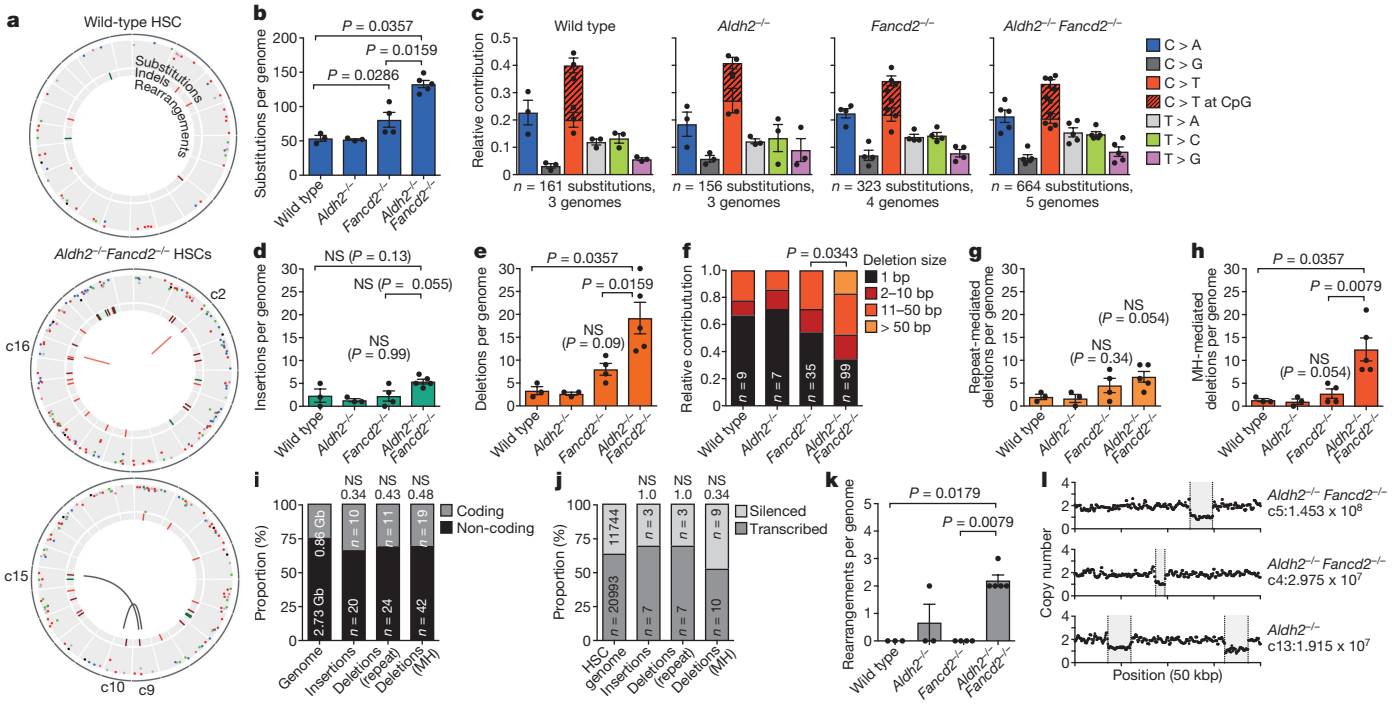

Figure 5 | Endogenous aldehydes mutate the HSC genome. a, Circos plots showing mutations in three HSCs. All HSCs are shown in Extended Data Fig. 4. **b**, **d**, **e**, **g**, **h**, **k**, Mutations of different classes per genome (*P* calculated by two-sided Mann–Whitney test; data shown as mean and s.e.m.; $n = 3$, 3, 4 and 5 HSC genomes, left to right). **b**, Number of substitutions. **c**, Point mutation classes in HSC genotypes. **d**, Number of insertions per genome. **e**, Number of deletions per genome. **f**, Distribution of the size of deletions (χ^2 test, *n* shows number of deletions). **g**, Number of repeat-mediated deletions per genome. **h**, Number of microhomology (MH)-mediated deletions per genome. **i**, **j**, Indels in *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ HSCs are randomly distributed: within or outside genes (**i**) (*P* calculated by hypergeometric distribution, *n* is number of indels), or between expressed or silenced genes (**j**) (*P* calculated by binomial distribution, *n* is number of indels). Numbers above columns, *P* values. **k**, Number of rearrangements per genome. **l**, Large copy-number losses in *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ and *Aldh2*$^{-/-}$ HSCs at the indicated locations.

provide the first whole-genome sequences obtained from single stem cells propagated *in vivo*. These stem cell genomes show that endogenous aldehydes induce a tapestry of inter-chromosomal changes that are mediated by mutagenic end-joining of DNA DSBs.

A p53 response removes aldehyde-damaged HSCs

Strikingly, most *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ HSCs failed to engraft (Fig. 4b). It is possible that HSCs that carry heavy DNA-damage burdens are eliminated, and selection pressure favours the survival of less damaged stem cells. It was therefore important to determine the mechanism of HSC loss in *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ mice, and if attenuated, it would be important to determine the mutagenic consequences. The p53 protein regulates the cellular response to DNA damage and, when activated, induces restorative processes or apoptosis. We found that *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ haematopoietic stem and progenitor cells (HSPCs) accumulated p53 and cleaved caspase-3, indicating that endogenous-aldehyde stress activates the p53 response (Extended Data Fig. 6a–c). Furthermore, we found that genetic ablation of p53 partially suppressed the acetaldehyde hypersensitivity of *Fancd2*-deficient splenic B cells and granulocyte/macrophage colony forming units (Extended Data Fig. 6d, e).

We therefore generated *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$*Trp53*$^{-/-}$ triple-knockout mice. The severe HSC depletion of *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ mice was completely rescued in the triple knockouts (Fig. 6a, b). In addition, the triple-knockout stem cells were functional, as the mice showed a complete rescue in the frequency of ST-HSCs upon bone-marrow transplantation (Extended Data Fig. 6f). Moreover, p53 deficiency fully restored the blood cytopenias of untreated *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ mice and made these mice more resistant to alcohol exposure (Extended Data Fig. 7). Notably, *Trp53* deletion did not rescue the embryonic lethality of *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ embryos (in *Aldh2*$^{-/-}$ mothers), suggesting that a different checkpoint might mediate developmental failure (Supplementary Information Table 2).

We reasoned that the rescue of haematopoiesis in *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$*Trp53*$^{-/-}$ mice must be occurring at the cost of genome integrity. Although the level of micronucleated NCEs in the blood of *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$*Trp53*$^{-/-}$ mice appeared similar to that of *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ mice (Extended Data Fig. 8a), we noticed a significant ($P = 0.0034$) increase in chromosome rearrangements in *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$*Trp53*$^{-/-}$ mice, as seen by M-FISH analysis of total bone marrow cells (Fig. 6c, Extended Data Fig. 8). However, neither of these analyses tell us whether genome stability is similarly compromised in *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$*Trp53*$^{-/-}$ HSCs. We therefore performed transplantation of single HSCs combined with whole-genome sequencing, as described earlier, and observed that p53 deficiency partially rescued the engraftment defect of *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ HSCs (Fig. 6d). Surprisingly, the genomes of *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$*Trp53*$^{-/-}$ stem cells did not carry a greater mutation burden compared to those of *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ HSCs (Fig. 6e, f). Indel and rearrangement calls were validated by targeted deep sequencing and PCR, respectively (Extended Data Figs 9, 10 and Methods). One plausible explanation for the lack of increased mutagenesis in *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$*Trp53*$^{-/-}$ HSCs is the possibility that the very small number of HSCs in *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ mice might have undergone more replicative cycles, thereby accruing a larger number of mutations. To address this, we quantified the frequency of 'clock' mutations (C to T at CpG sites), but this analysis showed no significant difference between *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$ and *Aldh2*$^{-/-}$*Fancd2*$^{-/-}$*Trp53*$^{-/-}$ HSCs (Fig. 6f). These results indicate that aldehyde-induced DNA damage induces p53 leading to HSC attrition, which is inconsistent with p53 being a negative regulator of Fanconi anaemia repair, as recently reported[27]. However, while *Trp53* deletion completely rescues HSC depletion, this does not occur at the expense of genome stability in blood stem cells.