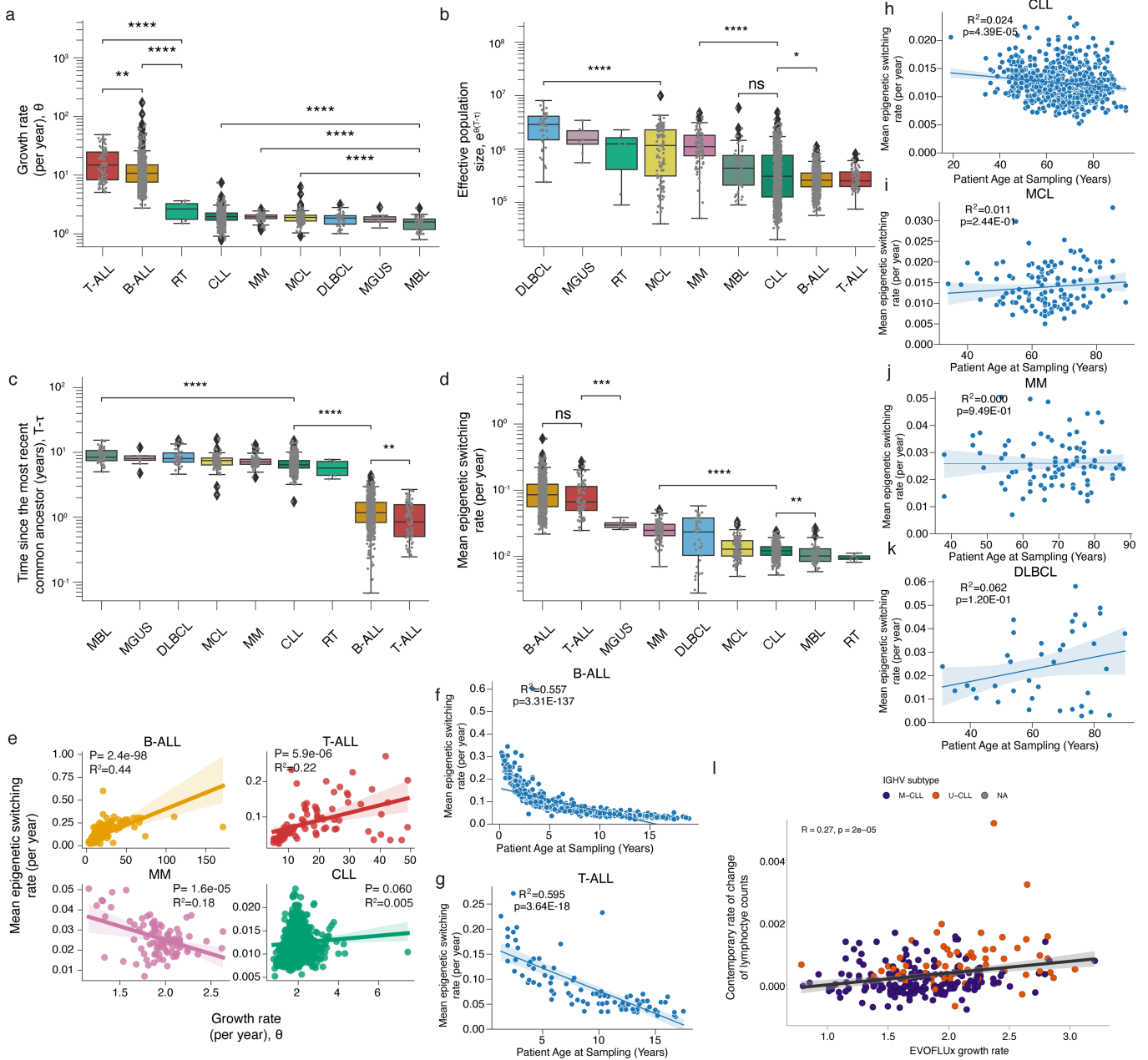
fCpGs in **c** as a function of the carrying capacity for the growth rate (**d**) and most recent common ancestor age (**e**). The dashed red line represents the ground truth parameter value. **f:** Histograms showing the fCpG methylation distributions of 10,000 synthetic fCpGs simulated under the heterogenous epigenetic switching model with varying switching rate standard deviation (Supplementary Information). **g-h:** EVOFLUX posterior median and 95% credible interval when run on a subset of 2,000 of the simulated fCpGs in **f** as a function of the switching rate standard deviation for the growth rate (**g**) and most recent common ancestor age (**h**).



Extended Data Fig. 5 | EVOFLUX inference is robust to missing data.

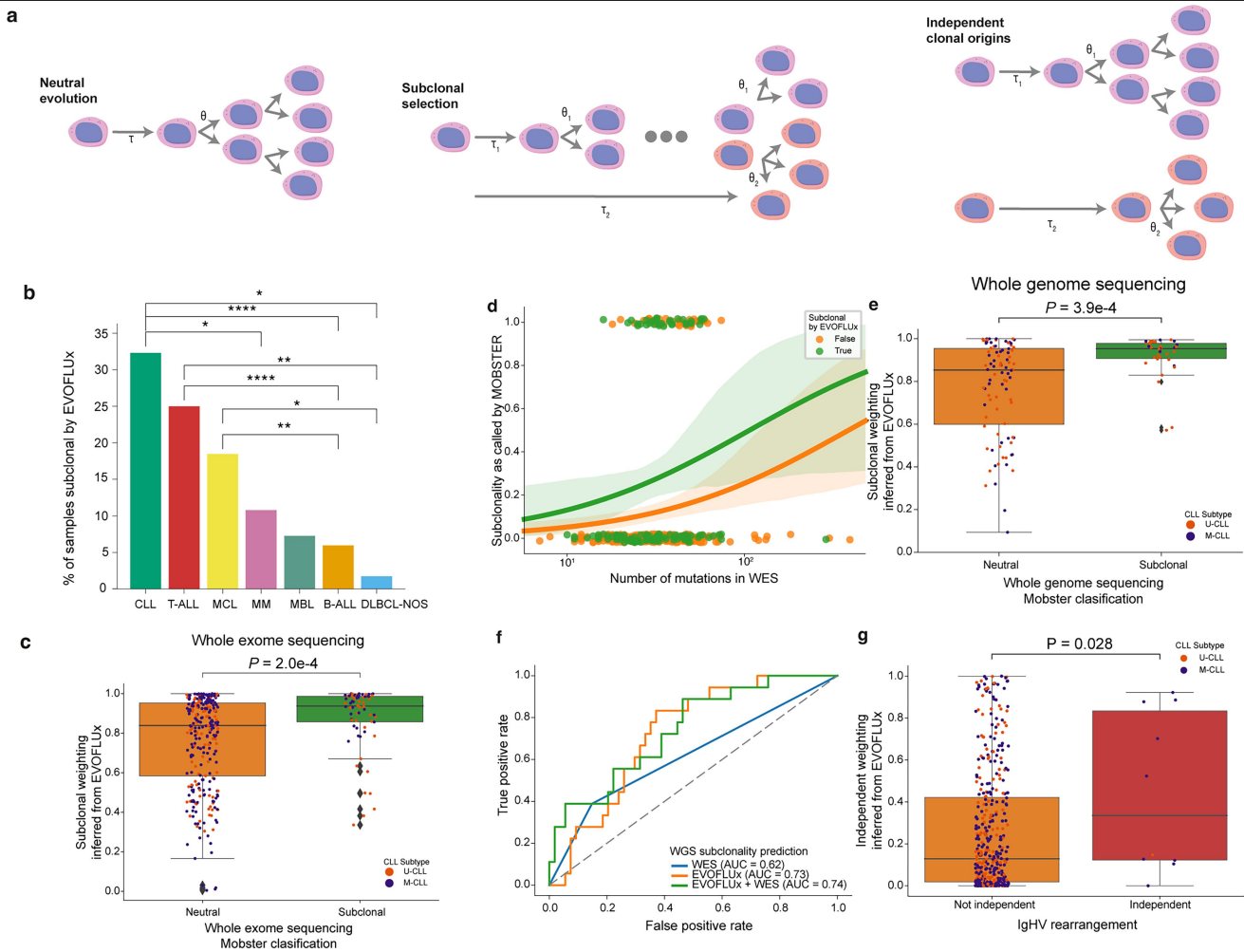
a: Example posterior resulting from running EVOFLUX on a CLL sample – a pairs plot showing the marginal (diagonal) and pairwise (off-diagonal) inferred posterior (orange) distributions, with the prior distributions overlaid (gray). Posteriors show marked tightening compared to the priors, demonstrating the parameters were well informed by the data. **b:** A histogram of the fCpG methylation distribution (blue) with the posterior predictive of the model fit overlaid (orange). **c-e:** Regression plots between the parameters inferred by running EVOFLUX on all 978 fCpGs (x-axis) vs just those fCpGs present on diploid regions (y-axis) in the CLL cohorts for the growth rate (c), the most

recent common ancestor age (d) and the effective population size (e). **f:** Example histograms showing the distribution of all 978 fCpGs (blue) in a CLL sample (SCLL-001), with a randomly downsampled subset of 20%, 40%, 60% and 80%. **g-i:** Plots showing the effect of the number of downsampled fCpGs included in the inference process against the inferred growth rate (g), time since the most recent common ancestor (h), and the effective population size (i). For each set of 10% increment, 10 replicate fCpG subsets were generated (grey dots) and the EVOFLUX inference repeated. Mean and standard error of the replicates represented with a blue dot and error bars respectively.



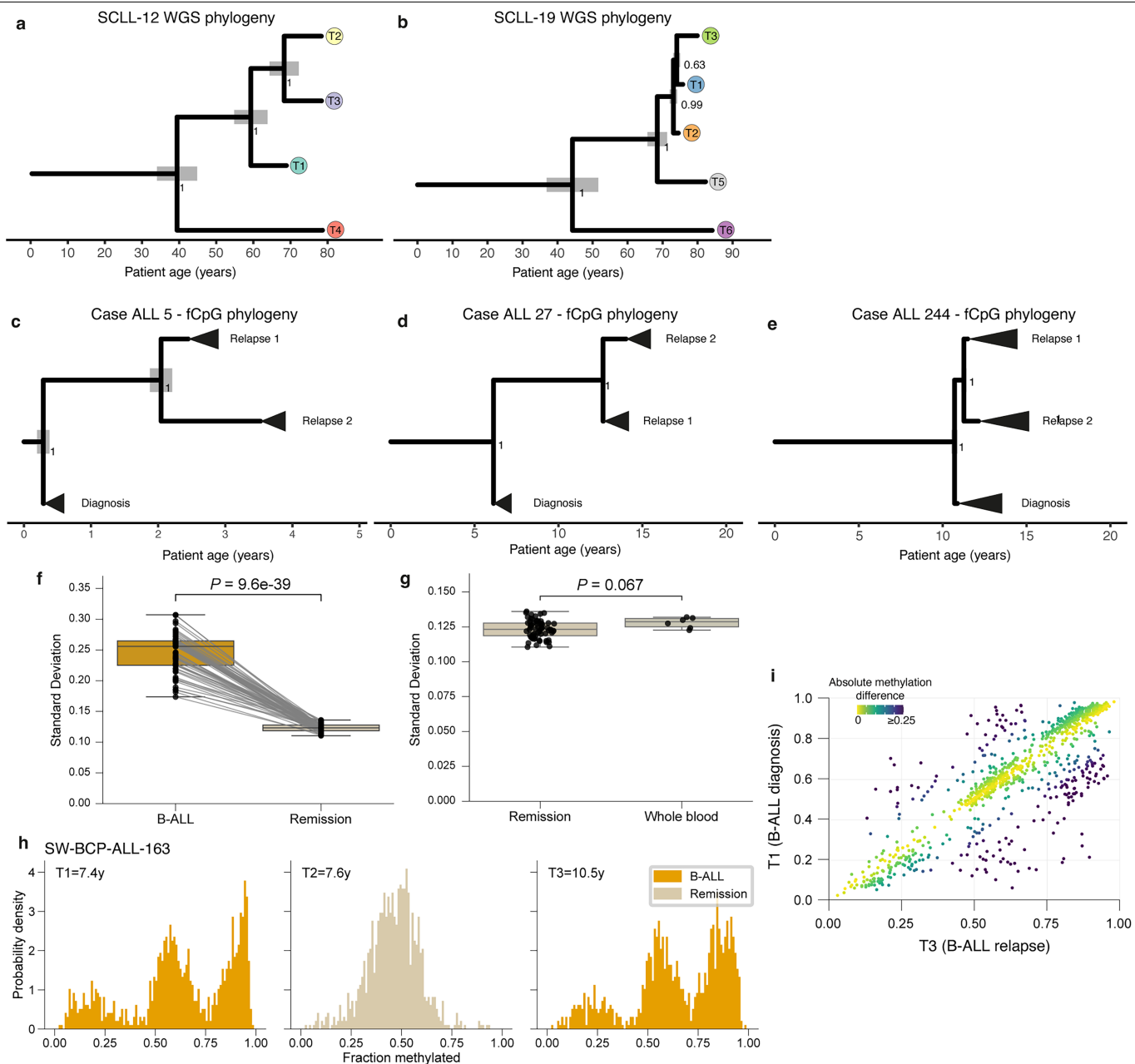
Extended Data Fig. 6 | EVOFLUX captures lymphoid cancers' evolutionary histories and methylation epimutation dynamics. a-d: Boxplots (whiskers extending to $\pm 1.5 \times \text{IQR}$) showing the distribution of inferred growth rate (a), effective population size (b), time since the most recent common ancestor (c), and mean epigenetic switching rate (i.e. mean of μ, ν, γ, ζ ; d) by disease. For interpretability, only a subset of the pairwise p values are annotated (Mann-Whitney U tests, hs correction). **e:** Linear regression between the growth rate and mean epigenetic switching rates separated by cancer types. There is a positive association in B-ALL ($P = 2.4e-98$, $R^2 = 0.44$) and T-ALL ($P = 5.9e-06$, $R^2 = 0.22$), a weak negative association in MM ($P = 1.6e-05$,

$R^2 = 0.18$) and no association in CLL ($P = 0.060$, $R^2 = 0.005$) or the other entities. **f-k:** Linear regressions between the patient age at sampling and the mean inferred epigenetic switching rate in B cell acute lymphoblastic leukaemia (B-ALL, f), T cell acute lymphoblastic leukaemia (T-ALL, g), chronic lymphocytic leukaemia (CLL, h), mantle cell lymphoma (MCL, i), multiple myeloma (MM, j) and diffuse large B cell lymphoma (DLBCL, k). **l:** Linear regression between the EVOFLUX inferred initial evolutionary growth rate and the estimate of a linear model of the number of historical lymphocyte counts with the sampling dates of patients with at least 10 sample timepoints before treatment ($P = 2e-5$, Supplementary Fig. 10).



Extended Data Fig. 8 | Inference and validation of subclonal selection using fCpG loci. a: Illustrations of three alternative evolutionary models: neutral evolution, in which the cancer population with a MRCA emerging at time τ grows exponentially at rate θ ; subclonal selection, in which an initial population emerging at time τ_1 growing at rate θ_1 is outcompeted by a fitter subclone emerging at time τ_2 with growth rate θ_2 ; and independent clonal origins, in which the fitter clone emerging at time τ_2 with growth rate θ_2 bears no clonal relationship to the initial clone. **b:** Bar chart comparing the fraction of cancer samples identified as subclonal by EVOFLUX across disease (pairwise chi-sq tests, holm-sidak (hs) correction). **c:** Boxplots comparing the distribution of subclonal weightings inferred by EVOFLUX in samples called as neutral vs under subclonal selection via whole exome sequencing (WES) data. **d:** Logistic regression between the probability of a cancer being identified as subclonal via running MOBSTER⁴³ on WES data and the number of mutations

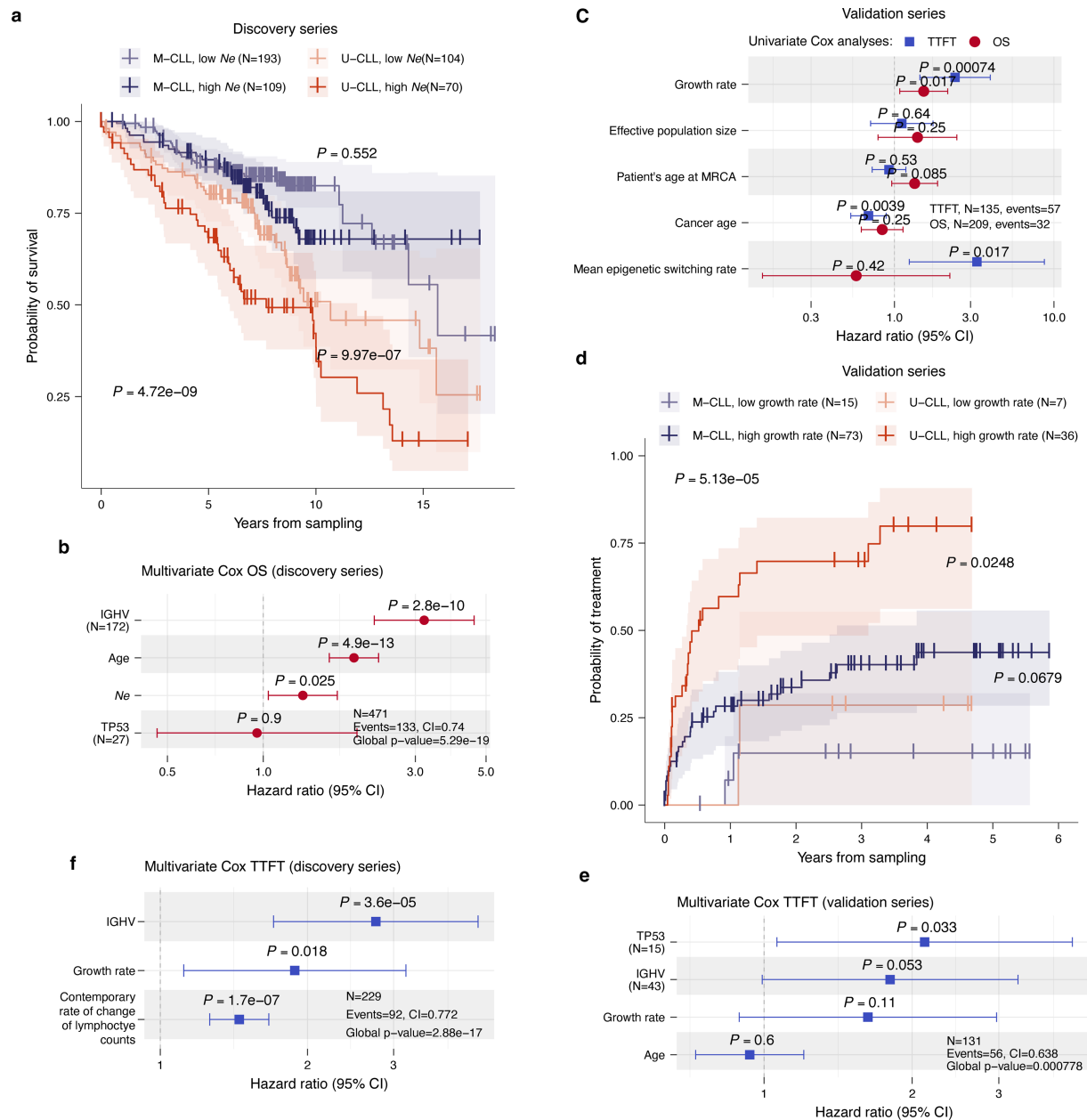
detected in the sample. Regression were run separately for samples also identified as subclonal/neutral via EVOFLUX (subclonal weighting > 95%). **e:** Boxplots comparing the distribution of subclonal weightings inferred by EVOFLUX in samples called as neutral vs under subclonal selection within whole genome sequencing (WGS) data. **f:** Receiver operating characteristic (ROC) curves showing the accuracy at predicting the subclonality of WGS data of competing logistic classifiers trained on just the EVOFLUX subclonality weighting, just the WES subclonality call and a model trained on both sources of data. **g:** Boxplots comparing the distribution of independent model weightings inferred by EVOFLUX in samples containing multiple IGHV rearrangements (independent cancers) vs those with only a single IGHV rearrangement detected via WGS, WES and/or RNA-seq⁴⁶ (likely a single clonal origin).



Extended Data Fig. 9 | Matched whole genome sequencing data validates

fCpGs as a phylogenetic character. **a-b:** Corresponding phylogenies reconstructed on matched WGS SNVs for CLL case 12 and 19 to Fig. 4a,b, respectively. **c-e:** The reconstructed phylogenies of the relationship between samples collected longitudinally in 3 individual ALL patients, annotated with the clinical classification of each sample. The black triangles represent the time that occurred since the most recent common ancestor, taken as the posterior median of $T - \tau$ from the single-sample EVOFLUX inferences. **f:** Paired boxplot showing the standard deviation of the fCpG methylation distributions of

B-ALL samples is greater than their matched remission sample ($p = 9.6e-39$, paired t-test). **g:** Comparison of the standard deviation of the fCpG methylation distributions of B-ALL patients in remission (i.e. no cancer cells present in blood) vs normal whole blood ($p = 0.067$, MW-U test). **h:** Example longitudinal samples from one patient showing the development of the fCpG distribution from diagnostic B-ALL, through remission and relapse. **i:** Scatterplots showing the marginal fCpG methylation distribution between diagnosis and relapse for the ALL case in **h**.



Extended Data Fig. 10 | Additional survival analyses. a: Kaplan-Meier curves comparing the OS between patients with high vs low (cut-off values identified using maxstat statistics) inferred effective population sizes (N_e) in the discovery cohort³³, separated by IGHV mutational status. **b** Multivariate Cox regression of the OS shows the N_e is significant when controlling for IGHV status, *TP53* alterations and age at sampling in the discovery cohort.

c: Univariate survival analysis of the time to first treatment (TTFT, blue) and overall survival (OS, red) in the validation CLL cohort^{28,29} for evolutionary variables inferred via EVOFLUX. Note this cohort contains a mixture of treated and untreated samples, of which only the untreated samples were included in the TTFT analysis. **d:** Kaplan-Meier curves comparing the TTFT between

patients with high vs low (cut-off values identified using maxstat statistics) inferred cancer growth rates in the validation cohort, separated by IGHV mutational status. **e:** Multivariate Cox regression of the effect of the cancer growth rate on the TTFT in the validation cohort, controlling for IGHV status, *TP53* alteration and age at sampling. **f:** Multivariate Cox proportional hazard regression on the TTFT for 229 CLL patients where longitudinal measurements of the lymphocyte numbers were available and therefore the contemporary growth rate could be estimated. The contemporary lymphocyte counts refer to the estimate of a linear model of the number of historical lymphocyte counts with the sampling dates of patients with at least 10 sample timepoints before treatment.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

No software was used to collect data.

Data analysis

Preprocessing of sample data and some analysis was performed using R (v4.3.1). This involved the use of the minfi package (v1.46.0), alongside IlluminaHumanMethylation450kanno.ilmn12.hg19 (version 0.6.1), IlluminaHumanMethylationEPICanno.ilm10b2.hg19 (version 0.6.0) and methylCIPHER (<https://github.com/MorganLevineLab/methylCIPHER>). Base core R and Bioconductor (version 3.1) packages as well as custom R scripts were used to mine, analyze and plot the data. ggplot2 (v3.5.2) was also used to plot data. Survival (v3.5-7), maxstat (v0.7-25), survminer (v0.4.9) and ggsurvfit (v0.3.1) R packages were used to analyze and display clinical associations.

Other data analysis was performed using Python (v3.7.4). Analysis employed the following package: numpy (v1.21.6), pandas (v1.3.5), scipy (1.7.3), statannotations (v0.4.4), statsmodels (v0.13.2) and scikit-learn (v0.23.2). The Bayesian programming language Stan was used to perform fCpG discretisation using a custom script. Pathway analysis was performed using gProfiler version e109_eg56_p17_1d3191d.

The EVOFLUX code employed to infer the evolutionary history of cancer samples from methylation array data is at <https://github.com/CalumGabbutt/evoflux>. Code to curate DNA methylation data and to perform clinical and additional bioinformatic analyses is available at <https://github.com/Duran-FerrerM/evoflux>. The phylogenetic method is available at <https://github.com/adamallo/PISCA>.

The EVOFLUX Python software contains a README.md listing the important packages, and a requirements.txt file indicating the appropriate versions. The main packages were numpy (basic maths functions), scipy (special maths functions), pandas (dataframes manipulation), dynesty (nested sampling - Bayesian inference), joblib (pickling large data files, preserving dynasty structure), notebook (Jupyter notebook), matplotlib (plotting), seaborn (plotting), arviz and corner (Bayesian plotting).

Phylogenetic inference on fCpGs was performed with the BEAST (v.1.8.4) framework (Drummond and Rambaut 2007; Drummond et al. 2012) using custom models implemented in PISCA v1.1. Phylogenetic inference on SNV data was performed using BEAST (v.1.10). Phylogenetic MCMC chains were assessed using Tracer v1.7.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

No new methylation bead array data was generated in the course of this study. The harmonised and filtered methylation matrix was deposited using Zenodo (<https://doi.org/10.5281/zenodo.15479736>), which we reference in the Methods.

Previously published DNA methylation data re-analysed in this study can be found under accession codes: B cells, EGAS00001001196; ALL, GSE56602, GSE49032, GSE76585, GSE69229; MCL, EGAS00001001637, EGAS00001004165; CLL, EGAD00010000871, EGAD00010000948, EGAD00010001975; MM, EGAS00001000841; DLBCL, EGAD00010001974.

CLL gene expression data is available EGAS00001000374 and EGAS00001001306. ChIP-seq datasets are available from Blueprint <https://www.blueprint-epigenome.eu/> under the accession EGAS00001000326.

Matched WES and WGS are available under accessions EGAS00000000092 and EGAD00001008954 respectively under controlled access.

Pathway analysis was run using the Gene Ontology: Biological Processes (GO:BP) release 2023-03-06 and the Human Protein Atlas (HPA) version 10.0 databases.

Matched Oxford nanopore long-read data were generated for 6 normal B cell and samples and 2 CLL/RT sample pairs. These data are in the process of being uploaded to EGA and the accession will be provided once the data has been processed and validated.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Where biological sex was annotated in the preexisting sample annotations, these annotations have been retained, but sex was not considered an important factor in this study.

Reporting on race, ethnicity, or other socially relevant groupings

Race and ethnicity information was not considered in this study.

Population characteristics

This is an analysis article using mostly epigenetic and clinical data reported in previous publications.

Recruitment

Individual studies from which data were obtained for this analysis article explain how patients were recruited. In general, all the studies recruited patients for methylation studies based on sample availability in their biobanks.

Ethics oversight

The study was approved by the ethics committees of the institutions that generated the data used in the present study.

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

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Sample size was not determined by calculations, but by sample availability and data quality.

Data exclusions

Cancer samples with tumor cell content below 60% were excluded, in line with our previous studies (Duran-Ferrer et al., 2020). Samples where the age of the patient upon sampling was missing were included for the purposes of fCpG selection and validation, but removed from the EVOFLUX inference of evolutionary dynamics.

Replication

Experimental replication was not feasible because this work is an observational, computational analysis of 1,976 publicly available tumour samples. Acquiring an additional cohort of patient samples and running additional methylation bead arrays would impose prohibitive cost,

time and ethical burden. Instead, robustness was ensured by testing that the clinical associations in CLL was observed in two independent cohorts and blinding of the researcher performing the EVOFLUX inference to the clinical outcome data. The detection of subclonal selection in whole exome sequencing data was validated in whole genome sequencing data.

Randomization Patient age and tumour purity was directly accounted for in the EVOFLUX inference model. The difference in inferred evolutionary parameters between molecular subtypes was found to exist even when tumour purity was controlled as a covariate in a multivariate linear regression ($p < 0.05$). The association between inferred evolutionary parameters and clinical outcomes were still significant when patient age, IGHV status and TP53 mutational status were controlled in a multivariate Cox regression model.

Blinding The study is a retrospective, fully computational re-analysis of anonymised public datasets, and therefore full blinding was not feasible. However, the inference of evolutionary parameters was performed blinded to the patients' clinical outcomes.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| n/a | Involvement |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants |

Methods

| n/a | Involvement |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Plants

| | |
|------------------------------|--|
| Seed stocks | <i>Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.</i> |
| Novel plant genotypes | <i>Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.</i> |
| Authentication | <i>Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.</i> |