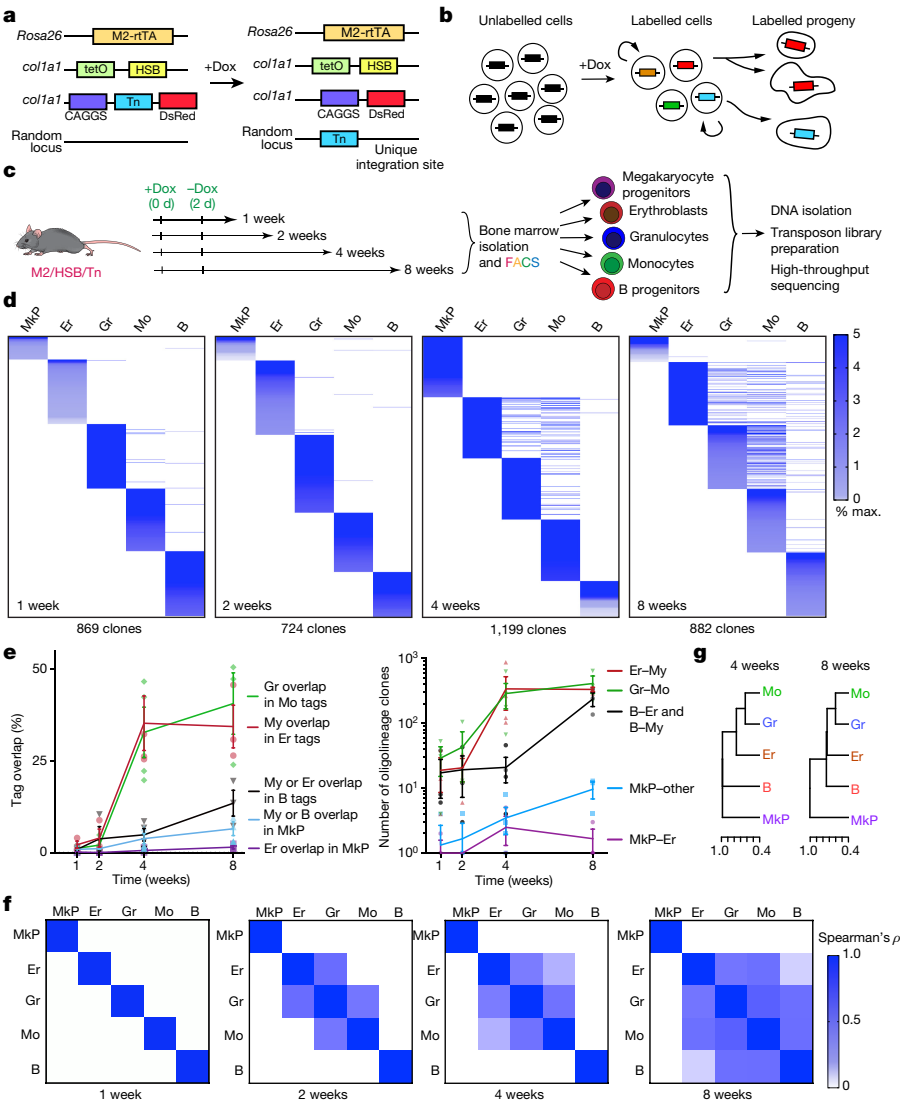

Figure 1 | Clonal analysis of haematopoietic lineage fates in the native bone marrow. **a**, M2/HSB/Tn (transcriptional activator M2/hyperactive Sleeping Beauty/transposon) mouse model. Addition of Dox induces random transposition of the transposon, and concomitant cell labelling with DsRed. The transposon insertion site is stable after removal of Dox. **b**, Transposon lineage tracing. Shared tags can be detected between a self-renewing progenitor stem cell and its progeny, or between two different mature cell populations. **c**, Experimental design. M2/HSB/Tn mice were labelled with Dox for 2 days and five blood lineages were isolated from bone marrow after different periods of time. Transposon insertion tag libraries were prepared and sequenced for each population. **d**, Alignment of transposon tags from different lineage-committed (Lin^+) blood cell populations in the bone marrow at 1–8 weeks. Tags are coloured by frequency in each lineage, and organized by rank. Each chart is representative of three independent experiments. MkP, megakaryocyte progenitors; Er, erythroblasts; Gr, granulocytes; Mo, monocytes; B, B-cell progenitors. **e**, Percentage of clonal overlap between designated lineage pairs (left), and quantification of total number of detected bi-/tri-lineage clones at 1–8 weeks (right). Abbreviation My refers to either granulocyte or monocyte lineage. Mean ± s.d. from three independent mice. **f**, Spearman's correlation coefficient (ρ) matrices for all Lin^+ tags at 1–8 weeks. Each matrix is the average from three independent experiments per time point. **g**, Hierarchical clustering of blood lineages using $(1 - \rho)$ as the distance measure (4 and 8 weeks after labelling).

MkP-producing MPP clones, MkP overlap was more lineage-restricted than any other lineage, even after eight weeks (MkP 67.8 ± 8.0% versus other 22.1 ± 4.6%; Fig. 2a, b and Extended Data Fig. 7b), indicating that at least a subset of MPPs is responsible for a stable restricted contribution to the megakaryocyte lineage.

Our analyses also provided relative quantitative information about the dynamics of lineage replacement by MPPs. For instance, the average clone size of MPP-derived erythromyeloid clones at eight weeks was 18.3 ± 7.7-fold larger than non-MPP-derived clones, suggesting a significant cellular amplification, in contrast to the B-cell progenitor lineage (1.2 ± 0.4-fold; Fig. 2c). In addition, we found that the erythroid lineage was replaced at the fastest rate, with at least 35% of all erythroblast reads overlapping with MPPs after just two weeks, from just a handful of erythroid-committed MPPs (Fig. 2d, e). By comparison, the granulocyte/monocyte-producing MPPs achieved similar levels of replacement only after two months. Considering that our analysis cannot measure the contribution of MPP clones that disappear from the MPP pool (that is, by cell death or differentiation), our results probably underestimate the overall MPP contribution.

To provide further insight into the heterogeneity and hierarchy of the haematopoietic stem cell (HSC)/MPP compartment, we sorted subsets within these populations using previously described surface markers and interrogated their single-cell gene expression landscape using inDrop (Fig. 3a–c)[9,17]. Louvain–Jaccard clustering analysis of transcriptomes resulted in 12 reproducibly distinct clusters (Fig. 3b). Most analysed cells (78.9% of all subsets combined) fitted into one of three major clusters that we labelled as unprimed ('C1', 'C2', 'C3') on the basis of the lack of expression of lineage-restricted gene signatures (Supplementary Table 2 and Extended Data Figs 8 and 9). We also identified several primed clusters (21.1% of HSCs/MPPs) that formed branches defined by progressive expression of genes associated with lineage commitment (Fig. 3b–d, right). Predictably, cells indexed as long-term (LT)-HSCs and MPP1s (also known as short-term (ST)-HSCs) mostly fitted into the 'C1' (67.9%) and 'C2' (78.3%) clusters, respectively. By contrast, other MPP subsets displayed different degrees of heterogeneity. MPP2s contained the largest proportion of primed cells (59.3%), and MPP4s the least (13.2%) (Fig. 3c, d). MPP2s comprised a larger number of erythroid-primed (18.7%) and megakaryocyte-primed (21.9%) cells, whereas MPP3s contained a larger number of granulocyte/monocyte-primed cells (20.8%) (Fig. 3c, d and Extended Data Fig. 8b). Using transposon tracing, we confirmed that MPP2s presented a preference for MkP production, and generated less multilineage output (5 ± 5% of all active clones) within the first week, where their immediate progeny is likely to be measured, compared with MPP3s and MPP4s (40.17 ± 11.4%) (Fig. 3e, f). Analysis of tags not arising from upstream progenitors at four weeks revealed similar findings (Fig. 3g, h). On the contrary, MPP4s produced most lympho-erythromyeloid and multilineage clones (Fig. 3h) and preferentially overlapped with MPP1/ST-HSCs, suggesting that at least a fraction of MPP4s represent direct activated progeny of MPP1/ST-HSCs (Fig. 3i). Combined, our data support the notion that a functional hierarchy, consisting of progenitors at varying degrees of lineage priming, exists already within HSCs/MPPs.