Genomic uracil occurence and the role of dUTPase in transgenic models

Short Summary of PhD Thesis

Hajnalka Laura Pálinkás

Genome Metabolism and Repair Research Group

Institute of Enzymology

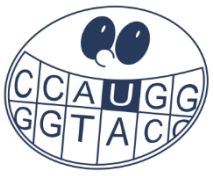
Research Centre for Natural Sciences

Doctoral School of Multidisciplinary Medical Science

Faculty of Medicine

University of Szeged

Supervisor: Beáta G. Vértessy PhD, DSc



Szeged

2020

List of scientific publications

This thesis is based on the following peer-reviewed publications:

I. Róna G, Scheer I, Nagy K, Pálinkás HL, Tihanyi G, Borsos M, Békési A, Vértessy BG. Detection of uracil within DNA using a sensitive labeling method for *in vitro* and cellular applications. Nucleic Acids Res. 2016; 44(3):e28. doi: 10.1093/nar/gkv977. **IF: 9.202**

II. Pálinkás HL, Rácz GA, Gál Z, Hoffmann OI, Tihanyi G, Róna G, Gócza E, Hiripi L, Vértessy BG. CRISPR/Cas9-Mediated Knock-Out of dUTPase in Mice Leads to Early Embryonic Lethality. Biomolecules. 2019; 9(4). pii: E136. doi: 10.3390/biom9040136. **IF: 4.694**

Cumulative impact factor of publications directly related to the thesis: 13.896

Preprint manuscript directly related to the subject of the thesis:

III. Pálinkás HL, Békési A, Róna G, Pongor L, Tihanyi G, Holub E, Póti Á, Gemma C, *et al*. Genome-wide alterations of uracil distribution patterns in human DNA upon chemotherapeutic treatments. bioRχiv preprint server, 2020. doi: https://doi.org/10.1101/2020.03.04.976977

Additional peer-reviewed publications, not included in this thesis:

IV. Szabó JE, Németh V, Papp-Kádár V, Nyíri K, Leveles I, Bendes AÁ, Zagyva I, Róna G, Pálinkás HL, *et al*. Highly potent dUTPase inhibition by a bacterial repressor protein reveals a novel mechanism for gene expression control. Nucleic Acids Res. 2014; 42(19):11912-20. doi: 10.1093/nar/gku882. **IF: 9.112**

V. Róna G, Pálinkás HL, Borsos M, Horváth A, Scheer I, Benedek A, Nagy GN, Zagyva I, Vértessy BG. NLS copy-number variation governs efficiency of nuclear import--case study on dUTPases. FEBS J. 2014; 281(24):5463-78. doi: 10.1111/febs.13086. **IF: 4.001**

VI. Periyasamy M, Singh AK, Gemma C, Kranjec C, Farzan R, Leach DA, Navaratnam N, Pálinkás HL, Vértessy BG, *et al*. p53 controls expression of the DNA deaminase APOBEC3B to limit its potential mutagenic activity in cancer cells. Nucleic Acids Res. 2017; 45(19):11056-11069. doi: 10.1093/nar/gkx721. **IF: 11.561**

VII. Molnár P, Marton L, Izrael R, Pálinkás HL, Vértessy BG. Uracil moieties in *Plasmodium falciparum* genomic DNA. FEBS Open Bio. 2018; 8(11):1763-1772. doi: 10.1002/2211-5463.12458. **IF: 1.05**

Cumulative impact factor of publications which are not included in the thesis: 25.724

1. Introduction

The faithful maintenance and transmission of genetic information is an essential task for all known living organisms. However, besides of the well-known Watson-Crick base pairs, DNA sequence also contains a large variety of modified bases and DNA defects. In my research, I focus on uracil, one of the most frequently occurring non-canonical base in DNA, implicated in several physiological processes. Two independent pathways may lead to the presence of uracil in DNA: thymine replacing misincorporation and hydrolytic cytosine deamination. The incorporation of uracil instead of thymine does not cause mutation, since uracil will base pair with adenine in the same way as thymine. Uracil misincorporation depends on the cellular deoxyuridine triphosphate / deoxythymidine triphosphate (dUTP/dTTP) ratio which is normally kept extremely low by two key enzymes: thymidylate synthase (TS) and deoxyuridine triphosphatase (dUTPase) [1,2]. Briefly, dUTPase can efficiently decrease dUTP level by hydrolizing it to deoxyuridine monophosphate (dUMP), from which TS synthetises deoxythymidine monophosphate (dTMP), the precursor of dTTP. Oxidative deamination of cytosines incorporated into the double helix, is one of the most common spontaneous DNA-damaging processes. The emerging uracil:guanine mismatch can cause a stable point mutation if no DNA repair is performed until the next replication cycle, as adenine will be incorporated into the DNA instead of guanine (leading to C:G to T:A transition) [3].

Thymidylate (deoxythymidine monophosphate, dTMP) biosynthesis is essential in proliferating cells, as it is the only source of dTTP synthesis which is one of the precursors of DNA synthesis. The bottleneck of thymidylate synthesis is dUMP, whose methylation is catalyzed by TS into dTMP. Since maintenance of proper dNTP pool and thereby genomic integrity are essential for cellular function and survival especially in actively replicating cells such as cancer cells, TS is a common target for chemotherapeutic agents [1]. Fluoropyrimidines (5-fluorouracil (5FU) and 5-fluoro-2'-deoxyuridine (5FdUR)) belong to the oldest and most-effective chemotherapeutics. The exact mechanism of their toxicity is not yet clear but they have a complex mode of action involving: i) irreversible inhibition of the enzyme TS by their metabolite 5-fluoro-dUMP (5FdUMP) and subsequent nucleotide pool imbalance (depletion of dTTP levels and elevated dUTP/dTTP ratio); and ii) misincorporation of dUMP and 5FdUMP into DNA by polymerases [4]. Antifolate derivatives inhibit the folate cycle or act directly by forming complex with TS (such as raltitrexed (RTX)), resulting in decreased dTTP and increased dUTP levels [4].

Since dUMP has been shown to be a key intermediate molecule for thymidylate biosynthesis, dUTPase has remarkable function prior to the action of TS enzyme. Most eukaryotic organisms

have two isoforms of dUTPase arising from alternative splicing of the mRNA due to different promoter and 5' exon usage [5]. dUTPase isoforms in vertebrates including *Mus musculus* and *Homo sapiens,* are translocated into the nucleus (nDut) or into the mitochondria (mDut). The enzyme dUTPase was found to be essential in *Escherichia coli* (*E. coli*) [6], *Saccharomyces cerevisiae* (*S. cerevisiae*) [7], *Caenorhabditis elegans* (*C. elegans*) [8], *Trypanosoma brucei* (*T. brucei*) [9], *Arabidopsis thaliana* [10] and *Mycobacterium smegmatis* [11], where the deficiency of dUTPase led to cell death. Additionally, silencing of dUTPase resulted in a lethal phenotype in early pupal stages of *Drosophila melanogaster,* whereas under physiological conditions, larvae, pupae and imago maintain greatly elevated uracil content of DNA [12,13]. Importantly, the simultaneous depletion of the uracil-DNA N-glycosylase, Ung1 has been shown to rescue the lack of dUTPase in *E. coli* [14], *S. cerevisiae* [15] and *C. elegans* [16], and these double mutant organisms incorporate uracil into their DNA with a high frequency. In mammalian cell line experiments, dUTPase disruption was only achieved so far by RNA silencing where it has been found that the expression level of dUTPase significantly affects the response to chemotherapy targeting TS. On the one hand, overexpression of dUTPase induced resistance to 5FdUR in human cancer cells [17,18]. On the other hand, silencing or inhibition of dUTPase increased responses to TS inhibitors leading to dUTP pool increment, replication defects and cytotoxicity [19,20,21,22]. Recently, a genetic study has been reported where a monogenic syndrome with diabetes and bone marrow failure in human patients could be associated with a missense mutation in dUTPase gene [23]. Apparent lack of genetic polymorphisms for dUTPase in the human population implies that deficiency may lead to severe consequences which are incompatible with life.

The appearing DNA lesions are primarily subjected for post-replication base excision DNA repair (BER) and secondly the mismatch repair (MMR) pathway. Uracil-DNA glycosylases (UDGs) are responsible for the specific recognition of the non-canonical deoxynucleotides such as uracil in DNA and they cleave the β-N-glycosidic bond between deoxyribose and the erroneous base [24,25]. Mammals possess several UDGs, among them UNG, SMUG1, TDG, MBD4 are well characterized in the literature. The uracil-DNA glycosylase (UNG) protein is the most active enzyme in the removal of uracil from DNA in human cells [25,26,27]. UNG isoforms can specifically excise uracil from both double-stranded (dsDNA) and single-stranded DNA (ssDNA) regardless to its origin. The created abasic (apurinic/apyrimidinic (AP)) site is then recognized by the AP-endonuclease (APE). After the removal of the remaining sugar fragment, ssDNA undergoes gap filling by a DNA polymerase and sealing of the nick by a

DNA ligase. However, under either high dUTP/dTTP ratio or total thymine deprivation, uracils presumably will be reincorporated during repair synthesis. Transformation of uracil-excision repair into a hyperactive futile cycle may lead to numerous AP sites, single- and double-stranded DNA breaks (DSBs), stalled replication forks and other potentially yet undiscovered consequences which trigger the so-called thymine-less cell death [28,29,30].

Accumulating evidence has been presented that uracils also appear in DNA under normal physiological conditions and may have key impact in highly diverse processes. Long-term stabilization of uracilated DNA can be achieved by the efficient inactivation or absence of UDG enzymes involved in the main uracil-excision repair system. It has been shown that a number of prokaryotes lack both genes for dUTPase and UNG (*dut-, ung-* genotype) [31] which is expected to result in uracil-containing DNA (U-DNA). However, even more numerous genomes lack the dUTPase gene but possess the UNG gene (*dut-, ung+* genotype) threatening genetic stability [31]. Specific protein inhibitors of UNG were identified in *Bacillus subtilis* bacteriophages PBS1/PBS2 and Φ29 encoding UGI [32] and p56 [33], respectively. Intriguingly, genome of these specific bacteriophages contains almost completely uracil instead of thymine, similarly to the *Yersinia enterocolitica* ΦR1–37 [34] and the general *Staphylococcace* S6 [35] phages where potential UNG inhibitors are still undiscovered. Not only phages have their own UNG inhibitor, but a similar DNA mimic protein was also identified in the genome of *S. aureus* itself, termed SAUGI [36], encoded on a mobile genetic element. Another striking example is human immunodeficiency virus (HIV) which does not encode its own dUTPase or UNG. The elevated uracil content of HIV DNA was suggested to protect the viral genome against suicidal autointegration, thereby promote favorable integration process into chromosomal DNA and increase viral infectivity [37]. Surprisingly, among eukaryotes there are also organisms or tissues lacking dUTPase or UNG enzymes, namely high genomic uracil level (200–2000 uracil/million bases) was found in larvae, pupae and imago of *Drosophila melanogaster* [12]. The induction of uracil-excision repair at well-defined uracil sites can also serve important cellular functions. Uracil arisen by the performance of activation-induced cytosine deaminase (AID) is a normal intermediate in variable regions of immunoglobulin (Ig) genes as part of somatic hypermutation (SHM), gene conversion (GC) and class-switch recombination (CSR) processes in adaptive immunity [25,38].

Since genomic uracil occurrence was implicated as a key regulator factor in different fields of biology, various methods were developed to gain qualitative and quantitative information on uracil levels for *in vitro* studies. Each method (such as mass spectrometry (MS); UNG-ARP

assay; real-time polymerase chain reaction (PCR) based techniques) has limitations and they do not allow *in situ* cellular detection of genomic uracil residues. Furthermore, no adequate antibody has yet been described for uracil moieties, in contrary to other non-orthodox DNA bases. Since uracil residues also escape detection by standard sequencing methods, scientists have to develop new approaches and methods.

2. Aims

Considering the wide range of physiological and pathological conditions where U-DNA level is altered, expanding our knowledge about genomic uracil occurrence including causes and consequences is of particular importance.

Towards this end, novel, reliable detection methods providing either qualitative or quantitative knowledge of U-DNA metabolism are essential. Developing and using a catalytically inactive UNG2 (ΔUNG) sensor protein, we aimed to elaborate a dot-blot based U-DNA measurement technique, and thereby quantify genomic uracil level in different biological samples. My goal was to prove that our U-DNA sensor-based method allows *in situ* cellular detection of uracils in DNA through an immunocytochemical approach.

Afterwards, I aimed to further develop this specific U-DNA sensor construct to enable microscopic visualization of genomic uracil residues in complex eukaryotic cells. After experimental validation of the optimized ΔUNG-based sensor and U-DNA labeling method, I wished to investigate the nascent genomic uracil distribution, using confocal and super-resolution microscopy. I also wished to develop an adequate U-DNA immunoprecipitation method to gain position and sequence-specific information by next-generation sequencing.

In another aspect, genomic uracil accumulation highly depends on the preventive function of dUTPase, which supports low cellular dUTP/dTTP ratio serving as an immunity against uracil incorporation into DNA. Until now, dUTPase function was only studied with gene silencing in cell cultures but gene knock-out strategy has not been published. Therefore, we aimed to investigate the physiological role of dUTPase *in vivo*, applying CRISPR/Cas9-mediated gene editing in mouse model. Importantly, I wished to reveal whether knocking-out of dUTPase could result in viable offspring or dUTPase deficiency may be incompatible with mammalian life.

3. Materials and methods

HCT116, MEF and 293T cells were grown in McCoy's 5A (Gibco, Life Technologies), DMEM/F12 HAM (Sigma) and DMEM (Gibco) medium, respectively; supplemented with 50 μg/ml Penicillin-Streptomycin (Gibco) and 10% fetal bovine serum (FBS) (Gibco). Cells were cultured at 37°C in a humidified incubator with 5% CO_2 atmosphere. Schneider S2 cells were grown in Schneider Insect Medium (Sigma) with the same supplements and were kept in a 26°C incubator.

For dot-blot based U-DNA measurements, genomic DNA isolated from CJ236 (*dut*-, *ung*-) *E. coli* strain served as a uracil standard. The amount of the inert carrier DNA (ultrapure salmon sperm DNA (Invitrogen)), was kept constant during the two-third dilution series. The different DNA samples were spotted onto a prewetted positively charged nylon membrane (Amersham Hybond-Ny+; GE Healthcare) using a vacuum driven microfiltration apparatus (Bio-Dot, BioRad). Immobilization was performed with 2 h of incubation at 80°C. Membrane was blocked by incubation in blocking buffer ETBS-T (25 mM Tris-HCl, pH = 7.4, 2.7 mM KCl, 137 mM NaCl, 1 mM EDTA, 0.05% Tween-20) containing 100 μg/ml salmon sperm DNA, 5% non-fat milk powder and 10 mM β-mercaptoethanol. Membrane was incubated with the 3xFlag-ΔUNG construct (18.1 μg/ml) in blocking buffer overnight at 4°C. After several washing steps with ETBS-T, anti-Flag M2 antibody (Sigma) was added for 1 h (1:2000, in ETBS-T with 5% non-fat milk powder). After washing the membranes, horseradish peroxidase coupled secondary antibody was applied. Immunoreactive bands were visualized by enhanced chemiluminescence reagent (GE Healthcare) and images were captured by a BioRad ChemiDoc™ MP Imaging system. Densitometry was done using ImageJ software.

Immunofluorescence staining of uracil residues in bacterial DNA was done using XL1-Blue (*dut*+, *ung*+) *E. coli* and CJ236 (*dut*-, *ung*-) *E. coli* cells. Cells were fixed with Carnoy's fixative (ethanol:acetic acid:chloroform = 6:3:1) for 20 min at 4°C. Rehydration was performed as following: washing with 1:1 ethanol:PBS, 3:7 ethanol:PBS and PBS containing 0.05% Triton X-100 (PBST) for 5 min. Cells were washed once with GTE buffer (50 mM glucose, 20 mM Tris, pH = 7.5 and 10 mM EDTA), and resuspended in GTE buffer with 10 mg/ml lysozyme (Sigma). The suspension was applied onto poly-L-lysine coated cover glasses and cells were left to air-dry. Cells were washed with PBST and were blocked in blocking buffer (5% BSA in PBST) for 15 min. Uracil residues were visualized by applying 4.64 μg/ml of the Flag-ΔUNG-DsRed construct in blocking buffer, overnight at 4°C. After several washing steps with PBST, anti-FLAG M2 antibody was applied (1:2000) for 1 h. FLAG epitope was visualized

by applying Alexa 488 conjugated secondary antibody (1:1000, Molecular Probes). Cells were counterstained with 1 μg/ml DAPI (4',6-diamidino-2-phenylindole, Sigma) nucleic acid stain. To test our uracil-staining method in mammalian cell lines, MEF cells were transfected in a 6-well plate with 4 μg of normal pEGFP-N1 (purified from XL1-Blue cells) or uracil-rich pEGFP-N1 vector (purified from CJ236 cells) and 12 μl FuGENE HD (Promega) reagent according to the manufacturer's recommendation. Staining of extrachromosomal plasmid DNA in MEF cells was done as described above with minor modifications. After fixation and rehydration, epitope unmasking was done by applying 1 M HCl, 0.5% Triton X-100 for 15 min. After neutralization with 0.1 M $Na_2B_4O_7$ (pH = 8.5) for 5 min, blocking was done for 1 h in blocking buffer: 200 μg/ml salmon sperm DNA, 5% fetal goat serum (FGS), 3% fetal bovine serum albumin (BSA) and 0.05% Triton X-100 in PBS. Labeling was obtained applying FLAG-ΔUNG-DsRed, followed by treatments with anti-FLAG M2 primary and secondary antibodies. To gain FLAG-ΔUNG-SNAP construct, SNAP encoding sequence from pSNAPf was PCR amplified. The fragments digested by KpnI and XhoI were ligated into the KpnI/XhoI sites of the plasmid construct FLAG-ΔUNG-DsRed (in a pET-20b vector). Primers were synthesized by Sigma-Aldrich and the constructs were verified by sequencing at Microsynth Seqlab GmbH. All UNG constructs were expressed in the *E. coli* BL21(DE3) *ung*- 151 strain and purified using Ni-NTA affinity resin (Qiagen). For comparison of FLAG-ΔUNG-DsRed and FLAG-ΔUNG-SNAP sensor constructs, staining of plasmid aggregates was done.

Generation of UGI-expressing stable HCT116 cell line was carried out by retroviral transduction. Briefly, 293T cells (1.5×10^6 cells in T25 tissue culture flasks) were transfected with 1.5 μg pLGC-hUGI/EGFP, 0.5 μg pCMV-VSV-G envelope and 0.5 μg pGP packaging plasmids using Lipofectamine 3000 reagent (Invitrogen) according to the manufacturer's recommendation. The supernatant, containing lentiviral particles was filtered through a 0.45 μm filter (Merck Millipore) 36 h after the transfection. Successfully transduced HCT116 cells were collected by fluorescence activated cell sorting (FACS) for GFP-positive cells using a BD FACSAria III Cell sorter (BD Biosciences). For comparison, HCT116 cells were also transfected transiently with FuGENE HD. UGI-expressing cells were treated with 20 μM 5FdUR (Sigma) or 100 nM RTX (Sigma) for 48 h before fixation for immunocytochemistry or collecting them for genomic DNA (gDNA) purification. For dot-blot measurements or U-DNA-Seq, DNA of eukaryotic HCT116 cells (as well as bacterial samples) was purified using the Quick-DNA™ Miniprep Plus Kit (Zymo Research) using the recommendations of the manufacturer and elution was carried out in nuclease-free water.

For immunofluorescent staining of genomic uracil residues, HCT116 cells stably expressing UGI were seeded onto 24-well plates containing cover glasses or onto µ-Slides (ibidi GmbH). In case of dSTORM imaging, coverslips were coated with poly-D-lysine before seeding. Sub-confluent cultures of cells were fixed using 4% paraformaldehyde (PFA, pH = 7.4 in PBS) or Carnoy's fixative for 15 min. After washing or rehydration steps, epitope unmasking was done by applying 2 M HCl, 0.5% Triton X-100 for 30 min. After neutralization, cells were incubated in blocking solution I (TBS-T (50 mM Tris-HCl, pH = 7.4, 2.7 mM KCl, 137 mM NaCl, 0.05% Triton X-100) containing 5% non-fat dried milk for 15 min, followed by incubation in blocking buffer I supplemented with 200 µg/ml salmon sperm DNA for 45 min. Uracil residues were visualized by applying 4 µg/ml of the FLAG-ΔUNG-SNAP construct for 1 h in blocking buffer I with 200 µg/ml salmon sperm DNA. After several washing steps, cells were labelled with 2.5 µM SNAP-Surface Alexa Fluor 546 or 647 (NEB) for 20 min in blocking buffer II (5% FGS, 3% BSA and 0.05% Triton X-100 in PBS), and optionally counterstained with 1 µg/ml DAPI. Confocal images were acquired on a Zeiss LSCM 710 or a Leica TCS SP8 STED 3X microscope. For STED images 660 nm STED laser was used for depletion. The same image acquisition settings were applied on each corresponding sample to allow comparison. Fluorescence images were processed using ZEN software.

For U-DNA immunoprecipitation, 12 µg of gDNA was sonicated into fragments ranging between 100 and 500 basepairs with a BioRuptor (Diagenode). 25% of the samples was saved as input, and the remaining DNA was resuspended in the following IP buffer: 30 mM Tris-HCl, pH = 7.4, 140 mM NaCl, 0.01% Tween-20, 1 mM 0.5 mM EDTA, 15 mM β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 5 mM benzamidine. Immunoprecipitation was carried out with 15 µg of 1xFLAG-ΔUNG construct for 2.5 h at room temperature with constant rotation. Anti-FLAG M2 agarose beads (Sigma) were equilibrated in IP buffer and then added to the IP mixture for 16 h at 4°C with constant rotation. Beads were resuspended in elution buffer containing 1% sodium dodecyl sulphate (SDS), 0.1 M $NaHCO_3$. Elution of uracil sensor protein binding U-DNA was done by vortexing for 5 min with an additional incubation for 20 min with constant rotation. After centrifugation (13000 rpm for 3 min), supernatant was incubated with 150 µg/ml RNAse A (Epicentre) for 30 min, followed by the addition of 500 µg/ml Proteinase K (Sigma) for 1 h at 37°C for removal of residual RNA and proteins. Immunoprecipitated DNA was purified with NucleoSpin Gel and PCR Clean-up Kit (MACHEREY-NAGEL GmbH & Co. KG) according to the manufacturer's instructions. Enrichment of uracil in the DNA samples was examined by dot-blot assay, then input and enriched U-DNA samples were subjected to next-generation sequencing (NGS).

For dUTPase knock-out experiments, the T7 single-guide RNA (sgRNA) and the Cas9 mRNA were obtained from Sigma-Aldrich. CRISPR/Cas9 efficiency was tested in MEF cells by transfection with 2.5 μg Cas9 mRNA, 250 ng target sgRNA and Lipofectamine™ 3000 Transfection Reagent (ThermoFisher Scientific) in 6-well plates. 24 h after transfection, cells were maintained in a fresh medium for 24 h, and then the genomic DNA was extracted with a MasterPure™ DNA Purification Kit (Epicentre). After DNA amplification, Cel 1 cleavage assay was performed using the Transgenomic SURVEYOR Mutation Detection Kit.

All FVB/N mice used in the experiments was produced and maintained in the Animal Care Facility at the Agricultural Biotechnology Institute, National Agricultural Research and Innovation Centre (NAIK). Genotyping of the heterozygous founder animals was carried out by amplifying the CRISPR target sites from genomic DNA, and the fragments were cloned into SalI/EcoRI sites of pBluescript SK (+) vector. Individual bacterial colonies were purified with NucleoSpin Plasmid DNA Purification Kit, followed by sequencing. Off-target analysis was also performed using the online predictor CCTop. The genotypes of mice were determined by PCR of the total genomic DNA extracted from mouse tails or from embryos. Samples were dissolved in a DNA lysis buffer (0.1 M Tris-HCl, pH = 7.4, 0.2 M NaCl, 5 mM EDTA, 0.2% SDS) and DNA was extracted with phenol-chloroform. Genotyping from blastocysts was performed by semi-nested PCR, and DNA fragments were visualized by 1% agarose gel electrophoresis.

For blastocyst outgrowth assay, embryos were flushed out from the uteri of pregnant mice at 3.5 dpc in M2 medium. Blastocysts were individually cultured on 0.1% gelatin-coated, 12-well dishes, in KO-DMEM ES cell culture medium supplemented with 1000 U/mL LIF and 20% FBS, in 5% CO_2 at 37°C for four days. Outgrowths were photographed daily, and on the fourth day of culture, outgrowths were removed and genotyped by PCR. For western blot analysis, embryos at 10.5 dpc were dissected from pregnant mice, then washed with PBS and resuspended in a lysis buffer (20 mM HEPES, pH = 7.5, 420 mM NaCl, 1 mM EDTA, 2 mM dithiothreitol, 25% glycerol). Homogenization was assisted with vortex until the tissue was sufficiently disrupted. After centrifugation the supernatant samples were boiled with SDS buffer at 95°C for 5 min. Total proteins were resolved on a 12% polyacrylamide gel and transferred to the PVDF membrane (Immobilon-P, Merck Millipore). Membranes were blocked with 5% non-fat dried milk in TBS-T (25 mM Tris-HCl, pH = 7.4, 140 mM NaCl, 3 mM KCl, 0.05% Tween-20) for 1 h at 4°C and were developed against dUTPase (1:2000, Sigma) or α-actin (1:1000, Sigma) for loading control. Next steps were the same as used for immunoreactive visualization in dot-blot assays. Densitometry was done using BioRad Image Lab™ 6.0.

4. Results

4.1. Uracil-DNA detection applying catalytically inactive UNG constructs in *in vitro* and cellular studies

We developed a new dot-blot based method for relative quantification of genomic uracil content using a catalytically inactive, flag-tagged ΔUNG construct. After primary and secondary antibody labelling, U-DNA levels can be calculated *via* interpolation to a standard calibration curve, which could be obtained using uracilated DNA derived from (*dut-*, *ung-*) *E. coli* cells. Combination of *ung-* genetic background (using different bacterial strains or *Drosophila* S2 cell line) and treatment with well-known drugs targeting *de novo* thymidylate biosynthesis pathways (MTX, RTX, 5FdUR) led to significantly increased genomic uracil level compared to non-treated cells. Then, we proceeded to measure U-DNA level alterations using a human colon carcinoma cell line (HCT116), and found that similar drug treatments on their own do not lead to elevated U-DNA level. To allow direct capture of the nascently appearing uracil moieties in DNA, inhibition of the major uracil-recognizing factor, UNG was essential. Using dual treatment (UGI expression and 5FdUR drug addition) we could obtain and detect significantly elevated U-DNA content, as compared to non-treated cells. Staining of genomic uracils was achieved either indirectly *via* immunocytochemistry against the Flag epitope tag or directly *via* the fluorescent DsRed signal in CJ236 (*dut-*, *ung-*) cells. According to our expectations, using the same labeling procedure in uracil-free XL1-Blue (*dut+*, *ung+*) *E. coli* cells did not result in any visible signal (Figure 1). In the uracil-rich milieu, U-DNA staining colocalized with the signal of DAPI used to counterstain DNA.

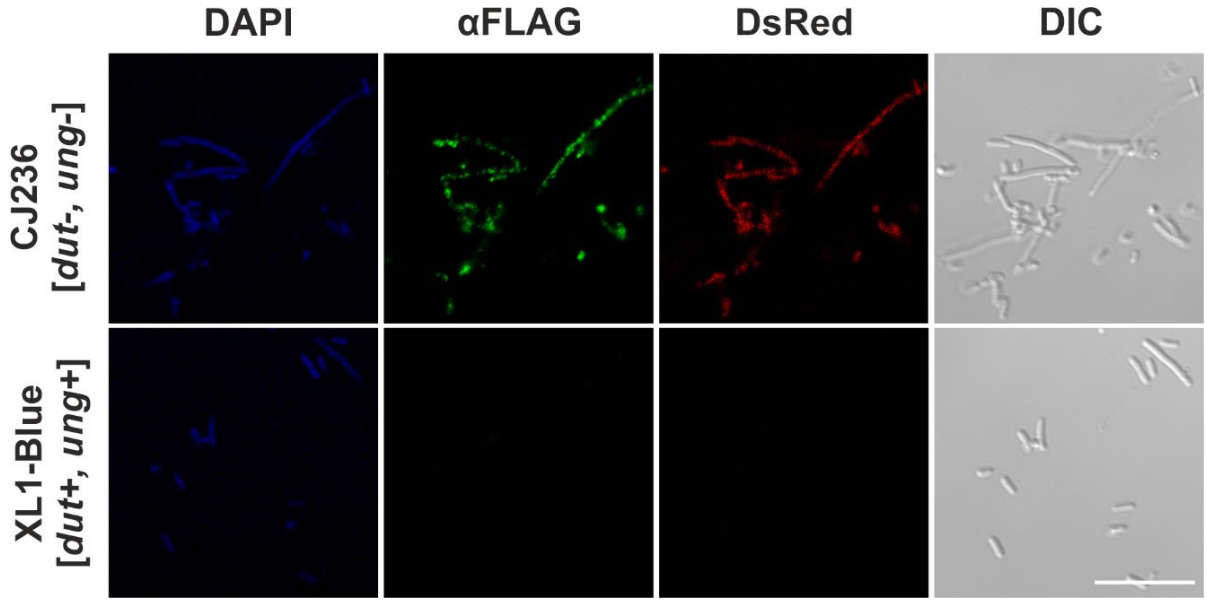


Figure 1. [39] Visualization of genomic uracil using the Flag-ΔUNG-DsRed sensor construct in uracil-containing CJ236 (*dut-, ung-*) *E. coli* cells as compared to uracil-free XL1-Blue (*dut+, ung+*) *E. coli* cells. Differential interference contrast (DIC) image was also acquired. Scale bar represents 10 μm.

Using *ung-/-* MEF cells for transfection with uracil-rich or normal plasmids, the herein presented U-DNA sensor protein succesfully detected cellular uracil-enriched plasmid aggregates colocalizing with DAPI staining.

4.2. Promising applications of ΔUNG-based U-DNA labeling techniques for analysis of uracil distribution pattern in human genomic DNA

We further sought to adapt the ΔUNG-based U-DNA labeling approach to visualize genomic uracil residues in complex eukaryotic cells. A versatile labeling technique could be achieved by cloning a SNAP-tag in the C terminal end of ΔUNG, yielding Flag-ΔUNG-SNAP. Comparison of Flag-ΔUNG-SNAP and Flag-ΔUNG-DsRed constructs demonstrated that the novel sensor shows also reliable U-DNA detection. Afterwards, an UNG-inhibited, human-derived, model cell line was established by retroviral transduction of human codon optimized UGI into HCT116 cells. Using the generated UGI-expressing cell line, drug treatment with 5FdUR or RTX elevated U-DNA level; and the Flag-ΔUNG-SNAP construct was succesfully applied for *in situ* staining of genomic uracil in these cells. U-DNA staining could be carried out with cross-linking PFA fixation of cells, instead of Carnoy fixation. Therefore, our staining procedures may be suitable for multi-color imaging in colocalization studies in the future. Genomic uracil distribution pattern was also investigated by super-resolution microscopy using appropriate SNAP substrates for either STED or dSTORM applications. A characteristic pattern could be observed in drug (5FdUR or RTX) treated cells, that differed from U-DNA signal of non-treated, UGI-expressing cells, as revealed by dSTORM imaging (Figure 2). Uracil staining showed signal enrichment at the nuclear membrane and areas surrounding the nucleoli in the case of drug-treated cells.

To gain sequence-specific information of genomic uracil distribution, I developed the U-DNA-Seq method. Briefly, U-DNA immunoprecipitation was carried out applying the Flag-tagged, catalytically inactive ΔUNG sensor in purified and fragmented gDNA, followed by pull-down with anti-FLAG beads before sequencing. Successful enrichment of U-DNA could be confirmed by dot-blot assay for 5FdUR or RTX treated, UGI-expressing cells. Then, input DNA and enriched U-DNA samples could be subjected to library preparation and next-generation sequencing (NGS).

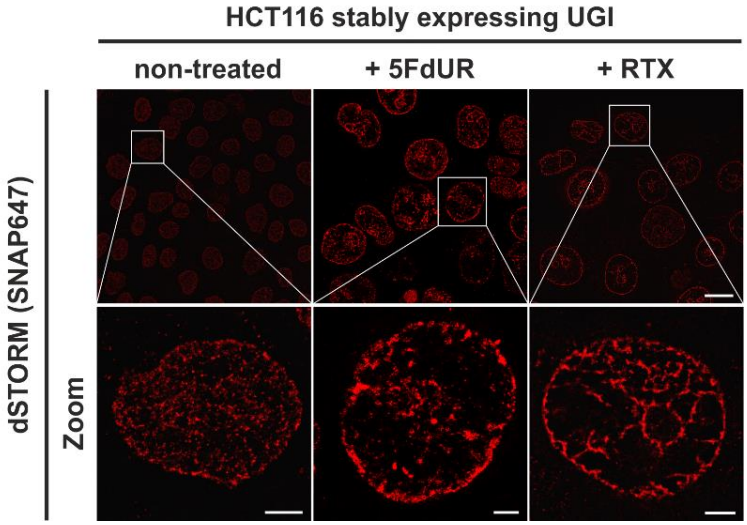


Figure 2. The Flag-ΔUNG-SNAP sensor enables super-resolution detection of genomic uracil by super-resolution microscopy. dSTORM imaging was performed on non-treated or drug-treated (5FdUR or RTX) HCT116 cells stably expressing UGI. Scale bar represents 10 μm for whole images and 2 μm for zoomed sections.

4.3. A critical investigation of the physiological role of dUTPase in a mammalian model, using CRISPR/Cas9-mediated gene editing

To investigate the effects of dUTPase deficiency *in vivo*, specific knock-out of *dut* gene was addressed by CRISPR/Cas9-mediated gene editing strategy. After validation of the chosen sgRNA and Cas9 mRNA by *in vitro* surveyor assay on cotransfected MEF cells, CRISPR/Cas9 system was used to produce dUTPase knock-out mice. Two founder animals (#2 and #4) were generated harboring distinct CRISPR/Cas9 editing events, namely six bp deletion and one base substitution (D6, M1) in mice #2; and 47 bp deletion (D47) in mice #4. CRISPR/Cas9-induced D47 mutation resulted in frameshift mutation on *dut* gene and consequently early stop codons. Further experiments were conducted on later generations of the mouse #4 with D47 mutation. Genotyping analysis using semi-nested PCR of numerous generations derived from heterozygous D47 crossings revealed that deficiency of dUTPase leads to embryonic lethality. Upon investigation of dissected embryos at different developmental stages, we found homozygous *dut* -/- mutant embryos only in blastocyst stage (3.5 dpc), but not at later stages, as shown in Table 1. Differences in the development of blastocysts with altered *dut* status were analysed in *in vitro* cultures. In *dut* -/- embryos, outgrowth of both ICM and TE regions were significantly reduced as compared to either *dut* +/+ WT or *dut* +/- variants (Figure 3), indicating developmental defects in the absence of dUTPase. In contrast, no significant difference between *dut* +/- and *dut* +/+ embryos were detected neither in blastocyst outgrowth assay, nor visually on dissected embryos at later embryonic stages (8.5 dpc, 9.5 dpc, 10.5 dpc).

DNA source	Genotype			Resorbed	No. Total
	+/+	+/-	-/-		
Postnatal	21	42	0	-	63
10.5 dpc	3	5	0	3	11
9.5 dpc	5	5	0	0	10
8.5 dpc	10	5	0	5	20
3.5 dpc	11	13	7	NA	31

Table 1. [40] Genotype analysis referring to *dut* gene after *dut* +/- D47 intercrossings. This table summarizes all the animals and embryos carrying *dut +/+; dut* +/- or *dut -/-* genotype, and also resorbed embryos with unknown genotype. NA, not applicable.

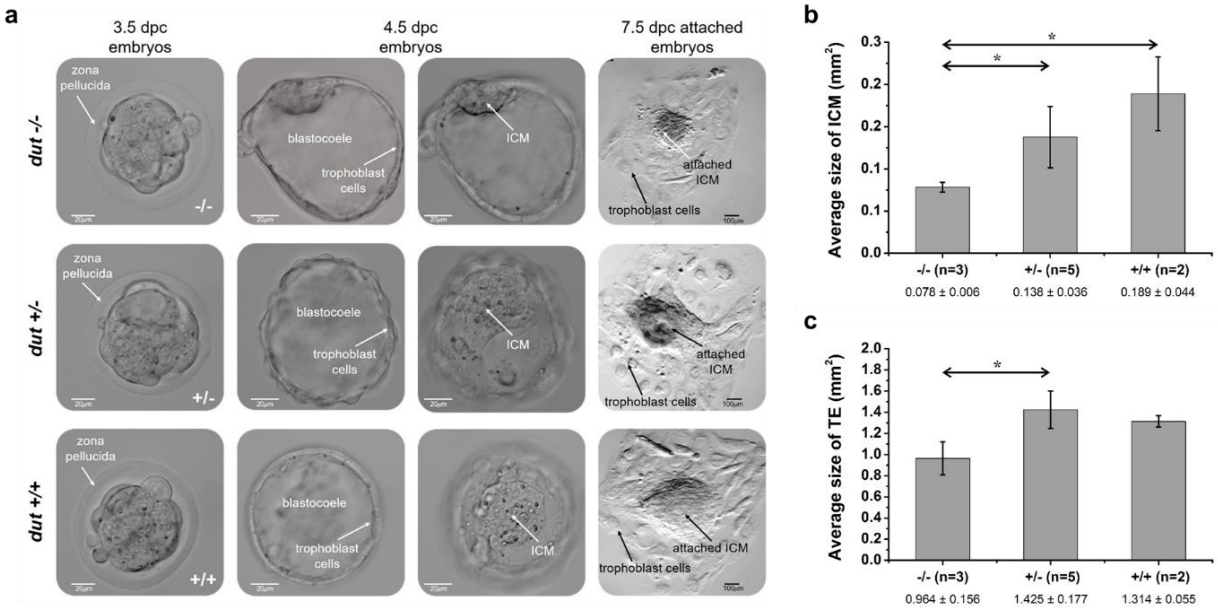


Figure 3. [40] ICM outgrowth assay for blastocysts produced by D47 heterozygous crosses. **(A)** Representative phase contrast images of *dut -/-*; *dut +/-*; *dut +/+* embryos in *in vitro* culture. The first column shows the isolated embryos at 3.5 dpc. The second and third columns represent the attached embryos at 4.5 dpc, focusing on the trophoblast cells or the ICM in the blastocoele. In the last column, outgrowths at 7.5 dpc are seen. Scale bar represents 20 μm, or 100 μm for images in the last column. Average size of ICM **(B)** and trophectoderm (TE) **(C)** were quantified for embryos at 7.5 dpc. Error bars indicate standard deviation. Sample size is also provided with n number for all the genotypes. * $p < 0.05$.

Examination by Western blots showed that *dut +/-* heterozygous embryos contain significantly reduced dUTPase protein level as compared to WT embryos (Figure 20). This difference might also apply for adult animals, but this needs further organ-specific experimental evidence. Regarding adulthood, progeny of D47 strains did not show any gross or fertility abnormalities, or systemic diseases during our study. Taken together, *dut +/-* heterozygosity and related potential suboptimal dUTPase protein level did not result in any visible phenotypic change in mice.

5. Discussion

In my PhD thesis, I have presented a novel, reliable detection method providing either qualitative or quantitative knowledge of U-DNA metabolism. It was demonstrated here that our dot-blot based U-DNA measurement technique relying on a catalytically inactive UNG2 (ΔUNG) sensor protein enables quantification of genomic uracil level in different biological samples. I also verified that our U-DNA sensor-based method allows *in situ* cellular detection of uracils in bacterial or extrachromosomal DNA through an immunocytochemical approach.

This specific U-DNA sensor construct and staining method was further developed, which allowed imaging of genomic uracil residues within DNA of complex eukaryotic cells using confocal as well as super-resolution microscopy. I also managed to establish a novel HCT116 cell line stably expressing the UNG-inhibitor UGI protein, by retroviral transduction. Drug (5FdUR or RTX) treatment of the UNG-inhibited cells resulted in significantly elevated U-DNA level, thus gaining an adequate model system for investigation of nascent genomic uracil distribution. To my best knowledge this is the first time that a labeling technique applicable for *in situ* microscopic detection of human genomic uracil residues, is presented. Super-resolution imaging by dSTORM revealed that drug treated cells showed different U-DNA pattern than non-treated cells. Namely, signal enrichment could be observed at the nuclear membrane and areas surrounding the nucleoli in the case of drug-treated cells. Colocalization studies are needed to be done in the future to explore potential correlations between U-DNA distribution pattern and localization pattern of histone markers or other chromatin factors. Importantly, I have also described an adequate DNA immunoprecipitation method (U-DNA-Seq) providing sequence-specific information of genomic uracil distribution by next-generation sequencing.

We also aimed to investigate the physiological role of dUTPase *in vivo*, applying CRISPR/Cas9-mediated gene editing in mouse model. Based on our results, we could conclude that deficiency of dUTPase leads to early embryonic lethality in mice. Homozygous *dut -/-* mutant embryos were only found in blastocyst stage, and we suggest that further developmental abnormalities may lead to lethality occuring in around or shortly after implantation. Considering the timing of the maternal-to-zygotic transition in mice, the first several duplication cycles of embryogenesis may occur independently on dUTPase expression. Since U-DNA occurrence is involved in the wide range of physiological and pathological processes, the herein presented *dut +/-* mouse model and early *dut -/-* embryos might be beneficial to gain new insights into the details of U-DNA metabolism.

6. Theses of the PhD dissertation

Theses of my PhD dissertation can be summarized as follows:

1. Application of a catalytically inactive, Flag-ΔUNG-DsRed construct uniquely allows *in situ* microscopic visualization of uracil residues within bacterial DNA, either indirectly *via* immunocytochemistry against the Flag tag or directly *via* the fluorescent DsRed signal. This labeling method has the potential to be further extended for detection of uracils within the highly complex chromatin of human cells.
2. The dUTPase encoding *dut* gene can be successfully targeted by CRISPR/Cas9-mediated gene editing to study the effects of dUTPase deficiency *in vivo* in mice.
3. Using CRISPR/Cas9 system in mice, only heterozygous *dut* +/- offspring could be achieved, while viable homozygous *dut* -/- offspring could never be found, implying that dUTPase deficiency lead to prenatally lethal phenotype.
4. Investigation of dissected embryos at different developmental stages, showed that homozygous *dut* -/- mutant embryos exist only in blastocyst stage, but not at later stages, suggesting early embryonic lethality in the absence of dUTPase in mice.
5. *In vitro* outgrowth assays demonstrated that both ICM and TE formation are significantly impaired in *dut* -/- blastocysts, indicating that lack of dUTPase may cause developmental defects leading to lethality around implantation into the uterus.
6. Examination by Western blots showed that *dut* +/- heterozygous embryos contain significantly reduced dUTPase protein level as compared to WT embryos, which difference might also apply for adult animals.

7. References

1. Grafstrom, R. H., Tseng, B. Y. & Goulian, M. The incorporation of uracil into animal cell DNA in vitro. *Cell* **15,** 131–40 (1978).
2. Vértessy, B. G. & Tóth, J. Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. *Acc. Chem. Res.* **42,** 97–106 (2009).
3. Vértessy, B. G. & Tóth, J. Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. *Acc. Chem. Res.* **42,** 97–106 (2009).
4. Grogan, B. C., Parker, J. B., Guminski, A. F. & Stivers, J. T. Effect of the thymidylate synthase inhibitors on dUTP and TTP pool levels and the activities of DNA repair glycosylases on uracil and 5-fluorouracil in DNA. *Biochemistry* **50,** 618–27 (2011).
5. Ladner, R. D., McNulty, D. E., Carr, S. A., Roberts, G. D. & Caradonna, S. J. Characterization of distinct nuclear and mitochondrial forms of human deoxyuridine triphosphate nucleotidohydrolase. *J. Biol. Chem.* **271,** 7745–51 (1996).
6. el-Hajj, H. H., Zhang, H. & Weiss, B. Lethality of a dut (deoxyuridine triphosphatase) mutation in Escherichia coli. *J. Bacteriol.* **170,** 1069–75 (1988).
7. Gadsden, M. H., McIntosh, E. M., Game, J. C., Wilson, P. J. & Haynes, R. H. dUTP pyrophosphatase is an essential enzyme in Saccharomyces cerevisiae. *EMBO J.* **12,** 4425–31 (1993).
8. Erdélyi, P. *et al.* Shared developmental roles and transcriptional control of autophagy and apoptosis in Caenorhabditis elegans. *J. Cell Sci.* **124,** 1510–8 (2011).
9. Castillo-Acosta, V. M., Estévez, A. M., Vidal, A. E., Ruiz-Perez, L. M. & González-Pacanowska, D. Depletion of dimeric all-alpha dUTPase induces DNA strand breaks and impairs cell cycle progression in Trypanosoma brucei. *Int. J. Biochem. Cell Biol.* **40,** 2901–13 (2008).
10. Dubois, E. *et al.* Homologous recombination is stimulated by a decrease in dUTPase in Arabidopsis. *PLoS One* **6,** e18658 (2011).
11. Pecsi, I. *et al.* The dUTPase enzyme is essential in Mycobacterium smegmatis. *PLoS One* **7,** e37461 (2012).
12. Muha, V. *et al.* Uracil-containing DNA in Drosophila: stability, stage-specific accumulation, and developmental involvement. *PLoS Genet.* **8,** e1002738 (2012).
13. Horváth, A., Békési, A., Muha, V., Erdélyi, M. & Vértessy, B. G. Expanding the DNA alphabet in the fruit fly: uracil enrichment in genomic DNA. *Fly (Austin).* **7,** 23–7 (2013).
14. Tye, B. K., Chien, J., Lehman, I. R., Duncan, B. K. & Warner, H. R. Uracil incorporation: a source of pulse-labeled DNA fragments in the replication of the Escherichia coli chromosome. *Proc. Natl. Acad. Sci. U. S. A.* **75,** 233–7 (1978).
15. Guillet, M., Van Der Kemp, P. A. & Boiteux, S. dUTPase activity is critical to maintain genetic stability in Saccharomyces cerevisiae. *Nucleic Acids Res.* **34,** 2056–66 (2006).
16. Dengg, M. *et al.* Abrogation of the CLK-2 checkpoint leads to tolerance to base-excision repair intermediates. *EMBO Rep.* **7,** 1046–51 (2006).
17. Canman, C. E. *et al.* Induction of resistance to fluorodeoxyuridine cytotoxicity and DNA damage in human tumor cells by expression of Escherichia coli deoxyuridinetriphosphatase. *Cancer Res.* **54,** 2296–8 (1994).
18. Canman, C. E., Lawrence, T. S., Shewach, D. S., Tang, H. Y. & Maybaum, J. Resistance to fluorodeoxyuridine-induced DNA damage and cytotoxicity correlates with an elevation of deoxyuridine triphosphatase activity and failure to accumulate deoxyuridine triphosphate. *Cancer Res.* **53,** 5219–24 (1993).
19. Koehler, S. E. & Ladner, R. D. Small interfering RNA-mediated suppression of dUTPase sensitizes cancer cell lines to thymidylate synthase inhibition. *Mol Pharmacol* **66,** 620–626 (2004).

20. Wilson, P. M., LaBonte, M. J., Lenz, H.-J., Mack, P. C. & Ladner, R. D. Inhibition of dUTPase induces synthetic lethality with thymidylate synthase-targeted therapies in non-small cell lung cancer. *Mol. Cancer Ther.* **11,** 616–28 (2012).
21. Merényi, G. *et al.* Cellular response to efficient dUTPase RNAi silencing in stable HeLa cell lines perturbs expression levels of genes involved in thymidylate metabolism. *Nucleosides. Nucleotides Nucleic Acids* **30,** 369–90 (2011).
22. Hagenkort, A. *et al.* dUTPase inhibition augments replication defects of 5-Fluorouracil. *Oncotarget* **8,** 23713–23726 (2017).
23. Dos Santos, R. S. *et al.* dUTPase (DUT) Is Mutated in a Novel Monogenic Syndrome With Diabetes and Bone Marrow Failure. *Diabetes* **66,** 1086–1096 (2017).
24. Kim, Y.-J. & Wilson, D. M. Overview of base excision repair biochemistry. *Curr. Mol. Pharmacol.* **5,** 3–13 (2012).
25. Krokan, H. E., Drabløs, F. & Slupphaug, G. Uracil in DNA--occurrence, consequences and repair. *Oncogene* **21,** 8935–48 (2002).
26. Pettersen, H. S. *et al.* UNG-initiated base excision repair is the major repair route for 5-fluorouracil in DNA, but 5-fluorouracil cytotoxicity depends mainly on RNA incorporation. *Nucleic Acids Res.* **39,** 8430–44 (2011).
27. Slupphaug, G. *et al.* Properties of a recombinant human uracil-DNA glycosylase from the UNG gene and evidence that UNG encodes the major uracil-DNA glycosylase. *Biochemistry* **34,** 128–38 (1995).
28. Khodursky, A., Guzmán, E. C. & Hanawalt, P. C. Thymineless Death Lives On: New Insights into a Classic Phenomenon. *Annu. Rev. Microbiol.* **69,** 247–63 (2015).
29. Kunz, C. *et al.* Base excision by thymine DNA glycosylase mediates DNA-directed cytotoxicity of 5-fluorouracil. *PLoS Biol.* **7,** e91 (2009).
30. Wilson, P. M., Danenberg, P. V, Johnston, P. G., Lenz, H.-J. & Ladner, R. D. Standing the test of time: targeting thymidylate biosynthesis in cancer therapy. *Nat. Rev. Clin. Oncol.* **11,** 282–98 (2014).
31. Kerepesi, C. *et al.* Life without dUTPase. *Front. Microbiol.* **7,** 1768 (2016).
32. Wang, Z. & Mosbaugh, D. W. Uracil-DNA glycosylase inhibitor gene of bacteriophage PBS2 encodes a binding protein specific for uracil-DNA glycosylase. *J. Biol. Chem.* **264,** 1163–71 (1989).
33. Serrano-Heras, G., Bravo, A. & Salas, M. Phage phi29 protein p56 prevents viral DNA replication impairment caused by uracil excision activity of uracil-DNA glycosylase. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 19044–9 (2008).
34. Kiljunen, S. *et al.* Yersiniophage phiR1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine. *Microbiology* **151,** 4093–102 (2005).
35. Uchiyama, J. *et al.* Intragenus generalized transduction in Staphylococcus spp. by a novel giant phage. *ISME J.* **8,** 1949–52 (2014).
36. Wang, H.-C. *et al.* Staphylococcus aureus protein SAUGI acts as a uracil-DNA glycosylase inhibitor. *Nucleic Acids Res.* **42,** 1354–64 (2014).
37. Yan, N., O'Day, E., Wheeler, L. A., Engelman, A. & Lieberman, J. HIV DNA is heavily uracilated, which protects it from autointegration. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 9244–9 (2011).
38. Kavli, B., Otterlei, M., Slupphaug, G. & Krokan, H. E. Uracil in DNA--general mutagen, but normal intermediate in acquired immunity. *DNA Repair (Amst).* **6,** 505–16 (2007).
39. Róna, G. *et al.* Detection of uracil within DNA using a sensitive labeling method for in vitro and cellular applications. *Nucleic Acids Res.* **44,** e28 (2016).
40. Pálinkás, H. L. *et al.* CRISPR/Cas9-Mediated Knock-Out of dUTPase in Mice Leads to Early Embryonic Lethality. *Biomolecules* **9,** (2019).