DNA Repair 25 (2015) 60–71

Contents lists available at ScienceDirect

DNA Repair

journal homepage: www.elsevier.com/locate/dnarepair

AID expression in B-cell lymphomas causes accumulation of genomic uracil and a distinct AID mutational signature

Henrik Sahlin Pettersen[a,c,d], Anastasia Galashevskaya[a], Berit Doseth[a], Mirta M.L. Sousa[a,e], Antonio Sarno[a], Torkild Visnes[a,1], Per Arne Aas[a], Nina-Beate Liabakk[a], Geir Slupphaug[a,e], Pål Sætrom[a,b], Bodil Kavli[a,*], Hans E. Krokan[a,*]

[a] Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway
[b] Department of Computer and Information Science, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway
[c] Liaison Committee between the Central Norway Regional Health Authority (RHA) and the Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[d] St. Olav's Hospital, Trondheim University Hospital, NO-7006 Trondheim, Norway
[e] The Proteomics and Metabolomics Core Facility (PROMEC) at NTNU, NO-7491 Trondheim, Norway

ARTICLE INFO

Article history:
Received 10 September 2014
Received in revised form 6 November 2014
Accepted 17 November 2014
Available online 24 November 2014

Keywords:
Activation-induced cytidine deaminase (AID)
Genomic uracil
B-cell lymphoma
Base excision repair
Kataegis
Mutational signature

ABSTRACT

The most common mutations in cancer are C to T transitions, but their origin has remained elusive. Recently, mutational signatures of APOBEC-family cytosine deaminases were identified in many common cancers, suggesting off-target deamination of cytosine to uracil as a common mutagenic mechanism. Here we present evidence from mass spectrometric quantitation of deoxyuridine in DNA that shows significantly higher genomic uracil content in B-cell lymphoma cell lines compared to non-lymphoma cancer cell lines and normal circulating lymphocytes. The genomic uracil levels were highly correlated with AID mRNA and protein expression, but not with expression of other APOBECs. Accordingly, AID knockdown significantly reduced genomic uracil content. B-cells stimulated to express endogenous AID and undergo class switch recombination displayed a several-fold increase in total genomic uracil, indicating that B cells may undergo widespread cytosine deamination after stimulation. In line with this, we found that clustered mutations (*kataegis*) in lymphoma and chronic lymphocytic leukemia predominantly carry AID-hotspot mutational signatures. Moreover, we observed an inverse correlation of genomic uracil with uracil excision activity and expression of the uracil-DNA glycosylases UNG and SMUG1. In conclusion, AID-induced mutagenic U:G mismatches in DNA may be a fundamental and common cause of mutations in B-cell malignancies.

1. Introduction

The only sources of uracil in DNA were previously thought to be misincorporation of dUMP during DNA replication and spontaneous deamination of DNA cytosine. The discovery of activation-induced cytidine deaminase (AID, also called AICDA) and several other APOBEC-family enzymes as probable DNA-cytosine deaminases introduced a third possible source (reviewed in [1]).

AID was first identified following induction of class switch recombination (CSR) in the CH12 mouse B-cell lymphoma cell line and initially thought to be an RNA-editing enzyme [2]. However, evidence that AID was a DNA mutator in *Escherichia coli* [3] and its functional interaction with uracil-DNA glycosylase UNG in adaptive immunity [4–6], indicated that AID is a DNA-cytosine deaminase. Later several of the other known APOBEC-family enzymes were also found to be DNA-cytosine deaminases *in vitro* [7,8]. DNA cytosine deamination by APOBEC-family enzymes is a natural event in both the adaptive and innate immune systems, through targeted deamination of immunoglobulin (Ig) genes by AID and deamination of viral DNA by APOBEC enzymes, respectively [7]. Despite their important physiological functions, these host defense mechanisms entail a high risk of potentially carcinogenic off-target genomic mutagenesis. Recent high-throughput sequencing of large numbers of human cancer genomes showed that mutations at cytosine residues, particularly C to T transitions, are the most prevalent

* Corresponding authors. Tel.: +47 72 57 30 74/+47 72 573221; fax: +47 72 57 64 00.
E-mail addresses: bodil.kavli@ntnu.no (B. Kavli), hans.krokan@ntnu.no (H.E. Krokan).
[1] Present address: Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-17121 Stockholm, Sweden.

http://dx.doi.org/10.1016/j.dnarep.2014.11.006

mutations in human cancer, highlighting enzymatic deamination of cytosine to uracil as a potential source of mutagenesis [9–11]. However, the actual uracil level in normal and various cancer genomes has remained elusive.

Here, a sensitive LC/MS/MS-based method for quantification of genomic 2′-deoxyuridine (dUrd) was applied to demonstrate that B-cell lymphoma cell lines contain several-fold increased levels of genomic uracil compared to normal human lymphocytes and non-lymphoma cell lines. Genomic uracil content correlated with AID protein expression but not with other APOBEC enzymes. In accordance with AID-generated uracil, we found that regions of clustered mutations (*kataegis*) in lymphoma and chronic lymphocytic leukemia (CLL) have a distinct AID-hotspot mutational signature. Importantly, we also show that uracil excision capacity and expression of the uracil-DNA glycosylases UNG and SMUG1 correlated negatively with genomic uracil levels and to some extent diminished the effect of AID. This study provides direct mechanistic evidence for genomic uracil accumulation due to enzymatic DNA cytosine deamination in human cancers.

2. Materials and methods

2.1. Primary cells, cell lines, cultivation, and reagents

Human cell lines HeLaS3 (ATCC CCL-2.2™), HEK293T (ATCC CRL-11268™), and U2OS (ATCC HTB-96™) were from ATCC. L428 (DSMZ ACC 197), DU145 (DSMZ ACC 261), KARPAS422 (DSMZ ACC 32), T24 (DSMZ ACC 376), DOHH2 (DSMZ ACC 47), SUDHL4 (DSMZ ACC 4956), JJN3 (DSMZ ACC 541), SUDHL5 (DSMZ ACC 571), SUDHL6 (DSMZ ACC 572 6), RAMOS (DSMZ ACC 603), RL (DSMZ ACC 613), DAUDI (DSMZ ACC 78 5), A431 (DSMZ ACC 91) were from DSMZ. OCILY3 was a gift from Dr. L.M. Staudt, Metabolism Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. Peripheral blood mononuclear cells (PMBCs) were purified from buffy coats from three healthy blood donors using the Lymphoprep™ (Progen) kit according to the manufacturer's protocol. Human B-lymphocytes were purified from buffy coats from three healthy blood donors using a negative selection kit from StemCell Technologies according to the manufacturer's protocol. HeLaS3, HEK293T, T24, A431, DU145, and U2OS cells were cultured in DMEM (4500 mg/l glucose) with 10% FCS, 0.03% L-glutamine, 0.1 mg/ml gentamicin and 2.3 μg/ml fungizone at 37 °C and 5% CO_2. DAUDI, DOHH-2, KARPAS, RAMOS, SU-DHL-4, SU-DHL-6, OCILY-3, L-428, RL, SU-DHL-5, and JJN3 cells were cultured in RPMI-1640 with 4500 mg/l glucose, 0.03% L-glutamine, Pen-Strep (1× final), 0.1 mg/ml gentamicin, and 2.3 μg/ml fungizone, and 20% heat inactivated (56 °C, 20 min) FCS at 37 °C and 5% CO_2. For quantitative rtPCR and uracil measurements cells were harvested at densities between 750 000 and 2 million cells/ml.

Cell doubling times for suspension cells were measured using a *Countess*® cell counter (Invitrogen) by two parallel daily measurements for three to five day periods from cell densities of 50 000–200 000 cells/ml to one to three million cells/ml. For adherent cells, doubling time was measured in 96 well plates (3–6 parallel wells; starting density 50 000 cells/ml) for a three day period by daily fluorescent measurement of resazurin (Sigma) metabolism according to the manufacturer's protocol. Doubling times were calculated by exponential regression.

SUDHL5 AID knockdown and control cells were made using Open Biosystem TransLenti Viral Packaging Mix, pTRIPZ AICDA shRNA (RHS4741-EG57379; vectors V2THS_58282, 58283, and 58319) or pTRIPZ non-silencing control vector according to the manufacturer's protocol. Briefly, lentiviruses were produced in HEK293T cells, and then supernatant from three consecutive days 48 h after HEK293T transfection were used to infect SUDHL5 cells. Infected SUDHL5 cells were amplified for another 48 h and then selected with 2 μg/ml puromycin for 30 days. Expression was induced with 1 μg/ml doxycycline.

CH12F3 AID-EYFP and EYFP stable transfectants, confocal microscopy, and stimulation experiments were described previously [12]. CH12F3 cells (2×10^6 cells/ml) were cultured in RPMI medium, with 10% heat-inactivated fetal calf serum, 0.03% L-glutamine, 50 μM β-mercaptoethanol, 1 mM Na-pyruvate, 0.1 mg/ml penicillin/streptomycin, 2.3 μg/ml fungizone, and 1.0 mg/ml G418. CH12F3 cells were stimulated to undergo class switch recombination by adding 10 ng/ml mouse recombinant IL-4 (Peprotech), 2 μg/ml anti-mouse CD40 monoclonal antibody (BD Biosciences) and 1 ng/ml human TGF-β1 (Peprotech) and harvested 48 h post stimulation for DNA and protein isolation. Western analysis of AID protein expression was performed using mouse anti-AID monoclonal antibody no. 39-2500, clone ZA001, 500 μg/ml (Invitrogen). Nuclear extracts from synchronized HeLa cells were prepared essentially as described [13,14].

2.2. RNA isolation and quantitative real-time PCR (qRT-PCR)

Total RNA for mRNA analysis was prepared using the mirVana miRNA isolation kit (Ambion) according to the manufacturer's instructions. RNA concentration and quality was measured on a NanoDrop ND-1000 UV–vis spectrophotometer. Total RNA (770 ng) was reverse transcribed for gene expression analysis using TaqMan reverse transcription reagents (Applied Biosystems). The following TaqMan gene expression assays (Applied Biosystems) were used: AID (Hs00757808_m1), UNG (Hs00422172_m1), SMUG1 (Hs04274951_m1), TDG (Hs00702322_s1), MBD4 (Hs00187498_m1), APOBEC1 (Hs00242340_m1), APOBEC2 (Hs00199012_m1), APOBEC3A (Hs00377444_m1), APOBEC3B (Hs00358981_m1), APOBEC3C (Hs00819353_m1), APOBEC3D (Hs00537163_m1), APOBEC3G (Hs00222415_m1), APOBEC3F (Hs01665324_m1), APOBEC3H (Hs00419665_m1), APOBEC4 (Hs00378929_m1), and GAPDH (Hs99999905_m1). Quantitative PCR was carried out on a Chromo4 (BioRad) real-time PCR detection system. Relative expression of mRNA was calculated by the ΔCt method using GAPDH as endogenous control. Regression analyses were done using GraphPad Prism where data were fitted by linear regression (log/linear(X) *vs.* log/linear(Y)) as indicated.

2.3. Quantification of uracil in DNA by LC/MS/MS

Genomic uracil was quantified as previously described [15]. Briefly, DNA was isolated by phenol:chloroform:isoamyl extraction, treated with alkaline phosphatase to remove free deoxyribonucleosides, and then enzymatically hydrolyzed to deoxyribonucleosides. Deoxyuridine (dU) was then separated from deoxycytidine (dC) by HPLC fractionation using a reverse-phase column with embedded weak acidic ion-pairing groups (2.1 mm × 150 mm, 5 μm, Primesep 200, SIELC technologies), using a water/acetonitrile gradient containing 0.1% formic acid. The dU fraction was finally analyzed by ESI-LC/MS/MS using a reverse phase column (2.1 mm × 150 mm, 3.5 μm, Zorbax SB-C18, Agilent Technologies), using a water/methanol gradient containing 0.1% formic acid on an API5000 triple quadrupole mass spectrometer (Applied Biosystems) in positive ionization mode. A small fraction of the hydrolyzed deoxyribonucleosides were quantified by LC/MS/MS in parallel and used to determine the amount of dU per 10^6 deoxyribonucleosides.

2.4. In vitro uracil DNA excision activity and complete BER assays

Standard UDG activity assay was performed as described [16]. Briefly, 20 μl reaction mixtures containing (final) 1.8 μM nick translated [^{3}H]-dUMP-labeled calf thymus DNA (U:A substrate), 1× UDG buffer (20 mM Tris–HCl, pH 7.5, 60 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.5 mg/ml BSA) and 1 μg whole cell extract were incubated at 30 °C for 10 min. Acid-soluble [^{3}H] uracil was quantified by scintillation counting. Whole cell extracts was prepared as described [17]. Oligodeoxynucleotide UDG assays were performed as described [16]. Briefly, double-stranded DNA substrates were generated by annealing 6-FAM-labeled oligonucleotides containing a centrally positioned uracil in an AID-hotspot (5′-CATAAAGAGUTAAGCCTGG-3′; Eurogentec) to complementary strands containing G opposite U. Activity was measured in 10 μl assay mixtures containing (final) 20 nM substrate, 1× UDG buffer and 0.4 μg cell extract, and incubated at 37 °C for 10 min. Reactions were stopped and AP-sites were cleaved by addition of 50 μl 10% piperidine followed by incubation at 90 °C for 20 min. Product and substrate were separated on PAGE, scanned on Typhoon Trio imager and quantified using ImageQuant TL software (GE healthcare).

BER assays were carried out essentially as described [14,17]. Briefly, 10 μg nuclear extract was incubated with 250 ng cccDNA (covalently closed circular DNA) substrates in final concentrations of 40 mM HEPES-KOH, 70 mM KCl, 5 mM $MgCl_2$, 0.5 mM DTT, 2 mM ATP, 20 μM dATP, 20 μM dGTP, 20 μM dTTP, 8 μM dCTP, 4.4 mM phosphocreatine, 62.5 ng/μl creatine kinase and 50 nCi/μl [α-^{32}P]dCTP in a final volume of 40 μl. Reactions were incubated for 25 min at 37 °C and stopped by addition of EDTA (18 mM final) and 6 μg RNaseA and incubated at 37 °C for 10 min followed by the addition of SDS (0.5% final) and 12 μg proteinase K. DNA was extracted by phenol/chloroform and precipitated in ammonium acetate/ethanol and digested with XbaI and HincII (New England Biolabs). Following 12% PAGE, bands were visualized and quantified using ImageQuant software (Fujifilm). We investigated relative contribution of SMUG1, TDG and UNG2 to the initiation of uracil repair by pre-incubating extracts with neutralizing antibodies to SMUG1 (0.11 μg/μl final concentration), UNG (0.3 μg/μl final concentration), and/or neutralizing anti-serum towards TDG (1:50 dilution) on ice for 30 min prior to the reaction.

2.5. Flow cytometric analysis of cell cycle

Cells were fixed in 70% methanol, washed twice with PBS, and then treated with 50 μl RNaseA (100 μg/ml in PBS) at 37 °C for 30 min prior to DNA staining with 200 μl propidium iodide (50 μg/ml in PBS) at 37 °C for 30 min. Cell cycle analyses were performed using a FACS Canto flow cytometer (BD-Life Science).

2.6. Sample preparation and targeted mass spectrometry

Cell pellets were resuspended in 1× packed cell volume in buffer I: 10 mM Tris–HCl pH 8.0, 200 mM KCl, 1× complete protease inhibitor, and 5× phosphatase-inhibitor cocktails I and II (Sigma-Aldrich), 10 μM suberoylanilide hydroxamic acid (SAHA) (Cayman Chemicals) and 0.05 μM, Ubiquitin Aldehyde (Biomol International LP) followed by addition of an equal final volume of buffer II: 10 mM Tris–HCl pH 8.0, 200 mM KCl, 10 mM EGTA, 10 mM $MgCl_2$, 40% glycerol, 0.5% NP40, 1 mM DTT, 1× complete protease inhibitor, and 5× phosphatase-inhibitor cocktails I and II (Sigma-Aldrich), 10 μM suberoylanilide hydroxamic acid (SAHA) (Cayman chemicals) and 0.05 μM, Ubiquitin Aldehyde (Biomol International LP) containing an endonuclease cocktail of 200 U Omnicleave (Epicenter Technologies), 2 U DNase I (Roche Inc.), 250 U Benzonase (EMD), 100–300 U micrococcal nuclease (Sigma-Aldrich), and 10 μg RNase A (Sigma-Aldrich) per 1 ml of buffer II. After resuspension, the lysates were incubated for 1.5 h at 4 °C in a roller. 50 μg protein of cell lysate pools consisting of 2–4 biological replicates from each cell line were incubated with 5 mM tris (2-carboxyethyl) phosphine (TCEP) for 30 min followed by alkylation with 1 μmol/mg protein of iodoacetamide for 45 min in the dark. Proteins were precipitated using a methanol–chloroform method as described [18], including another round of reduction and alkylation prior to overnight digestion with Trypsin (Promega) at 1:40 ratio (w/w, enzyme:protein) at 37 °C. Tryptic digests were dried out, resuspended in 0.1% formic acid and analyzed on a Thermo Scientific QExactive mass spectrometer operating in Targeted-MS2 mode coupled to an Easy-nLC 1000 UHPLC system (Thermo Scientific/Proxeon). Peptides (2 μg) were injected onto a Acclaim PepMap100 C-18 column (75 μm i.d. × 2 cm, C18, 5 μm, 100 Å) (Thermo Scientific) and further separated on a Acclaim PepMap100 C-18 analytical column (75 μm i.d. × 50 cm, C18, 3 μm, 100 Å) (Thermo Scientific). A 120 min method was used and consisted of a 300 nl/min flow rate, starting with 100% buffer A (0.1% Formic acid) with an increase to 5% buffer B (100% Acetonitrile, 0.1% Formic acid) in 2 min, followed by an increase to 35% Buffer B over 98 min and a rapid increase to 100% buffer B in 6 min, where it was held for 5.5 min. The solvent composition was quickly ramped to 0% buffer B, where it was subsequently held for 8 min to allow the column to equilibrate for the next run. The peptides eluting from the column were ionized by using a nanospray ESI ion source (Proxeon, Odense) and analyzed on the QExactive operating in positive-ion mode using electrospray voltage 1.9 kV and HCD fragmentation. Each MS/MS scan was acquired at a resolution of 35 000 FWHM, normalized collision energy (NCE) 28, automatic gain control (AGC) target value of 2×10^5, maximum injection time of 120 ms and isolation window 2 *m/z*.

All parallel reaction monitoring (PRM)-based targeted mass spectrometry methods were designed, analyzed, and processed using Skyline software version 2.5 [19]. *In silico* selection of proteotypic peptides was performed *via* Skyline using the *Homo sapiens* reference proteome available at www.uniprot.org to exclude non-unique peptides. Frequently modified peptides, such as those containing methionine, and peptides containing continuous sequences of R and K (*e.g.*, KR, RK, KK or RR) were avoided. However, when the inclusion of non-ideal peptides was necessary both unmodified and M-oxidized peptides as well as peptides containing a missed cleavage site were analyzed. Synthetic purified peptides (JPT Peptide Technologies) and tryptic digests from recombinant proteins were analyzed in a QExactive mass spectrometer. Information on retention time and fragmentation pattern of the top 2–6 ionizing tryptic peptides (2+ or 3+ charge states) for each protein were used to build a scheduled method with a retention time window of 5 min. The method was then used for peptide quantification in the cell lysate pools. A minimum of 2 peptides per protein was used for quantitative analysis except for APOBEC3F in which only one of the unique peptides tested was detectable in the samples. The sum of the integrated peak areas of the 3–5 most intense fragments was used for peptide quantification. Peptide areas for multiple peptides of the same protein were summed to assign relative abundance to that protein. The error bars represent the standard deviation of 3 technical replicates.

2.7. Bioinformatics analysis of DNA exome sequencing data

Kataegis regions and somatic mutations for CLL, B-Cell lymphoma, ALL, lung adenocarcinoma, and breast, liver, and pancreatic cancer were downloaded from the supplementary material of a published study [11]. The *kataegis* regions within specific cancer samples were provided as genomic coordinates into the human reference genome version 19 (hg19); the somatic mutations were provided as genomic coordinates in hg19 and nucleotide alterations. We used the following procedure to create mutational

signatures for the *kataegis* regions for each cancer type. First, for each *kataegis* region, its sample ID and genomic coordinates were used to identify the corresponding somatic mutations. Second, for each somatic mutation, the five nucleotides centered on the mutated nucleotide were retrieved from the genome sequence. Third, if the middle nucleotide within the retrieved sequence was a purine, the sequence was reverse-complemented such that all the mutations were represented by pyrimidines. Fourth, for each of the six possible single nucleotide mutations, the relative occurrence of each nucleotide at each position within the retrieved sequences was computed. These position-specific relative occurrences were the mutational signatures.

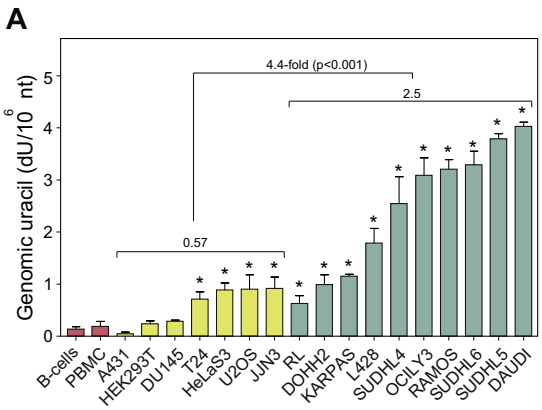
3. Results

3.1. High genomic uracil levels in B-cell lymphoma cells

To investigate whether uracil in the genome may be an important factor in lymphomagenesis, we measured genomic uracil in ten B-cell lymphoma cell lines, seven other human transformed cell lines and in lymphocytes from three healthy human blood donors (Fig. 1A). The origin and major characteristics of cell lines is displayed in Fig. 1B. We found as much as 72-fold variation in genomic uracil levels between the cell line with the highest uracil content (DAUDI, 4.03 deoxyuridines (dU) per 10^6 deoxyribonucleosides (nt)) and the cells with the lowest level of genomic uracil (A431, 0.056 dU/10^6 nt). Strikingly, all ten lymphoma cell lines and four of the other transformed cell lines had significantly ($P < 0.05$) elevated genomic uracil levels compared to genomic uracil in peripheral blood mononuclear cells (PBMC) from the mean value for three blood donors (0.19 dU/10^6 nt). We also measured genomic uracil in B-lymphocytes isolated from buffy coats from three healthy donors, using a kit for negative selection. The genomic mean uracil level in these was 0.14 dU/10^6 nt, with individual values of 0.07, 0.17 and 0.19 dU/10^6 nt, respectively. The mean value for the genomic uracil level in B-cell lymphoma cell lines (2.5 dU/10^6 nt) was 13-fold and 18-fold higher than in PBMC and primary B-cells, respectively. In addition it was significantly higher (4.4-fold, $P < 0.001$) than the mean for non-lymphoma cell lines (0.57 dU/10^6 nt). The B-cell lymphoma cell lines are likely to be exposed to enzymatic untargeted cytosine deamination by AID throughout the genome, since the total number of genomic uracils is in the range 3000–15 000 per haploid genome (this paper) and the density of genomic uracil in S_μ region of stimulated B-cells is only ~0.8 per kb [20]. Some of the non-lymphoma cancer cell lines had intermediate genomic uracil levels, clearly higher than normal peripheral blood lymphocytes, but lower than most of the B-cell lymphoma cell lines.

3.2. AID expression correlates with genomic uracil accumulation

AID has previously been shown to be expressed in several lymphoma subtypes [21–24] and AID/APOBEC family enzymes were suggested to contribute to mutational signatures in a number of cancers by deaminating cytosine to uracil in DNA [11]. We therefore investigated whether expression of AID and/or other APOBECs could explain the observed variation in genomic uracil levels in the cell line panel. We first measured mRNA expression of AID and all other APOBEC-family genes by quantitative rtPCR using GAPDH as reference gene (Fig. 2A). AID mRNA was detected in all 17 cell lines, although at highly variable levels, but not in the normal lymphocytes from blood donors. Furthermore, AID mRNA was substantially increased in lymphomas with high genomic uracil such that AID mRNA had a high positive correlation with genomic uracil ($R^2 = 0.70$, $P < 0.0001$). By contrast, APOBEC3B, -3D, -3F, and -3G mRNA content did not correlate with genomic uracil level although



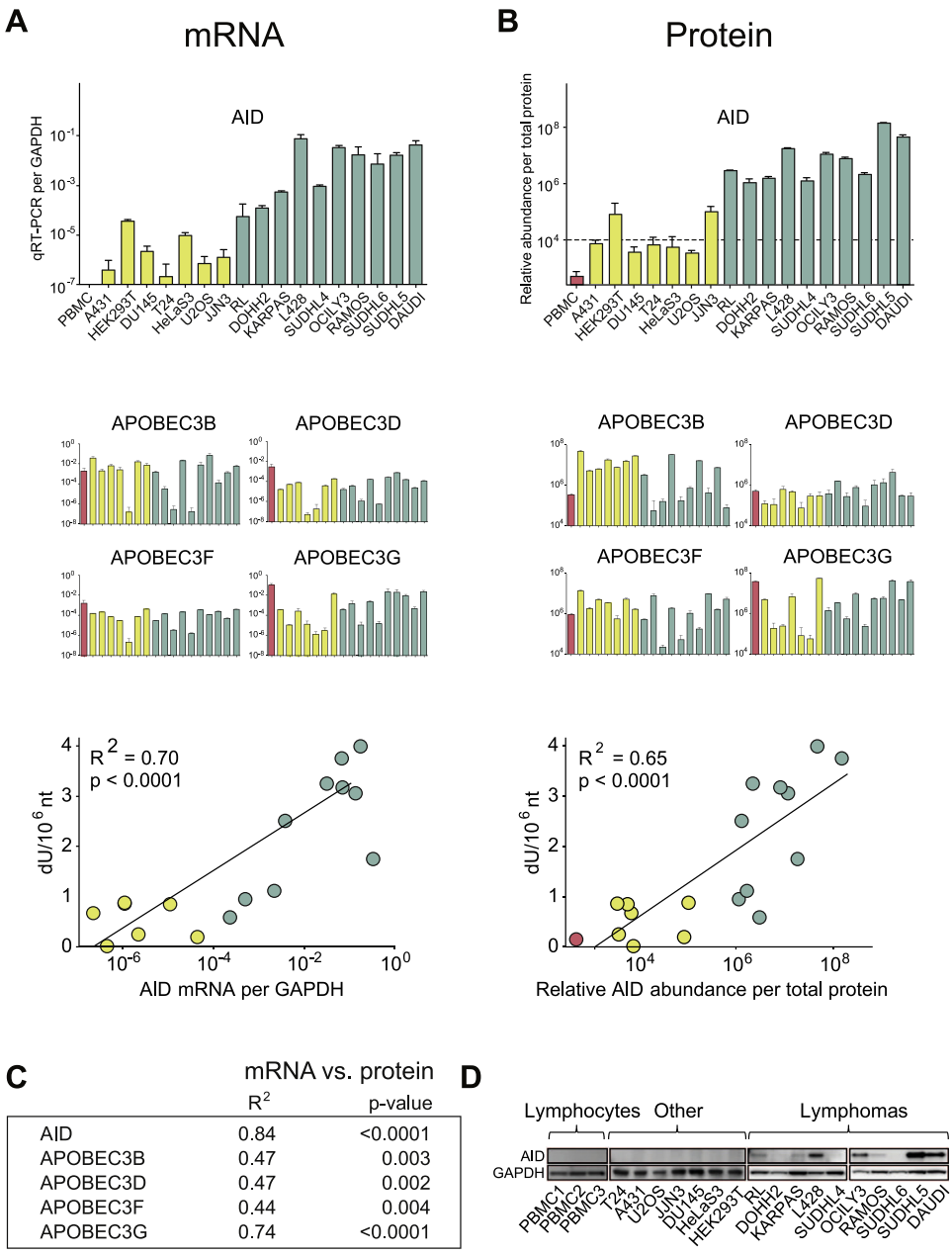
B

Name	Cell/tissue type	Disease
B-cells	B-lymphocyte/blood	Healthy blood donors
PBMC	White blood cells/blood	Healthy blood donors
A431	Epithelial/skin	Epidermoid carcinoma
HEK293T	Embryonic kidney	NA
DU145	Epithelial/prostate	Prostate cancer
T24	Epithelial/bladder	Bladder carcinoma
HeLaS3	Epithelial/cervix	Adenocarcinoma
U2OS	Osteocyte/bone	Osteosarcoma
JJN3	Plasma cell/bone marrow	Plasma cell leukemia
RL	B-cell/lymphatic	B-NHL
DOHH2	B-cell/lymphatic	B-NHL
KARPAS	B-cell/lymphatic	B-NHL
L428	B-cell/lymphatic	Hodgkin lymphoma
SUDHL4	B-cell/lymphatic	B-NHL
OCILY3	B-cell/lymphatic	B-NHL
RAMOS	B-cell/lymphatic	Burkitt lymphoma
SUDHL6	B-cell/lymphatic	B-NHL
SUDHL5	B-cell/lymphatic	B-NHL
DAUDI	B-cell/lymphatic	Burkitt lymphoma

Fig. 1. Genomic uracil levels in B-cell lymphoma-/non-lymphoma cell lines and on white blood cells from peripheral blood l. (A) Quantification of genomic uracil levels (dU/10^6 nt) by LC–MS/MS in lymphoma cell lines (green), non-lymphoma cell lines (yellow) and PBMCs or B-lymphocytes isolated from buffy coats from blood donors (red). Asterisk (*) signifies measurements significantly ($P < 0.05$) different from average genomic uracil levels in PBMC from three healthy blood donors (Student's *T*-test). Error bars represent mean and SD of at least two biological replicates. Cell lines within each group are ordered along the x-axis according to increasing genomic uracil levels. (B) Overview of cell lines, PBMCs and B-lymphocytes used in the study and their origin. B-NHL: B-cell non-Hodgkin lymphoma. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

they were expressed in all cell lines as well as in the normal lymphocytes (Fig. 2A). mRNA of the other APOBECs (APOBEC1, APOBEC2, APOBEC3A, and APOBEC) were detected only in some of the cell lines and mostly at very low levels (data not shown).

Although mRNA expression data is useful as a predictor of protein expression, it does not always correlate with the actual protein levels in the cells. Thus, we quantified AID and the APOBEC proteins by parallel reaction monitoring using a quadrupole-Orbitrap mass spectrometer (Fig. 2B). This is a highly selective method allowing quantification of many protein targets in a single sample [25,26]. In agreement with mRNA data, MS quantification revealed higher amounts of AID protein in lymphoma cells with increased genomic uracil (Fig. 2B, upper panel). Furthermore, similar to mRNA data (Fig. 2A, middle panel), APOBEC3B, -3D, -3F and-3G proteins were expressed in all cell lines (Fig. 2B, middle panel), while APOBEC1, APOBEC2, APOBEC3A, and APOBEC4 were not detectable or detected at very low levels (data not shown). In general, protein levels for AID and the APOBEC proteins (normalized to GADPH protein) correlated well with mRNA levels (Fig. 2C). As an additional control, we also quantified AID by western analysis, which yielded results similar to the MS analysis (Fig. 2D). Linear regression



	R^2	p-value
AID	0.84	<0.0001
APOBEC3B	0.47	0.003
APOBEC3D	0.47	0.002
APOBEC3F	0.44	0.004
APOBEC3G	0.74	<0.0001

Fig. 2. Expression of AID and APOBECs, and correlation with genomic uracil. Expression of AID and APOBEC3B, 3D, 3F, and 3G mRNAs measured by qRT-PCR (A) or protein by mass spectrometric quantification. (B) Lymphoma cell lines are shown in green, non-lymphoma cell lines in yellow, and PBMC in red. Cell lines within each group are ordered along the *x*-axis according to increasing genomic uracil levels, as in Fig. 1. mRNA levels have been normalized to GAPDH mRNA, and protein levels to MS signal counts per total injected protein. Note that mRNA and protein expression data are in log-scale. Regression plots of genomic uracil (dU/10^6 nt) *vs.* AID mRNA and protein levels are presented in the lower panels in Fig. 2A and B, respectively. (C) Table of correlation coefficients between mRNA and protein expression for AID and other APOBECs. (D) Western analysis of AID protein expression with GAPDH shown as a loading control. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Regression analysis of genomic uracil levels (linear) *vs.* AID and APOBEC protein expression (log) normalized to total protein. Bold green indicates significant positive correlation (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

	All cell lines including PBMC		B-cell lymphoma cell lines		Non-lymphoma cell lines	
	R^2	*P*-value	R^2	*P*-value	R^2	*P*-value
AID	**0.65**	**<0.0001**	**0.42**	**0.04**	0.00	0.97
APOBEC3B	0.10	0.2089	0.00	0.84	0.00	0.98
APOBEC3D	0.12	0.17	0.00	0.88	0.02	0.79
APOBEC3F	0.01	0.67	0.08	0.44	0.32	0.18
APOBEC3G	0.12	0.14	0.30	0.09	0.00	0.98

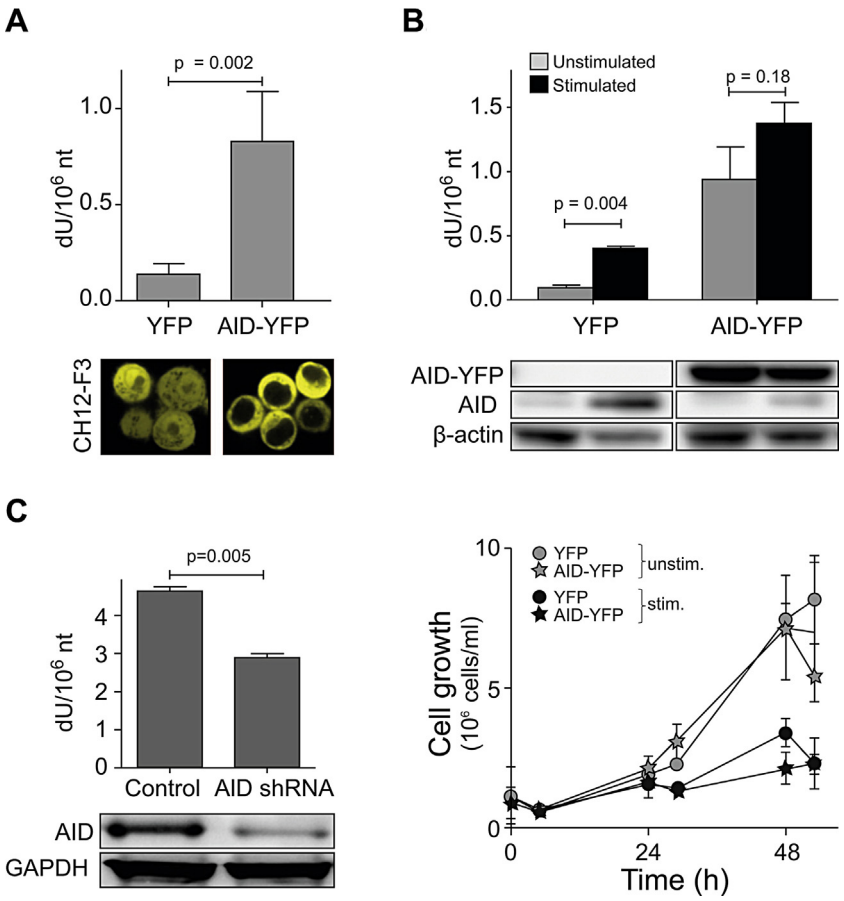


Fig. 3. Genomic uracil levels after stimulation of endogenous AID expression, AID-YFP overexpression, and AID knockdown. (A) Genomic uracil levels in DNA isolated from mouse lymphoma cells (CH12F3) stably transfected with AID-YFP or YFP, and confocal microscopy showing subcellular distribution of AID-YFP fusion protein or YFP. (B) Genomic uracil levels and cell growth of CH12F3 YFP cells and CH12F3 AID-YFP cells prior to stimulation and 48 h after being stimulated to undergo class switch recombination using mouse recombinant IL-4, CD40 monoclonal antibody and hTGF-β (upper panel) and western blots from one representative experiment showing AID protein expression levels and β-actin as loading control (middle panel). The lower panel shows cell growth of stimulated and unstimulated cells. Graphs represent mean and SD calculated from at least two biological replicates. *P*-values were calculated by a two-tailed Student's *T*-test. (C) Genomic uracil levels in SUDHL5 lymphoma cells stably transfected with AID-shRNA and control. Western blots shows AID protein expression levels with GAPDH as a loading control.

analysis of AID western signals against MS quantitation of AID protein revealed almost perfect correlation ($R^2 = 0.95$). Importantly, AID expression significantly correlated with genomic uracil also at the protein level ($R^2 = 0.65$, $P < 0.0001$) (Fig. 2B, and Table 1), and thereby seemed to account for a large part of the variation in genomic uracil between the cell lines. The correlation was still valid when including only the B-cell lymphoma cell lines in the regression analysis (Table 1). No significant correlations were observed between the other APOBEC proteins and genomic uracil (Table 1). Thus, AID was the only APOBEC-family member that correlated with genomic uracil in the human cancer cell lines examined here.

3.3. *AID expression causes several-fold increases in genomic uracil*

To investigate whether AID expression significantly increases the overall level of genomic uracil in an otherwise isogenic background, we used stable transfectants of the mouse B-cell lymphoma cell line CH12F3 expressing AID-YFP fusion protein, or YFP as control [27]. AID is mostly localized in the cytoplasm (Fig. 3A), but is actively imported into the nucleus where it may access the genome [12]. We found that the cells expressing AID-YFP displayed an almost six-fold higher level of genomic uracil compared to the YFP control (Fig. 3A). When appropriately stimulated, CH12F3 cells increase endogenous AID expression and have capacity to undergo CSR. Thus next, we investigated whether stimulation of these cell lines also increased the level of genomic uracil. A clear induction of AID and a four-fold increase in genomic uracil were observed in stimulated CH12F3-YFP cells already after 48 h (Fig. 3B, upper panel). An increase in genomic uracil was observed in the stimulated AID-YFP expressing cells as well, although this was not significant, probably due to the high constitutive expression AID-YFP. Importantly, the increase in genomic uracil observed after stimulation could not be ascribed to increased replicative misincorporation of dUMP due to higher proliferation rate because stimulated CH12F3 cells actually have reduced proliferation (Fig. 3B, lower panel). Finally, we examined the effect of knocking down AID using a lentiviral AID shRNA expressing vector. For this experiment, we used the human B-cell lymphoma cell line SUDHL5, which exhibited high constitutive AID expression (Fig. 2B and D). We found that a 60% knockdown of AID reduced genomic uracil level by 38% ($P = 0.005$; Fig. 3C). Taken together these results strongly support the view that enzymatic cytosine deamination is the major source of genomic uracil in AID-expressing cells.

3.4. *Uracil-DNA repair capacity is inversely correlated with genomic uracil levels*

Genomic uracil is predominantly repaired by base excision repair (BER), which is mainly initiated by the uracil-DNA glycosylase encoded by the *UNG* gene [16]. We have previously shown that UNG deficiency in human and mouse cells results in a several-fold

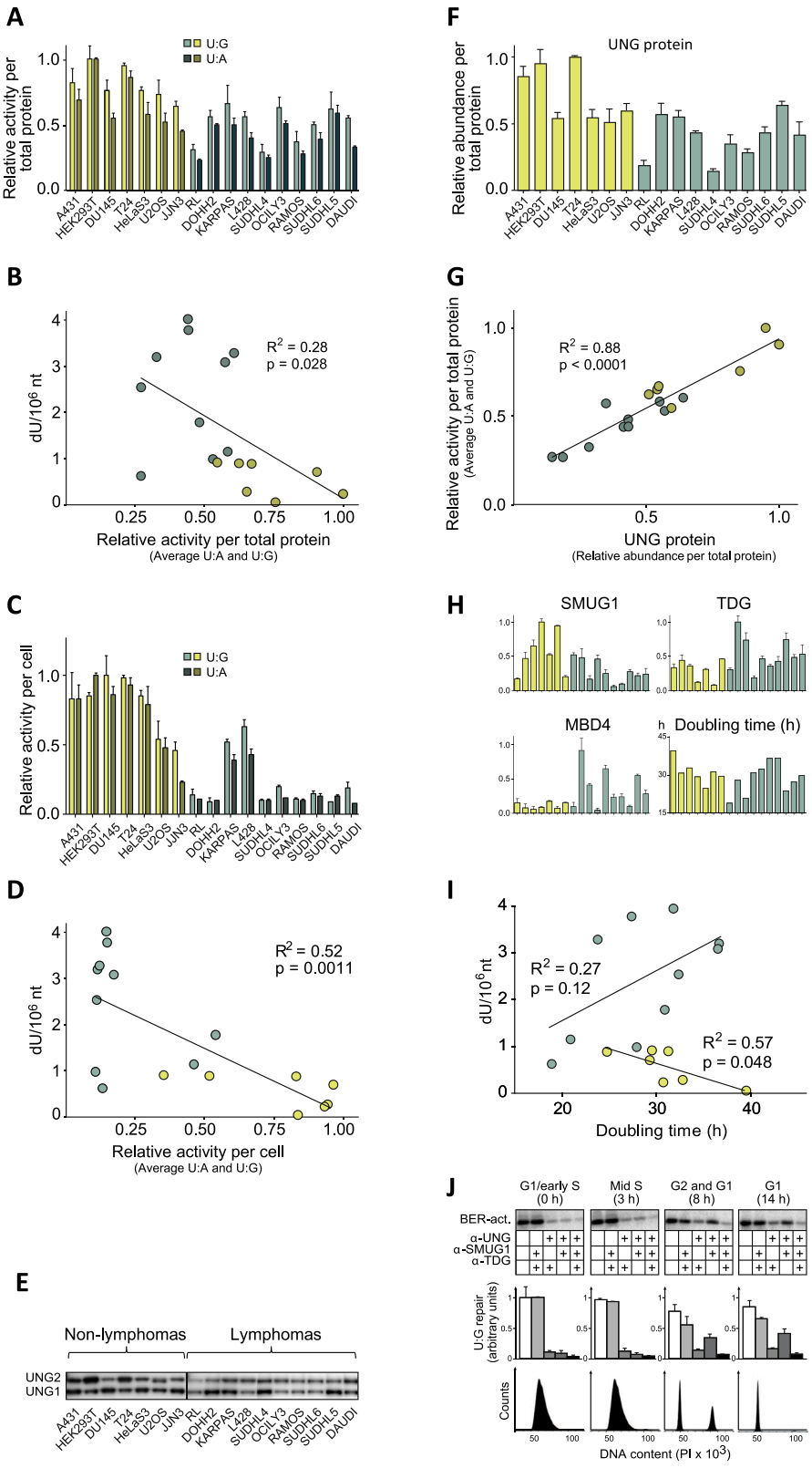


Fig. 4. Uracil excision activity, expression of uracil DNA glycosylases, and correlation with genomic uracil levels. Note that in all bar graphs cell lines are ordered according to increasing genomic uracil levels in lymphoma cell lines (green) and non-lymphoma cell lines (yellow), and *Y*-axes are normalized so that maximum activity or maximum protein abundance equals 1. Bars and error bars represent mean and SD of three biological replicates. (A) Relative uracil excision activity from an AID-hotspot sequence-oligomer containing uracil in U:G context (cleavage assay) and from a nick-translated DNA containing uracil in U:A context (^{3}H-uracil release assay), as indicated by color codes. Activity was normalized to total protein. (B) The corresponding correlation between genomic uracil and activity per total protein. (C) Relative uracil excision activity normalized to activity per cell, and (D) the corresponding correlation with genomic uracil with activity per cell. (E) Western blot of UNG2 and UNG1 in non-lymphoma and lymphoma cell lines. (F) Relative abundance of MS-quantified UNG protein per total protein; (G) Correlation plot of average uracil excision activity *vs.* relative abundance of MS quantified UNG protein. (H) Relative abundance of MS quantified DNA glycosylases SMUG1, TDG and MBD4 and cell doubling times of cell lines; (I) Correlation plot of genomic uracil content *vs.* doubling times of non-lymphoma cell lines and lymphoma cell lines. (J) Contribution of UNG, SMUG1 and TDG through the cell cycle measured by

increase in genomic uracil [15]. The other uracil-DNA glycosylases, *i.e.* SMUG1, TDG, and MBD4, are thought to be quantitatively less important contributors, at least in proliferating cells [16,28,29]. Furthermore, the DNA repair machinery has been shown to protect against AID-induced mutagenesis [30–32]. Therefore, we measured uracil excision activity of cell free extracts prepared from all cell lines against oligodeoxyribonucleotides with uracil in a U:G context. In addition, we measured [^{3}H]-uracil release from calf thymus DNA having uracil in a U:A context. The two different assays gave similar activity profiles (Fig. 4A). Regression analysis of uracil-excision activity (relative to protein content in the cell extracts) against genomic uracil content in the cells demonstrated a negative correlation (Fig. 4B), which is significant ($P < 0.05$), although weak. We also calculated relative uracil excision activity per cell since the glycosylases are predominantly nuclear enzymes and the cells tested vary in size and nucleus-to-cytoplasm ratios (Fig. 4C). Using these activity values, a stronger correlation with genomic uracil level was observed (Fig. 4D).

The *UNG* gene encodes both nuclear UNG2 and mitochondrial UNG1, having identical catalytic domains but specific N-terminal domains. These isoforms are differently regulated from two promoters [33,34]. Since activity assays measure total activity, we analyzed the isoforms by western blots. Nuclear UNG2, which is the isoform relevant for repair of genomic uracil, was expressed in all cell lines and accounted for approximately half of total UNG in most cell lines (Fig. 4E). UNG enzymes are the most active of the glycosylases, at least *in vitro*. However, each glycosylase with its specific or complementary role may exert a significant impact on the total level of genomic uracil *in vivo*. We therefore quantified all the uracil-DNA glycosylases at protein level by MS. The relative abundance of quantified UNG protein (UNG1 and UNG2) (Fig. 4F) correlated strongly with total uracil excision activity (Fig. 4G), in accordance with its presumed major role in uracil repair. Similar to the uracil excision activity, UNG protein per cell also correlated inversely with genomic uracil level when all cell lines were included in the regression analysis (Table 2). Furthermore, quantified SMUG1 protein (Fig. 4H) correlated negatively with genomic uracil, although more weakly. Surprisingly, however, SMUG1 was the only glycosylase that correlated with genomic uracil when only the B-cell lymphoma group was analyzed (Table 2). In addition, the AID/SMUG1 protein ratio displayed significantly higher correlation with genomic uracil in the B-cell lymphoma group ($R^2 = 0.65$) compared to AID alone ($R^2 = 0.42$). No significant correlations were found for TDG or MBD4 proteins and genomic uracil (Fig. 4H) when analyzed separately (Table 2) or in combination with AID or other glycosylases.

3.5. Cell doubling time, genomic uracil content and repair capacity in cell cycle phases

In cells that do not express AID, one would predict that genomic uracil from misincorporation of dUMP during replication should result in increased genomic uracil in cells with short doubling time, as suggested previously [35]. Indeed, we observed a significant inverse relationship between genomic uracil and doubling time in non-lymphoma cancer cells ($R^2 = 0.57$; $P = 0.048$; Fig. 4I). Furthermore, since AID has been shown to act in the G_1 phase of the cell cycle [36–38], one would expect that the lymphoma cell lines with long doubling times might have higher genomic uracil levels than those with shorter doubling time. However, we did not find a significant positive correlation with doubling time ($R^2 = 0.27$; $P = 0.12$), although the curve was apparently different from that of the non-lymphoma cell lines (Fig. 4I).

As mentioned above, we found an inverse correlation between genomic uracil and both total uracil excision capacity, and with SMUG1 and UNG protein levels. Nuclear UNG2 expression peaks during G_1/S-phase transition and during S-phase and is expressed at a lower level in late S-phase, G_2 and early G_1 [13,39]. In contrast, TDG is mainly expressed in the G_1 phase of the cell cycle [13,39]. Thus, TDG might have a role in counteracting untargeted generation of U:G mismatches by AID in G_1, although correlation studies did not give indications of this. SMUG1 is not cell cycle regulated [40] and may contribute in all cell cycle phases, but is a rather slow acting enzyme [16]. To explore the relative contribution of the uracil-DNA glycosylases in *in vitro* complete BER of uracil in different parts of the cell cycle, we synchronized HeLa cells by double thymidine block [13], prepared nuclear extracts from the different cell cycle phases (monitored by flow cytometry) and applied an assay for complete BER of U:G mismatches in DNA [14,17,41]. To examine UNG, SMUG1 and TDG separately, we used a combination of neutralizing antibodies against UNG, SMUG1 and TDG. UNG was found to be by far the major contributor to initiate BER in the G1/S transition and in the S phase. Total repair capacity in G2 and G1 was somewhat lower than in the S phase, but UNG remained a major contributor to the initial step in BER-process, although contributing only 1.5–1.7 more than TDG. SMUG1 contributed in all cell cycle phases, but to a minor degree (Fig. 4J). Thus, a role of TDG and SMUG1 in BER of U:G mismatches in the G1 phase, and a smaller role in the S-phase would seem likely from our *in vitro* data. The contribution of the different uracil-DNA glycosylases during the cell cycle is likely to be similar in other human cell lines, including B cell lymphoma cell lines, although this has not been formally demonstrated.

3.6. Lymphomas and CLL carry a distinct AID-hotspot mutational signature in kataegis regions

Large scale genome sequencing of cancers has produced the novel observation that several cancers carry localized hypermutation, named *kataegis*, in small regions that are also associated with genomic rearrangements. The mutational signatures observed in most cancer types with *kataegis* (acute lymphoblastic leukemia (ALL), lung adenocarcinomas, breast, pancreas, and liver cancer) suggest an association with APOBEC3 enzymes, with a strong enrichment of C to T transitions and C to G transversions at TCA/T sequence contexts [11]. As mentioned, these *kataegis* patterns might be different from those found in lymphomas and CLL [11], though this was not explored in detail in their comprehensive paper. We therefore reanalyzed these exome sequencing data from *kataegis* regions of lymphomas and CLL and compared them to *kataegis* regions in cancers with typical APOBEC signatures (Fig. 5). The preferred sequence for C to T mutation in *kataegis* regions of B-cell lymphomas and CLL revealed a target sequence that overlap with the known AID hotspot motif (WRCY W = A/T, R = purine, Y = pyrimidine). The general mutational pattern for C to T transitions in lymphomas and CLL was AG**C**T, rather than T**C**A/T for the other cancer types with *kataegis* (Fig. 5). This strongly implicates AID-induced genomic uracil formation in the development of localized hypermutation in B-cell malignancies, in accordance

an *in vitro* assay for complete BER of a single uracil in a defined U:G context. HeLa cells were synchronized by double thymidine block, and harvested after 0, 3, 8, and 14 h representing G1/early S-phase, mid S-phase, G1 and G2 phase, and G1 phase, respectively, as shown by flow cytometric confirmation of cell cycle distribution in the top row. The contribution of each uracil DNA glycosylase was measured by using neutralizing antibodies to UNG, SMUG1, or TDG as indicated. Note that the column size values in the panels are directly comparable since they are generated from the same gel using the same substrate. The data points represent mean of independent triplicate experiments. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Regression analysis of genomic uracil levels (linear) *vs.* expression of uracil-DNA repair glycosylases (linear) normalized either to total protein or to total protein per cell. Bold red indicates significant negative correlation (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Per total protein

	All cell lines		B-cell lymphoma cell lines		Non-lymphoma cell lines	
	R^2	*P*-value	R^2	*P*-value	R^2	*P*-value
UNG	**0.24**	**0.05**	0.01	0.82	0.23	0.28
SMUG1	**0.28**	**0.03**	**0.41**	**0.04**	0.13	0.43
TDG	0.05	0.35	0.02	0.69	0.13	0.41
MBD4	0.07	0.27	0.02	0.7	0.05	0.63

Per cell

	All cell lines		B-cell lymphoma cell lines		Non-lymphoma cell lines	
	R^2	*P*-value	R^2	*P*-value	R^2	*P*-value
UNG	**0.42**	**0.005**	0.05	0.52	0.20	0.31
SMUG1	**0.28**	**0.03**	0.16	0.24	0.06	0.6
TDG	0.22	0.06	0.14	0.27	0.32	0.17
MBD4	0.00	0.94	0.05	0.55	0.00	0.88

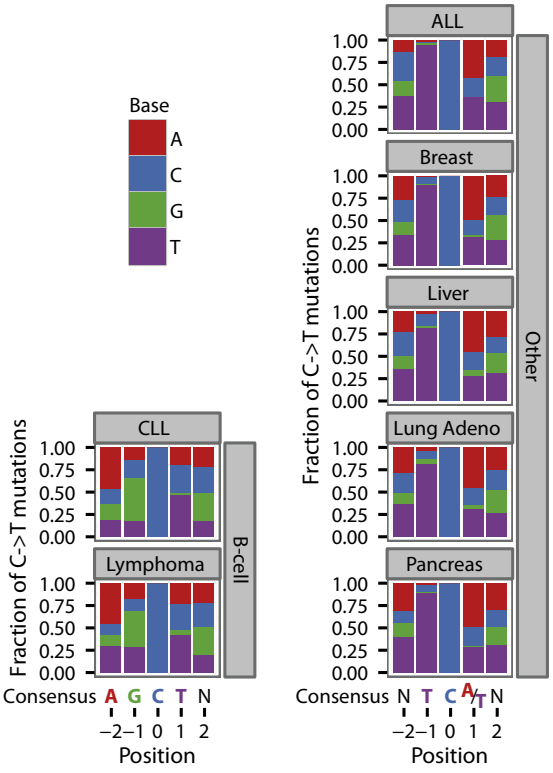


Fig. 5. Sequence context of C to T transitions in *kataegis* regions of lymphomas. Sequence analyses are based on exome sequencing data obtained from [11]. Sequence context of C to T transitions in *kataegis* regions of lymphomas ($n = 21$; 1102 single mutation sites) and CLL ($n = 15$; 290 single mutation sites) showing an AID-hotspot consensus sequence (—AG**C**TN—), where *N* represents no significant difference between A, T, C or G. Comparative analyses of cancers with known APOBEC signatures in *kataegis* regions showing an APOBEC consensus signature (—NT**C**ATN—), from ALL ($n = 1$; 153 single mutation sites), breast ($n = 67$; 5021 single mutation sites), liver ($n = 15$; 175 single mutation sites), lung adenocarcinoma ($n = 20$; 2024 single mutation sites), and pancreas ($n = 11$; 439 single mutation sites).

with our genomic uracil measurements and the published associations between AID and lymphomas [21–24,42–44] and CLL [45,46]. Moreover, these 122 lymphoma *kataegis* regions mapped to 70 distinct 100 kb blocks on 16 chromosomes, further supporting that enzymatic cytosine deamination by AID is not restricted to the S_μ region but occurs genome-wide.

4. Discussion

A major finding in our study is that AID expression is apparently a predominant source of genomic uracil in B-cell lymphoma cell lines. The LC–MS/MS method used quantifies genomic uracil as 2′-deoxyuridine in DNA [15]. The contribution of AID in this process was not solely made plausible by correlations, but also demonstrated by physiological induction of the endogenous AID gene, overexpression of recombinant AID, as well as knockdown of AID by shRNA. We feel that these results provide convincing evidence of dC to dU conversion *in vivo* by AID, which was considered missing in a recent review [47]. Furthermore; we found that mutational signatures in *kataegis* regions in human B-cell malignancies carry a distinct AID signature, strongly supporting the concept that AID is a DNA cytosine deaminase that, when mistargeted cause mutations and eventually B-cell malignancies. The increased genomic uracil is in general agreement with a recent report on relative increases in genomic uracil in B-cell lymphoma cell lines expressing AID, using an indirect genomic uracil-quantification method [48]. Evidence for targeted generation of uracil in Ig-genes has been obtained using a ligation-mediated PCR approach [20,49]. AID is normally only expressed in activated germinal center B-cells [2,50] and at low but detectable levels in early developing B-cells in the bone marrow [51]. This is apparently a risky process because AID strongly promotes the generation of germinal center-derived lymphomas [22,52,53], in which off-target activity of AID may contribute to point mutations and translocations during lymphomagenesis [31,54,55].

Recently, high-throughput sequencing of complete human cancer genomes and exomes revealed distinct mutational signatures compatible with mutagenesis by APOBEC-family enzymes in several common human cancers. This suggests that enzymatic off-target deamination of DNA-cytosine to uracil might be a major cause of mutation in human cancers [9–11]. However, direct evidence from measurements of uracil in the cancer genomes has largely been missing. Importantly, we found that endogenous AID-induction in CH12F3 mouse B-cells increases genomic uracil four-fold, from approximately 750 to 3000 uracils per genome already after 48 h. It is unlikely that this substantial increase can be confined to target regions in the Ig genes. Therefore, the increase in genomic uracil levels following endogenous AID expression indicates that even transiently induced AID expression during CSR causes widespread cytosine deamination. This is also in accordance with the mutational AID signatures found at many regions in human B-cell malignancy genomes. We did not observe correlation of genomic uracil with expression of other APOBEC-family

members. This does not rule out these as significant mutators in cancer cells, particularly since we only examined seven non-lymphoma cell lines. Low levels of enzymatic cytosine deamination may be overshadowed by dUMP misincorporation and spontaneous cytosine deamination. In addition, the strong effect of AID in B-cell lymphomas may obscure contribution of other APOBEC enzymes. A contribution from APOBECs may become significant over time and help drive transformation from normal cell to cancer cell, as indicated by mutational signatures [11,56].

Although AID expression levels correlated with variation in genomic uracil in the cells we tested, our results indicate that additional factors may modulate genomic uracil levels. The most obvious factor would be uracil repair capacity, which varies considerably between cell lines, and dUMP incorporation. We have previously shown that UNG is a rate-limiting factor in complete *in vitro* BER of genomic uracil [14] although UNG and SMUG1 may have complementary roles in uracil repair [16,29,57] and in the prevention of mutagenesis [58]. Studies on *UNG*$^{-/-}$ cells have documented an important function for UNG in keeping genomic uracil levels low [15]. However, the complete absence of any BER factor is a dramatic and rare event, whereas several-fold variation is rather common, at least in transformed cells. Earlier work demonstrated that AID-induced mutagenesis was counteracted by UNG, which initiates U:G DNA repair [31]. Our data showed that UNG and SMUG1 protein levels both correlated inversely with genomic uracil, with UNG showing the strongest correlation across all cells, while only SMUG1 correlates significantly in the lymphoma cell lines. Consequently, these results indicate that BER protein levels do affect genomic uracil. These results do not in themselves, however, necessarily reveal the relative importance of individual glycosylases for *in vivo* BER. We therefore made an effort to analyze the role of the glycosylases independently, using an assay for complete BER based on nuclear extracts from synchronized HeLa cells and a plasmid containing a single uracil. The results indicated that overall, UNG is the main contributor in initiating BER of uracil, at least in HeLa cells. However, SMUG1 and TDG may contribute significantly in G_1 (and G_2), which is also the time when AID is most active.

It is thought that U:G mismatches arising from AID in Ig genes and U:G from spontaneous deamination are processed by different mechanisms. Indeed, in order for SHM and CSR to be successfully carried out, canonical uracil DNA repair may be locally suppressed. One factor contributing to this may be transcription factor E2A, which induces AID [59,60], but represses both UNG-expression and its binding to relevant regions in the Ig genes [60]. Furthermore, p53 is actively reduced in germinal center B cells, presumably to allow mutagenic processing required for antibody maturation [61]. Although complex, the evidence that AID may drive carcinogenesis is well supported. In mice, AID expression was shown to be required for translocations between Ig loci and proto-oncogenes, a hallmark of several human B-cell lymphomas [62]. In contrast, AID knockout mice have fewer translocations [63] and accumulate fewer mutations in genes linked to B cell tumorigenesis [31]. AID expression has also been shown to confer a mutator phenotype in established lymphomas [42–44], but the role of AID in cancer progression remains unsettled [23,64,65]. Interestingly, AID expression has been reported in numerous cancers of non-B-cell origin, including breast, prostate, stomach, liver, and lung [66]. It would be interesting to investigate whether aberrant AID expression also confers high genomic uracil levels in these cancers. Interestingly, *Ung*$^{-/-}$ mice have roughly a 20-fold higher frequency of B-cell lymphoma compared with wild-type mice, but no apparent increase in other cancer types [67,68]. A straightforward explanation for this observation would be that SMUG1 and TDG together with MMR may compensate for UNG-deficiency in most tissues, but not in B-cells expressing AID, due to their increased genomic uracil levels.

A central role for AID-induced mutagenesis in lymphomas is also indicated by the AID-hotspot signature in the *kataegis* regions of a random selection of all lymphomas and CLLs (Fig. 5). We find that the *kataegis* AID-hotspot signature is not limited to lymphomas, but is also present in CLL, which overlaps with the category small lymphocytic lymphoma. Indeed, AID expression as cause of an ongoing mutator phenotype has been suggested for both lymphomas [42–44] and CLL [45,46]. Interestingly, progression of established cancers through expression of AID was also demonstrated in other blood cell cancers, such as ALL [69] and chronic myelogenous leukemia (CML), in which AID expression may lead to fatal lymphoblastoid crisis [70]. Thus, AID may be involved in development and progression of B-cell malignancies, and possibly only in late stage progression of other blood cell malignancies. This would be in agreement with the lack of an overall AID signature in ALL, as observed in our study.

In conclusion, we have provided strong evidence that AID is a DNA-cytosine deaminase that due to persistent expression causes accumulation of genomic uracil in B-cell lymphoma cell lines, as well as AID mutational signatures in human B-cell malignancies. Other factors, including expression levels for uracil-DNA glycosylases and cell doubling time, may modulate genomic uracil levels, but AID levels remain the strongest predictor.

Conflict of interest statement

There are no conflicts of interest.

Funding

This work was supported by The Norwegian Cancer Association (Grant id.: 576160), the Research Council of Norway (Grant id's: 191408; HEK) and 205316; PS), the Svanhild and Arne Must Fund for Medical Research, the Liaison Committee between the CentralNorway Regional Health Authority (RHA) and the Norwegian University of Science and Technology (NTNU) (Grant id.: 46047800; HSP) and Norwegian University of Science and Technology.

References

[1] B. Kavli, M. Otterlei, G. Slupphaug, H.E. Krokan, Uracil in DNA—general mutagen, but normal intermediate in acquired immunity, DNA Repair 6 (2007) 505–516.
[2] M. Muramatsu, V.S. Sankaranand, S. Anant, M. Sugai, K. Kinoshita, N.O. Davidson, T. Honjo, Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells, J. Biol. Chem. 274 (1999) 18470–18476.
[3] S.K. Petersen-Mahrt, R.S. Harris, M.S. Neuberger, AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification, Nature 418 (2002) 99–103.
[4] J. Di Noia, M.S. Neuberger, Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase, Nature 419 (2002) 43–48.
[5] K. Imai, G. Slupphaug, W.I. Lee, P. Revy, S. Nonoyama, N. Catalan, L. Yel, M. Forveille, B. Kavli, H.E. Krokan, H.D. Ochs, A. Fischer, A. Durandy, Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination, Nat. Immunol. 4 (2003) 1023–1028.
[6] C. Rada, G.T. Williams, H. Nilsen, D.E. Barnes, T. Lindahl, M.S. Neuberger, Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice, Curr. Biol.: CB 12 (2002) 1748–1755.
[7] S.G. Conticello, The AID/APOBEC family of nucleic acid mutators, Genome Biol. 9 (2008) 229.
[8] R.S. Harris, S.K. Petersen-Mahrt, M.S. Neuberger, RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators, Mol. Cell. 10 (2002) 1247–1253.
[9] J. Zhang, V. Grubor, C.L. Love, A. Banerjee, K.L. Richards, P.A. Mieczkowski, C. Dunphy, W. Choi, W.Y. Au, G. Srivastava, P.L. Lugar, D.A. Rizzieri, A.S. Lagoo, L. Bernal-Mizrachi, K.P. Mann, C. Flowers, K. Naresh, A. Evens, L.I. Gordon, M. Czader, J.I. Gill, E.D. Hsi, Q. Liu, A. Fan, K. Walsh, D. Jima, L.L. Smith, A.J. Johnson, J.C. Byrd, M.A. Luftig, T. Ni, J. Zhu, A. Chadburn, S. Levy, D. Dunson, S.S. Dave, Genetic heterogeneity of diffuse large B-cell lymphoma, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 1398–1403.
[10] C. Greenman, P. Stephens, R. Smith, G.L. Dalgliesh, C. Hunter, G. Bignell, H. Davies, J. Teague, A. Butler, C. Stevens, S. Edkins, S. O'Meara, I. Vastrik, E.E. Schmidt, T. Avis, S. Barthorpe, G. Bhamra, G. Buck, B. Choudhury, J. Clements, J.

Cole, E. Dicks, S. Forbes, K. Gray, K. Halliday, R. Harrison, K. Hills, J. Hinton, A. Jenkinson, D. Jones, A. Menzies, T. Mironenko, J. Perry, K. Raine, D. Richardson, R. Shepherd, A. Small, C. Tofts, J. Varian, T. Webb, S. West, S. Widaa, A. Yates, D.P. Cahill, D.N. Louis, P. Goldstraw, A.G. Nicholson, F. Brasseur, L. Looijenga, B.L. Weber, Y.E. Chiew, A. DeFazio, M.F. Greaves, A.R. Green, P. Campbell, E. Birney, D.F. Easton, G. Chenevix-Trench, M.H. Tan, S.K. Khoo, B.T. Teh, S.T. Yuen, S.Y. Leung, R. Wooster, P.A. Futreal, M.R. Stratton, Patterns of somatic mutation in human cancer genomes, Nature 446 (2007) 153–158.

[11] L.B. Alexandrov, S. Nik-Zainal, D.C. Wedge, S.A. Aparicio, S. Behjati, A.V. Biankin, G.R. Bignell, N. Bolli, A. Borg, A.L. Borresen-Dale, S. Boyault, B. Burkhardt, A.P. Butler, C. Caldas, H.R. Davies, C. Desmedt, R. Eils, J.E. Eyfjord, J.A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilicic, S. Imbeaud, M. Imielinski, N. Jager, D.T. Jones, D. Jones, S. Knappskog, M. Kool, S.R. Lakhani, C. Lopez-Otin, S. Martin, N.C. Munshi, H. Nakamura, P.A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J.V. Pearson, X.S. Puente, K. Raine, M. Ramakrishna, A.L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T.N. Schumacher, P.N. Span, J.W. Teague, Y. Totoki, A.N. Tutt, R. Valdes-Mas, M.M. van Buuren, L. van't Veer, A. Vincent-Salomon, N. Waddell, L.R. Yates, I. Australian Pancreatic Cancer Genome, I.B.C. Consortium, I.M.-S. Consortium, I. PedBrain, J. Zucman-Rossi, P.A. Futreal, U. McDermott, P. Lichter, M. Meyerson, S.M. Grimmond, R. Siebert, E. Campo, T. Shibata, S.M. Pfister, P.J. Campbell, M.R. Stratton, Signatures of mutational processes in human cancer, Nature 500 (2013) 415–421.

[12] Y. Hu, I. Ericsson, K. Torseth, S.P. Methot, O. Sundheim, N.B. Liabakk, G. Slupphaug, J.M. Di Noia, H.E. Krokan, B. Kavli, A combined nuclear and nucleolar localization motif in activation-induced cytidine deaminase (AID) controls immunoglobulin class switching, J. Mol. Biol. 425 (2013) 424–443.

[13] L. Hagen, B. Kavli, M.M. Sousa, K. Torseth, N.B. Liabakk, O. Sundheim, J. Pena-Diaz, M. Otterlei, O. Horning, O.N. Jensen, H.E. Krokan, G. Slupphaug, Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA, EMBO J. 27 (2008) 51–61.

[14] T. Visnes, M. Akbari, L. Hagen, G. Slupphaug, H.E. Krokan, The rate of base excision repair of uracil is controlled by the initiating glycosylase, DNA Repair 7 (2008) 1869–1881.

[15] A. Galashevskaya, A. Sarno, C.B. Vagbo, P.A. Aas, L. Hagen, G. Slupphaug, H.E. Krokan, A robust, sensitive assay for genomic uracil determination by LC/MS/MS reveals lower levels than previously reported, DNA Repair 12 (2013) 699–706.

[16] B. Kavli, O. Sundheim, M. Akbari, M. Otterlei, H. Nilsen, F. Skorpen, P.A. Aas, L. Hagen, H.E. Krokan, G. Slupphaug, hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup, J. Biol. Chem. 277 (2002) 39926–39936.

[17] M. Akbari, M. Otterlei, J. Pena-Diaz, P.A. Aas, B. Kavli, N.B. Liabakk, L. Hagen, K. Imai, A. Durandy, G. Slupphaug, H.E. Krokan, Repair of U/G and U/A in DNA by UNG2-associated repair complexes takes place predominantly by short-patch repair both in proliferating and growth-arrested cells, Nucleic Acids Res. 32 (2004) 5486–5498.

[18] T.S. Batth, J.D. Keasling, C.J. Petzold, Targeted proteomics for metabolic pathway optimization, Methods Mol. Biol. 944 (2012) 237–249.

[19] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, M.J. MacCoss, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments, Bioinformatics 26 (2010) 966–968.

[20] R.W. Maul, H. Saribasak, S.A. Martomo, R.L. McClure, W. Yang, A. Vaisman, H.S. Gramlich, D.G. Schatz, R. Woodgate, D.M. Wilson, P.J. 3rd, Gearhart, Uracil residues dependent on the deaminase AID in immunoglobulin gene variable and switch regions, Nat. Immunol. 12 (2011) 70–76.

[21] J. Greeve, A. Philipsen, K. Krause, W. Klapper, K. Heidorn, B.E. Castle, J. Janda, K.B. Marcu, R. Parwaresch, Expression of activation-induced cytidine deaminase in human B-cell non-Hodgkin lymphomas, Blood 101 (2003) 3574–3580.

[22] L.A. Smit, R.J. Bende, J. Aten, J.E. Guikema, W.M. Aarts, C.J. van Noesel, Expression of activation-induced cytidine deaminase is confined to B-cell non-Hodgkin's lymphomas of germinal-center phenotype, Cancer Res. 63 (2003) 3894–3898.

[23] I.S. Lossos, R. Levy, A.A. Alizadeh, AID is expressed in germinal center B-cell-like and activated B-cell-like diffuse large-cell lymphomas and is not correlated with intraclonal heterogeneity, Leukemia 18 (2004) 1775–1779.

[24] L. Pasqualucci, R. Guglielmino, J. Houldsworth, J. Mohr, S. Aoufouchi, R. Polakiewicz, R.S. Chaganti, R. Dalla-Favera, Expression of the AID protein in normal and neoplastic B cells, Blood 104 (2004) 3318–3325.

[25] S. Gallien, E. Duriez, C. Crone, M. Kellmann, T. Moehring, B. Domon, Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer, Mol. Cell. Proteomics: MCP 11 (2012) 1709–1723.

[26] A.C. Peterson, J.D. Russell, D.J. Bailey, M.S. Westphall, J.J. Coon, Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics, Mol. Cell. Proteomics: MCP 11 (2012) 1475–1488.

[27] Y. Hu, I. Ericsson, B. Doseth, N.B. Liabakk, H.E. Krokan, B. Kavli, Activation-induced cytidine deaminase (AID) is localized to subnuclear domains enriched in splicing factors, Exp. Cell Res. 322 (2014) 178–192.

[28] H.E. Krokan, M. Bjoras, Base excision repair, Cold Spring Harbor Perspect. Biol. 5 (2013) a012583.

[29] H.S. Pettersen, O. Sundheim, K.M. Gilljam, G. Slupphaug, H.E. Krokan, B. Kavli, Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms, Nucleic Acids Res. 35 (2007) 3879–3892.

[30] A. Yamane, W. Resch, N. Kuo, S. Kuchen, Z. Li, H.W. Sun, D.F. Robbiani, K. McBride, M.C. Nussenzweig, R. Casellas, Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes, Nat. Immunol. 12 (2011) 62–69.

[31] M. Liu, J.L. Duke, D.J. Richter, C.G. Vinuesa, C.C. Goodnow, S.H. Kleinstein, D.G. Schatz, Two levels of protection for the B cell genome during somatic hypermutation, Nature 451 (2008) 841–845.

[32] M.G. Hasham, N.M. Donghia, E. Coffey, J. Maynard, K.J. Snow, J. Ames, R.Y. Wilpan, Y. He, B.L. King, K.D. Mills, Widespread genomic breaks generated by activation-induced cytidine deaminase are prevented by homologous recombination, Nat. Immunol. 11 (2010) 820–826.

[33] H. Nilsen, M. Otterlei, T. Haug, K. Solum, T.A. Nagelhus, F. Skorpen, H.E. Krokan, Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene, Nucleic Acids Res. 25 (1997) 750–755.

[34] H. Nilsen, K.S. Steinsbekk, M. Otterlei, G. Slupphaug, P.A. Aas, H.E. Krokan, Analysis of uracil-DNA glycosylases from the murine Ung gene reveals differential expression in tissues and in embryonic development and a subcellular sorting pattern that differs from the human homologues, Nucleic Acids Res. 28 (2000) 2277–2285.

[35] S. Andersen, T. Heine, R. Sneve, I. Konig, H.E. Krokan, B. Epe, H. Nilsen, Incorporation of dUMP into DNA is a major source of spontaneous DNA damage, while excision of uracil is not required for cytotoxicity of fluoropyrimidines in mouse embryonic fibroblasts, Carcinogenesis 26 (2005) 547–555.

[36] E.C. Ordinario, M. Yabuki, R.P. Larson, N. Maizels, Temporal regulation of Ig gene diversification revealed by single-cell imaging, J. Immunol. 183 (2009) 4545–4553.

[37] G. Sharbeen, C.W. Yee, A.L. Smith, C.J. Jolly, Ectopic restriction of DNA repair reveals that UNG2 excises AID-induced uracils predominantly or exclusively during G1 phase, J. Exp. Med. 209 (2012) 965–974.

[38] C.E. Schrader, J.E. Guikema, E.K. Linehan, E. Selsing, J. Stavnezer, Activation-induced cytidine deaminase-dependent DNA breaks in class switch recombination occur during G1 phase of the cell cycle and depend upon mismatch repair, J. Immunol. 179 (2007) 6064–6071.

[39] U. Hardeland, C. Kunz, F. Focke, M. Szadkowski, P. Schar, Cell cycle regulation as a mechanism for functional separation of the apparently redundant uracil DNA glycosylases TDG and UNG2, Nucleic Acids Res. 35 (2007) 3859–3867.

[40] J. Pena-Diaz, S.A. Hegre, E. Anderssen, P.A. Aas, R. Mjelle, G.D. Gilfillan, R. Lyle, F. Drablos, H.E. Krokan, P. Saetrom, Transcription profiling during the cell cycle shows that a subset of Polycomb-targeted genes is upregulated during DNA replication, Nucleic Acids Res. 41 (2013) 2846–2856.

[41] M. Akbari, J. Pena-Diaz, S. Andersen, N.B. Liabakk, M. Otterlei, H.E. Krokan, Extracts of proliferating and non-proliferating human cells display different base excision pathways and repair fidelity, DNA Repair 8 (2009) 834–843.

[42] M.S. Hardianti, E. Tatsumi, M. Syampurnawati, K. Furuta, K. Saigo, Y. Nakamachi, S. Kumagai, H. Ohno, S. Tanabe, M. Uchida, N. Yasuda, Activation-induced cytidine deaminase expression in follicular lymphoma: association between AID expression and ongoing mutation in FL, Leukemia 18 (2004) 826–831.

[43] C. Bodor, A. Bognar, L. Reiniger, A. Szepesi, E. Toth, L. Kopper, A. Matolcsy, Aberrant somatic hypermutation and expression of activation-induced cytidine deaminase mRNA in mediastinal large B-cell lymphoma, Br. J. Haematol. 129 (2005) 373–376.

[44] A.J. Deutsch, A. Aigelsreiter, P.B. Staber, A. Beham, W. Linkesch, C. Guelly, R.I. Brezinschek, M. Fruhwirth, W. Emberger, M. Buettner, C. Beham-Schmid, P. Neumeister, MALT lymphoma and extranodal diffuse large B-cell lymphoma are targeted by aberrant somatic hypermutation, Blood 109 (2007) 3500–3504.

[45] H. McCarthy, W.G. Wierda, L.L. Barron, C.C. Cromwell, J. Wang, K.R. Coombes, R. Rangel, K.S. Elenitoba-Johnson, M.J. Keating, L.V. Abruzzo, High expression of activation-induced cytidine deaminase (AID) and splice variants is a distinctive feature of poor-prognosis chronic lymphocytic leukemia, Blood 101 (2003) 4903–4908.

[46] F. Palacios, P. Moreno, P. Morande, C. Abreu, A. Correa, V. Porro, A.I. Landoni, R. Gabus, M. Giordano, G. Dighiero, O. Pritsch, P. Oppezzo, High expression of AID and active class switch recombination might account for a more aggressive disease in unmutated CLL patients: link with an activated microenvironment in CLL disease, Blood 115 (2010) 4488–4496.

[47] A.S. Yousif, A. Stanlie, N.A. Begum, T. Honjo, Opinion: uracil DNA glycosylase (UNG) plays distinct and non-canonical roles in somatic hypermutation and class switch recombination, Int. Immunol. 26 (2014) 575–578.

[48] S. Shalhout, D. Haddad, A. Sosin, T.C. Holland, A. Al-Katib, A. Martin, A.S. Bhagwat, Genomic uracil homeostasis during normal B cell maturation and loss of this balance during B cell cancer development, Mol. Cell. Biol. 34 (2014) 4019–4032.

[49] B. Roche, A. Claes, F. Rougeon, Deoxyuridine triphosphate incorporation during somatic hypermutation of mouse VkOx genes after immunization with phenyloxazolone, J. Immunol. 185 (2010) 4777–4782.

[50] E.E. Crouch, Z. Li, M. Takizawa, S. Fichtner-Feigl, P. Gourzi, C. Montano, L. Feigenbaum, P. Wilson, S. Janz, F.N. Papavasiliou, R. Casellas, Regulation of AID expression in the immune response, J. Exp. Med. 204 (2007) 1145–1156.

[51] J.H. Han, S. Akira, K. Calame, B. Beutler, E. Selsing, T. Imanishi-Kari, Class switch recombination and somatic hypermutation in early mouse B cells are mediated by B cell and Toll-like receptors, Immunity 27 (2007) 64–75.

[52] L. Pasqualucci, G. Bhagat, M. Jankovic, M. Compagno, P. Smith, M. Muramatsu, T. Honjo, H.C. Morse, M.C. Nussenzweig III, R. Dalla-Favera, AID is required for germinal center-derived lymphomagenesis, Nat. Genet. 40 (2008) 108–112.

[53] A. Kotani, N. Kakazu, T. Tsuruyama, I.M. Okazaki, M. Muramatsu, K. Kinoshita, H. Nagaoka, D. Yabe, T. Honjo, Activation-induced cytidine deaminase (AID)

promotes B cell lymphomagenesis in Emu-cmyc transgenic mice, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 1616–1620.
[54] O. Hakim, W. Resch, A. Yamane, I. Klein, K.R. Kieffer-Kwon, M. Jankovic, T. Oliveira, A. Bothmer, T.C. Voss, C. Ansarah-Sobrinho, E. Mathe, G. Liang, J. Cobell, H. Nakahashi, D.F. Robbiani, A. Nussenzweig, G.L. Hager, M.C. Nussenzweig, R. Casellas, DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes, Nature 484 (2012) 69–74.
[55] I.A. Klein, W. Resch, M. Jankovic, T. Oliveira, A. Yamane, H. Nakahashi, M. Di Virgilio, A. Bothmer, A. Nussenzweig, D.F. Robbiani, R. Casellas, M.C. Nussenzweig, Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes, Cell 147 (2011) 95–106.
[56] M.B. Burns, L. Lackey, M.A. Carpenter, A. Rathore, A.M. Land, B. Leonard, E.W. Refsland, D. Kotandeniya, N. Tretyakova, J.B. Nikas, D. Yee, N.A. Temiz, D.E. Donohue, R.M. McDougle, W.L. Brown, E.K. Law, R.S. Harris, APOBEC3B is an enzymatic source of mutation in breast cancer, Nature 494 (2013) 366–370.
[57] H. Nilsen, K.A. Haushalter, P. Robins, D.E. Barnes, G.L. Verdine, T. Lindahl, Excision of deaminated cytosine from the vertebrate genome: role of the SMUG1 uracil-DNA glycosylase, EMBO J. 20 (2001) 4278–4286.
[58] Q. An, P. Robins, T. Lindahl, D.E. Barnes, C –> T mutagenesis and gamma-radiation sensitivity due to deficiency in the Smug1 and Ung DNA glycosylases, EMBO J. 24 (2005) 2205–2213.
[59] J. Hauser, N. Sveshnikova, A. Wallenius, S. Baradaran, J. Saarikettu, T. Grundstrom, B-cell receptor activation inhibits AID expression through calmodulin inhibition of E-proteins, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 1267–1272.
[60] A. Wallenius, J. Hauser, P.A. Aas, A. Sarno, B. Kavli, H.E. Krokan, T. Grundstrom, Expression and recruitment of uracil-DNA glycosylase are regulated by E2A during antibody diversification, Mol. Immunol. 60 (2014) 23–31.
[61] R.T. Phan, R. Dalla-Favera, The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells, Nature 432 (2004) 635–639.
[62] R. Kuppers, R. Dalla-Favera, Mechanisms of chromosomal translocations in B cell lymphomas, Oncogene 20 (2001) 5580–5594.
[63] Y. Dorsett, D.F. Robbiani, M. Jankovic, B. Reina-San-Martin, T.R. Eisenreich, M.C. Nussenzweig, A role for AID in chromosome translocations between c-myc and the IgH variable region, J. Exp. Med. 204 (2007) 2225–2232.
[64] M. Leuenberger, S. Frigerio, P.J. Wild, F. Noetzli, D. Korol, D.R. Zimmermann, C. Gengler, N.M. Probst-Hensch, H. Moch, M. Tinguely, AID protein expression in chronic lymphocytic leukemia/small lymphocytic lymphoma is associated with poor prognosis and complex genetic alterations, Mod. Pathol. 23 (2010) 177–186 (an official journal of the United States and Canadian Academy of Pathology, Inc).
[65] K. Willenbrock, C. Renne, M. Rottenkolber, W. Klapper, M. Dreyling, M. Engelhard, R. Kuppers, M.L. Hansmann, B. Jungnickel, The expression of activation induced cytidine deaminase in follicular lymphoma is independent of prognosis and stage, Histopathology 54 (2009) 509–512.
[66] A. Orthwein, J.M. Di Noia, Activation induced deaminase: how much and where? Semin. Immunol. 24 (2012) 246–254.
[67] S. Andersen, M. Ericsson, H.Y. Dai, J. Pena-Diaz, G. Slupphaug, H. Nilsen, H. Aarset, H.E. Krokan, Monoclonal B-cell hyperplasia and leukocyte imbalance precede development of B-cell malignancies in uracil-DNA glycosylase deficient mice, DNA Repair 4 (2005) 1432–1441.
[68] H. Nilsen, G. Stamp, S. Andersen, G. Hrivnak, H.E. Krokan, T. Lindahl, D.E. Barnes, Gene-targeted mice lacking the Ung uracil-DNA glycosylase develop B-cell lymphomas, Oncogene 22 (2003) 5381–5386.
[69] T.A. Gruber, M.S. Chang, R. Sposto, M. Muschen, Activation-induced cytidine deaminase accelerates clonal evolution in BCR-ABL1-driven B-cell lineage acute lymphoblastic leukemia, Cancer Res. 70 (2010) 7411–7420.
[70] L. Klemm, C. Duy, I. Iacobucci, S. Kuchen, G. von Levetzow, N. Feldhahn, N. Henke, Z. Li, T.K. Hoffmann, Y.M. Kim, W.K. Hofmann, H. Jumaa, J. Groffen, N. Heisterkamp, G. Martinelli, M.R. Lieber, R. Casellas, M. Muschen, The B cell mutator AID promotes B lymphoid blast crisis and drug resistance in chronic myeloid leukemia, Cancer Cell 16 (2009) 232–245.