Single-cell SNP-DNA sequencing precisely maps genotoxic events in CRISPR-edited primary cells

Julian Boutin [1,2]†, Sabrina Fayet[1]†, Victor Marin[1,2]†, Camille Bergès[1,3], Maude Riandière[1], Jérôme Toutain[1,3], Isabelle Lamrissi-Garcia[1], Chloé Thibault[1], David Cappellen[1,4], Sandrine Dabernat[1,2], Arthur Poulet[1], Maëla Francillette[5], Nathalie Droin[5,6], François Moreau-Gaudry[1,2]†, Aurélie Bedel[1,2]†

1 Univ Bordeaux, INSERM, UMR 1312, BRIC (Bordeaux Institute of Oncology) 146 Rue Léo Saignat, Case 10, F-33076 Bordeaux, France.

2 CHU de Bordeaux, Biochemistry Laboratory, F-33000 Bordeaux, France.

3 CHU de Bordeaux, Medical genetic laboratory, F-33000 Bordeaux, France.

4 CHU de Bordeaux, Tumor biology and tumor bank laboratory, F-33000 Bordeaux, France.

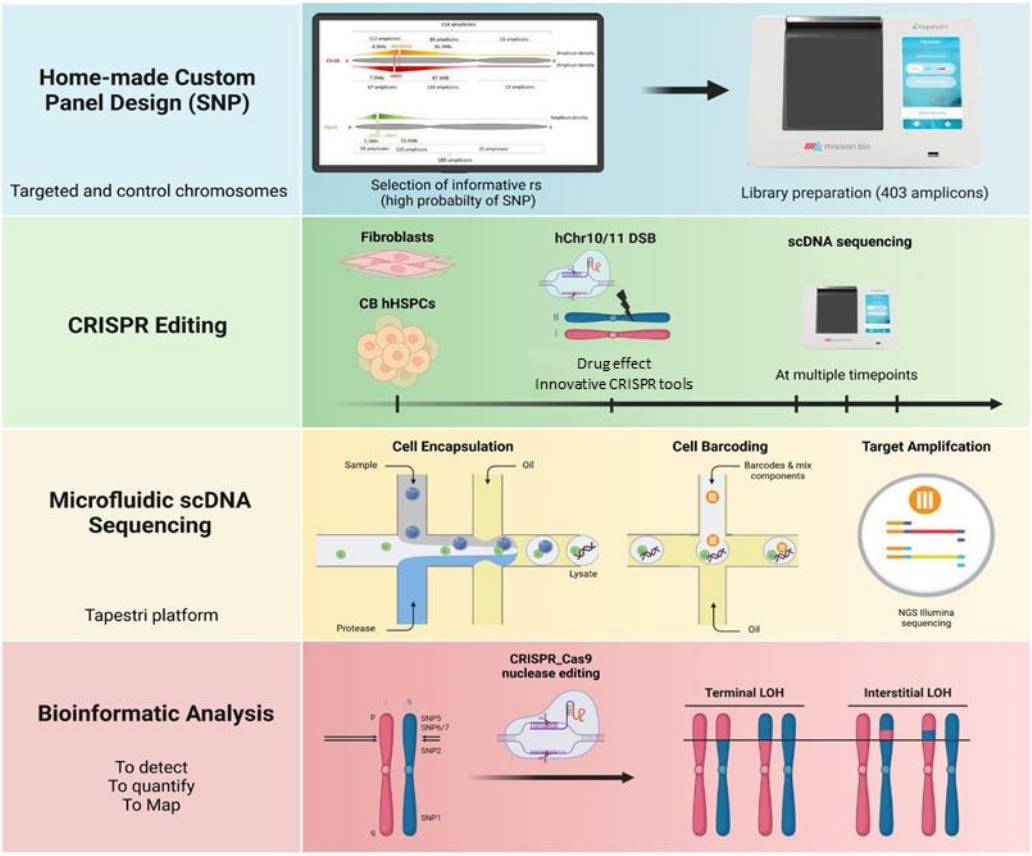
5 INSERM US23, CNRS UAR 3655, AMMICa, Genomic platform, Gustave Roussy Cancer Center, F-94805 Villejuif, France.

6 INSERM U1287, Gustave Roussy Cancer Center, F-94805 Villejuif, France.

† equally contributed to this work

Corresponding author: Aurélie Bedel. e-mail: aurelie.bedel@u-bordeaux.fr

Graphical abstract



Abstract

Genome editing by CRISPR-Cas9 is promising for genetic disease and cancer gene therapy. However, safety concerns are still present, particularly the ON-target genotoxicity for protocols using nucleases. Quality control of edited cells before and after graft is mandatory, especially to assay megabase-scale genomic rearrangements induced at the targeted *locus*. These unintended events are fortunately rare but potentially deleterious. Classical PCR-based bulk approaches do not detect them or underestimate their frequency. Single-cell approaches are promising but RNA sequencing only estimates copy number variation of the genome. Here, we propose single-cell DNA sequencing to accurately evaluate and monitor CRISPR-mediated genotoxicity in primary cells (human fibroblasts and hematopoietic stem/progenitor cells). We designed a homemade panel using single nucleotide polymorphisms to detect, map and characterize induced-losses of heterozygosity (terminal, interstitial, copy-loss and copy-neutral). This innovative approach revealed intense genotoxicity linked to the double strand break. The risk was associated with DNA repair modulators in particular DNA-PKcs inhibitor AZD7648. Importantly, the CDK4/6 inhibitor palbociclib prevented these rearrangements in hematopoietic stem/progenitor cells. This work strongly suggests that single-cell DNA sequencing should be routinely implemented in clinical applications before CRISPR-edited cell infusions.

Keywords: single-cell DNA sequencing, CNV, Loss of heterozygosity (LOH), chromosomal instability, gene therapy, CRISPR-Cas9, genotoxicity.

Introduction

CRISPR-Cas9 is promising for gene therapy[1-4]. The editing process yields a heterogeneous population with precise genome editing (PGE) and more frequently non-conservative insertions/deletions (InDels). Numerous CRISPR-based clinical trials are ongoing and the first treatments are already marketed. However, the editing process linked to targeted Double-strand break (DSB) can induce unpredictable outcomes with a genotoxic risk and potential edited-cell phenotype modifications. Targeted DSB repairs can induce ON-target genotoxicity. It encompasses kilobase-scale interstitial deletions[5], megabase-scale terminal deletions[6], copy-neutral loss-of-heterozygosity (CN-LOH) notably with altered imprinted genes expression[7], or even whole-chromosome losses[8] and chromothripsis[9-11]. Importantly, these genomic rearrangements are observed in cell lines and adult primary cells, such as hematopoietic stem/progenitor cells (HSPCs) and CAR-T cells and even in embryos[8,12-14].

Bulk analysis of alleles is classically used to monitor CRISPR genotoxicity but this global approach lacks sensitivity and can mask heterogeneity among cells. Moreover, short-read PCR-based technologies do not amplify alleles with large deletions. Long-read sequencing allows the detection of kilobase-scale deletions but not larger genomic rearrangements. Manual genomic single-cell analysis by cell cloning is laborious, time-consuming and not always possible, especially for primary cells and for techniques that require important material such as SNP/CGH-array and optic genomic mapping. Single-cell-RNA sequencing was recently proposed to measure CRISPR-mediated aneuploidies and chromosomal truncations[8]. Based on individual cell analyses, it increases the sensitivity to detect genomic by-products but is an indirect approach based on transcriptomic modulation and cannot easily detect CN-LOH.

Single-cell-DNA sequencing (scDNA-seq) should be more appropriate to analyze the genome integrity but is challenged by only two DNA copies in each cell. Microfluidic scDNA-sequencing is already used to monitor cancer intra-tumoral heterogeneity, clonal evolution and residual disease[15-16]. scDNA-seq technologies recently enabled the investigation of chromosome instability in tumors, revealing aneuploidies specific to individual cells within the same tumor[17]. By using a droplet-based targeted scDNA-seq solution, it was possible to sequence hundreds of *loci* across thousands of cells. Here, we used micronuclei detection (MN), LOH detection by an in-house Fluorescence-Assisted Megabase-scale Rearrangements Detection (FAMReD)[18] and scDNA-seq to assay CRISPR-Cas9 ON-target genotoxicity at the single-cell level in clinically relevant cells. Importantly, by combining SNP and copy number variations (CNV) analyses, we demonstrate for the first time that scDNA-seq is a powerful

assay to measure chromosomal instability, in particular CN-LOHs with high sensitivity. We revealed and mapped strong ON-target genotoxicity in human primary cells. We then demonstrated in HSPCs the powerful effect of palbociclib in preventing genotoxicity. Our sensitive tools uncovered AZD7648, a potent and selective DNA-PKcs inhibitor routinely used as an HDR booster, as highly genotoxic. We also used these accurate systems to assess HDR-booster effects and observed a high-genotoxicity for AZD7648, a potent and selective DNA-PKcs inhibitor. Our data show that scDNA-seq beats the limits of the current methods, including scRNA-seq

RESULTS

Design of the scDNA-seq library and method validation for LOH quantification.

To detect genotoxicity induced by CRISPR-Cas9, including CN- LOHs, we designed a home-made scDNA-seq library, focusing on SNPs in Chr10/Chr11. We identified frequent SNPs (Fig.1a). We then designed probes to obtain 403 amplicons/cell with higher concentration near the cut sites (Fig.1b). Amplicons on Chr10 served as control in the event of Chr11 targeting and vice-versa. After CRISPR-Cas9-editing, each amplicon of each cell was sequenced to analyze genome integrity, and LOH presence, with allelic losses (AL) of identified SNPs.

For method validation, we created a positive control cell line LOH^{+}. We used FAMReD, based on the heterozygous *UROS*$^{+/-}$ hFFs[18] (Fig.1c). By targeting *ABRAXAS2*, 1Mb centromeric to *UROS*, we observed a slight increase in the proportion of fluorescent cells 15 days after editing (0.1 to 0.2%), suggesting the occurrence of rare LOH at *UROS locus*. Next, we sorted the fluorescent cells (Fig.1c), and analyzed them by scDNA-seq. We selected five SNPs located along the Chr10 (Fig.1d). As expected, more than 99% of cells had a LOH from the CRISPR cut-site to the telomere. We identified a unique LOH event, with an allele comprising the *UROS* mutation (LOH10qI), validating the method at the allele level (Fig.1d). Importantly, the non-targeted, Chr11 remained fully heterozygous.

scDNA-seq and global UMAP reveals frequent ON-target genotoxicity after double-cut in *HBG1/2* promoters

To address genotoxicity in HSPCs, we edited cells with the gRNA-68 targeting the *HBG1/2* promoters as areported[19]. Then, we performed scDNA-seq to evaluate genome rearrangements using Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction and cluster identification (Fig.2a). We identified the present SNPs in the non-edited HSPCs. Using principal component analysis (PCA) and UMAP, we clustered 5161 edited cells, revealing three populations based on 8 components and 88 SNPs (Fig.2b). Most of cells were heterozygous for all sequenced SNPs (orange) and two minor red and green populations appeared (2.4% and 2.2% for LOH11pI and LOH11pII respectively). In these

cells, 18 SNPs located between the *HBG1/2* cut sites and the Chr11p telomere exhibited AL (pink or blue, Fig.2b right). Identification of CNVs by read-depth showed abnormal ploidy of the chr11p extremity in both LOH populations (Fig.2c). Reduced ploidy (blue) corresponded to extra-large deletion spanning from the cut site to the telomere. In contrast, ploidy of the non-targeted Chr10 remained normal. Altogether, UMAP-based SNP analysis revealed high-genotoxicity with 4.6% of megabase-scale truncations of Chr11p in HSPCs.

scDNA-seq with focus SNP analysis confirms high-terminal copy-loss LOH and reveals interstitial LOH.

To further explore in-depth megabase-scale LOHs telomeric to the cut sites (5 Mb tLOH), we focused on five amplicons containing SNPs with a high sequencing rate (>60% of cells), along the targeted chromosome (Fig.3a). scDNA-seq revealed a high-percentage of cells (7.4%) with tLOH (SNP 3-5-AL), on day 4 (D4). The similar distribution of both repaired alleles, red or blue, demonstrates the robustness of the methods. We then looked for interstitial LOHs (iLOH) by changing SNP3/4 to SNP6/7, 12 and 13.6 kb from the cut site respectively (Fig. 3b). We confirmed the presence of 7.4% of cells with tLOH observed with UMAP. We observed two new cell genome profiles, 2% of cells had AL at SNP6/7 but not at the telomeric SNP5. These iLOH were equally distributed between red and blue alleles (Fig.3b). The overall percentage of LOH was about 9%. No LOH was found on the non-targeted Chr10 confirming the induction of the events by CRISPR-Cas9 DSB. To determine the nature of the i/tLOH (CN or copy-loss (CL)), we analyzed the CNVs (Fig.3c). In cells with tLOH11pI and tLOH11pII, the calculated LOH ploidies were 1.18 and 1.25 respectively, suggesting that around 80% of LOHs were due to a terminal deletion (Fig.3c) and a minority of LOHs were CN. In cells with iLOH11pI and iLOH11pII, the ploidies were 1.21 and 1.44 respectively, suggesting a majority of interstitial deletion (Fig.3c). This show that ploidy analysis allowed a more precise mapping of the iLOHs, from the cut site to *AC104389.16* (43Kb). In summary, we demonstrated the usefulness of SNP-based scDNA-seq to decipher and precisely map the genotoxicity of CRISPR-Cas9 in primary cells.

Palbociclib prevents ON-target genotoxicity in HSPCs

Due to the genotoxicity detected using scDNA-seq (Fig.2/3), we tested palbociclib[18], a CDK4/6 inhibitor, to prevent the occurrence of CRISPR-mediated LOH in clinically relevant HSPCs. First, we validated the dose and exposure duration of palbociclib to transiently block cells in the G0/G1 phases (Fig.4a). Just after thawing, HSPCs are in G0/G1 phases (99%). Whereas cells in the G0/G1 phase tended to disappear after 12h, most of them remained in the G0/G1 phase with palbociclib (88% vs 62% without palbociclib, Fig.4b). The effect was enhanced with a 48h-palbociclib exposure (96% of cells in G0/G1 phase vs 64% without palbociclib) and reversed after removal of palbociclib (Fig.4b, 68% of cells in G0/G1 phase with 48h-palbociclib and a 48h-clearance vs 67% without palbociclib).

We edited *HBG1/2* in presence of palbociclib by collected the cells on D4 and performed scDNA-seq. The total LOH frequency dropped from 9.2% to 2.9% or 2.1% when using palbociclib 24h or 48h after transfection showing for the first time that palbociclib is efficient to protect HSPCs from ON-target genotoxicity. Interestingly, analysis of the LOH profile shows that palbociclib preferably reduced tLOH from 6.95% to 1.44% and 1%, corresponding to a 5 and 7-fold decrease (Fig.4c). We analyzed the ploidy of each LOH ("cGH like" scheme in Extended Data Fig.1). The proportion of deletion and CN-LOH was not modified with palbociclib at 24h (ploidy around 1.3, with 70% of deletion for both conditions). With palbociclib at 48h, the number of events was too low to calculate ploidy (Fig.4d). The decrease in iLOH was more modest (2.25% to 1.45% and 1.1%) (Fig.4c) with identical ploidy (Fig.4d). Therefore, it is possible to prevent LOH by controlling the cell cycle in HSPCs by using palbociclib.

AZD7648 increases genotoxicity in hFFs and HSPCs

Pharmacological modulators "HDR boosters" have been proposed to improve the HDR-editing rate. To evaluate the impact of three HDR boosters on CRISPR-induced genotoxicity on hFFs and HSPCs. We exposed the cells to XL413, an inhibitor of CDC7 and to two DNA-PKcs inhibitors, NU7441 and AZD7648, during HDR editing, with a ssODN template (Extended Data Fig.2a).

First, the "HDR-booster" genotoxicity was quantified using a cytometry-based MN assay. We edited the *UROS locus* in HSPCs and measured the MN rate on D2. AZD7648 or XL413 exposure strongly increased MN formation (up to 3% of MN/nuclear vs 1% in edited HSPCs), showing an increased genotoxicity (Extended Data Fig.2b).

We also edited *ABRAXAS2* in hFFs. To combine short and long-term genotoxicity analyses, we performed MN analysis (D4) and megabase-scale LOH detection by FAMReD (D18). The three "HDR-boosters" tended to increase MN, AZD7648 being the only one to reach a statistically significant effect (Extended Data Fig.2c). AZD7648 was also the only one to induce a significant HDR boost (Extended Data Fig.2d). On D18, both DNA-PKs inhibitors increased LOHs (Extended Data Fig.2e). AZD7648 reached the highest MN rate, LOH frequency and HDR rate.

To confirm the AZD7648-induced genotoxicity, we edited the *HBG1/2* on Chr11p in primary cell-types. AZD7648 dramatically increased MN in HSPCs and hFFs on D4 (reaching to 4.5% in HSPCs and 2.6% in hFFs Extended Data Fig.2f). On D18, AZD7648 promoted LOH (detected by FAMReD) in p53-proficient and deficient hFFs (Extended Data Fig.2g). Because FAMReD only detects 50% of the events, the true long-term LOH is the double (14% of edited cells). Finally, we used DAPI-imaging to confirm MN formation and other nuclear abnormalities such as nuclear buddings, hallmarks of the elimination of amplified DNA and/or DNA repair complexes (Extended Data Fig.2h). Taken together, these data demonstrate that HDR-boosters unevenly induce MN and LOH formation. AZD7648, induces short and

long-term genotoxicity with an immediate cellular impact (nuclear abnormalities) and persistent extra-large LOHs.

AZD7648 delays but increases the onset of genotoxicity in hFFs

To map and determine the nature of LOHs (interstitial or terminal) in the presence or absence of AZD7648, we performed scDNA-seq. After *ABRAXAS2* editing using CRISPR-Cas9 and ssODN template, with/without AZD7648, we still obtained the same three profiles: a majority of cells with SNPs heterozygosity persistence along the Chr10 ("WT"), and two abnormal populations with allele loss of 3 SNPs 3/4/5, either from each parental chromosome (named "LOH10qI" and "LOH10qII", Fig.5a). In "LOH10qI and II" cells, AL began just after the cut site and continued to the Chr10q telomere, revealing tLOH in 1.48% of cells on D2 (Fig.5b). To characterize induced-LOHs, we studied CNVs in conditions with more than five cells in the LOH10qI and II profiles. Most of these LOHs were associated with low ploidy (1.07 and 1.37 corresponding to 93% and 63% of deletions in LOH10qI and LOH10qII respectively, Fig.5c). Unexpectedly, in the presence of AZD7648, we did not find LOH on D2 (Fig. 5b). We monitored the genotoxic events of the same edited cells on D20. At this time and without AZD7648, most of the tLOHs had disappeared (from 1.48% to 0.23%). By contrast, AZD7648 had drastically increased tLOHs (1.37%, a 6-fold increase relative to without AZD7648 Fig.5b, D20), confirming the previous D18 FAMReD results and highlighting the delay in LOH appearance. These LOH were mostly CN (ploidy 1.68 and 1.59, Fig.5c). "cGH like" schemes are shown in Extended Data Fig.3. No LOH was found on the non-targeted Chr11, even with AZD7648 (not shown). These data revealed the long-term genotoxicity of AZD7648 in hFFs which appeared over time. We hypothesized that DNA-PKcs inhibition during a DSB could affect the cell cycle. The occurrence of a DSB concomitantly to AZD7648 and NU7441 exposure forced hFFs to accumulate into the G0 phase at D2 (17% of G0 phase cells with CRISPR compared to 25% with NU7441 (Extended Data Fig.4) and 30% with AZD7648, Fig.5d). These results could explain the delay in the onset of genotoxicity.

AZD7648 increases terminal CN-LOHs in HSPCs

To further characterize the genotoxicity of AZD7648 in HSPCs, we investigated genome editing outcomes. We designed time-course scDNA-seq experiments to track genotoxicity at various time points. These cells were edited in *UROS* in the presence or the absence of AZD7648. By nanopore NGS we confirmed a higher HDR rate with AZD7648 (63.6% *vs* 21.7%) and as expected a modification of the InDels profile (InDels <5 pb due to NHEJ mostly disappeared in favor of larger InDels due to MMEJ, Extended Data Fig.5). HSPCs were cultured and collected on D4,7 and 14 for scDNA-seq. We found three cell populations after editing: "WT" without LOH, tLOHs "LOH 10qI" and "LOH 10qII" (Fig.6a). On D4, CRISPR induced the presence of 1.98% of rearranged cells with tLOH ("LOH 10qI and LOH 10qII"),

confirming the ON-target genotoxicity of CRISPR, even after one DSB (Fig.6b). AZD7648, on D4, significantly increased these tLOHs (7.32%, 3.5-fold relative to CRISPR without AZD7648). Kinetic analysis showed that the frequency of tLOH persisted at 1.52% without AZD7648 and at a higher-level with AZD7648 (6.95 and 7.51% on D7 and D14 respectively, Fig. 6b). No LOH was found on the non-targeted Chr11, even with AZD7648.

To characterize the induced-LOHs, we studied the CNVs. In all conditions with or without AZD7648, and whatever the moment, most of the LOHs were CN (ploidy around 2, Fig.6c). The ploidy pattern along the chromosome is shown in Extended Data Fig.6.

DISCUSSION

Our custom panel for scDNA-seq enabled the precise mapping of the edited Chr10/11. CNV analysis and robust SNPs sequencing allowed to determine the frequency and nature of LOH with high-sensitive detection of CN and CL-LOH (around 0.1%). To assess genomic outcomes, we considered all the SNPs, or focused only on five discriminant SNPs. We obtained comparable tLOH percentages (5% and 7% respectively targeting *HBG1/2* promoters in HSPCs). However, global SNPs analysis did not detect any iLOH, probably because they were rarer. Our focused method enabled us to detect and count the heterogenous ON-target genotoxicity involving several DSB repair pathways and their kinetics. We then used this more focused method for the study. We were able to identify the heterogenous ON-target genotoxicity due to the implication of several DSB repair pathways, and its time-course. Notably, CN-LOH are not detectable by FISH, cGH-array and complicated with scRNA-seq but can be easily detected by scDNA-seq if the amplicons encompass SNPs. This approach can be a key innovative tool to screen pharmacological drugs modulating chromosomal genotoxicity.

It is mandatory to understand the genotoxicity associated with CRISPR-Cas9 editing. Spontaneous LOHs occur infrequently in human cells (<0.01%)[20]. The high-frequency of *in-vitro* CN/CL-LOH starting at the cut site, interstitial or reaching the telomere described in this work cannot be at random. The genotoxicity in HSPCs was more severe with *HBG1/2* than with *UROS* editing, with a higher frequency of tLOH or iLOH mostly associated with deletions probably that were due to two DSBs in homologous sequences[19]. Thanks to the early analysis (D4), we found, for the first time, kilobase/megabase-scale deletions (5.2Mb encompassing 169 genes) in HSPCs, which raises safety concerns. More studies will be required to decipher the time-course of the unwanted genomic by-products and determine whether they disappear in HSPCs like rearranged CAR-T cells[8]. In the absence of telomerase, extra-large CL, including telomeres, are hard to tolerate for primary cells. Recently, Jasin et al reported that neo-telomere can be added[21]. If a neo-telemere cannot be added, we hypothesize

that such events could be rescued by a CN-LOH. Targeting two loci in *HBB* and *HBD* in HSPCs[22], we previously observed large CN-LOHs by subcloning and long-term amplification (D30) that could induce paternal uniparental disomy with bi-allelic paternal imprinting pattern resulting in transcriptional modifications in the *H19/IGF2* parental genomic imprinting-center[7]. Extended *in-vivo* experiments (HSPCs graft in immunodeficient mice) will be required to evaluate (i) the graft capability of these rearranged edited HPSCs, (ii) the long-term phenotypic consequences, and (iii) the putative selective advantage/clonogenicity of LOH clones. It is a critical issue although these events are rare, given that millions of edited cells need to be transplanted in clinical practices.

We studied the risk associated with DNA-repair modulators. Several "HDR-boosters" can improve HDR, such as XL413[23] and DNA-PKcs inhibitors (NU7441 and AZD7648), with impressively increased PGE[24-27] or targeted integration enhancer[28,29]. scDNA-seq showed that AZD7648 is risky since it induces ON-target megabase-scale genotoxicity, with an increase in tLOH in two primary cell types targeting two different *loci*. We observed a delay in the onset of genotoxicity with AZD7648 at D2, making it necessary to wait a few days before analyzing cells.

Our scDNA-seq results were confirmed by a rapid increase in MN and other nuclear abnormalities, which are the hallmarks of chromosomal missegregation[30]. If the chromosomal DSB-mediated Cas9 is not repaired before cell division, the telomeric fragment lacking a functional centromere can missegregate, forming a micronucleus that can lead to instability[9]. Corn *et al* very recently reported the genotoxicity of AZD7648[31] (translocations and kilobase-scale deletions in HSPCs) and Jasin et al reported increased LOHs using the inhibitors of PolQ/DNA-PKcs[21]. This is in agreement with Kosicki and al[32] describing rearrangement in NHEJ-deficient mouse ES cells. We now demonstrate that AZD7648 increases mega-base-scale CN-LOHs on D4, 7 and 14 in HSPCs. By blocking rapid synapse formation by NHEJ machinery at the DSBs, DNA-PKcs inhibition could increase the persistence of DSBs, thereby dissociating the ends from their partners and risking of long-term extra-large LOH. LOHs persisted after 2 weeks in HSPCs and after 3 weeks in fibroblasts treated with AZD7648, without selective disadvantage. This genotoxicity linked to DNA repair pathway modulation by drugs, is compatible with their use in oncology to sensitize cancer cells to chemo/radiotherapies/PARP inhibitors. These treatments induce a catastrophic genomic instability[27,33-37]. The error-prone NHEJ is above all a conservative way to maintain chromosomal integrity.

We also demonstrated the accuracy of scDNA-seq to evaluate protocols that mitigate ON-target genotoxicity. For the first time, we show that palbociclib drastically reduces LOHs in HSPCs. Further *in-vivo* studies are required to test its innocuity in preserving the stemness of the edited cells graft. Our results agree with (i) our previous publication showing that blocking G0/G1 prevents LOH in fibroblasts[18], (ii) a recent paper[12] showing that non-proliferating CAR-T editing is safer than proliferating T-cell editing, and (iii) a paper reporting that low HSPC cycling during editing prevented

genotoxicity[38]. These recent data confirmed the role of the cell cycle in the induction of genotoxicity during editing, and prove the power of scDNA-seq to screen the safest protocols. Alternatively, new DSB-less CRISPR tools (e.g. base, prime and click editors[39-43]) are thought to be less genotoxic. However, recent data reported unintended ON-target genotoxicity using base/prime editing in HSPCs[44,45]. scDNA-seq could therefore become an option to assess their safety.

Our panel can be used to map kilobase/megabase-scale rearrangements. However, it cannot capture the whole genome picture. A recent custom panel[17] measure genome wide rearrangement detected whole chromosome aneuploidies. However, it could not reliably detect aneuploidies in shorter regions. Further optimizations of such panels combining local and whole genome coverage will allow the precise global analysis of genotoxicity.

In conclusion, inferring SNP analysis and CNV alteration from scDNA-seq data is an unbiased high-throughput method to identify the heterogeneity of nuclease-induced chromosomal aberrations. Our discovery of a high-rate of persistent LOHs, in particular with AZD7648, highlights potential oncogenic risks but proposes a safer solution by using palbociclib. scDNA-seq should be implemented as a quality control step before the transplantation of CRISPR-Cas9-edited cells

Figure legends

Fig.1: scDNA-seq design to assay CRISPR-induced ON-target genotoxicity in Chr10 and 11 and method validation for LOH quantification. a, Design of home-made scDNA-seq library in Chr10 and 11 focusing on frequent SNPs. **b,** Mapping of chromosomes 10 and 11 by amplicon density. **c, upper panel,** hFAMRed system based on *UROS*$^{+/-}$ hFFs. hFFs switched from non-fluorescent to fluorescent, in the event of CRISPR-induced 10q telomeric LOH. *ABRAXAS2* targeting by CRISPR-Cas9 induced fluorescent cells 15 days after editing (0.1 to 0.2%). Fluorescent cells were sorted and cultured for 2 months before analysis by scDNA-seq and Mosaic 3 software. **d, Top Left panel,** Chr10 representation encompassing *UROS* and the five SNPs used to detect LOHs by scDNA-seq. Top **Right panel,** scDNA-seq confirmed the presence of only 1 LOH profile (LOH10qI) in fluorescent cells = fluo+. **Lower panel,** SNP genomic position, SNP rs (reference SNP cluster ID), and distance to the DSB are indicated.

Fig.2: UMAP clustering and LOH detection following *HBG1/2* promoter targeting. a, Experimental workflow. *HBG1/2* promoters were targeted using CRISPR-Cas9, followed by single-cell DNA sequencing (scDNA-seq) analysis on day 4. Dimensionality reduction was performed using principal component analysis (PCA) to capture major variance components, followed by uniform manifold approximation and projection (UMAP) for non-linear embedding. Clustering was performed using

hierarchical density-based spatial clustering of applications with noise (HDBSCAN) with a minimum cluster size of 22, enabling the identification of distinct genomic subpopulations. **b**, UMAP projection and SNP-based classification of single-cells. **Left panel**, UMAP representation of 5,161 analyzed cells (99.47% retained), with three distinct profiles: no LOH (orange), LOH11pI (red), and LOH11pII (green). **Right panel**, Convolved heatmap displaying all heterozygous SNPs in cell population on chr11 diversity across groups. **c,** CNV analysis confirming LOH events. Averages of read depth of each amplicon were calculated in non-rearranged cells ("WT") and used as a reference for a ploidy of 2. **Left panel**, heatmap representation of CNV across samples, highlighting copy-loss LOH regions (blue). **Right panel**, chromosomal copy number profiles for LOH11pI (red box) and LOH11pII (green box), showing a reduction in ploidy at the targeted Chr11 compared to the control Chr10. The position of the DSB is indicated.

Fig.3: *HBG1/2* promoter targeting by CRISPR induces a wide range of LOHs in HSPCs. a, upper panel, HSPCs cells from cord blood (CB) were transfected with Cas9 nuclease RNP and gRNA targeting *HBG1/2* promoters. Five SNPs along the Chr11 were chosen, one on the non-targeted region 11q (SNP1), one in Chr11p between the centromere and the cut site *HBG1/2* (SNP2) and three telomeric to the cut site at 0.3, 2.8 and 4.6 Mb, (SNP 3-5). scDNA-seq revealed the presence of 3 LOH profiles (no LOH = WT, tLOH I (LOH11pI, red allele), tLOH II (LOH11pII, blue allele). % of LOH is indicated at day 4 after transfection. **lower panel**: SNP genomic position, SNP rs, and distance to the DSB are indicated. **b, upper panel,** A novel set of SNPs was chosen using two interstitial SNPs (6 and 7) on the telomeric side. scDNA-seq revealed the presence of 2 novel interstitial LOH profiles (iLOH). % of LOH is indicated at day 4 after transfection. **lower panel,** SNP genomic position, SNP rs, and distance to the DSB are indicated. **c, left panel**, Ploidies and CGH -like representation (CNV analyses) along the targeted Chr11 for the 5 profiles (WT, LOH11pI, LOH 11pII, iLOH11pI and iLOH11pII), and of the non-targeted Chr10 (negative control). Ploidies are calculated by the estimated ploidy averages of each amplicon between the cut site and the telomere **right panel,** cumulative histogram of copy-neutral and copy-loss LOHs in each profile.

Fig.4: Palbociclib prevents genotoxicity in HSPCs. a, experimental design, HSPCs cells were edited with Cas9 RNP and gRNA targeting *HBG1/2* promoters with or without palbociclib exposure (1 µM, overnight before transfection and for 1 or 2 days after transfection). **b, Left panel,** illustrative histogram of initial cell cycle and after 12h and 48h of cell culture with or without palbociclib. **Right panel,** representative FACS analyses of cell cycles with or without palbociclib in HSPC medium at 12 hours or 48 hours. **c,** % of LOHs (interstitial and terminal) at day 4 in each condition. **d**, ploidy for each LOH type.

Fig. 5: AZD 7648 delays but increases long-term ON-target genotoxicity targeting *ABRAXAS2* in hFFs. A, upper panel, hFFs were transfected with Cas9 RNP, ssODN template and gRNA targeting *ABRAXAS2* with or without AZD7648 exposure (1 µM, 18 h before transfection and for 2 days after transfection). Five SNPs along the Chr10 were chosen. scDNA-seq revealed the presence of 3 LOH profiles (no LOH = WT, tLOH I (LOH10qI, red allele), tLOH II (LOH10qII, blue allele). % of LOH is indicated at day 2 and day 20 after transfection. **lower panel**: SNP genomic position, SNP rs, and distance to the DSB are indicated. **b,** % of LOH is indicated at day 2 and day 20 after transfection with or without CRISPR-Cas9 and AZD7648 exposure. Percentages of LOHs in HSPCs were compared by the Chi-square test. **c,** % of copy-neutral and copy-loss LOHs for experimental conditions with more than 5 cells by LOH profile. **d, upper panel,** illustrative histogram of cell cycle analysis at day 2 in hFFs, with or without AZD7648. DMSO used as negative control. **lower panel,** illustrative cytometry graphs

Fig. 6: Time course of AZD7648 effect on ON-target genotoxicity targeting *UROS* in HSPCs. a, upper panel, HSPCs were transfected with Cas9 RNP, ssODN template and gRNA targeting *UROS* with or without AZD7648 (1 µM, 18 h before transfection and for 2 days after transfection). 5 SNPs along the Chr10 were chosen. scDNA-seq revealed the presence of three LOH profiles (no LOH = WT, tLOHI (LOH10qI, red allele), tLOH II (LOH10qII, blue allele). % of LOH is indicated at day 4 day 7 and day 14 after transfection. **lower panel**: SNP genomic position, SNP rs, and distance to the DSB are indicated. **b,** % of LOH is indicated at day 4 day 7 and day 14 after transfection with or without CRISPR-Cas9 and AZD7648 exposure. Percentages of LOHs in HSPCs were compared by the Chi-square test. **c,** % of copy-neutral and copy-loss LOHs for experimental conditions with more than 5 cells per LOH profile.

Extended Data

Extended Data Fig. 1: Ploidy analysis and cGH-like representation of LOH+ HSPC populations targeting *HBG1/2* promoters in presence of palbociclib

Extended Data Fig. 2: HDR boosters unevenly induce MN and LOH formation.

Extended Data Fig. 3: Ploidy analysis and cGH-like representation of LOH+ hFF populations after *ABRAXAS2* targeting in presence or absence of AZD7648

Extended Data Fig. 4: Cell cycle phase distribution in hFFs targeted or not in *ABRAXAS2* in presence of ssODN template and HDR boosters

Extended Data Fig. 5: AZD efficiency in HSPCs monitored by NGS analysis

Extended Data Fig. 6: Ploidy analysis and cGH-like representation of LOH+ HSPC populations after *UROS* targeting in presence or absence of AZD7648 at day 4, day 7 and day 14.

Extended Data Table 1: list of selected amplicons in Chr10 and Chr11

Acknowledgements

Tapestri equipments from Mission Bio was acquired at Gustave Roussy Institute thanks to MyProbe ANR-17-RHUS-0008. This work was supported by the French national research agency (ANR) (grants ANR-21-CE18-0002 and ANR-PRME-23-CE52 and by la Fondation pour la Recherche Médicale (FRM, grant EQU202403018062). C.B. was supported by funding from INSERM (Poste accueil INSERM). C.T. is supported by AFM telethon (#24581). We thank Ivan Lukic and Stephane Collinet from Mission Bio for their advice and technical support for Mission Bio software. We thank the FACSility cytometry platform at Bordeaux University (TBMCore, Bordeaux, France). We thank Sandrine Hamon, Patrice Dutrinus and Stephanie Lannelongue for administrative assistance and financial management.

Author contribution

V.M, S.F., J.B., A.B., F.M.G. design, analysis of data and drafting of the manuscript. **I.L-G, S.F., M.R**. **C.B.**: cell culture, CRISPR editing, FACS experiments. **C.T.**: nanopore sequencing and analysis; **N.D**., **M.F.**: scDNA-seq experiments. **A.P**.: cell cycle protocol development. **V.M., J.B.**: bioinformatics analysis; **J.T., J.B.**: FISH experiments. **S.D**., **J.B.**, **D.C**.: helpful discussion and revision of the manuscript. **A.B., F.M.G**.: supervision, funding acquisition, analyzing data, writing the manuscript, final approval of the manuscript. All authors edited and approved the final manuscript.

Disclosures

S.F., A.B., and F.M.G declared a patent application EP23305760 filed on May 13th 2023 and claims the use of palbociclib to prevent genotoxicity induced by nuclease. The remaining authors declare no competing interests.

References

1. Rouet P., Smih F. & Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl Acad. Sci. USA* **91**, 6064–6068 (1994).
2. Jasin M., Haber J. E. The democratization of gene editing: Insights from site-specific cleavage and double-strand break repair. *DNA Repair* **44**, 6–16 (2016).

3. Doudna J. A., Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
4. Cong L., Ran F.A., Cox D. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
5. Kosicki M., Tomberg K., Bradley A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol* **36**, 765–771 (2018).
6. Cullot G., Boutin J., Toutain J., et al. CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat. Commun* **10**, 1136 (2019).
7. Boutin J., Rosier J., Cappellen D., et al. CRISPR-Cas9 globin editing can induce megabase-scale copy-neutral losses of heterozygosity in hematopoietic cells. *Nat Commun* **12**, 4922 (2021).
8. Nahmad A.D., Reuveni E., Goldschmidt E., et al. Frequent aneuploidy in primary human T cells after CRISPR-Cas9 cleavage. *Nat Biotechnol* **40**, 1807-1813 (2022).
9. Leibowitz M.L., Papathanasiou S., Doerfler P.A., et al. Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *Nat Genet* **53**, 895-905 (2021).
10. Papathanasiou S., Markoulaki S., Blaine L.J. et al. Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing. *Nat Commun* **12**, 5855 (2021).
11. Boutin J., Cappellen D., Rosier J. et al. ON-target adverse events of CRISPR-Cas9 nuclease: more chaotic than expected. *CRISPR J* **5**, 19–30 (2022).
12. Tsuchida C.A., Brandes N., Bueno R., et al. Mitigation of chromosome loss in clinical CRISPR-Cas9-engineered T cells. *Cell* **21**, 4567-4582 (2023).
13. Alanis-Lobato G., Zohren J., McCarthy A. et al. Frequent loss of heterozygosity in CRISPR-Cas9-edited early human embryos. *Proc Natl Acad Sci U S A* **118**,e2004832117 (2021).
14. Zuccaro M.V., Xu J., Mitchell C., et al. Allele-Specific Chromosome Removal after Cas9 Cleavage in Human Embryos. *Cell* **183,**1650-1664 (2020).
15. Ruff D.W., Dhingra D.M., Thompson K., et al. High-Throughput Multimodal Single-Cell Targeted DNA and Surface Protein Analysis Using the Mission Bio Tapestri Platform. *Methods Mol Biol* **2386**, 171-188 (2022).
16. Borsi E., Vigliotta I., Poletti A., et al. Single-Cell DNA Sequencing Reveals an Evolutionary Pattern of CHIP in Transplant Eligible Multiple Myeloma Patients. *Cells* **13**, 657 (2024).
17. Mays J.C., Mei S., Kogenaru M. et al. KaryoTap Enables Aneuploidy Detection in Thousands of Single Human Cells. *bioRxiv* doi: 10.1101/2023.09.08.555746, 29 september 2024, pre-print: Not peer-reviewed (2024).
18. Cullot G., Boutin J., Fayet S., et al. Cell cycle arrest and p53 prevent ON-target megabase-scale rearrangements induced by CRISPR-Cas9. *Nat Commun* **10**, 4072 (2023).

19. Sharma A., Boelens J.J., Cancio M., et al. CRISPR-Cas9 Editing of the HBG1 and HBG2 Promoters to Treat Sickle Cell Disease. *N Engl J Med* **389**, 820-832 (2021).
20. Davis L., Khoo K.J., Zhang Y. et al. POLQ suppresses interhomolog recombination and loss of heterozygosity at targeted DNA breaks. *Proc Natl Acad Sci U S A* **117**,22900-22909 (2020).
21. Regan S.B., Medhi D., White T.B. et al Megabase-scale loss of heterozygosity provoked by CRISPR-Cas9 DNA double-strand breaks. bioRxiv. doi: 10.1101/2024.09.27.615517. September 2024, pre-print: Not peer-reviewed (2024).
22. Antoniani C., Meneghini V., Lattanzi A. Induction of fetal hemoglobin synthesis by CRISPR/Cas9-mediated editing of the human β-globin locus. *Blood* **131**, 1960–1973 (2018).
23. Wienert B., Nguyen D.N., Guenther A. et al. Timed inhibition of CDC7 increases CRISPR-Cas9 mediated templated repair. *Nat Commun* **11**, 2109 (2020).
24. Wimberger S., Akrap N., Firth M., et al. Simultaneous inhibition of DNA-PK and PolΘ improves integration efficiency and precision of genome editing. *Nat Commun* **14,** 4761 (2023).
25. Cloarec-Ung FM, Beaulieu J, Suthananthan A et al. Near-perfect precise on-target editing of human hematopoietic stem and progenitor cells. *Elife* **12**, RP91288 (2024).
26. Robert F., Barbeau M., Éthier S., et al. Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med* **7**, 93 (2015).
27. Fok J.H.L., Ramos-Montoya A., Vazquez-Chantada M. et al. AZD7648 is a potent and selective DNA-PK inhibitor that enhances radiation, chemotherapy and olaparib activity. *Nat Commun* **10**, 5065 (2019).
28. Selvaraj S., Feist W.N., Viel S., et al. High-efficiency transgene integration by homology-directed repair in human primary cells using DNA-PKcs inhibition. *Nat Biotechnol* **42**, 731-744 (2024).
29. Stack J.T., Rayner R.E., Nouri R. et al. DNA-PKcs inhibition improves sequential gene insertion of the full-length CFTR cDNA in airway stem cells. *Mol Ther Nucleic Acids* **35**, 102339 (2024).
30. Podrimaj-Bytyqi A., Borovečki A., Selimi Q., et al. The frequencies of micronuclei, nucleoplasmic bridges and nuclear buds as biomarkers of genomic instability in patients with urothelial cell carcinoma. *Sci Rep* **8**,17873 (2018).
31. Cullot G., Aird E.J., Schlapansky M.F. et al. Genome editing with the HDR-enhancing DNA-PKcs inhibitor AZD7648 causes large-scale genomic alterations. *Nat Biotechnol* (2024).
32. Kosicki M., Allen F., Steward F. et al. Cas9-induced large deletions and small indels are controlled in a convergent fashion. *Nat Commun* **13**, 3422 (2022).
33. Laroche-Clary A., Josensi C., Derieppe M.A. et al. Selective DNA-PK Inhibition Enhances Chemotherapy and Ionizing Radiation Activity in Soft-Tissue Sarcomas. *Clin Cancer Res* **30**, 629-637 (2024).

34. Deng L., Yang L., Zhu S. et al. I Identifying CDC7 as a synergistic target of chemotherapy in resistant small-cell lung cancer via CRISPR/Cas9 screening. *Cell Death Discov* **9**, 40 (2023).
35. Liu S., Deng P., Yu Z., et al. CDC7 Inhibition Potentiates Antitumor Efficacy of PARP Inhibitor in Advanced Ovarian Cancer. *Adv Sci* (Weinh) **11**, e2403782 (2024).
36. Wang J., Sadeghi C.A., Frock R.L. DNA-PKcs suppresses illegitimate chromosome rearrangements. *Nucleic Acids Res* **52**,5048-5066 (2024).
37. Hung K.L., Yost K.E., Xie L. et al. ecDNA hubs drive cooperative intermolecular oncogene expression. *Nature* **600**,731-736 (2021).
38. Zeng, J., Nguyen, M. A., Liu, P. et al. Gene editing without ex vivo culture evades genotoxicity in human hematopoietic stem cells. *Cell Stem Cell* **32**,191-208.e11 (2025).
39. Komor, A. C., Kim, Y. B., Packer, M. S., et al. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature.* **533**, 420–424 (2016).
40. Gaudelli, N. M. et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).
41. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet* **19**, 770–788 (2018).
42. Anzalone, A.V. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
43. Ferreira da Silva, J., Tou, C.J., King, E.M. *et al.* Click editing enables programmable genome writing using DNA polymerases and HUH endonucleases. *Nat Biotechnol* (2024).
44. Fiumara, M., Ferrari, S., Omer-Javed, A. *et al.* Genotoxic effects of base and prime editing in human hematopoietic stem cells. *Nat Biotechnol* **42**, 877–891(2024).
45. Huang M.E., Qin Y., Shang Y. et al. C-to-G editing generates double-strand breaks causing deletion, transversion and translocation. *Nat Cell Biol* **26**, 294–304 (2024).
46. Bloh K., Kanchana R., Bialk P., et al. Deconvolution of Complex DNA Repair (DECODR): Establishing a Novel Deconvolution Algorithm for Comprehensive Analysis of CRISPR-Edited Sanger Sequencing Data. *CRISPR J* **4**,120-131 (2021).

Ethics statement

Our research complies with all relevant ethical regulations. Cord blood Human $CD34^+$ HSPCs were obtained in collaboration with EFS Nouvelle Aquitaine-Limousin (24SPL 02 agreement for the cessation

of between EFS and Bordeaux University), according to the hospital's ethical institutional review board and with the mother's informed consent.

Cell culture

Human foreskin fibroblasts immortalized with hTERT (hFFs) were from ATCC® (CRL 4001, BJ-5ta). They were partially inactivated for *UROS* using a ribonucleoprotein (RNP) made of Cas9 protein complexed with a gRNA targeting *UROS* exon 4[18]. hFFs were maintained in DMEM, high-glucose (4.5 g.L-1), L-Glutamine (1 g.L-1) and pyruvate (Gibco® by Life-technologies™, Carlsabad, CA, USA, catalog n°31966047) supplemented with 10% fetal bovine serum FBS (Eurobio, Les Ulis, France, Catalog #CVFSVF00-01), 100 U/mL penicillin and 100µg/mL streptomycin (Gibco® by Life-technologies™, Carlsabad, CA, USA; catalog #15070063), 10 µg/mL ciprofloxacin (Biogaran™, Colombes, France) and 0.5 µg/mL amphotericin B (Sigma-Aldrich®, Saint Louis, MO, USA, catalog #A2942).

Human $CD34^+$ HSPCs were isolated from the cord blood. Briefly, mononuclear cells were isolated by Ficoll gradients. HSPCs were purified according to the manufacturer's instructions (hCD34-Positive Selection kit II, from Stem Cell Technologies, Vancouver, BC, Canada, catalog #17865) and purity was analyzed by flow cytometry using phycoerythrin (PE)-conjugated anti-CD34 antibody (clone 561, Biolegend,San Diego, CA, USA). Cryopreserved HSPCs were thawed and cultured in expansion medium consisting in StemSpan SFEM (Stem Cell Technologies, catalog #09600) supplemented with Flt3-L (100 ng/mL, catalog #GMP300-19-50UG), SCF (100 ng/mL, catalog #GMP300-07-50UG), hTPO (100 ng/mL, catalog #GMP300-18-50UG), (all from Peprotech, Cranbury, NJ, USA), 100 U/mL penicillin, and 100 µg/mL streptomycin (Gibco® by Life-technologies™, Carlsabad, CA, USA, catalog n°15070063), 10 µg/mL ciprofloxacin (Biogaran™, Colombes, France) and 0.5 µg/mL amphotericin B (Sigma-Aldrich®, Saint Louis, MO, USA, catalog n°A2942).

All cells were grown at 37°C with 5% CO_2 levels. All cell lines were tested for mycoplasma.

To increase HDR, we added 3 "HDR boosters" in the medium for editing : NU7441 (2.5 µM, 18 h before transfection and for 2 days after transfection, from Selleckchem, Houston, US, catalog #S2638), AZD7648 (1 µM, 18 h before transfection and for 2 days after transfection, from MedChemExpress, Monmouth Junction, New Jersey, US, catalog #HY-111783) or XL413 (20 µM, for 2 days after transfection, no cell exposure performed before transfection, from Selleckchem, Houston, US, catalog #S7547). Dimethyl sulfoxide (DMSO, from Sigma-Aldrich®, Saint Louis, MO, USA, catalog #D2438), used to dissolve these compounds was used as a control.

To prevent genotoxicity, cells were synchronized in G0/G1 phase by incubation with palbociclib (1 µM, Sigma Aldrich, Saint Louis, MO, USA, catalog #PZ0383) 18 h before and 24h or 48h (for scDNA-seq) after RNP transfection.

Editing tools

The crRNA for the *ABRAXAS2* and *UROS* targets (Ch10q) was designed using CHOP-CHOP software (chopchop.cbu.uib.no) (*ABRAXAS2* 5'-GAGATCCTCCCACTCGATGG-3' and *UROS* exon 4: 5'-GGAAGCAGCAGAGTTATGTT-3'). The crRNA chosen for editing of the *HBG1* and *HBG2* promoters (Chr11p) is an already published gRNA (gRNA-68)[19]. The different components of the CRISPR-Cas9 system were combined to form a RNP (all ordered at IDT, Integrated DNA Technologies, Coralville, IA, USA). For this, 17 µg of Alt-R® S.p. HiFi Cas9 Nuclease V3 (abbreviated hereafter as Cas9, Integrated DNA Technologies, Coralville, IA, USA, catalog #1081059) and 6.5 µM of sgRNA, were mixed and then incubated for 10-20 minutes at room temperature before electroporation. For HDR booster experiments, 6.5 µM of ssODN template was added (*ABRAXAS2* 5'-TCTGCTTCTTCATCCTAGTTCACTCACTTTGCCCCCAGCTGCCT**CCATCGAGTGGGAGG*AA*GAGCTCGA**CAGAG CCTGTTCACTCCTAGtcatttatgcaacagacatcaatggaaacat-3'). Finally, 3.9 µM of Alt-R Cas9 Electroporation Enhancer solution (Integrated DNA Technologies, Coralville, IA, USA, catalog #1075916) was added to the RNP complex to improve electroporation efficiency.

Cells were transfected by electroporation using the Nucleofector 4D AMAXA electroporation system (Lonza®, Bale, Switzerland). In brief, 2.10^5 to 3.10^5 depending on the cell type (HSPCs and hFFs) were resuspended in P3 Primary Cell Line 4D-Nucleofector® (Lonza, Bale, Switzerland, catalog #V4XP-3024) and added to the RNP complex. Then cells were nucleofected using DO-100 and CZ-167 programs respectively.

Editing efficiency

Genomic DNA of edited cells, and their associated controls, was extracted using Nucleospin® Tissue (Macherey-Nagel®, Düren, Germany, catalog #740952.250) according to the manufacturer's protocol. The genomic region flanking the expected cut-site was amplified by PCR (HotStarTaq Plus DNA polymerase, Qiagen®, Venlo, Netherlands, catalog #203605) with adequate primers (*ABRAXAS2* targeting F: 5'-CCTGGCTCACTCACTCTGTG-3', R: 5'-AGGCCAGTTCAAACGCTCTT-3'; *UROS* targeting F: 5'-TAGTTCCAGGCACATAGTAAGCAC-3', R: 5'-TCCCAAGGCAGAGTCTGTGA-3'; *HBG1/2* promoters targeting F: 5'- ACGGCTGACAAAAGAAGTCCT -3', R: 5'-AGCCTTGTCCTCCTCTGTGA-3'). PCR products were purified with Nucleospin® Gel and PCR Clean-up (Macherey-Nagel®, Düren, Germany, catalog #740609.50S). Sanger sequencing was done on purified PCR products and sequenced by LIGHTRUN

(GATC Biotech, Konstanz, Germany). Sanger sequencing data were analyzed using ICE v2 CRISPR Analysis tool (ICE) software (Synthego, Redwood City, CA, USA) and DecodR v3[46]. Purified PCR products from non-edited cells were used as control chromatograms.

To validate AZD7648 efficiency in HSPCs, we used Nanopore NGS. Briefly, at least 5 days after editing, genomic DNA was extracted from cell pellets using Nucleospin® Tissue (Macherey-Nagel®, Düren, Germany, catalog #740952.250) according to the manufacturer's protocol. The genomic region flanking the edited site was amplified by High-Fidelity PCR (Phusion High-Fidelity PCR Master Mix HF Buffer, New England Biolabs, Ipswich, MA, USA, catalog #M0531S) with adequate primers. Sequencing library was prepared using Rapid Barcoding Kit 96 V14 (Oxford Nanopore Technologies, Oxford, United-Kingdom, catalog #SQK-RBK114.96) according to the manufacturer's protocol. Sequencing was then performed using a MinION Mk1B device and a Flow Cell R10.4.1 (Oxford Nanopore Technologies, Oxford, United-Kingdom, catalog #FLO-MIN114) to reach at least 2000 qualify-reads per sample (Q score > 8). After sequencing, super-accurate basecalling was performed using MinKNOW (v24.02.6) and fastq were aligned on human reference genome GrCH38 (GCA_0000001405.15) using Epi2Me (v5.1.9) and wf-alignement pipeline (v1.1.2). Mission bio. Then, BAM were analyzed using a homemade bioinformatic pipelines in Python (v3.8) using pysam (v0.16). Briefly, it calculates the proportion of each nucleotide for each selected position to determine the HDR-editing rate per base returning these data in xlsx format (pandas v1.1.3). https://github.com/victormar1/Nanopore_toolset.

Micronuclei analysis by cytometry

The MicroFlow In Vitro-250/50 Kit (Litron Laboratories, Rochester, NY, USA, catalog #INVITRO250/50) was used for all MN assays, following the manufacturer's instructions. Briefly, cells were stained and lysed following the protocol described in the kit. Following harvest, a photo-activated dye Dye 1, Ethidium monoazide (EMA) was used to stain apoptotic and necrotic cells then the cells were washed. Subsequently, healthy chromatin was stained with the nucleic acid Dye 2, SYTOX Green. Cells could be stored up to 24 hours at ambient temperature, protected from light, or refrigerated for up to 72 hours, prior to analysis. Using a BD Accuri C6 Plus (BD Biosciences, Franklin Lakes, NJ, USA), flow cytometric analysis distinguished 10.000 to 20.000 real MN (SYTOX Green only) from nuclei, based on their weaker DNA-associated fluorescence signal.

MN analysis by nuclear fluorescent imaging

Cytogenetic examination was carried out on interphase nuclei on HSPCs after fixation by a solution of 3:1 methanol/glacial acetic acid fixative. DAPI (Sigma Aldrich, Saint Louis, MO, USA, catalog #MBD0015) staining was performed onto a microscopic slide and MN examination was evaluated

under a microscope Leica DM5500 B (Leica microsystem, Wetzlar, Germany) with the 63x objective lens (i.e. 630x magnification). 300 cells were visualized for each condition. Blind count of micronuclei with a medical cytogeneticist (J.T. in author list).

LOH analysis by FAMReD system (cytometry)

FAMReD (Fluorescence-Assisted Megabase-scale Rearrangements Detection) is a cytometry-based system to detect induced LOHs in Chr10q with or without *TP53* inactivation[18]. On day 15 post-editing of hFFs *UROS*$^{+/-}$, 0.3 mM of 5-ALA (Sigma Aldrich, Saint Louis, MO, USA, catalog #A7793) were added to culture media. After 16 h of exposure (overnight), cells were washed twice with DPBS 1X (Gibco, by Life technologies, Carlsabad, CA, USA, catalog #14190144) and put back in fresh media. Upon LOH, loss of *UROS* can be detected by the appearance of fluorescence due to porphyrin accumulation. Following 8 hours of clearance, fluorescent cells were quantified by flow cytometry using a BD Accuri C6 Plus (BD Biosciences, Franklin Lakes, NJ, USA). UV-sensitive porphyrins were excited at 488 nm and the emitted wavelength was approximately 667 nm, detected by the PE-Cy5A PMT channel.FL-1 is a control green-fluorescent channel used to exclude auto-fluorescent cells. The system only detects 50% of the events, when the cells become UROS $^{-/-}$. The true LOH rate is double.

Cell cycle

To evaluate impact of palbociclib and HDR boosters on hFFs and HSPCs cell cycle, cells were harvested and washed with DPBS (Gibco® by Life-technologiesTM, Carlsabad, CA, USA, catalog #14190144) with 2% FBS (Eurobio, Les Ulis, France, catalog #CVFSVF00-01) and EDTA 2mM (Éthylènediamine-tétraacétate disodium dihydrate, Euromedex, Souffelweyersheim, France, catalog #EU0007). Cells were fixed and permeabilized using the True-Nuclear Transcription Factor Buffer Set following the manufacturer's instructions. True-nuclear Fixation buffer (1X, from True-Nuclear Transcription Factor Buffer Set, 424401, BioLegend, San Diego, California, USA, catalog #424401) was added and incubated for 45 min in the dark at ambient temperature. Subsequently, a true-nuclear 1X Perm Buffer (from the same set, BioLegend, San Diego, California, USA, catalog #424401) was used twice to wash. Cells were then stained with a PE mouse anti-Ki67 antibody (BD Biosciences, Franklin Lakes, NJ, USA, catalog #556027) for 30 min in the dark at ambient temperature and washed with true-nuclear 1X Perm Buffer (from the same set, BioLegend, San Diego, California, USA, catalog n°424401). Pellet was harvested in DPBS with FBS 2% and EDTA 2mM and resuspended in a 2µg/mL Hoechst solution (bisBenzimide H 33342 trihydrochloride, Sigma-Aldrich®, Saint Louis, MO, USA, catalog B2261). Samples were stored at +4°C overnight before analysis performed on a BD LSRFortessa™ Cell Analyzer (BD Biosciences, Franklin

Lakes, NJ, USA). Data were analyzed with BD FACSDiva™ Software (BD Biosciences, Franklin Lakes, NJ, USA).

Design of custom Tapestri panel for sequencing

Panel comprises 403 probes mainly across the 10 & 11 human autosomes (Extended Data Table 1). To identify candidate target regions for the panel, we used the dbSNP (NCBI dbSNP Build 155, hg19) and considered only variant SNPs with a major allele frequency between >0.5 and <0.7 without discrepancy on certain ancestry group (eg global MAF = 0,5 and <0,5 in African and Caucasian ancestry in GnomAd v4). *In Silico* validation of the SNP design was assessed by testing it on the CORIELL (GIAB consortium, sample NA12878, HG001) using bcftools (v1.18). The objective of >30% informative SNPs was largely achieved with 63%. All candidate SNPs were submitted to the Tapestri Panel Designer to generate a panel design and ensured that the designed probes targeted the candidate SNPs and had similar GC contents. Support for the custom panel design and synthesis of the panel was provided by Mission Bio using v3 Tapestri Chemistry (Mission Bio, South San Francisco, CA, USA).

Single-Cell DNA-seq

Single-cell DNA-seq was performed using the Tapestri platform (Mission Bio, South San Francisco, CA, USA) according to the manufacturer's specifications with kit v3. Briefly, cryopreserved human cells were gently thawed, washed, and quantified using a Countess II cell counter (ThermoFisher Scientific, Waltham, MA, USA). The cells were then diluted to a concentration of 2,000 to 3,000 cells per µL in the cell buffer. Next, 35 µL of cell suspension was loaded onto a microfluidics cartridge and cells were encapsulated on the Tapestri instrument followed by cell lysis with protease digestion followed by heat inactivation using a thermal cycler. The cell lysate was reintroduced into the cartridge and processed such that each cell possessed a unique molecular barcode. Amplification of the custom targeted DNA regions was carried out by incubating the barcoded DNA emulsions in a thermocycler following Mission Bio's specifications. Emulsions were broken, DNA digested and purified with AMPure XP beads (Beckman Coulter, Brea, CA, USA, catalog #A63881). The beads were pelleted and washed and DNA libraries were generated. Final libraries were purified with AMPure XP beads. All libraries were sized and quantified using an Agilent Fragment Analyzer and pooled for sequencing on an Illumina NovaSeq6000 S1 flowcell with 2 x 150bp at Gustave Roussy.

Tapestri Single Cell DNA analysis

Single-cell DNA sequencing data were obtained from primary cells and analyzed using the Tapestri Mosaic pipeline and customized versions of Mission Bio's available jupyter workflows (Mission Bio, San

Francisco, CA, USA). Quality control steps included the removal of low-quality or missing genotype cells as per the pipeline's automated filters. Variants were annotated using the default settings provided by the manufacturer in the Tapestri Mosaic suit.

For HSPCs edited in *HBG1/2* promoters, clustering method using principal component analysis (PCA) and UMAP, based on 88 heterozygous SNPs on chromosome 11 (Chr11) was performed. The optimal number of principal components (PCs) was determined through an elbow plot analysis, identifying the inflection point where the explained variance ratio begins to plateau. This transition, occurring around 8 PCs (determined via visual inspection), represents the balance between maximizing data variance retention and minimizing noise contribution. UMAP were then generated using the selected PCs as input, employing the Euclidean distance metric with default Tapestri Mosaic parameters. Clustering was carried out in an unsupervised manner using the *hdbscan* algorithm.

In case of selection of 5 SNPs, variant annotation manufacturer settings were not modified. LOHs analysis was performed by adjusting the max ADO score at 0.2 and the Minimum clone size at 0.01. Only heterozygous SNPs (0.40<AF<0.60) present in at least 60% of the cells were retained.

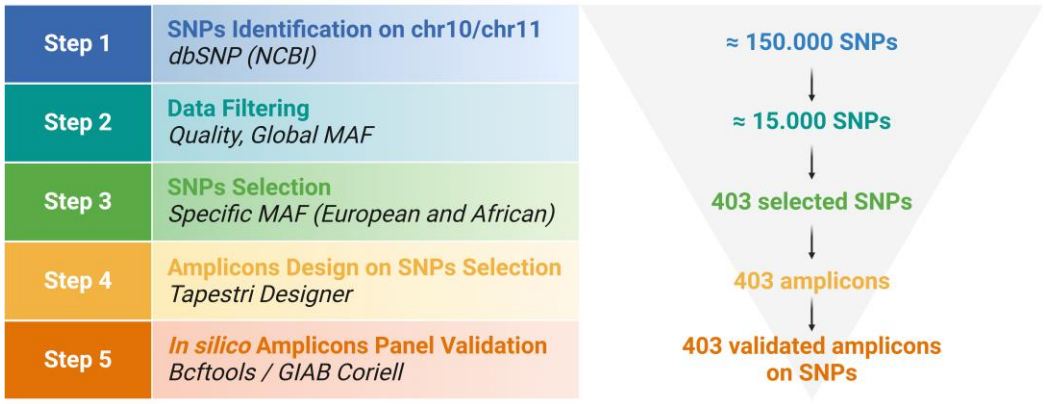
SNP profiles were retained if they contained at least 6 cells. The CNVs analysis was only performed in the LOH subtypes with at least 6 cells. The filters were set up to 50 for amplicon completeness, 10 for amplicon read depth and 10 for mean cell read depth. Ploidy was calculated by averaging the ploidies obtained for all genes affected by LOH.

Statistical analysis and reproducibility

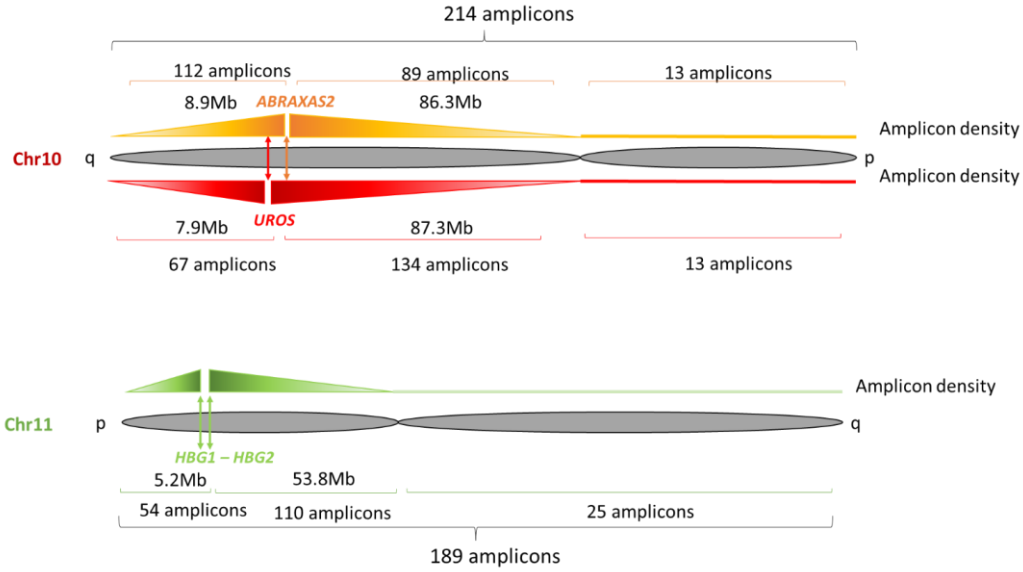
When possible, experiments were conducted at least 3 times independently. No statistical method was used to predetermine sample size. No data were excluded from the analyses. The experiments were not randomized. DAPI analysis was performed blindly. Statistical significance was inferred when necessary. Exact distinct and independent experiments size is indicated in each legend (*n*). Graph Pad Prism 10 software (GraphPad, Boston, MA, USA) was used for statistical analysis. Results are presented as mean ± SD. The parametric *T* test (two-sided) was used when distribution was Gaussian/normal (Shapiro–Wilk test). The non-parametric Mann–Whitney test (one-sided) was used to compare two groups for extended data 2d. One-way ANOVA (two-sided), complemented with the unprotected Fisher's Least Significant Difference test, was used to compare more than two groups. Percentages of LOHs in HSPCs were compared by the Chi-square test.

Fig.1

a

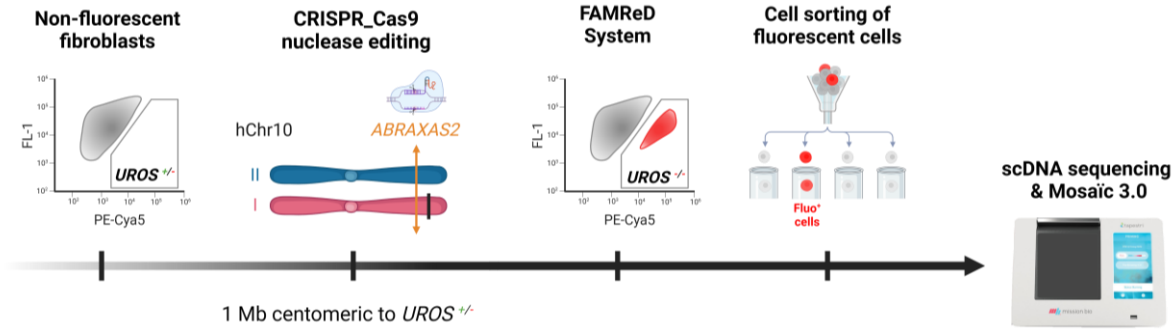


b

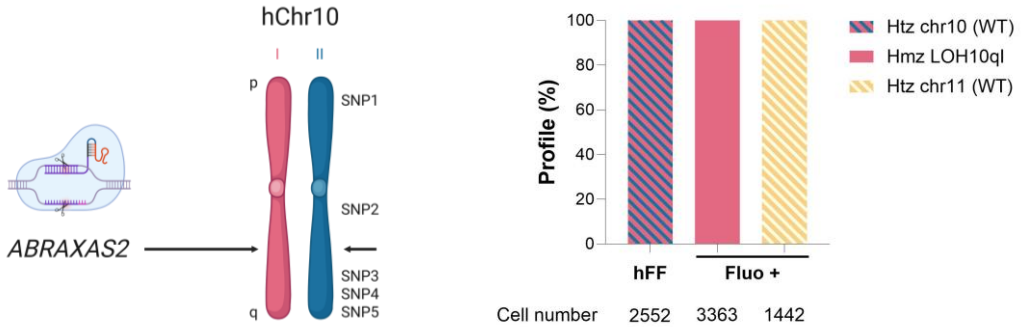


c

FAMReD system



d



Legend	Hg19 Genomic Position	SNP rs	Distance to DSB	WT	LOH 10ql	LOH 10qII
1	OLAH:chr10:15106441:A/G	rs10906818	111 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)
2	WDFY4:chr10:50110232:A/G	rs7069142	76.4 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)
3	ADAM12:chr10:127823151:T/G	rs1278329	1.3 Mb	Htz (T/G)	Hmz (G)	Hmz (T)
4	DOCK1:chr10:129168150:T/G	rs12570609	2.7 Mb	Htz (T/G)	Hmz (G)	Hmz (T)
5	TCERG1L:chr10:132984320:T/C	rs2918100	6.5 Mb	Htz (T/C)	Hmz (C)	Hmz (C)

Fig.2

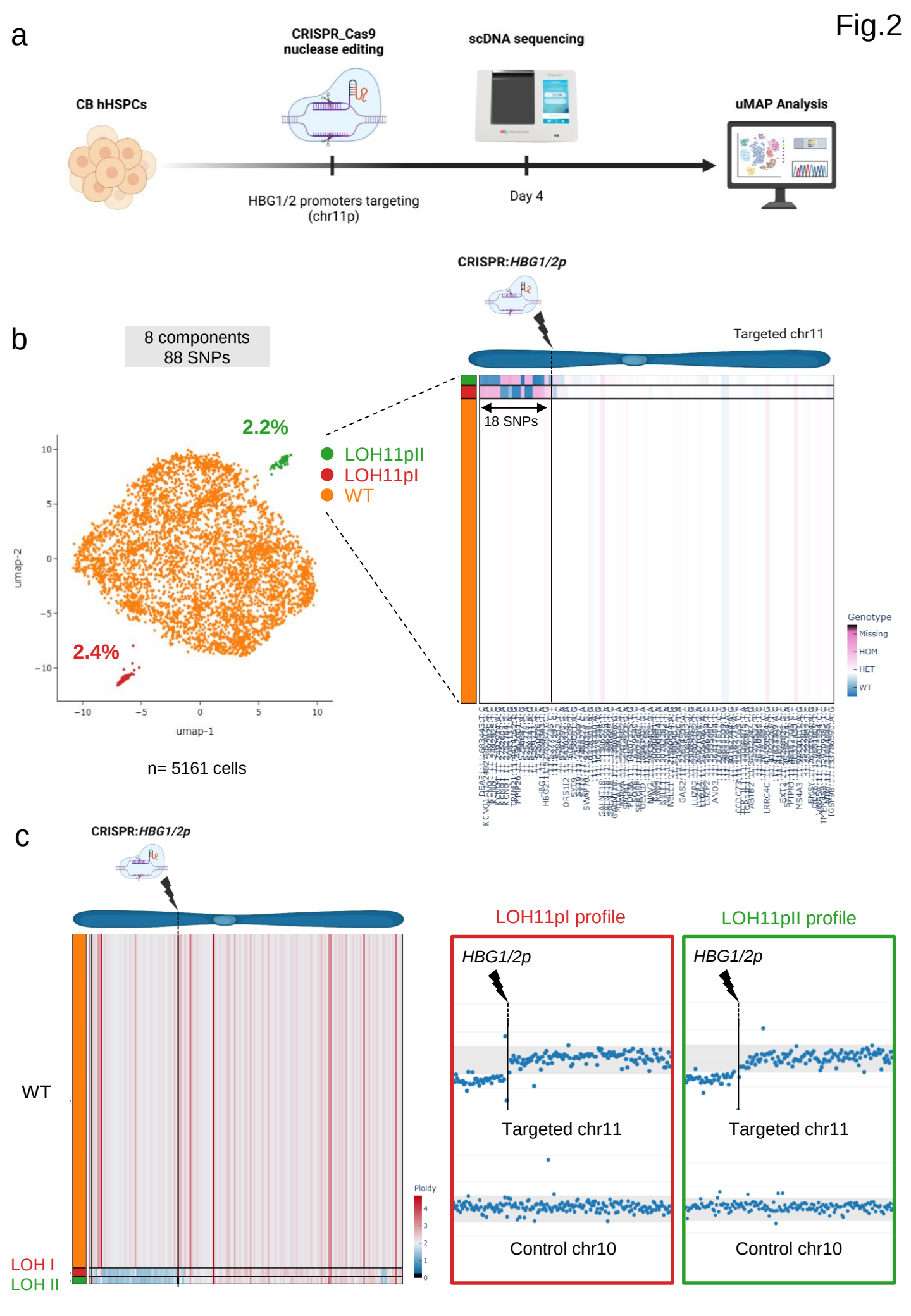
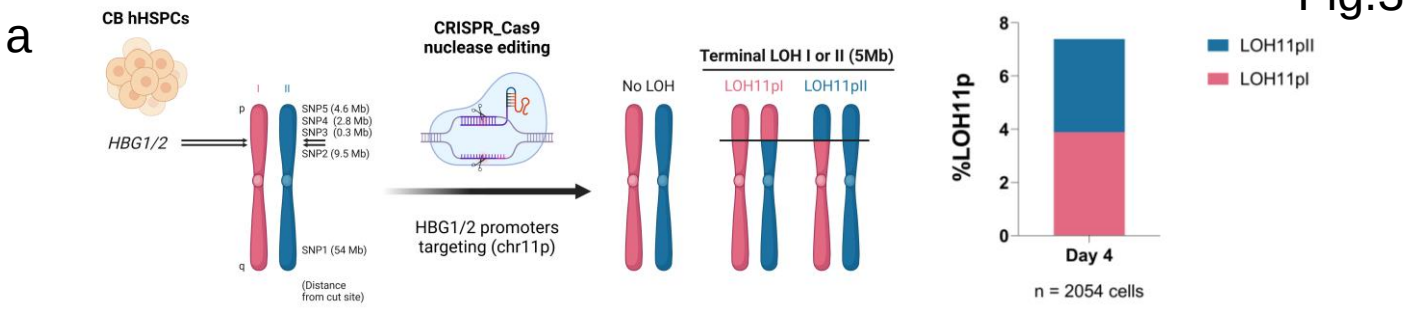

a
CRISPR_Cas9 nuclease editing
scDNA sequencing
CB hHSPCs
uMAP Analysis
HBG1/2 promoters targeting (chr11p)
Day 4
b
CRISPR:HBG1/2p
8 components
88 SNPs
Targeted chr11
2.2%
2.4%
LOH11pII
LOH11pI
WT
umap-2
umap-1
18 SNPs
Genotype
Missing
HOM
HET
WT
n= 5161 cells
c
CRISPR:HBG1/2p
LOH11pI profile
LOH11pII profile
HBG1/2p
HBG1/2p
WT
LOH I
LOH II
Ploidy
Targeted chr11
Targeted chr11
Control chr10
Control chr10

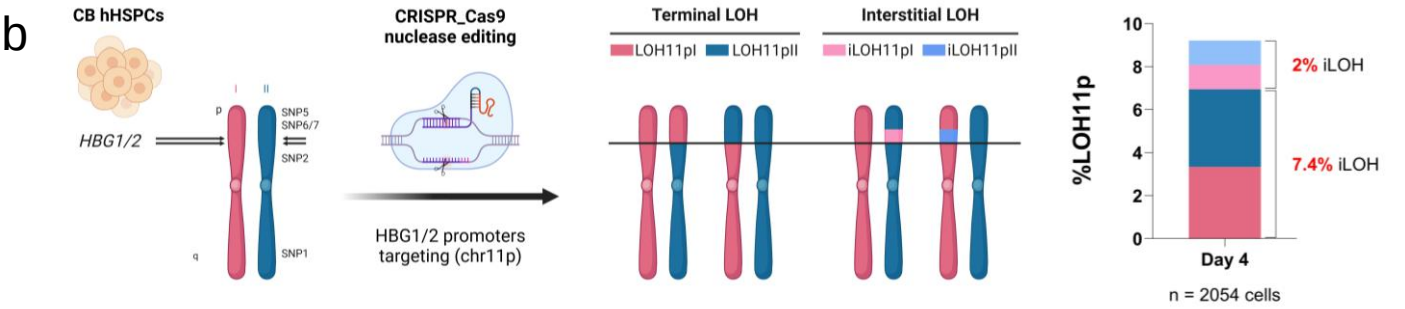
Fig.3

a



Legend	Hg19 Genomic Position	SNP rs	Distance to DSB	WT	LOH 11pI	LOH 11pII
1	MS4A3:chr11:59830103:A/G	rs481950	54 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)
2	PDE3B:chr11:14807375:T/G	rs2122942	9.5 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)
3	MMP26:chr11:4962001:T/C	rs2596002	0.3 Mb	Htz (T/C)	Hmz (T)	Hmz (C)
4	KCNQ1:chr11:2249879:T/C	rs11604319	2.8 Mb	Htz (T/C)	Hmz (C)	Hmz (T)
5	DEAF:chr11:683443:T/C	rs9666056	4.6 Mb	Htz (T/C)	Hmz (C)	Hmz (T)

b



Legend	Hg19 Genomic Position	SNP rs	Distance to DSB	WT	LOH 11pI	LOH 11pII	iLOH 11pI	iLOH 11pII
1	MS4A3:chr11:59830103:A/G	rs481950	54 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)	Htz (A/G)	Htz (A/G)
2	PDE3B:chr11:14807375:T/G	rs2122942	9.5 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)	Htz (A/G)	Htz (A/G)
6	chr11:5258741:A/G	rs16912026	12 kb	Htz (A/G)	Hmz (G)	Hmz (A)	Hmz (G)	Hmz (A)
7	chr11:5257778:G/A	rs3759074	13.6 kb	Htz (G/A)	Hmz (A)	Hmz (G)	Hmz (A)	Hmz (G)
5	DEAF:chr11:683443:T/C	rs9666056	4.6 Mb	Htz (T/C)	Hmz (T)	Hmz (C)	Htz (T/C)	Htz (T/C)

c

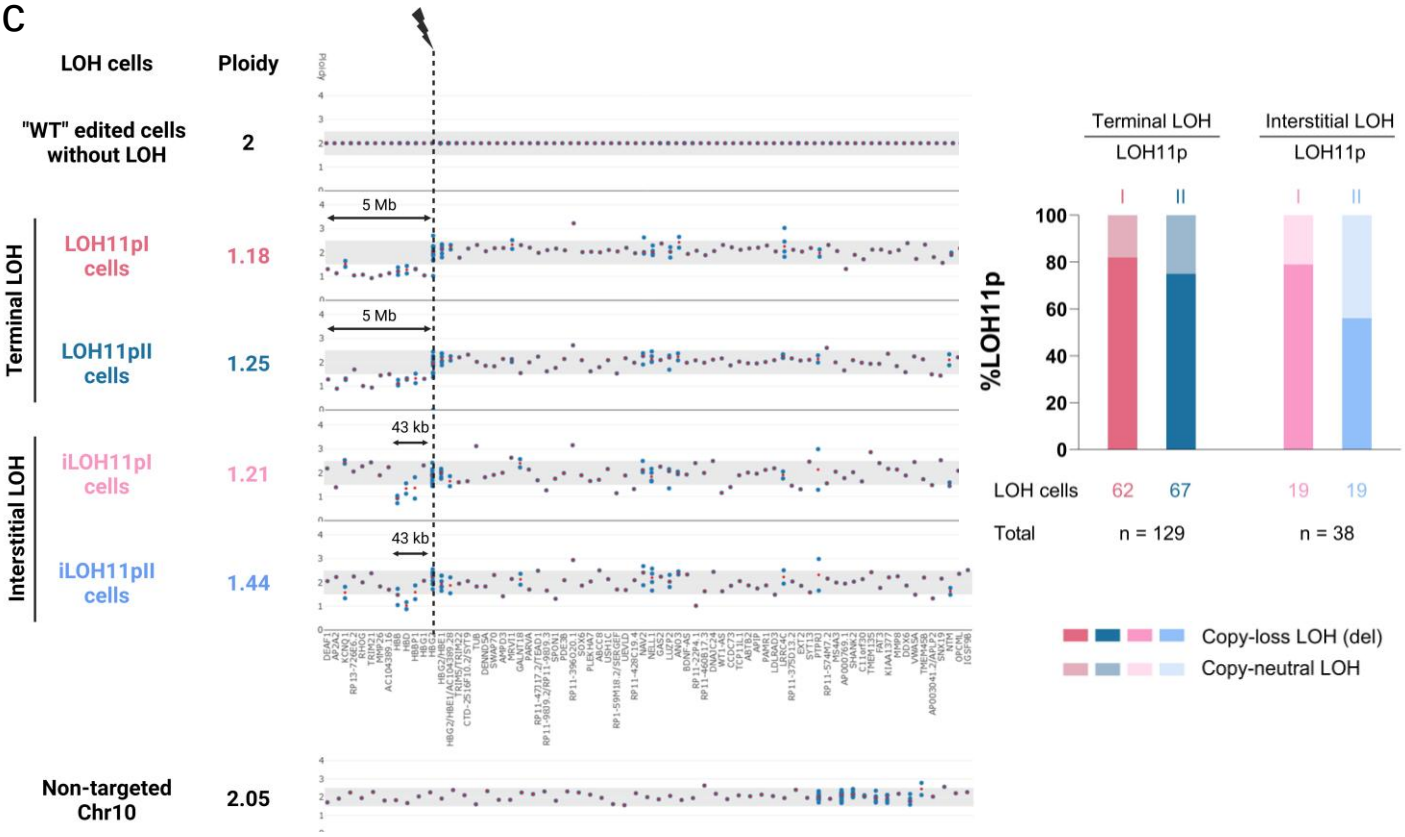
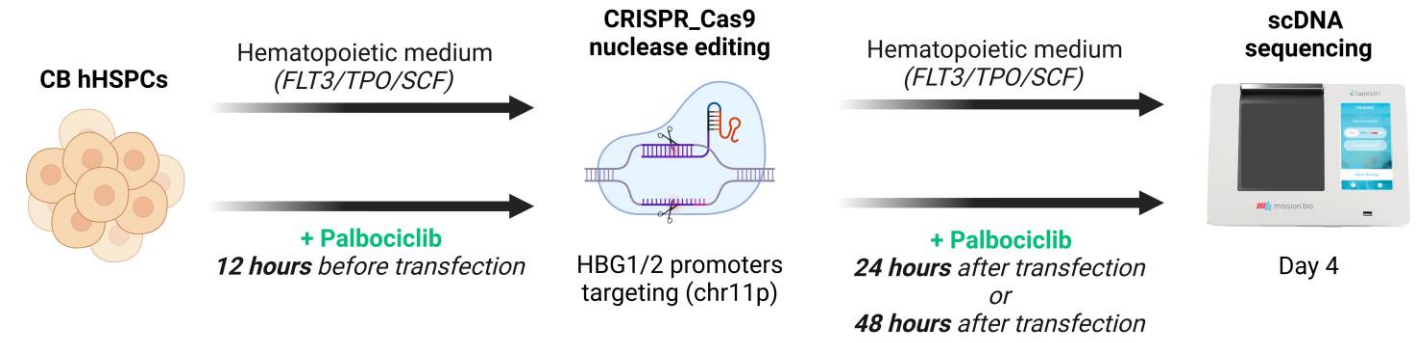
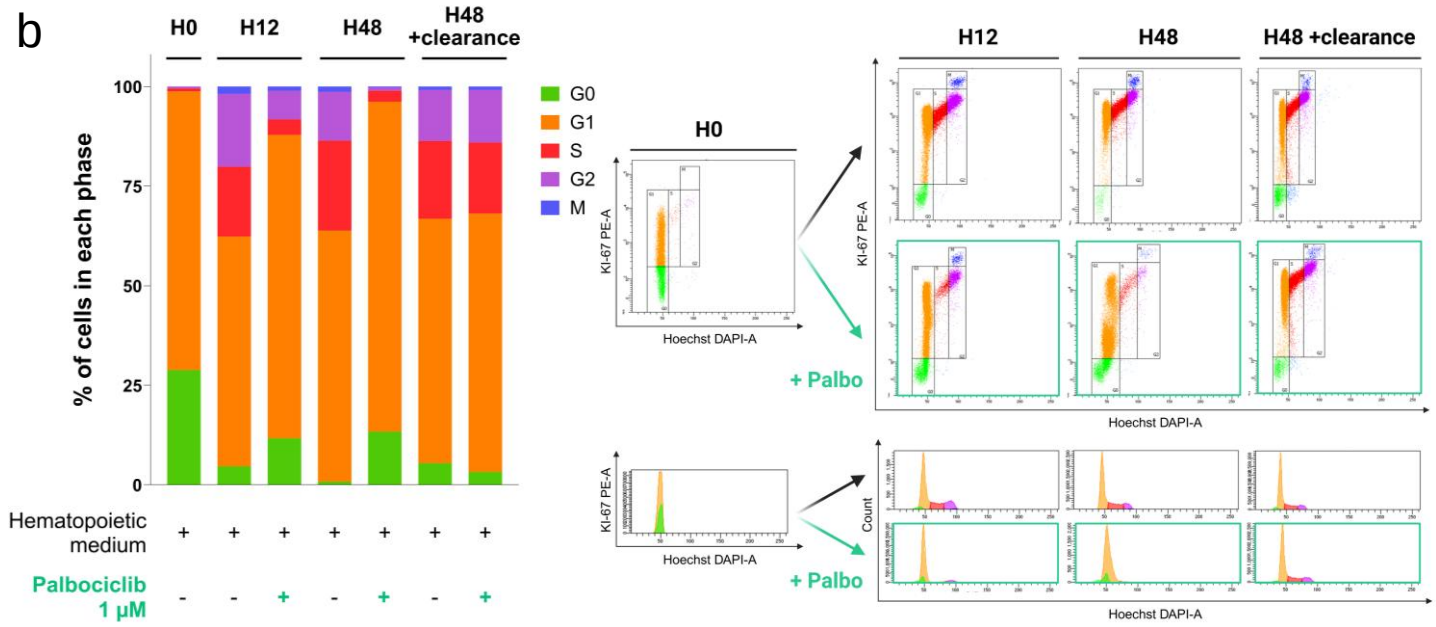


Fig. 4

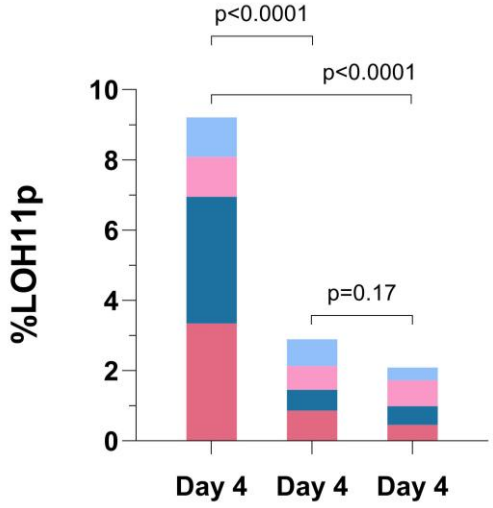
a



b



c



Cas9:sgHBG1/2	+	+	+
Palbociclib	-	24H	48H
Total cells	1855	2562	1112
LOH cells	167	73	23

d

Terminal LOH

LOH11pI LOH11pII

	Day 4	Day 4 Palbociclib 24H	Day 4 Palbociclib 48H
LOH11pI	ploidy = 1.18 n = 62	ploidy = 1.24 n = 22	ND n = 5
LOH11pII	ploidy = 1.25 n = 67	ploidy = 1.27 n = 15	ploidy = 1.84 n = 6
Average ploidy	*1.2*	*1.3*	-

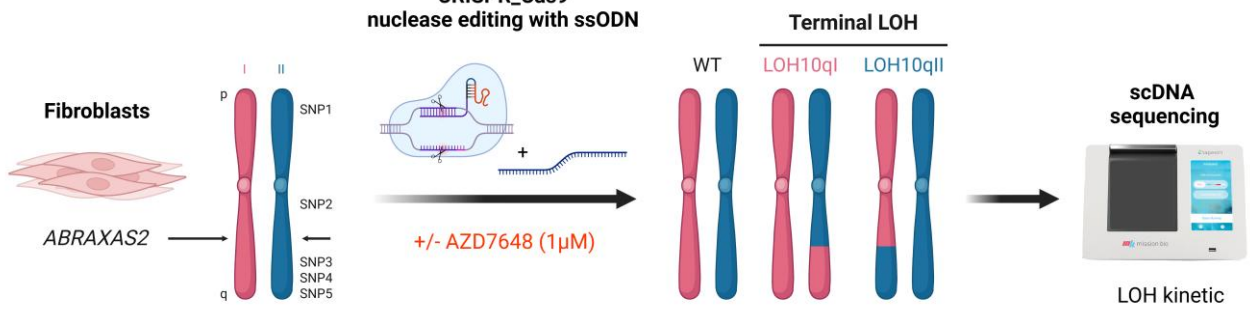
Interstitial LOH

iLOH11pI iLOH11pII

	Day 4	Day 4 Palbociclib 24H	Day 4 Palbociclib 48H
iLOH11pI	ploidy = 1.21 n = 19	ploidy = 1.21 n = 17	ploidy = 1.21 n = 8
iLOH11pII	ploidy = 1.44 n = 19	ploidy = 1.52 n = 19	ND n = 4
Average ploidy	*1.3*	*1.4*	-

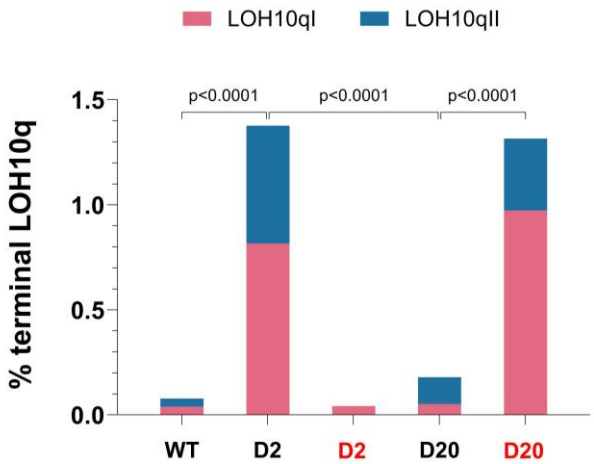
Fig.5

a



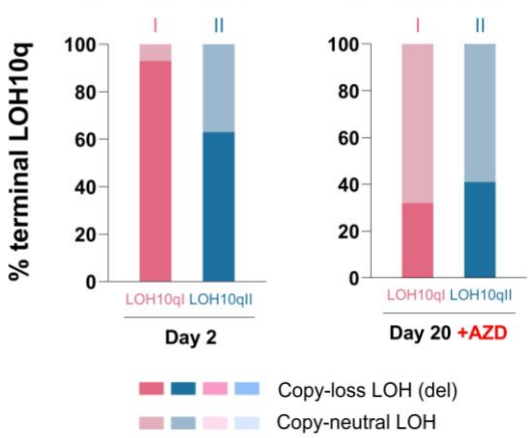
Legend	Hg19 Genomic Position	SNP rs	Distance to DSB	WT	LOH 10qI	LOH 10qII
1	OLAH:chr10:15106441:A/G	rs10906818	111 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)
2	WDFY4:chr10:50110232:A/G	rs7069142	76.4 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)
3	ADAM12:chr10:127823151:T/G	rs1278329	1.3 Mb	Htz (T/G)	Hmz (G)	Hmz (T)
4	DOCK1:chr10:129168150:T/G	rs12570609	2.7 Mb	Htz (T/G)	Hmz (G)	Hmz (T)
5	TCERG1L:chr10:132984320:T/C	rs2918100	6.5 Mb	Htz (T/C)	Hmz (C)	Hmz (C)

b



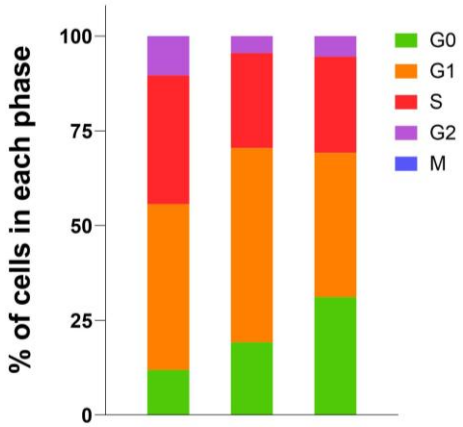
	WT	D2	D2	D20	D20
Cas9:sgABRAXAS2	-	+	+	+	+
ssODN	-	+	+	+	+
AZD7648	-	-	+	-	+
Total cells	2591	1961	2378	3900	5249
LOH cells	2	27	1	7	59

c



d

Cell cycle analysis



Cas9:sgABRAXAS2	-	+	+
ssODN	-	+	+
DMSO 1µM	-	+	-
AZD-7648 1µM	-	-	+

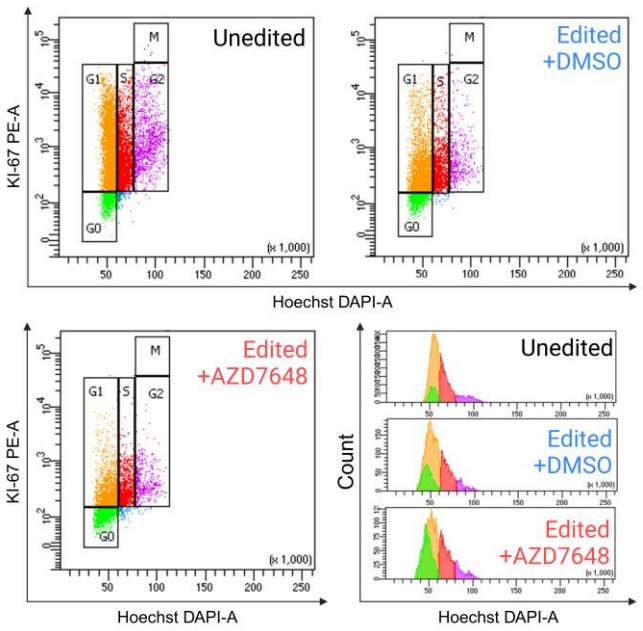
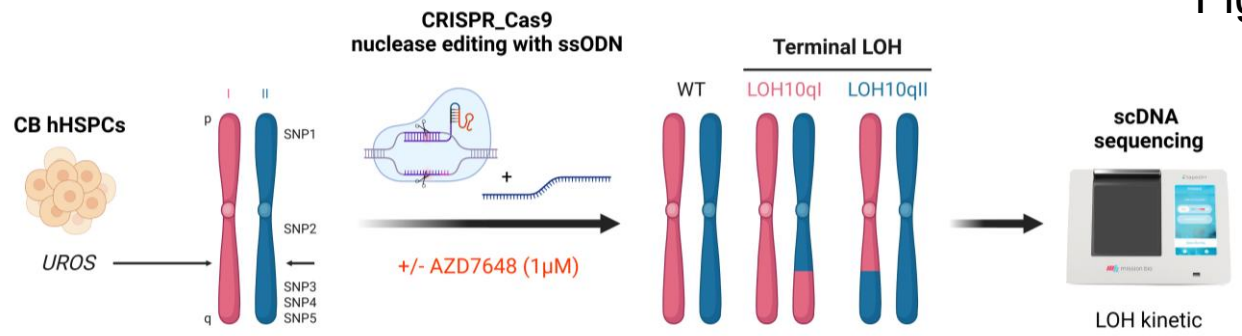


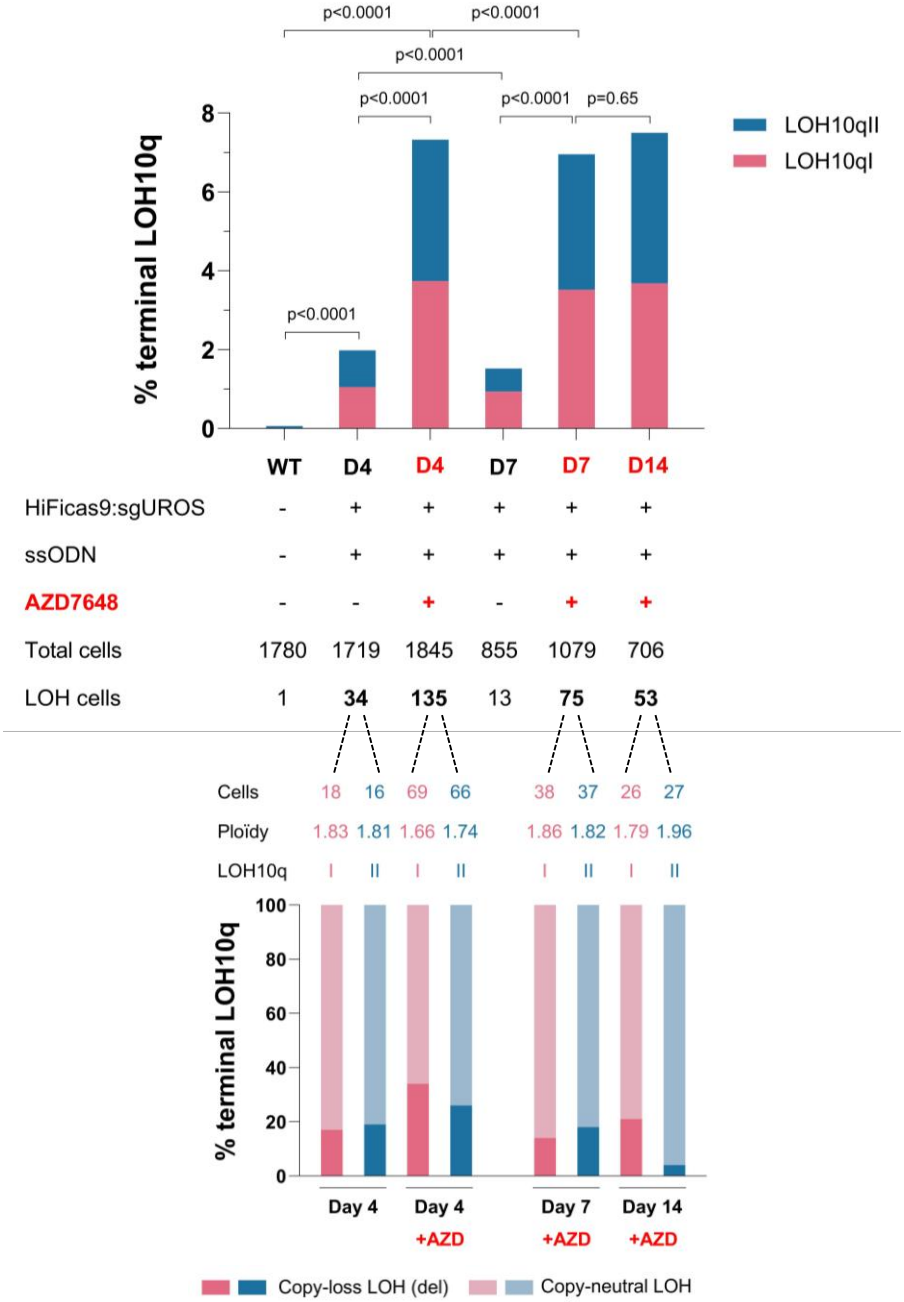
Fig.6

a



Legend	Hg19 Genomic Position	SNP rs	Distance to DSB	WT	LOH 10qI	LOH 10qII
1	CDC123:chr10:12292344:A/G	rs10951	115 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)
2	WDFY4:chr10:50110232:A/G	rs7069142	76.4 Mb	Htz (A/G)	Htz (A/G)	Htz (A/G)
3	ADAM12:chr10:127823151:T/G	rs1278329	0.44 Mb	Htz (T/G)	Hmz (G)	Hmz (T)
4	DOCK1:chr10:129168150:T/G	rs12570609	1.6 Mb	Htz (C/G)	Hmz (C)	Hmz (G)
5	TCERG1L:chr10:132984320:T/C	rs2918100	5.5 Mb	Htz (G/A)	Hmz (G)	Hmz (A)

b



c