Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The type of mismatch repair deficiency shapes dinucleotide repeat expansion and WRN inhibitor sensitivity in colorectal cancer

Cathy Saab
Paris-Saclay University / CNRS / Gustave Roussy
Emmanuelle Despras
Sorbonne University / INSERM
Louisa Bekkouche
Paris-Saclay University / CNRS / Gustave Roussy
Marine Guedet
Paris-Saclay University / CNRS / Gustave Roussy
Gwendolyn Renaudin
Sorbonne University / INSERM
Aurelie Siret
Inserm
Cody Feys
Sorbonne University / INSERM
Nathalie Droin
Institut Gustave Roussy
Florence Renaud
Sorbonne University / INSERM
Alex Duval
Sorbonne University / INSERM
Gabriel Matos-Rodrigues
Gabrieleduardo.DE-MATOS-RODRIGUES@gustaverou
Paris-Saclay University / CNRS / Gustave Roussy

Article

Keywords:

Posted Date: September 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10851996/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Even though cancer genetics was revolutionized by the discovery of microsatellite instability (MSI) in colorectal cancer (CRC) about thirty years ago, this phenotype is yet to be fully characterized[1–3]. Here, we used long-read sequencing to uncover the landscape of repeat instability in primary CRC. We found that long TA repeats, which underlie the formation of alternative DNA structures that break upon inactivation of WRN helicase, are a common feature of MSH2-, MLH1-, PMS2-, but not MSH6-deficient primary CRCs. Dinucleotide expansion is exacerbated in PMS2-deficient tumors, where most dinucleotide motifs are prone to expand. Importantly, acquisition of resistance to WRN inhibition in MSI cells is not driven by overcoming DNA structure formation. Altogether, our data uncover the full landscape of repeat instability in CRC where dinucleotide expansions are unlikely to be overcome by acquired resistance to WRN-inhibitor therapy, defining a novel actionable genetic feature of CRC.

INTRODUCTION

In the early 1990's the field of cancer genetics was revolutionized by studies exploring the instability of microsatellites (also known as simple repeats or short tandem repeats) in colorectal carcinomas[1–3]. This genomic feature, characterized by small insertion-deletions in mono- and dinucleotide repeats in microsatellite regions called microsatellite instability (MSI), arises from inactivation of MMR genes *MSH2*, *MSH6*, *MLH1*, or *PMS2* and has been utilized for diagnostic purposes ever since[1,2,4–7]. Approximately 5% of all tumors present MSI, including ± 15% of colorectal tumors[8–11], in which about 20% derive from Lynch syndrome patients, a heritable cancer predisposition disease caused by a germline mutation in one of the four MMR genes (*MSH2*, *MSH6*, *MLH1*, or *PMS2)*, while the other 80% are sporadic and driven either by bi-allelic epigenetic silencing of *MLH1* or two-hits somatic mutations affecting *MLH1, MSH2, MSH6 or PMS2*. MMR-deficient (MMRD) tumors show an overall increase in the mutational burden, and this feature has been explored in the clinic, as MSI/MMRD tumors are one of the best responders to immune checkpoint inhibitor (ICI)-based immunotherapy[5,12,13]. Indeed, the mutational burden found in MSI/MMRD tumors is associated with the formation of neo-peptides and an increased lymphocytic cell infiltrate. This preexisting immune microenvironment is thought to be activated by immunotherapy with ICIs[14–16]. Recent clinical trials revealed a 5-year overall survival rate of more than 50% in MSI/MMRD metastatic CRC patients treated with ICIs[17–19], while trials are also underway to test these molecules in patients with localized CRC. However, even though immunotherapy has provided an important improvement in patient prognosis, resistance has been reported[20], and new approaches are needed to further improve treatment options.

In the past few years, particular attention has been given to the function of the helicase Werner (WRN) after the discovery of its essential function for the survival of many MSI/MMRD cancer cells[21–25]. Studies performed in cancer cell lines have shown that expansions of TA repeats are found in MSI/MMRD colon cancer cells sensitive to WRN[26,27]. However, expansions have unknown frequency in MSI primary tumors from patients, and their relationship with the type of MMR deficiency is still poorly

understood, although it has been suspected that these expansions might not be systematically observed in MMR-deficient cells in *in vitro* models. MSI-related expanded TA dinucleotide repeats are prone to form alternative DNA structures called DNA cruciforms[26,27]. The current data support a model in which cruciform-like alternative DNA structures must be resolved by the DNA helicase WRN; otherwise, they are processed into DNA double-strand breaks (DSB), which cause chromosomal shredding and cell death[26,27]. Thus, based on these mechanistic studies performed in cancer cell lines, the TA expansions, not inactivation of MMR-repair *per se*, render MSI cancer cells sensitive to WRN inactivation. This biological model lacks validation from analysis in primary tumors, as the frequency with which MSI/MMRD CRCs and other tumor types carry TA expansions is still unknown. This is of particular importance given that highly promising innovative WRN inhibitors have been developed, and clinical trials testing the efficiency of WRN inhibition in MSI/MMRD patients are underway[28,29].

RESULTS

Analysis of TA expansions is challenged by the fact that long and pure TA repeats are refractory to short-read sequencing[25], which has been the technology used in most pan-cancer whole-genome sequencing studies. This technical limitation makes re-analysis of the available short-read sequencing data unsuitable for identifying TA expansions, which is exemplified by a recent pan-cancer genome-wide repeat expansions analysis that did not find TA expansions associated with MSI/MMRD tumors[30]. To address this issue, we have analyzed a cohort of 39 primary CRC (**Supplementary Table 1**) tumors and adjacent non-tumoral tissues using ONT-Nanopore long-read whole-genome sequencing. From the 39 CRC tumors analyzed, 23 were MSI, and 16 were MSS. Among the MSI tumors, 5 tumors were MLH1-deficient, 6 were MSH2-deficient, 3 were PMS2-deficient, 4 were MSH6-deficient, and for the remaining 5 tumors, previously published[31] genome sequences were used, and the MMR defect status was not available (N/A). By analyzing the structural variants (SV) in CRC, we uncovered that the presence of large insertions is a genomic feature that distinguishes most of the MSI/MMRD CRCs from the MSS ones (Fig. 1a), as we observed that 19 out of 23 MSI samples show more than 500 insertions (15bp or larger). No MSS tumor had more than 300 insertions. Importantly, only MSI MSH6-deficient tumors showed insertion levels similar to MSS tumors. Closer inspection of the identity of the insertions revealed that most insertions are formed by pure (≥ 80% of a motif) dinucleotide repeats (Fig. 1b), except tumor C553 (pie chart - Fig. 1b), for which the insertions are mainly composed of other sequence motifs and do not predominantly contain sequences with a high dinucleotide content. Considering all MSI tumors, TA repeats were the most frequently found dinucleotide insertions, with a median insertion size varying from 59 to 221 base pairs (**Supplementary Fig. 1a-d**). MLH1 and MSH2-deficient tumors showed a similar phenotype, with pure TA sequences accounting for over 90% of insertions (Fig. 1b). Interestingly, PMS2-deficient tumors had a different spectrum of dinucleotide insertions, and ± 30% of insertions were found in dinucleotide repeats other than TA, with median insertion length varying from 49 to 60.5 base pairs (**Supplementary Fig. 1e**). The GC/CG were the only dinucleotide motif insertions not found in PMS2-deficient tumors, in agreement with the low proportion of annotated GC/CG microsatellites in the hg38 genome (**Supplementary Fig. 2a, b**). PMS2-deficient tumor insertions were likely expansions of

already existing repeats instead of *de novo* events, as most of them overlapped with already annotated microsatellites (**Supplementary Fig. 2c**). MSH6-deficient tumors showed virtually no dinucleotide insertions (Fig. 1b). The number of the most recurrent insertions (TA) was not associated with the age of the patients (Fig. 1c), sex (male/female), or disease type (sporadic/Lynch syndrome) (Fig. 1d), indicating that large insertions are mainly defined by MMR gene deficiency.

Next, we sought to define the genomic features of TA insertions given the potential significance of these markers for determining WRN inhibitor sensitivity in MSI tumors. Consistent with the notion that the insertions observed in CRC long-read sequencing are repeat expansions, 18329 genomic regions containing TA insertions in MLH1-deficient, MSH2-deficient, and PMS2-deficient tumors, 16014 overlapped with the annotated TA repeats, accounting for 20% of the annotated TA microsatellites. These regions were hereafter named “expanded TA repeats”. Among TA insertions, 3320 were unique to a given tumor, with no common regions found for all tumors, and only one region overlapped between 16 of the 19 samples analyzed (**Supplementary Fig. 3a**). Given that previous studies have reported the existence of expanded TA microsatellites in the tumor suppressor gene *DCC* in CRC MSI cells[24], we asked whether specific genomic regions related to tumorigenesis were prone to show expansions. We observed no enrichment in genomic regions (intergenic, exonic, intronic, or promoter regions), oncogenes, tumor suppressor genes (TSG), or genes related to CRC (**Supplementary Fig. 3b-d**), indicating that the TA expansions are not overtly selected to influence tumor-related genes. Instead, we found that expanded TA repeats are particularly long (Fig. 2b). Compared to non-expanded TA repeats, expanded TA repeats have a medium size about 7 times longer in the annotated hg38 genome (Fig. 2b). Given that TA repeats are refractory to short-read sequencing, we confirmed these observations by analyzing three different whole genomes sequenced via ONT-Nanopore sequencing derived from the 1000 genomes project [32]. The long-read sequencing data echoed our findings in the hg38 genome, with a similar size difference between expanded and non-expanded TA repeats (Fig. 2b). It is noteworthy that the overall medium size of the TA repeats was larger in long-read sequencing data, with a medium gain of 8bp in the expanded TA repeats and only 1bp in non-expanded TA repeats, compared to the annotated sequences in the hg38 genome (Fig. 2b). In essence, TA repeats are larger than in the annotated genome with the expanded TA repeats being particularly prone to show insertions even in healthy individuals. Furthermore, whole-genome long-read sequencing of MSI cell lines confirmed our observation from CRC samples, as the MSH6-deficient HCT15 cell line shows a similar median size of TA repeats compared to healthy individuals, whereas in the HCT116 cell line, size specifically increased at expanded TA repeats (Fig. 2c).

It was previously shown that MSI CRC cell lines show expanded TA repeats forming alternative DNA structures that require the WRN helicase for their resolution. These cruciform-like structures are susceptible to *in situ* digestion with purified, recombinant MUS81-EME1, a structure-specific nuclease that cleaves cruciform structures[27,33,34] or the single-stranded DNA (ssDNA)-specific S1-nuclease. In this last method, sites of recurrent S1-cleavage are revealed by analysis of *in situ* DSB ends at single-nucleotide resolution via END-seq [26]. We observed that expanded TA repeats are particularly prone to

form alternative DNA structures scored by S1-END-seq in MSI cells but not in MSS ones (**Supplementary Fig. 4**). Importantly, this is not an overall feature of MSI cells, as we observed that expanded TA repeats are larger and more prone to form secondary DNA structures in HCT116 cells (MLH1-deficient) than in HCT15 cells (MSH6-deficient) (Fig. 2c, d). Therefore, a subset of inherently large TA repeats is prone to expand and form DNA structures in MSH2-, MLH1-, or PMS2-deficient CRC, but not in MSH6-deficient contexts. In addition, expanded TA repeats defined in our long-read sequencing from CRC allowed for the prediction of the sites of DNA breakage upon WRN-inactivation in MSI cells. Indeed, analysis of DNA double-stranded breaks via END-seq in cells treated with the non-covalent WRN inhibitor HRO761 or DMSO for 48 hours showed the enrichment of DNA double-strand breaks in expanded TA repeats in MLH1-deficient, but not in MSH6-deficient cells (Fig. 3a, b). This feature was associated with the lack of cell toxicity in HCT15 (MSH6-deficient) cells treated with the WRN inhibitor (Fig. 3c), establishing a direct link between the expanded TA repeats found in CRC and the sensitivity of cell lines to WRN inhibition.

Resistance to therapy can be inherent or acquired. Our data show that MSH6-deficient CRC are inherently resistant to WRN inhibition, likely because they lack expansion of TA repeats. However, resistance to WRN inhibition can be acquired [35–38]. Thus, we then asked whether pressure selection of MSI CRC cells can be overcome by the reduction of structure formation in MSI tumors. To generate WRNi-resistant cells, WRNi-sensitive HCT116 cells were chronically treated with either the non-covalent HRO761 or the covalent VVD214 WRN inhibitor[39,40] using different strategies (Fig. 4a). Resistance to HRO761 and/or VVD214 was confirmed by clonogenic assays (Fig. 4b and **Supplementary Fig. 5a**). Using whole-exome sequencing, we identified heterozygous missense mutations in the *WRN* gene in all resistant cell lines, which were absent from the parental cell line (**Supplementary Table 2**). These variants targeted either residues lying in the flexible hinge between the RecA1 and RecA2 helicase domains (C727S, C729S, R732Q), close to the inhibitor binding site (T705A) (**Supplementary Fig. 5b**), or lying in the HRDC domain (G1207D), which has been predicted to contribute to VVD214 binding[36]. The C727S/C729S mutants ($WRNi^R$ #2 and #5) and, to a lesser extent, the R732Q mutant ($WRNi^R$ #1) displayed increased resistance to both HRO761 and VVD214, whereas the T705A ($WRNi^R$ #3) and G1207D ($WRNi^R$ #4) mutants were specifically resistant to VVD214, in agreement with previous findings[36–38] (Fig. 4b, **Supplementary Fig. 5a** and **Supplementary Table 2**). Analysis of WRNi-induced degradation of WRN protein corroborated these observations (**Supplementary Fig. 5c**).

Interestingly, we did not observe changes in the formation of cruciform-like structures in WRNi-resistant cell lines (Fig. 4c). In agreement, WRNi-resistant cells remained sensitive to WRN knockdown (**Supplementary Fig. 5d**). This demonstrates that resistance to treatment likely arises from on-target WRN mutations that impair WRNi efficiency, but not from the selection of cells that lack structure formation at expanded TA repeats and become independent of WRN helicase activity.

Discussion

Our analysis of colorectal cancer (CRC) revealed a distinct subset of microsatellite loci characterized by expanded TA dinucleotide repeats, which are substantially longer than their non-expanded counterparts. In MSI cell lines, these expanded repeats exhibit an increased tendency to adopt alternative DNA conformations that become vulnerable to DNA breakage following WRN helicase inactivation. Moreover, investigation of MSI cells that acquired resistance to WRN-targeted therapy showed that resistance is unlikely to arise through suppression of structure formation at expanded repeats. Collectively, these findings provide a comprehensive view of repeat instability in CRC and provide definitive proof of TA dinucleotide expansions as a therapeutically relevant genetic vulnerability. Furthermore, acquired resistance to WRN inhibition arises from on-target mutations, instead of unlikely direct inhibition of DNA structure formation at expanded repeats.

Even though it is still unclear how insertions are created in dinucleotide repeats, our data provide mechanistic insight into how the MMR counteracts these events. MMR corrects errors occurring during DNA replication, including base-base mismatches and insertion/deletion loops, which predominantly result from replicative DNA polymerase slippage. The process begins with the recognition of these mismatches by MutS homolog complexes, followed by the recruitment of other repair proteins that degrade the newly synthesized incorrect DNA strand and fill in the gap[41,42]. Two main MutS complexes exist in mammalian cells: MutSα and MutSβ. MutSα, composed of MSH2 and MSH6, primarily detects base-base mismatches and small loops of one or two nucleotides[43]. The MutSβ complex, formed by the MSH2 and MSH3 heterodimer, specializes in recognizing larger insertion/deletion loops of up to 15 nucleotides[44]. We propose that MMR participates in resolving loop-outs formed in particularly large dinucleotide repeats. The structures are recognized by the MutSβ complex, and MutSα (MSH6) plays no role in inhibiting the expansions. In the absence of functional MutSβ, loop-outs stay unresolved, giving rise to an expansion-prone repair mechanism.

After mismatch or structure recognition, the MutS complex recruits MutL heterodimers to help process and repair the DNA. In mammalian cells, there are three MutL complexes involved in DNA mismatch repair: MutLα (MLH1-PMS2), which is the primary complex responsible for correcting replication errors by nicking the newly synthesized DNA strand to initiate repair; MutLβ (MLH1-PMS1), which has a less well-defined role and lacks known endonuclease activity, possibly acting in regulatory or specialized repair functions; and MutLγ (MLH1-MLH3), which has a primary role in meiotic recombination to ensure proper chromosome crossover and segregation, has endonuclease activity and a redundant or compensatory function with MutLα in somatic cells. We observed that PMS2-deficient tumors show a specific phenotype where most dinucleotides, not only TA repeats, show an increased propensity to expand. Two non-exclusive hypotheses can explain the phenotype observed in PMS2-deficient tumors: 1) PMS2 plays an independent direct role in the stability of non-TA-dinucleotide repeats, and its deficiency leads to their expansions; 2) In the absence of PMS2, other MutL complexes compensate for the lack of PMS2, but they are prone to induce expansions in TA and non-TA-dinucleotide repeats. Observations in mouse models might indicate a stronger role for the second hypothesis. Yao and colleagues[45] analyzed three CA:TG repeats and reported that *Pms2* knockout mice show a particular

tendency for expansions in these dinucleotide repeats. This feature was not observed in *Mlh1* KO and was abolished in *Pms2* and *Mlh3* double knockout mice, suggesting that *Mlh3* participates in the induction of CA:TG insertions[45]. Importantly, repeat expansions in CRC tumors are mechanistically distinct from the known repeat expansion diseases, which have been extensively studied over the past three decades[41,46]. As in these diseases, the MMR pathway is recognized as the primary molecular mechanism responsible for inducing repeat expansions[41,46,47].

Our study opens new avenues to the understanding of DNA repeat instability with irrefutable clinical implications. Based on our findings, future studies should aim to use the diagnosis of TA repeat expansion as well as the status of MMR deficiency as key parameters to define the patients that should be targeted with WRN inhibitors.

METHODS

Cell culture

HCT116 and KM12 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% foetal calf serum, 100 U/ml penicillin and 100 µg/ml streptomycin at 37°C under a 5% CO_2 atmosphere. HCT15 cells were cultured in RPMI 1640 (Gibco) supplemented with 10% foetal bovine serum, 100 U/ml penicillin and 100 µg/ml streptomycin at 37°C under a 5% CO_2 atmosphere.

Clonogenic assay

1 to 2 thousand cells were seeded in 6-well plates 24 hours before treatment at the indicated doses. Cells were grown for 9 to 10 days and fixed in 20% methanol, 0.5% crystal violet. Plate images were acquired on iBright 1500 imaging system (Thermo Scientific) using the universal mode. Surviving fraction was assessed as total colony area normalized to the DMSO-treated condition using FIJI software. The IC_{50} values were estimated with a four-parameter fit using GraphPad Prism 11.

Cell viability assay

30 thousand cells per well were seeded in six-well plates and treated with HRO761 or vehicle DMSO after 24 hours. Following a 5-day incubation period, cell viability was assessed using trypan blue exclusion and an automated cell counter.

Transfection

For RNA interference, siRNAs targeting WRN (On-TARGET plus SMART pool #L-010378-00-0005, siWRN) and a non-targeting control pool (#D-001810-10-05, siNT) were purchased from Dharmacon. 150 thousand cells were seeded in 6-well plates and transfected with 20 nM siRNAs using Lipofectamine RNAiMAX reagent (Thermo Scientific #13778030), according to the manufacturer's instructions.

Inhibitors

Stock solutions of WRN inhibitors (WRNi) HRO761 (Selleck #SE-E1908-5MG) and VVD214 (MedChemExpress #HY-158116) were made at 10 mM in DMSO and stored at -80°C.

Establishment and characterization of WRNi-resistant cells: To establish WRNi-resistant HCT116 cells, different strategies were used. Pulse protocol (HCT116 WRNi^{R} #1): cells were treated for 3 days with increasing doses of HRO761 (from 15 nM to 10 µM) followed by 3–4 days (or until the cells recovered) in drug-free medium. 3 additional pulses at 10 µM were performed at the end of the dose escalation. The full procedure took 3 months. Continuous protocol (HCT116 WRNi^{R} #2–5): cells were continuously treated with WRNi until reaching no apparent cell toxicity (i.e. detached cells). Conditions used were 350 nM VVD214 for 39 days (WRNi^{R} #3), 500 nM VVD214 for 47 days (WRNi^{R} #4), or 200 nM WRNi for 15 days followed by 2 µM for 25 days (WRNi^{R} #2 selected with HRO761 and WRNi^{R} #5 with VVD214).

Whole-exome sequencing of WRNi^{R}-resistant cells

Genomic DNA was extracted with Qiamp DNA mini kit (Qiagen #51306) and quantified by Qubit® dsDNA HS Assay kit. 300 ng were sent to the Novogene company, which performed library preparation, hybridization capture (Agilent SureSelect V6 60 Mb), captured library enrichment by PCR, whole-exome sequencing (Illumina NovaSeq X Plus PE150, average sequencing depths > 100x), and bioinformatic analysis. Variant calling was performed with GATK HaplotypeCaller using hg38 as the human reference genome and annotated with ANNOVAR using RefSeq and Gencode databases for genes and region annotations. The annotated files were filtered to retrieve *WRN* variants using a depth criterion of at least 10 total reads. Variants already present in the parental HCT116 cell line were further excluded to identify WRNi-resistant cell-specific mutations (**Supplementary Table 2**). Of note, an intronic deletion variant was identified in WRNi^{R} #2, #3, and #5 cells, which was not found in parental or WRNi^{R} #1 and #4 cells. However, given that these three resistant models showed differential sensitivity to HRO761 and VVD214 inhibitors, this variant was considered unlikely to be causative of WRNi resistance.

For WRNi^{R} #1, #2, #3, and #5 cells, mutations and heterozygosity were confirmed by Sanger sequencing of WRN cDNA. RNA was extracted using the RNeasy mini kit (Qiagen #74106). RT-PCR was performed to amplify the region coding for the WRN helicase domain (codons 484 to 911) using a High-Capacity cDNA reverse transcription kit (Applied #4368813) and Qiagen Taq DNA polymerase (#201207) with forward primer GGCACGGTAGAACCAACTCA and reverse primer TTTGTCTCCTACATCTGCTAGAAT[36] synthesized by Eurofins. Sanger sequencing of PCR products was performed by Eurofins.

Western blot: Whole-cell extracts were prepared in RIPA buffer (Thermo Scientific #89901) supplemented with Halt protease inhibitor cocktail (Thermo Scientific #78410), according to the manufacturer's instructions. Proteins were quantified with Pierce BCA protein assay kit (Thermo Scientific #23225), heat denatured in NuPAGE LDS sample buffer with Sample reducing agent (Thermo Scientific #NP0008 and #NP0009), separated by SDS-PAGE, and transferred onto PVDF membranes. Membranes were blocked in 5% milk in PBS 0.1% Tween20 and incubated with primary antibodies overnight at 4°C, followed by incubation with HRP-coupled secondary antibodies for 1 h at room

temperature. The following primary antibodies were used: mouse monoclonal anti-WRN 8H3 (Cell Signaling #4666, 1/1,000 dilution), mouse monoclonal anti-αtubulin clone DM1A (Sigma-Aldrich #T9026, 1/1,000 dilution). Detection was carried out using SuperSignal West Pico PLUS substrate (Thermo Scientific #34580) on iBright 1500 imager. Signals were quantified with FIJI software.

Samples and patients

MSI CRC and adjacent normal tissue samples were obtained from a retrospectively enrolled cohort of patients who underwent surgical resection of CRC at Hôpital Saint-Antoine and for whom frozen samples were available (N = 19). MSI/MMRd status was identified at diagnosis using PCR and immunohistochemistry (IHC) gold standards. Samples were selected based on IHC staining of MMR proteins in order to include MLH1-deficient (MLH1d) and MSH2-deficient (MSH2d) tumors, as well as tumors with isolated loss of MSH6 (MSH6d) or PMS2 (PMS2d) protein expression. Samples were stored at -80°C until DNA extraction.

Tissue recovery and DNA extraction

Genomic DNA from frozen tumor and adjacent normal colic mucosa was extracted using the AllPrep DNA/RNA Mini kit (Qiagen #80204) according to the manufacturer's instructions. DNA was quantified by Quant-iT dsDNA broad range assay (Thermo Scientific #Q33130). MSI status was confirmed retrospectively by Pentaplex PCR and whole-exome sequencing using MSICare algorithm[48].

END-seq and S1-END-seq

For END-seq and S1-END-seq, between 7–10 million cells were collected and embedded in 1% agarose plugs. Samples were processed following previously described protocols[26,49,50]. For S1-END-seq, cell-containing plugs were digested with proteinase K (1 hour at 50°C, followed by 7 h at 37°C), washed once with plug washing buffer (WB − 10 mM Tris, pH 8.0, 50 mM EDTA, diluted in nuclease free water), washed three times with Tris-EDTA (TE − 10 mM Tris, pH 8, 1 mM EDTA, in nuclease-free water) and then treated with RNase A for 1 h at 37°C. Following RNase treatment, plugs were washed three times with WB, once with elution buffer (EB − 10 mM Tris, pH 8.0, in nuclease-free water) and equilibrated with two 15-minute washes in S1 nuclease buffer (40 mM sodium acetate pH 4.5, 300 mM NaCl, 2 mM ZnSO4). S1 digestion was performed using 200 U of S1 nuclease (ThermoFisher) diluted in 100 µl S1 nuclease buffer per plug at 37°C for 30 min. The reaction was stopped by adding EDTA to a final concentration of 10mM. Finally, plugs were processed using the standard END-seq protocol.

END-Seq and S1-END-seq libraries were performed by Novogene using 150-bp paired-end Illumina runs. For each experimental condition, the corresponding control (vehicle or untreated) library was sequenced in parallel on the same flow cell. Final sequencing data were processed using the bcl2fastq software.

Five publicly available S1-END-seq datasets (BioProject PRJNA841652) were retrieved from the NCBI Sequence Read Archive (SRR19364486, SRR19364488, SRR19364489, SRR19364490, SRR19364491).

Oxford Nanopore Technologies (ONT) whole-genome sequencing: Genomic DNA from cell lines using the DNeasy Blood & Tissue Kit (Qiagen #69504) following the manufacturer's instructions. Quality of the purified DNA was evaluated using a TapeStation genomic DNA (gDNA) ScreenTape assay (Agilent #6057). Genomic DNA was quantified using a Qubit Fluorometer (Thermo Fisher) and size average was determined using Fragment Analyzer (Agilent). Prior to library preparation, gDNA was mechanically fragmented to an average size of approximately 10 kb using Covaris g-TUBEs (Covaris, Woburn, MA, USA) according to the manufacturer's recommendations. Whole-genome sequencing libraries were prepared using Ligation Sequencing Kit V14 (SQK-LSK114, ONT, Oxford, UK). Briefly, 1µg to 2µg of fragmented gDNA underwent DNA repair and end-preparation using the NEBNext Companion Module for Oxford Nanopore Technologies, followed by magnetic bead purification. Sequencing adapters supplied with the kit were ligated to the repaired DNA fragments, and the final library was purified according to the manufacturer's protocol. Only flow cells with at least 5,000 pores at the initial scan were used for sequencing. Sequencing libraries were loaded onto R10.4.1 PromethION flow cells (FLO-PRO114M) and sequenced on a PromethION P-48 instrument (ONT, Oxford, UK) at Gustave Roussy using MinKNOW and the super-accuracy base-calling model (SUP) selected with m5C and hm5C analysis. Data summary of the long-read sequencing can be found in **Supplementary Table 3**. 5 MSI (C538, C543, C553, C577, C586) and 16 MSS samples used in this study were obtained from the publicly available dataset previously reported[31]. The data are available in the Genome Sequence Archive (GSA) China National Center for Bioinformation (CNCB) under accession number: HRA002638. Additional clinical and sequencing information can be found in Supplementary Tables 1 and 3[31]. Long-read Nanopore whole-genome sequencing from GM20752, HC00733, and NA19036 reference genomes is publicly available https://data.humanpangenome.org/raw-sequencing-data.

Nanopore whole-genome sequencing alignment

POD5 files from PromethION were used to perform base calling with the Dorado basecaller (v0.8.1). The resulting unaligned BAM files were trimmed with Dorado trim and filtered by read length (≥ 1000bp) and quality score (≥ 10). The BAM files are aligned to a reference genome using Dorado aligner. Next, read quality was controlled before alignment is performed using FastQC, FastQ Screen (after uBAM files conversion to FASTQ using Samtools), and NanoPlot (directly on the BAM files). The mapping control quality is generated using Qualimap, Mosdepth, and Samtools on the aligned BAM files.

Structural variant calling

Germline structural variant calling is performed with Sniffles2 or cuteSV, and somatic SV, including large insertions, is executed using NanomonSV. SV annotation is done with AnnotSV. SV annotations were subsequently analyzed and quantified using R-studio (v4.5.2). Insertion sequences were analyzed to identify dinucleotide repeat motifs with a purity ≥ 80%, calculated as the proportion of the insertion sequence composed of the dinucleotide repeat.

Data visualization and quantification

The dinucleotide repeat location and size were extracted from the simple repeat mask annotation from UCSC (hg38). The expanded TA repeats corresponded to TA insertions overlapping with annotated TA/AT simple repeat regions. Genomic annotation of expanded and non-expanded TA repeats was defined using ChIPseeker and the TxDb.Hsapiens.UCSC.hg38.knownGene annotation database. Genes associated with TA repeats were intersected with the OncoKB cancer gene database to identify oncogenes and tumor suppressor genes. CRC-associated genes were identified using KEGG colorectal cancer pathway (hsa05210). R-studio (v4.5.2) was used for sequencing-based data analysis. The barplots, pie charts, and boxplots were generated using ggplot2. The S1-END-seq heatmaps were produced using EnrichedHeatmap. Statistical analyses are specified in the figure legends.

Supplementary Fig. 1. Median of insertion length measured in bp in MSI/MMRD tumors. (a) TA repeat length in N/A tumors. **(b)** TA repeat length in MSH2d tumors. **(c)** TA repeat length in MLH1d tumors. **(d)** TA repeat length in PMS2d tumors. **(e)** Dinucleotide repeat length in PMS2d tumors.

Supplementary Fig. 2. Insertions in MSI/PMS2D tumors. (a) Distribution of insertions by dinucleotide repeat type. **(b)** Proportion of dinucleotide repeats in hg38 genome. **(c)** Proportions of annotated or non-annotated dinucleotide expanded repeats using hg38 genome as reference.

Supplementary Fig. 3. Genomic characterization of expanded TA repeats. (a) Number of overlapping TA insertions between tumors. **(b)** Genomic location of non-expanded and expanded TA repeats. **(c)** Proportion of cancer-related genes (oncogenes/TSGs) overlapping with non-expanded and expanded TA repeats. **(d)** Proportions of CRC-related genes overlapping with non-expanded and expanded TA repeats.

Supplementary Fig. 4. **S1-END-seq reveals formation of cruciform-like structure at expanded TA repeats in MSI cells but not in MSS cells.** S1-END-seq heatmaps comparing the S1-END-seq signal in reads per million (RPM) related to 500 bp from the center of expanded TA repeats (top panel) and non-expanded TA repeats (bottom panel) in MSS (SW620, SW837) and MSI (RKO, KM12, SW48) cell lines.

Supplementary Fig. 5. **Characterization of WRN inhibitor-resistant cell lines. (a)** The IC_{50} values were estimated from the data shown in Fig. 4b using a variable slope four-parameter fit (left table). For the estimation of the IC_{50} values of the parental HCT116 cell line, survival curves after low doses of WRN inhibitors were used (middle and right panels, n = 3). **(b)** Identification of WRN mutations in $WRNi^R$ cell lines by whole-exome sequencing. The images show read alignments on the hg38 reference genome using the IGV software, confirming that the identified variants were absent from the parental HCT116 cell line (see also **Supplementary Table 2**). **(c)** The table recapitulates the identified CDS mutations and their consequences at the protein level. Of note, $WRNi^R$ #2 and #5 shared the same two mutations, both on the same allele, although the two cell lines were established independently. Protein structures were visualized using UCSF ChimeraX software based on PDB structure 8PFO and PDB structure 7GQU, representing the interaction between WRN and HRO761 and WRN and VVD214, respectively. WRN protein is represented in gray in both panels, while interacting molecules are shown in turquoise for HRO761 (left panel) and yellow for VVD214 (right panel). Residues mutated in resistant cell lines are

highlighted in dark blue. Predicted contact sites between WRN and molecules are indicated by red dots. Interactions between residues separated by fewer than 90 amino acids in the primary sequence were excluded from the contact analysis. **(d)** WB analysis of WRN protein after treatment with 5 μM HRO761 or VVD214 for 24 hours. The same volume of DMSO was used for mock-treated conditions. α-tubulin was used as a loading control. Histograms showed the quantification of WRN normalized to α-tubulin and represented as a percentage of DMSO-treated cells from 2 independent experiments. **(e)** Clonogenic survival assay of HCT116 and $WRNi^R$ cell lines following depletion of WRN by RNA interference. Representative images of crystal violet-stained colonies are shown. WB analysis of WRN protein confirmed the efficiency of the siRNA directed against *WRN* (siWRN). α-tubulin was used as a loading control.

Declarations

ETHICS DECLARATIONS

The authors declare that they have no competing interests or conflicts of interest.

This study was approved by our institutional review board/ethic committee (Sorbonne University), and informed consent was obtained. Research has been performed in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

GM-R conceived the study. GM-R, A.D., C.S., L.B., and E.D. designed the experiments and data analyses. C.S. and E.D. performed most of the experimental work, with assistance from M.G., G.R., and A.S. L.B. conducted most of the bioinformatic analyses with support from GM-R. C.F. performed the exome sequencing analysis of WRN-resistant cell lines. N.D. established the nanopore sequencing protocols at the Gustave Roussy Functional Genomics Platform. F.D. and A.D. provided the tumor samples. GM-R supervised the study and secured funding. GM-R wrote the manuscript, with input and feedback from all authors.

DATA AVAILABILITY

All sequencing data will be made publicly available upon publication.

ACKNOWLEDGEMENTS

We thank Dr. Patricia Kannouche for support in the development of the project and feedback on the manuscript. We thank Nomenahosandratra Rabearivelo, Marine Aglave, Marc Deloger, the members of the SIRIC-Epicure team, and Maëla Francillette from the genomic platform for their help in the development of Nanopore sequencing and its analysis. ONT PromethION-48 acquisition with financial support from ITMO Cancer of Aviesan within the framework of the 2021-2030 Cancer Control Strategy, on funds administered by Inserm and co-funded by Gustave Roussy IHU PRISM (National PRecISion

Medicine Center in Oncology), managed by the French National Research Agency under the France 2030 program (ANR-23-IAHU-0002) to Nathalie Droin.The computational resources of Gustave Roussy cluster, Flamingo, were used for High-Performance Computing and data analysis. The laboratory of Gabriel Matos-Rodrigues was funded by the SIRIC-Epicure (INCa-DGOS-Inserm-ITMO Cancer_18002), an ATIP-Avenir award with co-funding from la Ligue Nationale Contre le Cancer, an AFM-Telethon trampoline award, Université Paris-Saclay, a PJA1 grant from the Association pour la Recherche contre le Cancer (ARC – PJA1), and a start-up package from Gustave Roussy Foundation. The laboratory of Dr. Alex Duval was funded by grants from SIRIC CURAMUS and la Ligue Nationale Contre le Cancer (labeled team) and an Alliance Sorbonne université Emergence grant held by Emmanuelle Despras. Cody Feys was supported by a doctoral fellowship from la Ligue Nationale Contre le Cancer.

References

1. Peinado, M.A., Malkhosyan, S., Velazquez, A. & Perucho, M. Isolation and characterization of allelic losses and gains in colorectal tumors by arbitrarily primed polymerase chain reaction. *Proc Natl Acad Sci U S A* 89, 10065–10069 (1992).
2. Ionov, Y., Peinado, M.A., Malkhosyan, S., Shibata, D. & Perucho, M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363, 558–561 (1993).
3. Thibodeau, S.N., Bren, G. & Schaid, D. Microsatellite instability in cancer of the proximal colon. *Science* 260, 816–819 (1993).
4. Boland, C.R. & Goel, A. Microsatellite instability in colorectal cancer. *Gastroenterology* 138, 2073–2087 e2073 (2010).
5. Johannet, P., Rousseau, B., Aghajanian, C., Foote, M.B. & Diaz, L.A., Jr. Therapeutic targeting of mismatch repair-deficient cancers. *Nat Rev Clin Oncol* (2025).
6. Buhard, O., *et al.* Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. *J Clin Oncol* 24, 241–251 (2006).
7. Suraweera, N., *et al.* Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* 123, 1804–1811 (2002).
8. Bonneville, R., *et al.* Landscape of Microsatellite Instability Across 39 Cancer Types. *JCO Precis Oncol* 2017(2017).
9. Middha, S., *et al.* Reliable Pan-Cancer Microsatellite Instability Assessment by Using Targeted Next-Generation Sequencing Data. *JCO Precis Oncol* 2017(2017).
10. Landre, T. & Des Guetz, G. Microsatellite instability-high status as a pan-cancer biomarker for immunotherapy efficacy. *Cancer Immunol Immunother* 74, 122 (2025).
11. Vilar, E. & Gruber, S.B. Microsatellite instability in colorectal cancer-the stable evidence. *Nat Rev Clin Oncol* 7, 153–162 (2010).

12. Le, D.T., *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* 372, 2509–2520 (2015).
13. Le, D.T., *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409–413 (2017).
14. Dolcetti, R., *et al.* High prevalence of activated intraepithelial cytotoxic T lymphocytes and increased neoplastic cell apoptosis in colorectal carcinomas with microsatellite instability. *Am J Pathol* 154, 1805–1813 (1999).
15. Schwitalle, Y., *et al.* Immune response against frameshift-induced neopeptides in HNPCC patients and healthy HNPCC mutation carriers. *Gastroenterology* 134, 988–997 (2008).
16. Roudko, V., *et al.* Shared Immunogenic Poly-Epitope Frameshift Mutations in Microsatellite Unstable Tumors. *Cell* 183, 1634–1649 e1617 (2020).
17. Andre, T., *et al.* Nivolumab plus Ipilimumab in Microsatellite-Instability-High Metastatic Colorectal Cancer. *N Engl J Med* 391, 2014–2026 (2024).
18. Andre, T., *et al.* Pembrolizumab versus chemotherapy in microsatellite instability-high or mismatch repair-deficient metastatic colorectal cancer: 5-year follow-up from the randomized phase III KEYNOTE-177 study. *Ann Oncol* 36, 277–284 (2025).
19. Andre, T., *et al.* Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. *N Engl J Med* 383, 2207–2218 (2020).
20. Ratovomanana, T., *et al.* Prediction of response to immune checkpoint blockade in patients with metastatic colorectal cancer with microsatellite instability. *Ann Oncol* 34, 703–713 (2023).
21. Chan, E.M., *et al.* WRN helicase is a synthetic lethal target in microsatellite unstable cancers. *Nature* 568, 551–556 (2019).
22. Kategaya, L., Perumal, S.K., Hager, J.H. & Belmont, L.D. Werner Syndrome Helicase Is Required for the Survival of Cancer Cells with Microsatellite Instability. *iScience* 13, 488–497 (2019).
23. Lieb, S., *et al.* Werner syndrome helicase is a selective vulnerability of microsatellite instability-high tumor cells. *Elife* 8(2019).
24. Behan, F.M., *et al.* Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* 568, 511–516 (2019).
25. Picco, G., *et al.* Werner Helicase Is a Synthetic-Lethal Vulnerability in Mismatch Repair-Deficient Colorectal Cancer Refractory to Targeted Therapies, Chemotherapy, and Immunotherapy. *Cancer Discov* 11, 1923–1937 (2021).
26. Matos-Rodrigues, G., *et al.* S1-END-seq reveals DNA secondary structures in human cells. *Mol Cell* 82, 3538–3552 e3535 (2022).
27. van Wietmarschen, N., *et al.* Repeat expansions confer WRN dependence in microsatellite-unstable cancers. *Nature* 586, 292–298 (2020).
28. Wainberg, Z.A. WRN Helicase: Is There More to MSI-H than Immunotherapy? *Cancer Discov* 14, 1369–1371 (2024).

29. Picco, G., *et al.* Novel WRN Helicase Inhibitors Selectively Target Microsatellite-Unstable Cancer Cells. *Cancer Discov* 14, 1457–1475 (2024).
30. Erwin, G.S., *et al.* Recurrent repeat expansions in human cancer genomes. *Nature* 613, 96–102 (2023).
31. Xu, L., *et al.* Long-read sequencing identifies novel structural variations in colorectal cancer. *PLoS Genet* 19, e1010514 (2023).
32. Schloissnig, S., *et al.* Structural variation in 1,019 diverse humans based on long-read sequencing. *Nature* 644, 442–452 (2025).
33. Ehmsen, K.T. & Heyer, W.D. Saccharomyces cerevisiae Mus81-Mms4 is a catalytic, DNA structure-selective endonuclease. *Nucleic Acids Res* 36, 2182–2195 (2008).
34. Gaillard, P.H.L., Noguchi, E., Shanahan, P. & Russell, P. The endogenous Mus81-Eme1 complex resolves Holliday junctions by a nick and counternick mechanism. *Mol Cell* 12, 747–759 (2003).
35. Fletcher, C.T., *et al.* Structural insights into WRN helicase reveal conformational states and opportunities for MSI-H cancer drug discovery. *Commun Biol* 9(2026).
36. Fowler, F., *et al.* Microsatellite instable cancer cells acquire on-target resistance mutations to WRN helicase inhibitors. *Mol Cancer Ther* (2026).
37. Picco, G., *et al.* On-target mutations confer resistance to WRN helicase inhibitors in Microsatellite Unstable Cancer Cells. *bioRxiv*, 2025.2012.2023.695783 (2025).
38. Orcholski, M.E., *et al.* Anticipating on-target resistance to WRN inhibitors in microsatellite unstable cancers. *bioRxiv*, 2026.2001.2022.700152 (2026).
39. Ferretti, S., *et al.* Discovery of WRN inhibitor HRO761 with synthetic lethality in MSI cancers. *Nature* 629, 443–449 (2024).
40. Baltgalvis, K.A., *et al.* Chemoproteomic discovery of a covalent allosteric inhibitor of WRN helicase. *Nature* 629, 435–442 (2024).
41. Iyer, R.R., Pluciennik, A., Napierala, M. & Wells, R.D. DNA triplet repeat expansion and mismatch repair. *Annu Rev Biochem* 84, 199–226 (2015).
42. Pecina-Slaus, N., Kafka, A., Salamon, I. & Bukovac, A. Mismatch Repair Pathway, Genome Stability and Cancer. *Front Mol Biosci* 7, 122 (2020).
43. Genschel, J., Littman, S.J., Drummond, J.T. & Modrich, P. Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. *J Biol Chem* 273, 19895–19901 (1998).
44. Palombo, F., *et al.* hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr Biol* 6, 1181–1184 (1996).
45. Yao, X., *et al.* Different mutator phenotypes in Mlh1- versus Pms2-deficient mice. *Proc Natl Acad Sci U S A* 96, 6850–6855 (1999).
46. Wang, N., *et al.* Distinct mismatch-repair complex genes set neuronal CAG-repeat expansion rate to drive selective pathogenesis in HD mice. *Cell* 188, 1524–1544 e1522 (2025).

47. Mouro Pinto, R., *et al.* In vivo CRISPR-Cas9 genome editing in mice identifies genetic modifiers of somatic CAG repeat instability in Huntington's disease. *Nat Genet* 57, 314–322 (2025).
48. Ratovomanana, T., *et al.* Performance of Next-Generation Sequencing for the Detection of Microsatellite Instability in Colorectal Cancer With Deficient DNA Mismatch Repair. *Gastroenterology* 161, 814–826 e817 (2021).
49. Canela, A., *et al.* DNA Breaks and End Resection Measured Genome-wide by End Sequencing. *Mol Cell* 63, 898–911 (2016).
50. Wong, N., John, S., Nussenzweig, A. & Canela, A. END-seq: An Unbiased, High-Resolution, and Genome-Wide Approach to Map DNA Double-Strand Breaks and Resection in Human Cells. *Methods Mol Biol* 2153, 9–31 (2021).

Figures

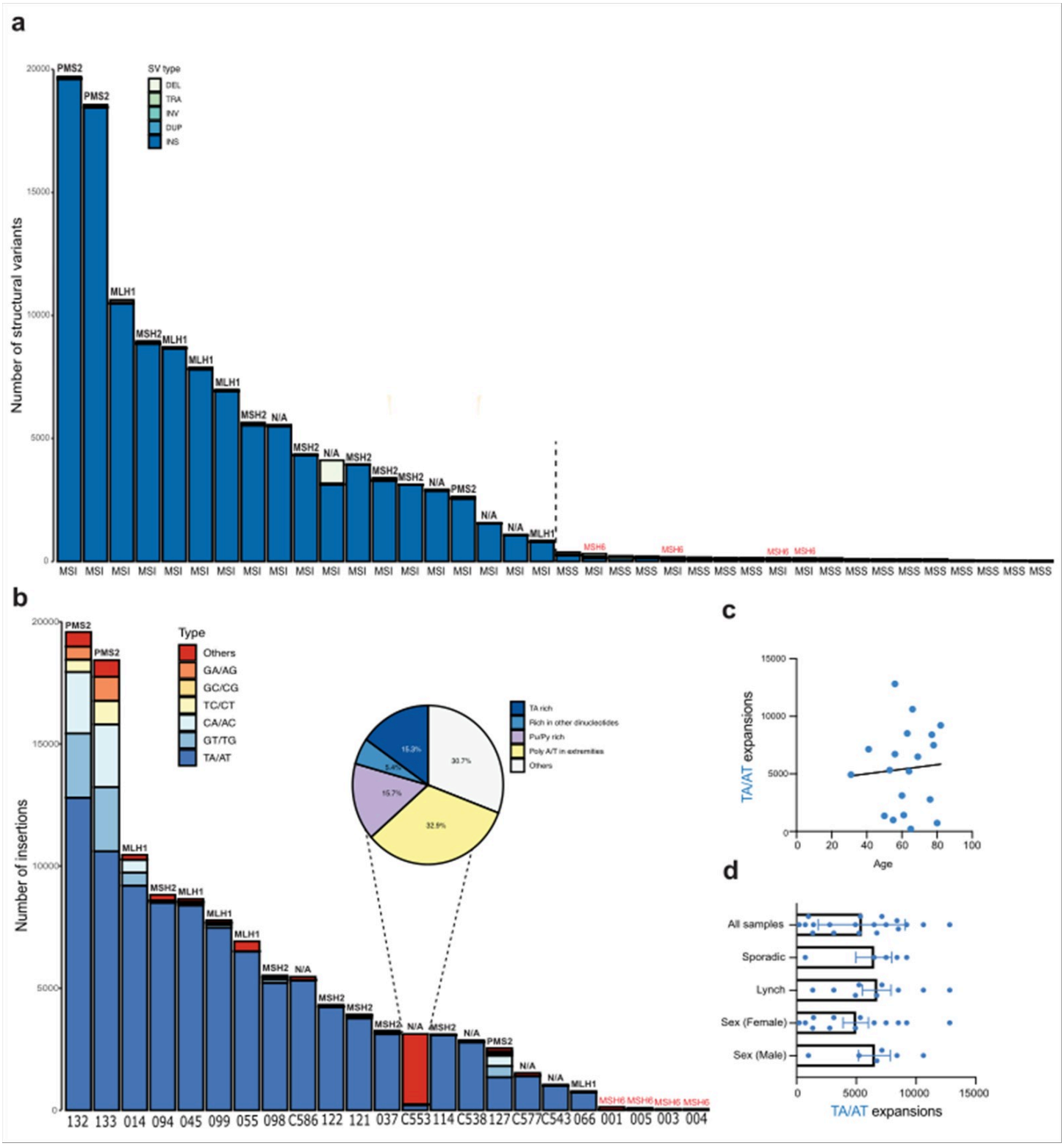

Figure 1

Structural variant analysis in MSI/MMRD and MSS CRC. (a) Number of structural variant types in the tumors. **(b)** Insertions representing dinucleotide pure repeats (≥80%). The “Others” category represents the remaining insertions. The pie chart represents the tumor C553 insertion composition. We filtered the sequences according to several additional criteria: a TA and dinucleotide content between 50% and 80%, a homopurine/homopyrimidine content between 70% and 100%, and the presence of a poly(A/T) tract at

either the 5' or 3' end. Sequences that did not fall into any of these categories were grouped into the “other” category. In tumor C553, the insertions consisted of approximately one-third (32.9%) sequences containing poly(A/T) ends that resembled transposable elements, 15.7% sequences with more than 70% homopurine/homopyrimidine content, 15.3% sequences containing more than 50% TA motifs, 5.4% sequences enriched in dinucleotides other than TA, and finally about one-third (30.7%) sequences harboring other unidentified motifs. **(c)** Relation between TA expansion number and patient age. **(d)** Relation between TA expansion number and patient sex/type of disease.

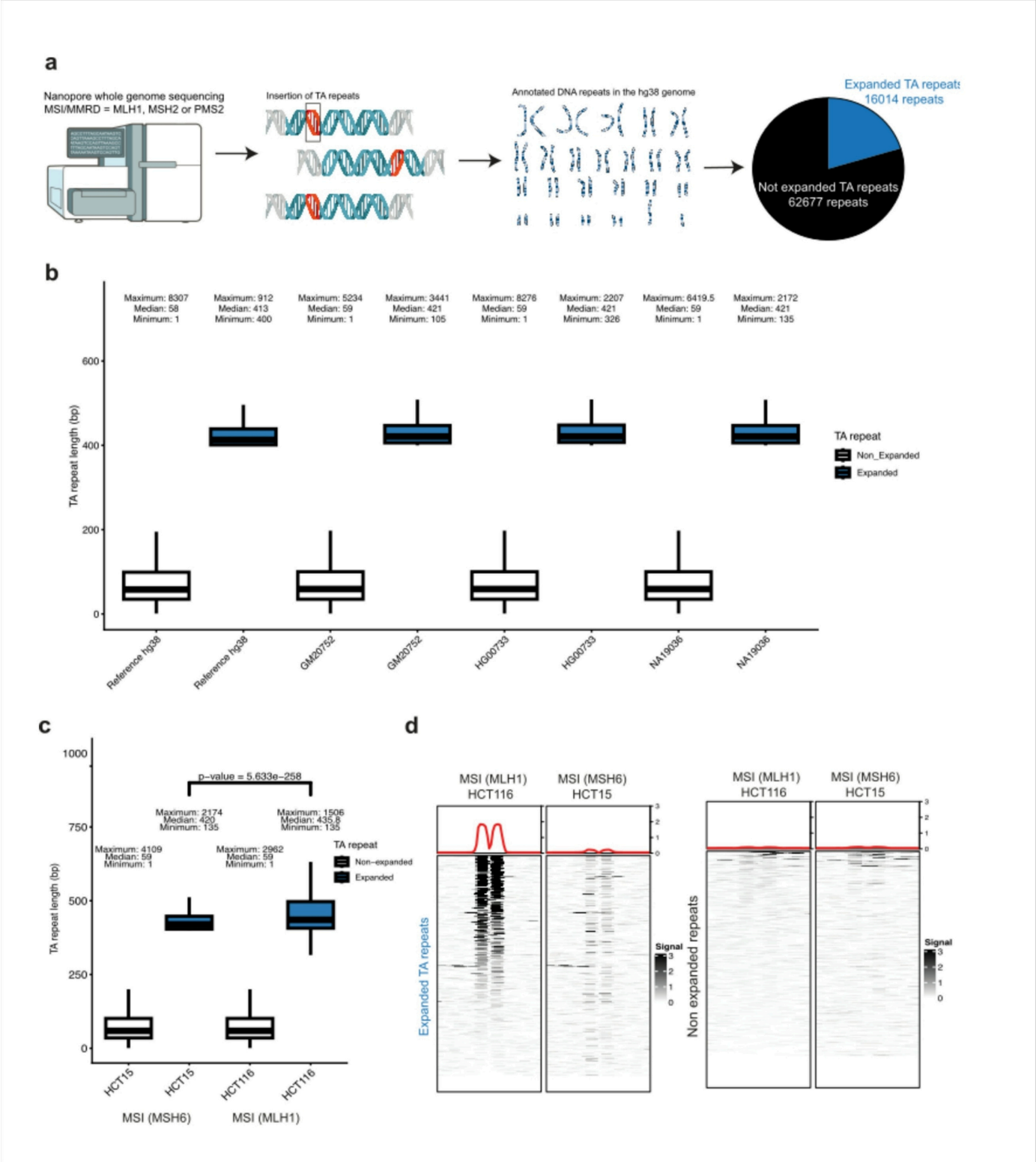

Figure 2

Expanded TA repeats are longer and prone to form alternative DNA structures. (a) Schematic model of expanded and non-expanded TA repeat characterization using hg38 genome. **(b)** TA repeat length measured in bp in expanded and non-expanded regions in the hg38 genome reference and in three different whole genomes sequenced using Nanopore sequencing. **(c)** Comparison between TA repeat length measured in bp in expanded and non-expanded regions in HCT15 (MSI/MSH6D) and HCT116 (MSI/MLH1D) cell lines. Wilcoxon test was assessed between the length of TA repeat expanded regions (p-value = 5.633^{e-258}). **(d)** S1-END-seq heatmaps comparing the S1-END-seq signal in reads per million (RPM) related to 500 bp from the center of expanded TA repeats (left panel) and non-expanded TA repeats (right panel) in HCT116 (MSI/MLH1D) and HCT15 (MSI/MSH6D) cell lines.

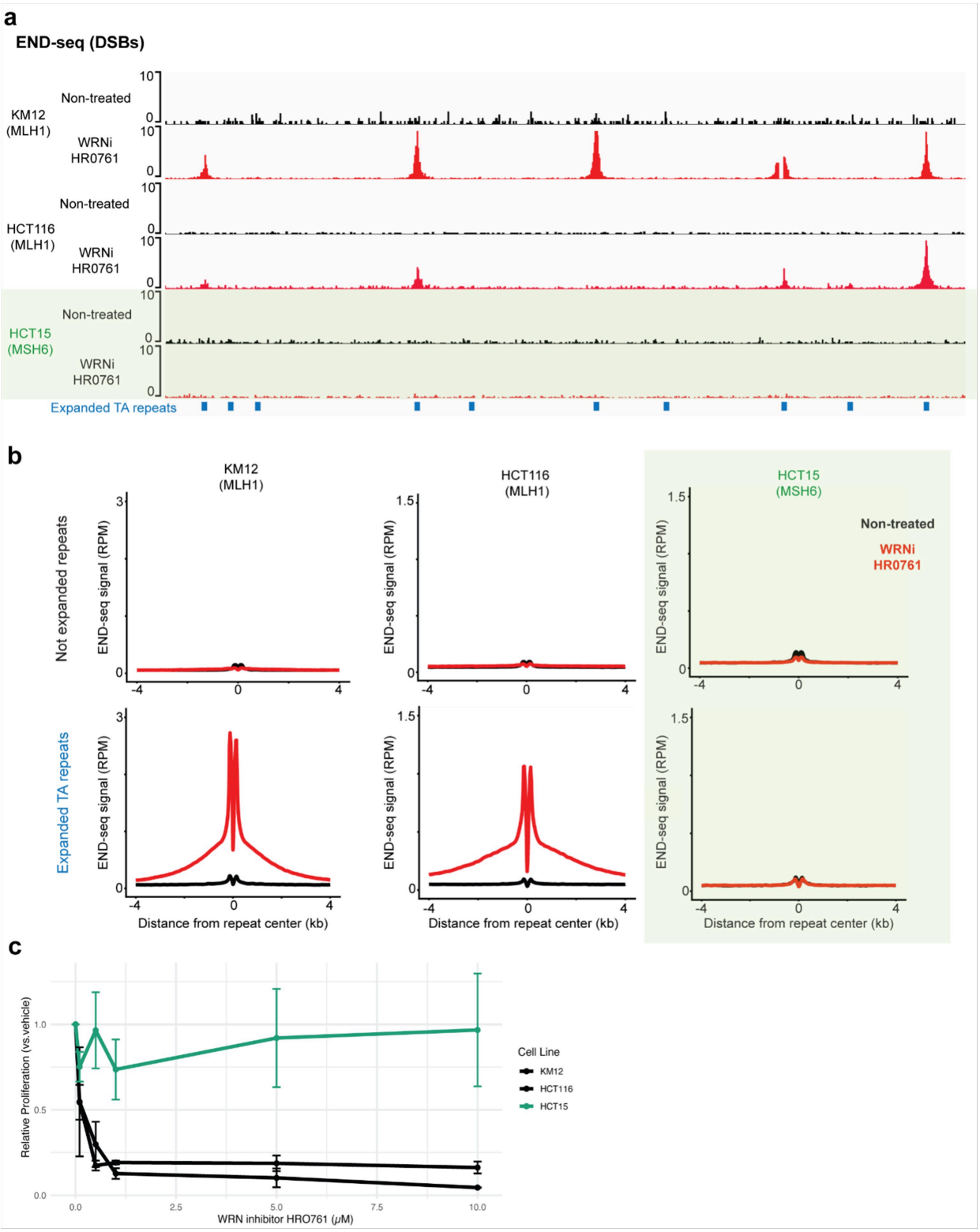

Figure 3

MMRD genotype-specific accumulation of breaks in expanded TA repeats and cytotoxicity upon WRN inhibition (a) Genome browser screenshots shown as RPM (reads per million) for END-seq in KM12, HCT116, and HCT15 colon cancer cell lines treated with HRO761 (red) or DMSO (black) for 48 hours. **(b)** Aggregate plot of END-seq signal (RPM – reads per million) at the center of the expanded TA repeats region or not expanded repeats region in KM12, HCT116, and HCT15 cell lines treated with HRO761 or

DMSO for 48 hours. **(c)** Cell viability assay evaluating the sensitivity of colon cancer cell lines KM12, HCT116, and HCT15 treated with increasing doses of HRO761 WRNi.

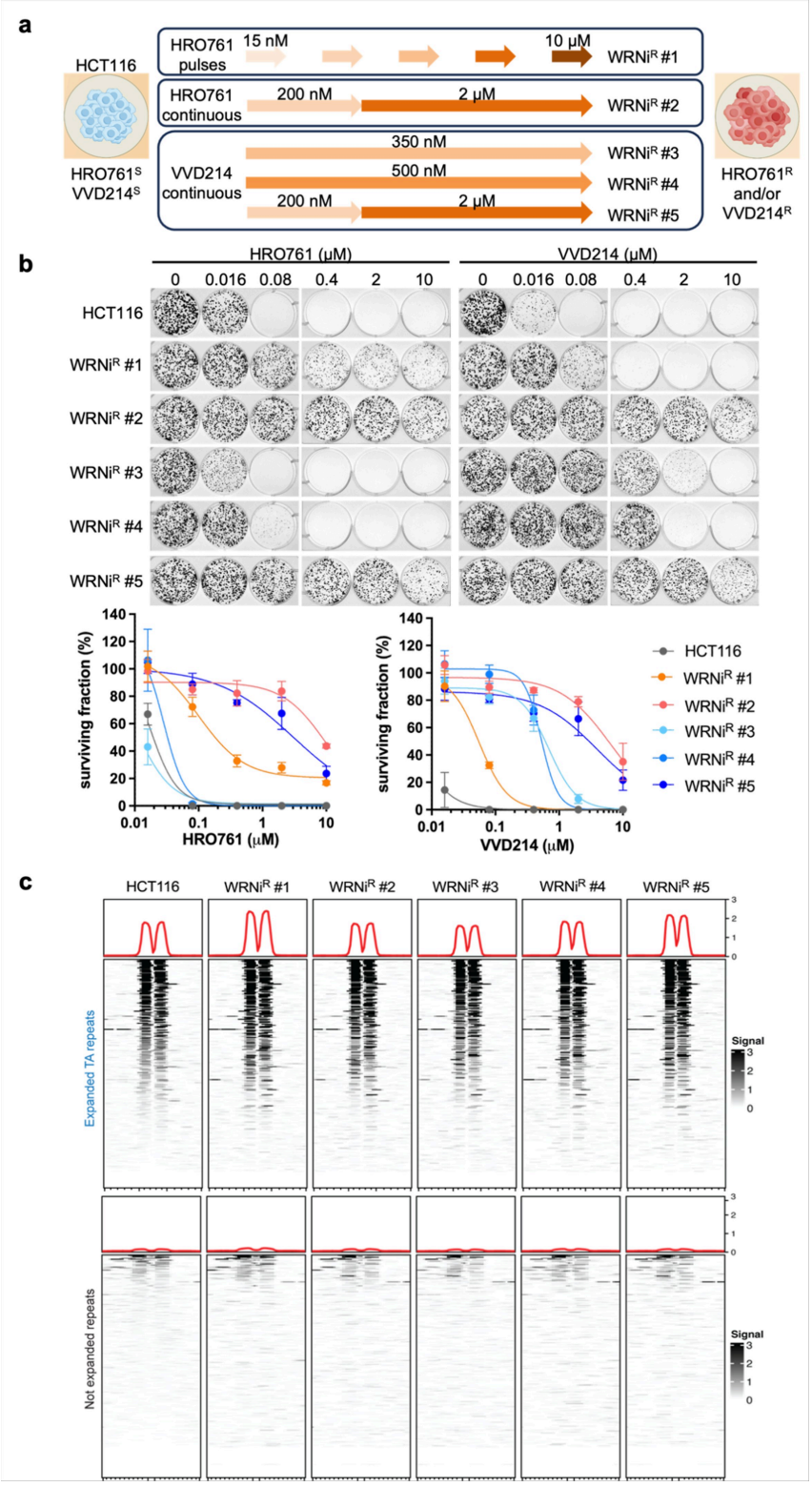

Figure 4

Expanded TA repeats form cruciform structures in WRN-inhibitor-resistant cell lines. (a) Schematic representation of the workflow used for the establishment of HCT116 cell lines resistant to HRO761 and

VVD214 molecules. Parental HCT116 cells were either treated with pulses of increasing doses of HRO761 (pulse protocol) or continuously exposed to WRN inhibitors (continuous protocol) until cells grew without apparent toxicity. **(b)** Clonogenic survival assay of HCT116 and WRNi^{R} cell lines following treatment with increasing concentrations of HRO761 and VVD214. Upper panels: representative images of crystal violet-stained colonies. Lower panels: quantification of the surviving fraction of HCT116 and WRNi^{R} cell lines treated with HRO761 or VVD214. **(c)** S1-END-seq heatmaps comparing the S1-END-seq signal in reads per million (RPM) related to 500 bp from the center of expanded TA repeats (top panel) and non-expanded TA repeats (bottom panel) in HCT116 cell lines.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuppTable1.xlsx
- SuppTable2.xlsx
- SuppTable3.xlsx
- Saabuncroppedblots.pdf
- SupplementaryFigures.docx