

The cytidine deaminase APOBEC3C has unique sequence and genome feature preferences

Grant W. Brown ^{1,2,*}

¹Department of Biochemistry, University of Toronto, 1 King's College Circle, Toronto, ON, Canada M5S 1A8

²Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, 160 College Street, Toronto, ON, Canada M5S 3E1

*Corresponding author: Department of Biochemistry and Donnelly Centre, University of Toronto, 160 College Street, Toronto, Canada M5S 3E1. Email: grant.brown@utoronto.ca

APOBEC proteins are cytidine deaminases that restrict the replication of viruses and transposable elements. Several members of the APOBEC3 family, APOBEC3A, APOBEC3B, and APOBEC3H-I, can access the nucleus and cause what is thought to be indiscriminate deamination of the genome, resulting in mutagenesis and genome instability. Although APOBEC3C is also present in the nucleus, the full scope of its deamination target preferences is unknown. By expressing human APOBEC3C in a yeast model system, I have defined the APOBEC3C mutation signature, as well as the preferred genome features of APOBEC3C targets. The APOBEC3C mutation signature is distinct from those of the known cancer genome mutators APOBEC3A and APOBEC3B. APOBEC3C produces DNA strand-coordinated mutation clusters, and APOBEC3C mutations are enriched near the transcription start sites of active genes. Surprisingly, APOBEC3C lacks the bias for the lagging strand of DNA replication that is seen for APOBEC3A and APOBEC3B. The unique preferences of APOBEC3C constitute a mutation profile that will be useful in defining sites of APOBEC3C mutagenesis in human genomes.

Keywords: mutagenesis; cytidine deaminase; mutation signature; humanized yeast; mutation accumulation; DNA replication stress

Introduction

The human APOBEC3 family comprises 7 paralogous genes that encode cytidine deaminases. APOBEC3 proteins target single-stranded DNA (ssDNA) and RNA to perform nonredundant functions in restricting the replication of viruses and the transposition of LINE elements (reviewed in Koito and Ikeda 2013; Harris and Dudley 2015; Pecori et al. 2022). More recently, several APOBEC3 proteins have been implicated in the mutagenesis of cancer genomes and in causing genome instability (reviewed in Mertz et al. 2022). Individual APOBEC3 proteins have diagnostic sequence and structural contexts for their preferred deamination sites, and the identification of these contexts has been instrumental in dissecting the contributions of individual APOBEC3 proteins to genome mutagenesis (Nik-Zainal et al. 2012; Roberts et al. 2012; Alexandrov et al. 2013; Burns et al. 2013a; Chan et al. 2015; Supek and Lehner 2017; Petljak et al. 2019, 2022; Buisson et al. 2019; Mas-Ponte and Supek 2020; Jakobsdottir et al. 2022; Sanchez et al. 2024).

Currently there is only a modest indication of the APOBEC3C deamination site consensus sequence. Early studies indicated that the target specificity of APOBEC3C differed from that of APOBEC3G (Harris et al. 2002). A preference for the TCA context (the underlined C being the deamination target) was noted in deamination of hepatitis B virus genomes by APOBEC3C (Khalfi et al. 2022), as was a preference for T or C at the −1 position relative to

the deaminated deoxycytidine, and C at +1 (Chen et al. 2021). A preference for T or C at −1 was also noted in deamination assays of Moloney leukemia virus, as was some preference for T or A at −2 (Langlois et al. 2005). Purified APOBEC3C preferred TCA and TCG contexts in vitro (Ito et al. 2017). Analysis of 67 mutations accumulated in a yeast strain expressing APOBEC3C indicated a preference for T at −1 and at −2 (Taylor et al. 2013). By contrast, the deamination site consensus sequences for APOBEC3A and APOBEC3B are well understood (Chan et al. 2015; McDaniel et al. 2020; Hou et al. 2021; Petljak et al. 2022). Both prefer a TCW context (Taylor et al. 2013; Hoopes et al. 2016), as is typical of the APOBEC3 family (with the exception of A3G (Schumacher et al. 2005; Liu et al. 2023)), but differ in their preference at the −2 position (Chan et al. 2015). Structural analyses have shed additional light on mechanisms of substrate preference, particularly at the central TC motif for APOBEC3A and APOBEC3B, and provide a basis for DNA conformation preferences (Kouno et al. 2017; Shi et al. 2017; Hou et al. 2021; Sanchez et al. 2024). Structural information has not yet extended to APOBEC3C, and existing data suggest that the preferred deamination sites of APOBEC3C likely differ from those of APOBEC3A and APOBEC3B.

Prevailing models suggest that 2 factors contribute to deamination of the nuclear genome by APOBEC3 proteins: (1) the intrinsic properties of each APOBEC3 protein, and (2) the features of the genome that could reveal substrates for deamination by APOBEC3s. Analyses in human cells, cancer genomes, and in

Received on 17 January 2024; accepted on 22 May 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of The Genetics Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

humanized yeast models have established connections to genome features that are preferred sites of APOBEC3 action. Two themes predominate. First, there is an association with DNA replication and with active transcription (Kazanov et al. 2015; Morganella et al. 2016; Chervova et al. 2021). Second, there is a preference for regions presumed to have more single-stranded DNA: the lagging strand of DNA replication, the nontranscribed strand of transcription units, DNA repair tracts, and the displaced DNA strand in R-loops (Haradhvala et al. 2016; Morganella et al. 2016; Septyarskiy et al. 2016; Mas-Ponte and Supek 2020; Sui et al. 2020; Chen et al. 2021; Chervova et al. 2021; Zhang et al. 2022; McCann et al. 2023). Whether these preferences are universal is unclear, and little is currently known about the genome features that might promote deamination by APOBEC3C.

Here I describe a humanized yeast model to define the deamination site consensus sequence and the preferred genome features for deamination by human APOBEC3C.

Methods

Yeast strains

Yeast strains (Supplementary Table 1 and Reagents Table) used in this study were derived from BY4741 (Brachmann et al. 1998) with gene corrections to improve sporulation and mitochondrial function (Harvey et al. 2018). Strains were constructed using standard yeast genetic and molecular cloning methods and were cultured under standard conditions (Dunham et al. 2015). The APOBEC3C wild-type and variant ORFs (C97S C100S and S188I) were assembled from PCR amplicons of codon-optimized synthetic DNAs containing the desired mutations. The NCBI Reference Sequence NP_055323.2 was used as the wild-type APOBEC3C amino acid sequence, and the codon-optimized wild-type APOBEC3C DNA sequence is in Supplementary Table 2. The gene expression transcription units were assembled from a hybrid pGPD14/15 promoter (Kotopka and Smolke 2020), the APOBEC3C open reading frame, sequences encoding the superFLAG epitope (Layton et al. 2019), and the tsynth27 transcription terminator (Curran et al. 2015), using Golden Gate assembly (Hamedirad et al. 2019). Transcription units were integrated into the *LYP1* locus, targeted with a CRISPR/Cas9 double-strand break specified by an sgRNA sequence (5'-CATAATAACGTCCAATAAAT) cloned in the plasmid pUB1306 (constructed by Gavin Schlissel in Jasper Rine's lab, and a kind gift from Elçin Ünal, UC Berkeley). APOBEC3 strains (wild type, C97S C100S, and S188I; strains GBY789, GBY793, and GBY805) were confirmed by PCR, by Sanger sequencing of PCR amplified gene loci, and by immunoblot analysis. *UNG1* was deleted by amplifying the *ung1Δ::kanMX* locus from genomic DNA purified from the *Saccharomyces cerevisiae* genome deletion project MATa *ung1Δ::kanMX* strain, followed by transformation of the resulting PCR product into DHY214 to yield strain GBY732. The *ung1Δ* was crossed into the APOBEC3C strain followed by identification of an *ung1Δ::kanMX lyp1::A3C* segregant (GBY819). *RAD5* was deleted by amplifying the *rad5Δ::kanMX* locus from genomic DNA purified from the *S. cerevisiae* genome deletion project MATa *rad5Δ::kanMX* strain, followed by transformation of the resulting PCR product into DHY213 to yield strain GBY810. Strains are available on request.

Immunoblot analysis

For validation of protein expression, strains were grown to mid logarithmic phase and 5 OD₆₀₀ units of cells were collected, diluted to a total volume of 10 ml in 10% trichloroacetic acid (TCA), incubated at room temperature for 15 minutes, then pelleted by centrifugation at 3,000 rpm for 5 minutes. Fixed cells

were washed with 1 ml of 1 M Tris-HCl pH 6.8, harvested, resuspended in 50 µl of 5X Laemmli sample buffer (0.25 M Tris-HCl pH 6.8, 0.5 M DTT, 10% SDS, 50% glycerol, 0.5% bromophenol blue), and lysed by vortexing with 0.5-mm glass beads for 10 minutes. An additional 50 µl 5X Laemmli sample buffer was added to the sample and the extract was incubated at 95°C for 5 minutes and microfuged briefly prior to gel loading. Immunoblots were stained with 0.2% Ponceau S in 3% TCA, destained with TBST, and blocked with 5% nonfat dry milk in TBST. APOBEC3C tagged with the superFLAG epitope was detected by immunoblotting with an anti-FLAG mouse antibody (Sigma anti-FLAG M2 antibody F3165, 1:5,000 dilution).

Mutation rate assays

Mutation rates at *CAN1* were measured using a Luria-Delbrück fluctuation test (Lang 2018), with modifications. For each fluctuation test, saturated cultures were diluted 1:10,000 in 10 ml of SD-arginine, and 96 cultures (30 µl) were grown in a flat-bottom 96-well plate for 48 hours at 30°C without shaking. After incubation, 8 cultures were pooled, diluted with dH₂O, and plated on fully supplemented SD to calculate N(t). For the remaining 88 cultures, 200 µl of SD-arginine containing 60 mg/l canavanine was added to each well. The plate was then sealed with an adhesive plate seal and incubated for 3 days at 30°C. Wells were scored for colony formation, and mutation rate was calculated according to the Poisson distribution (Supplementary Table 3). Strains assayed were DHY213, GBY789, GBY793, GBY805, GBY818, and GBY819.

Genetic interaction analysis

To assess genetic interactions between APOBEC3C alleles and *RAD5*, strains GBY789, GBY793, and GBY805 were each crossed to GBY810. Diploids of each were sporulated, and 20 tetrads of each were dissected on YPD. Plates were imaged after 48 hours at 30°C, and replica-plated to YPD + G418 to identify *rad5Δ* strains and to SD-lysine + S-(2-aminoethyl)-L-cysteine hydrochloride to identify APOBEC3C strains. Colony sizes were measured from the plate images using Balony (Young and Loewen 2013), and raw values are listed in Supplementary Table 4. The fitness of each genotype was calculated from the colony size values and expressed relative to the fitness of the wild-type strain (Supplementary Table 5). Genetic interaction scores were calculated as the deviation from expectation where expectation is the product of the single mutant fitness values (Supplementary Table 5).

Mutation accumulation

Strain GBY819, which expresses APOBEC3C in an *ung1Δ* background, was used for mutation accumulation. Eighteen independent lines were established from single colonies. Each line was streaked for single colonies on YPD and incubated at 30°C for 48 hours for 1 passage (approximately 25 generations). The single colony closest to the center of the plate was streaked for single colonies for each subsequent passage. After 15 passages (~375 generations) or 20 passages (~500 generations) the single colony closest to the center of the plate was inoculated into 2-ml YPD and incubated with shaking at 30°C for 48 hours. Genomic DNA was prepared from 0.5 ml of culture with the MasterPure Yeast DNA Purification kit (Biosearch Technologies) according to the manufacturer's instructions. Genomic DNA concentrations were measured using Qubit dsDNA Broad Range reagent (ThermoFisher Scientific).

Sequencing and variant calling

Sequencing libraries were prepared by tagmentation, using Nextera (Illumina) reagents scaled to small volume reactions

(Vonesch *et al.* 2021). Libraries were pooled and subjected to paired-end sequencing with 75-bp reads on an Illumina NextSeq. De-multiplexed raw sequencing reads are available from the NCBI Sequencing Read Archive, BioProject ID PRJNA938955.

Variants were called for each mutation accumulation line using the Snippy tool in Galaxy (<https://usegalaxy.org>). Paired reads were concatenated and treated as single-end reads, the reference genome was *sacCer3*, and minimum read depth of 10 and minimum variant proportion of 0.9 were required to call a variant. Variant data were read into R. Any variant that occurred more than once was presumed to be parental and was removed. Complex variants and multiple nucleotide variants were manually deconvoluted into single-nucleotide variants (SNVs). Variants mapping to the mitochondrial genome were removed. The resulting unique nuclear genome variants are tabulated in [Supplementary Table 6](#).

To make comparisons with APOBEC3A and APOBEC3B expressed in yeast, the data from (Hoopes *et al.* 2016) were acquired (NCBI SRA BioProject ID PRJNA307256 and PRJNA307256), and variants were called with the VarScan Somatic tool in Galaxy, using *sacCer3* as the reference genome, minimum coverage = 9, minimum supporting reads = 4, and minimum variant allele frequency = 0.4 (as these data were derived from diploid cells). Files (vcf) for SNVs were processed to retain heterozygous variants (TUMOR = 0/1) with FILTER = PASS, INFO = SOMATIC, and to remove variants that occurred more than once. SNVs mapping to the mitochondrial genome were removed. The resulting unique nuclear genome SNVs are tabulated in [Supplementary Tables 7](#) (APOBEC3A) and [8](#) (APOBEC3B).

SNVs resulting from deamination (C->T and G->A) were annotated and retained. SNVs in repetitive regions of the yeast genome were removed, using the definitions in Sui *et al.* (2020).

Mutation signatures

The 3 and 7 nucleotide contexts of each deamination SNV were extracted using the mutSignatures package (Fantini *et al.* 2020) in R. Seven nucleotide contexts were analyzed with the web implementation of pLogo (O'Shea *et al.* 2013) (<https://plogo.uconn.edu>), using S288C_reference_sequence_R64-3-1_20210421.fsa downloaded from SGD (<https://www.yeastgenome.org>) as the reference genome. Three nucleotide contexts were used to define the 96 trinucleotide mutation signatures for comparison to the COSMIC mutation signatures. Trinucleotide mutation signatures were corrected for trinucleotide frequency differences between the yeast and human genomes. The 96 trinucleotide signatures were plotted using the MutSignatures package (Fantini *et al.* 2020), and compared to the single-base substitution (SBS) signatures detailed in [table S21](#) from Degasperis *et al.* (2022), using the MutationalPatterns package (Blokzijl *et al.* 2018).

Mutation clustering

Clustering of deamination mutations was evaluated by calculating the distance between each pair of SNVs for each mutation accumulation line, yielding the intermutation distance for each pair. Each intermutation distance was annotated as strand-coordinated (both SNVs either C->T or G->A) or not. A previously established cluster definition (Mas-Ponte and Supek 2020) of intermutation distance < 500 bp was applied, with directly adjacent SNV pairs retained, and the likelihood of strand coordination was evaluated with a χ^2 test. The expected number of deamination clusters was determined by randomly broadcasting the SNVs for each chromosome for each mutation accumulation line across the relevant chromosome, adjusting to the nearest C or G coordinate, before repeating the intermutation distance calculation and

strand coordination evaluations. Randomization of SNVs was repeated 5 times to arrive at a median clustering expectation.

Transcription bias analyses

To evaluate enrichment of APOBEC3 SNVs at different genomic features, potential APOBEC3A/B/C deamination targets were defined as the set of TC and GA dinucleotides in the *sacCer3* reference genome and the coordinates of the C's and G's were extracted for use as the background reference set. The proportion of SNVs in genes and intergenic regions, and enrichment therein, was evaluated using the gene coordinates of the *sacCer3* reference genome. To evaluate enrichment near start codons of genes, distance to nearest ATG and the strand of the nearest ATG was annotated for each SNV. Start codon proximal SNVs (± 500 bp) were collected in 50-bp bins, and the proportion of SNVs on the non-transcribed strand was calculated for each bin. A similar analysis was performed for tRNA genes, using the coordinates of the 275 tRNA genes in the reference genome, extracted using YeastMine (Balakrishnan *et al.* 2012) (<https://yeastmine.yeastgenome.org>; accessed on 1/24/2023). Analysis of proximity to highly expressed genes utilized the gene expression table in Supplementary Dataset S3-4 from Sui *et al.* (2020), from which the highest and lowest 5% of expressed genes were extracted after removing genes whose expression was not detected.

Replication bias analyses

Replication timing and strand bias were evaluated by first calculating the distance between each SNV and the nearest early-firing DNA replication origin (Balint *et al.* 2015) or the nearest efficient origin (efficiency > 0.7) (McGuffee *et al.* 2013). As the results were highly similar, calculations from the (Balint *et al.* 2015) dataset were subjected to further analyses. SNVs within ± 7 kb of an early origin were collected in 700-bp bins for visualization, and enrichment for early origin proximity was evaluated by χ^2 tests. Replication strand bias was evaluated using early-firing origin annotations (Balint *et al.* 2015) without normalization of inter-origin distances, as the certainty of replication strand (leading or lagging) diminishes with the actual distance from an efficient early origin. Distances-to-origin were collected in 5-kb bins for 50 kb flanking each early origin and the proportion of C->T SNVs was calculated for each bin. C->T SNVs to the right of an origin and G->A SNVs to the left of an origin are defined as lagging strand template deaminations. As for the transcription bias analyses, APOBEC3 targets were defined as TC or GA dinucleotides.

Computational analyses and software

Statistical analysis, data manipulation, and data visualization were performed using the R Statistical Software (R Core Team 2021) in RStudio. Figures were assembled using Adobe Illustrator v28.1 using the Lost in Translation color palette (<https://www.instagram.com/filmandcolor/>).

Results

A humanized yeast platform to study APOBEC3C-mediated mutagenesis

To rigorously define the preferred sites of deamination by APOBEC3C, I took advantage of the approach used successfully for APOBEC3G (Schumacher *et al.* 2005), APOBEC3A, and APOBEC3B (Hoopes *et al.* 2016), and expressed APOBEC3C constitutively in the budding yeast *S. cerevisiae*. A yeast codon-optimized APOBEC3C open reading frame was flanked with a strong constitutive artificial promoter (Kotopka and Smolke 2020) and a synthetic terminator

(Curran et al. 2015) (Fig. 1a). The APOBEC3C transcription unit was then integrated in single copy in the *LYP1* gene to provide optimal stability. In addition to the wild-type APOBEC3C gene, a catalytic-dead variant encoding C97S and C100S (Horn et al. 2014) and a naturally occurring hyperactive allele encoding S188I (Duggal et al. 2013; Wittkopp et al. 2016) were also expressed.

Expression of the 3 APOBEC3C proteins was confirmed by immunoblot analysis (Fig. 1b). Wild-type and catalytic-dead proteins were expressed at similar levels, while the hyperactive allele was expressed at lower levels. To test whether APOBEC3C expressed in yeast was mutagenic, fluctuation analyses were performed to measure the forward mutation rate at the *CAN1* locus (Fig. 1c). As anticipated, APOBEC3C expression increased the mutation rate, from 2×10^{-7} to 26×10^{-7} . Expression of the hyperactive S188I variant further increased mutation rate to 52×10^{-7} . Expression of the catalytic-dead C97S/C100S mutant did not increase the mutation rate (1.8×10^{-7}). To test whether the mutagenesis by APOBEC3C was the result of cytidine deamination a deletion of the *UNG1* gene, which encodes the yeast uracil DNA glycosylase, was introduced into the APOBEC3C expression strain. Deletion of *UNG1* resulted in a large increase in mutation rate, to 213×10^{-7} , as would be expected when removal of uracils by base excision repair is prevented following cytidine deamination. As a final measure of APOBEC3C function in vivo in the humanized yeast system, a deletion of the *RAD5* gene was introduced into the APOBEC3C expression strains and measured cell fitness (Fig. 1, d–f). Deletion of *RAD5* results in accumulation of ssDNA in the nucleus (Gallo et al. 2019), and results in fitness defects when combined with expression of APOBEC3A or APOBEC3B (Hoopes et al. 2016). Although APOBEC3C expression in *rad5Δ* resulted in only a modest fitness defect (Fig. 1e), expression of the hyperactive S188I variant resulted in a fitness defect that was much greater than expectation (Fig. 1d; multiplicative model, comparing to *rad5Δ* and APOBEC3C expression alone). I conclude that APOBEC3C expressed in yeast is deaminating nuclear DNA, likely ssDNA, to produce mutations.

Mutation accumulation during APOBEC3C expression

To analyze the sites deaminated by APOBEC3C in vivo, I established 18 clonal lineages originating from the same *ung1Δ* APOBEC3C expression strain, propagated them independently for 375–500 generations to allow mutations to accumulate, and imposed a single cell bottle-neck every ~25 generations to minimize the effects of genetic drift (Zhu et al. 2014). Since the mutation rate of *ung1Δ* was 44-fold less than *ung1Δ* APOBEC3C, most mutations should be the result of APOBEC3C catalyzed deamination events. Genomic DNA from each strain was deep sequenced, and a total of 6,771 unique variants were identified (Fig. 2a). After removing variants that mapped to repetitive regions of the genome, there were 6,692 variants, the vast majority of which (6,685) were SNVs resulting from deamination (i.e. C to T transitions). The average number of deamination SNVs per mutation accumulation line was 370 (Fig. 2b). The mutation accumulation data reported for APOBEC3A and APOBEC3B expressed in *ung1Δ* yeast (Hoopes et al. 2016) were also analyzed to facilitate comparison with the APOBEC3C data, identifying 1,421 and 1,643 deamination SNVs, respectively (Fig. 2b). The average mutation rate across the 18 APOBEC3C mutation accumulation lines was 0.73×10^{-7} per base per generation (Fig. 2c), about 400-fold the wild-type rate (Zhu et al. 2014), and about half that observed upon APOBEC3B expression in yeast (Sui et al. 2020).

The APOBEC3C mutation signature differs from other APOBEC3 proteins

Using the APOBEC3C mutation accumulation data, along with the data reported by Hoopes et al. (2016) for APOBEC3A and APOBEC3B, I extracted the genome contexts for each unique deamination and identified over- and under-represented bases at positions flanking the deaminated deoxycytidine (Fig. 3a). As is typical of APOBEC3 proteins (with the exception of APOBEC3G) APOBEC3C showed a strong preference for T at the –1 position. Both A and G were under-represented at –1, while C was neutral. At –2, the preference for T indicated by Taylor et al. (2013) is apparent, as is a very strong under-representation of C and G. The motif derived for APOBEC3C extends 3 nucleotides upstream and downstream of the target C, with under-representation of G and C apparent at –3, +1, +2, and +3. Examination of the motif for APOBEC3C revealed clear differences from APOBEC3A and APOBEC3B (Fig. 3a). The motifs derived for APOBEC3A and APOBEC3B were similar to the TCW trinucleotide motif reported in the original analysis of the datasets (Hoopes et al. 2016), and were similar to the extended 5 nucleotide motifs reported by Chan et al. (2015) and to the preferences at –2 reported by Taylor et al. (2013). The APOBEC3C motif differs most at the –2 position, with an over-representation of T that is not observed for APOBEC3B and that is stronger than APOBEC3A. APOBEC3C also shows an under-representation of C at –2 that is not apparent for either APOBEC3A or APOBEC3B. Additionally, C is neutral at –1 in the APOBEC3C motif, whereas it is under-represented in the APOBEC3A and APOBEC3B motifs. The differences in deamination motifs between APOBEC3C, APOBEC3A, and APOBEC3B are also evident in the conventional triplet motif representation of their mutation signatures (Fig. 3b and Supplementary Fig. 1), where the high fractions of TCA and TCT variants seen for APOBEC3A and APOBEC3B are substantially diminished for APOBEC3C. The CCA, CCT, and CCC variants each approach 10% for APOBEC3C (9.4%, 7.7%, and 9.8%), while making up fewer than 3% of APOBEC3A variants, and fewer than 4% of APOBEC3B variants.

Having defined the mutational signature for APOBEC3C, I compared the trinucleotide frequencies in the APOBEC3C signature to the COSMIC (Alexandrov et al. 2020) and Signal (Degasperis et al. 2022) SBS mutational signatures (Fig. 3c). As anticipated, APOBEC3A and APOBEC3B both had the highest similarity to SBS2 (cosine similarity of 0.99 and 0.95), a signature attributed to APOBEC3A and APOBEC3B activities (Nik-Zainal et al. 2012; Chan et al. 2015; Petljak et al. 2022). By contrast, the APOBEC3C mutational signature was less similar to SBS2 (cosine similarity = 0.78) and is most similar to SBS100 (cosine similarity = 0.90), a signature of unknown etiology. APOBEC3C had a higher similarity to SBS11, SBS30, and SBS101 than either APOBEC3A or APOBEC3B. Similarities to the APOBEC3 signature SBS13 were absent as all mutation accumulation strains were deficient in *UNG1*, which is important for APOBEC3 to generate SBS13 (Petljak et al. 2022). Together, these data indicate that APOBEC3C generates a distinct mutational profile.

APOBEC3C mutations are found in clusters

APOBEC mutations in cancer genomes are associated with interesting patterns, where regional clustering of SNVs is observed, often with mutations occurring on the same DNA strand (Nik-Zainal et al. 2012). Two clustering patterns have been described: kataegis, where local strand-coordinated hypermutation events of tens of mutations occur over a span of kilobases (Nik-Zainal et al. 2012;

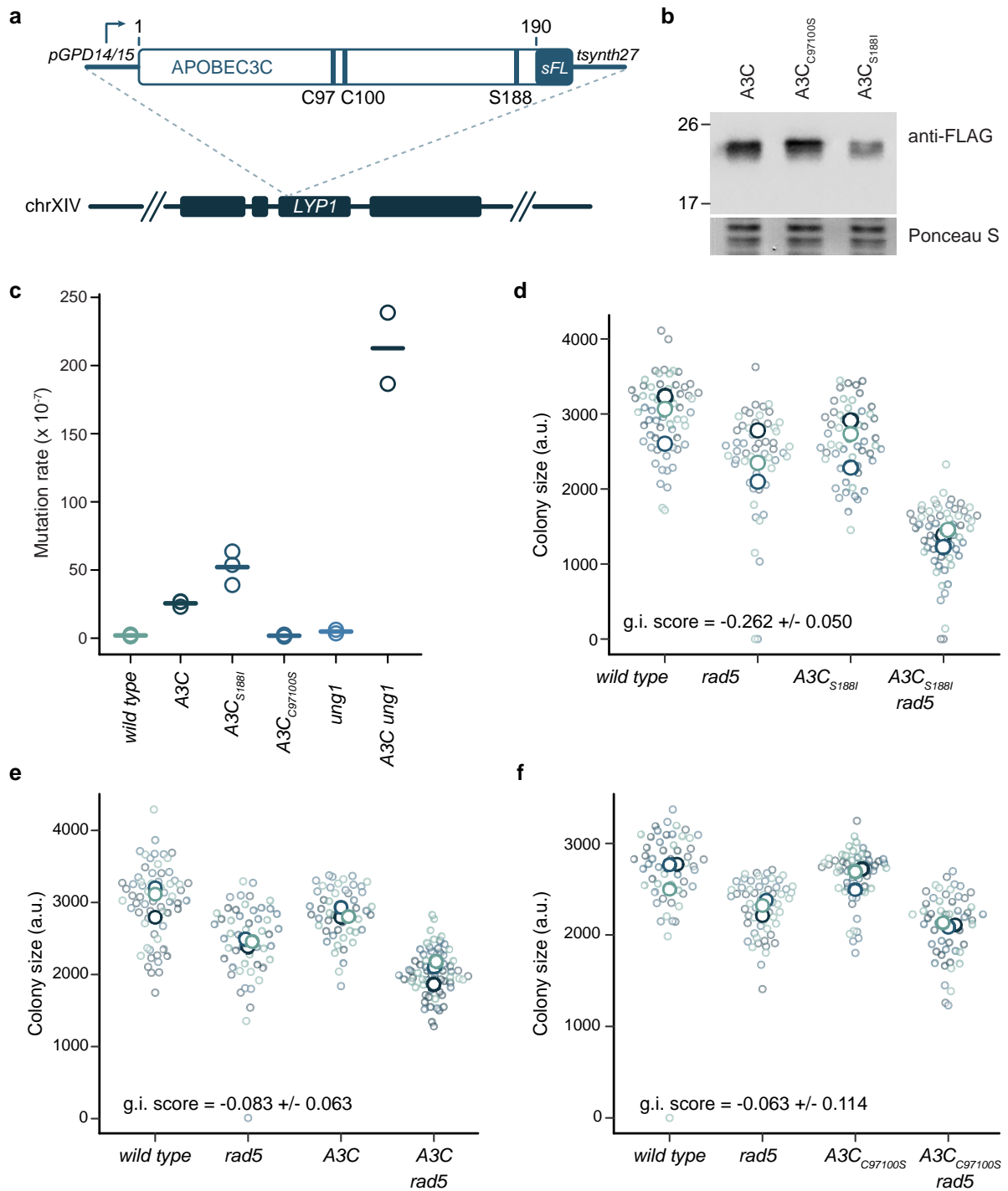


Fig. 1. Expression of human APOBEC3C in yeast is mutagenic. **a**) Schematic diagram of the APOBEC3C expression construct. The promoter (pGPD14/15) and terminator (tsynth27) are indicated, as are the positions of the superFLAG (sFL) epitope tag, the catalytic-dead mutation (C97/C100), and the activating mutation (S188). The integration site in the LYP1 gene is shown below. **b**) Immunoblot analysis of extracts of yeast strains expressing wild-type (A3C), catalytic-dead (A3C_{C97100S}), or hyperactive (A3C_{S188I}) alleles of APOBEC3C. The blot was probed with anti-FLAG antibody to detect the A3C proteins. The positions of molecular weight standards are indicated in kDa. An image of a region of the blot stained for total protein (Ponceau S) is shown as a loading control. **c**) Mutation rates at the CAN1 locus were measured for yeast strains expressing the indicated APOBEC3C alleles. Where indicated (*ung1Δ*), the UNG1 gene was deleted. P-values from 2-sided Student's t-tests are listed in [Supplementary Table 3](#), and were <0.05 for A3C and A3C_{S188I} vs wild type, and for A3C *ung1Δ* vs *ung1Δ*. *n* = 3 for UNG1 strains and *n* = 2 for *ung1Δ* strains. **d**) Colony size in arbitrary units (a.u.) was measured for the indicated strains following dissection of tetrads from *rad5Δ* × APOBEC3C_{S188I} crosses. Small circles show individual colony sizes (18–22 per genotype per replicate), and the large circles show the mean colony size for each of the 3 replicates. The genetic interaction (g.i.) score is indicated. **e**) Colony size in arbitrary units (a.u.) was measured for the indicated strains following dissection of tetrads from *rad5Δ* × APOBEC3C crosses. Small circles show individual colony sizes (13–27 per genotype per replicate) and the large circles show the mean colony size for each of the 3 replicates. The genetic interaction (g.i.) score is indicated. **f**) Colony size in arbitrary units (a.u.) was measured for the indicated strains following dissection of tetrads from *rad5Δ* × APOBEC3C_{C97100S} crosses. Small circles show individual colony sizes (15–23 per genotype per replicate) and the large circles show the mean colony size for each of the 3 replicates. The genetic interaction (g.i.) score is indicated.

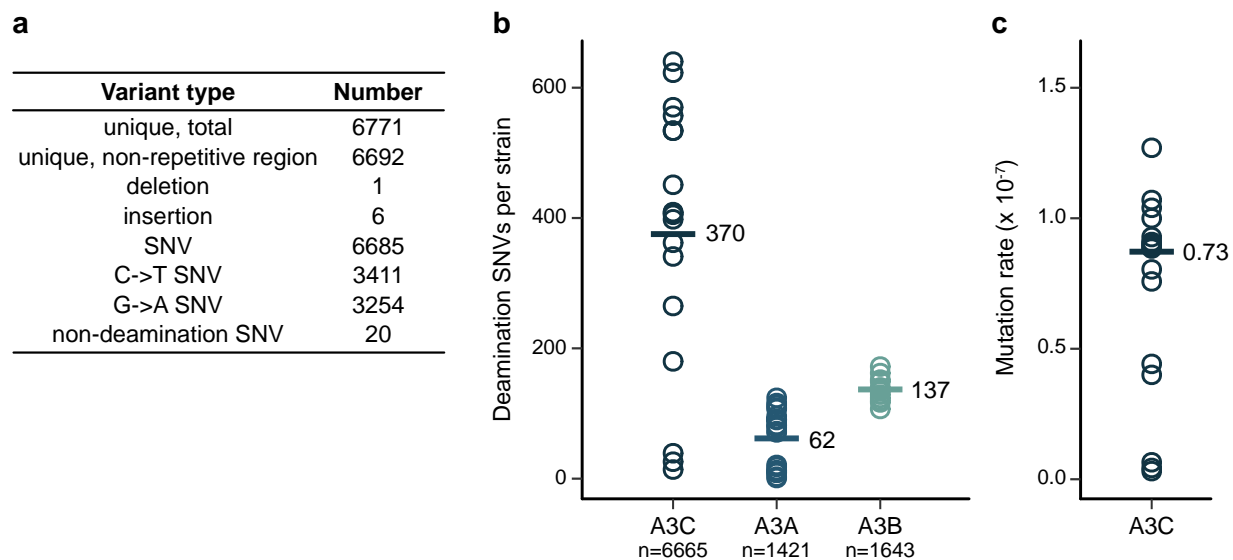


Fig. 2. Mutation accumulation in *ung1Δ* yeast cells expressing APOBEC3C. a) Summary of the variant types in the 18 mutation accumulation lines. The number of single-nucleotide variants (SNVs) and the type of SNV (C->T or G->A) are indicated. b) The number of SNVs due to deamination is plotted for each mutation accumulation line for *ung1Δ* yeast expressing APOBEC3C, 3A, or 3B. Horizontal bars indicate the means, and the total number of deamination SNVs for each mutation accumulation experiment is indicated. APOBEC3A and 3B data are from Hoopes et al. (2016). c) The mutation rate is plotted for each of the 18 APOBEC3C mutation accumulation lines. The horizontal bar indicates the mean.

Taylor et al. 2013), and omikli, where mutation clusters occur in pairs or triplets with a shorter intermutation distance (Mas-Ponte and Supek 2020). I measured the distances between APOBEC3C SNVs for each mutation accumulation line to assess the extent of clustering (Fig. 4a). Kataegis is expected to be rare in the APOBEC3C dataset as yeast models indicate that kataegis requires the action of uracil DNA glycosylase (Taylor et al. 2013), although kataegis in human cells does not require uracil DNA glycosylase (Petljak et al. 2022). The APOBEC3C yeast strains used here carry complete deletion of the uracil DNA glycosylase gene *UNG1*. Indeed, there were no mutation clusters that met a reasonable definition of kataegis (≥ 5 consecutive mutations with intermutation distances of ≤ 1 kb (Alexandrov et al. 2013; Petljak et al. 2019; Mas-Ponte and Supek 2020)). For APOBEC3C, 7.3% of SNV pairs had intermutation distances ≤ 500 bp (expectation = 1.9%), and therefore were defined as clustered (Supek and Lehner 2017; Mas-Ponte and Supek 2020). APOBEC3A and APOBEC3B caused clustered mutations at lower frequencies (1.1 and 1.3%, respectively, Fig. 4, b and c) that were still greater than expectation (0.3 and 0.7%, respectively). For APOBEC3C and APOBEC3B clustered mutations were more likely to occur on the same DNA strand than were non-clustered mutations (χ^2 test, $P = 1 \times 10^{-10}$ and $P = 0.009$, respectively). Strand coordination was particularly likely for APOBEC3C SNV pairs with the shortest intermutation distances (Fig. 4d). All strand-coordinated mutation clusters caused by APOBEC3C, of which there were 308, consisted of pairs or triplets, reminiscent of omikli (Mas-Ponte and Supek 2020). APOBEC3A clustered mutations were not more likely to be strand-coordinated than non-clustered mutations (χ^2 test, $P = 0.5$). Absence of strand coordination of APOBEC3A clustered mutations is consistent with the lack of processivity of APOBEC3A in vitro (Love et al. 2012). The presence of strand coordination for APOBEC3B and APOBEC3C is consistent with the processive nature of both deaminases, where deamination of more than 1 cytosine occurs in a single enzyme-substrate encounter (Adolph et al. 2017a, 2017b).

APOBEC3C targets both strands of RNA polymerase II and tRNA genes

Transcription can expose ssDNA substrates for ectopic action by cytidine deaminases in humanized yeast models (Taylor et al. 2014; Lada et al. 2015; Hoopes et al. 2016). There is some controversy as to whether exposure of ssDNA by transcription is also important for APOBEC3 action in its native context, as assessed by cancer genome sequencing (Kazanov et al. 2015; Mas-Ponte and Supek 2020; Chervova et al. 2021). Given the compact nature of the yeast genome, intergenic regions tend to comprise promoter elements and transcription start sites (TSS), with the latter being in close proximity to the start codon of the open reading frame (median ~40 bp (Lu and Lin 2019; Ronsmans et al. 2019)). I quantified deaminations across the start codon of open reading frames (± 500 bp) for APOBEC3C, APOBEC3A, and APOBEC3B (Fig. 5a). Both APOBEC3C and APOBEC3A showed deamination peaks in the TSS (-100 to -50 bp) and in the 5' region of the open reading frame ($+50$ to $+100$ bp). By contrast, APOBEC3B deaminations were not enriched in regions proximal to start codons (Fig. 5, a and b). I examined the strand biases of APOBEC3C and APOBEC3A deaminations that were proximal to start codons (Fig. 5c) and observed an interesting difference. APOBEC3A deaminations were biased to the nontranscribed strand, as previously noted (Hoopes et al. 2016). By contrast, APOBEC3C deaminations were biased to the transcribed strand, particularly within the first ~100 bp of the open reading frame (Fig. 5c).

APOBEC3B deaminations are enriched proximal to tRNA genes on the nontranscribed strand (Saini et al. 2017; Sui et al. 2020). I examined the propensity for APOBEC3C, APOBEC3A, and APOBEC3B to catalyze deaminations near tRNA genes (Fig. 5d). All 3 deaminases caused peaks of deamination in the first 75 bp of tRNA genes, as well as enrichments for tRNA gene-proximal deamination (Fig. 5e). APOBEC3B displayed the previously described bias for the nontranscribed strand (Saini et al. 2017; Sui et al. 2020), as did APOBEC3A (Fig. 5f). By contrast, APOBEC3C deaminations occurred in identical numbers on the transcribed and

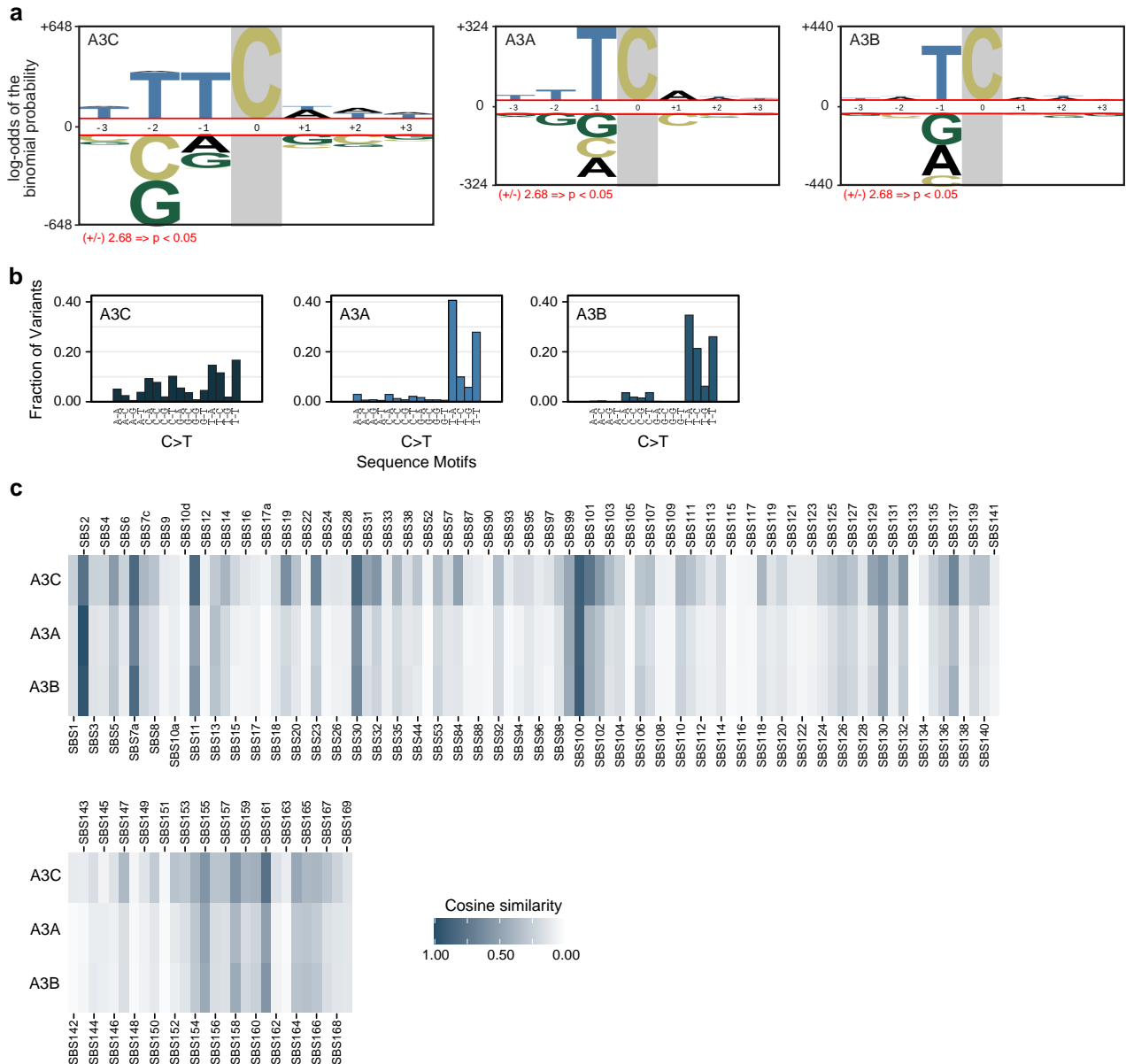


Fig. 3. The mutation signature of APOBEC3C is distinct from that of APOBEC3A or 3B. **a**) A probability logo visualization of the APOBEC3C, 3A, and 3B deamination sequence motifs. Nucleotides are scaled relative to their log-odds binomial probabilities, with over-represented nucleotides above the x-axis. Horizontal lines indicate the Bonferroni-corrected support threshold of $P < 0.05$, which is ± 2.68 in all cases. **b**) The trinucleotide mutational signatures are plotted for APOBEC3C, 3A, and 3B. Mutation types are grouped with only the 16 SNV types centered on C to T mutations shown. **c**) Heatmap showing the cosine similarity of the mutational signatures in (b) with the COSMIC and Signal single-base substitution (SBS) mutational signatures.

nontranscribed strands, resulting in a small bias for the transcribed strand relative to that of APOBEC3 targets. I infer that APOBEC3C lacks the strong bias for the nontranscribed strand that is seen for APOBEC3A and APOBEC3B.

Since tRNA genes are highly expressed (Warner 1999), I asked whether highly expressed RNA pol II genes were particularly susceptible to deamination by APOBEC3C (Fig. 5, g and i). APOBEC3C deaminations showed a clear peak at the TSS of highly expressed pol II genes (the highest 5% of genes (Sui et al. 2020); Fig. 5g), as did APOBEC3A deaminations (Fig. 5h). Deaminations by APOBEC3C and APOBEC3A were enriched by 3.9- and 3.5-fold near highly expressed genes (Fig. 5i), whereas APOBEC3B deaminations showed a more modest 1.5-fold enrichment (Fig. 5i and Supplementary Fig. 2a). None of the APOBEC3 deaminations were enriched near the lowest expressed genes (Supplementary Fig. 2b).

APOBEC3C lacks a DNA replication strand bias

APOBEC3 mutagenesis has been associated with early-replicating regions in human cancer genomes (Kazanov et al. 2015; Chervova et al. 2021). However, the association with replication timing might differ for clustered vs scattered mutations (Septyarskiy et al. 2016) and is not evident in all analyses (Morganella et al. 2016; Buisson et al. 2019). Given the extensive replication origin annotations for the yeast genome (Siow et al. 2012), as well as precise origin timing and efficiency datasets (Siow et al. 2012; McGuffee et al. 2013), humanized yeast systems are ideal for analyzing relationships between DNA replication and APOBEC3 mutagenesis. I first asked if APOBEC3 deaminations were enriched in regions proximal to early-firing DNA replication origins (Fig. 6, a–c, and Supplementary Fig. 3a), since the regions closest to early-firing

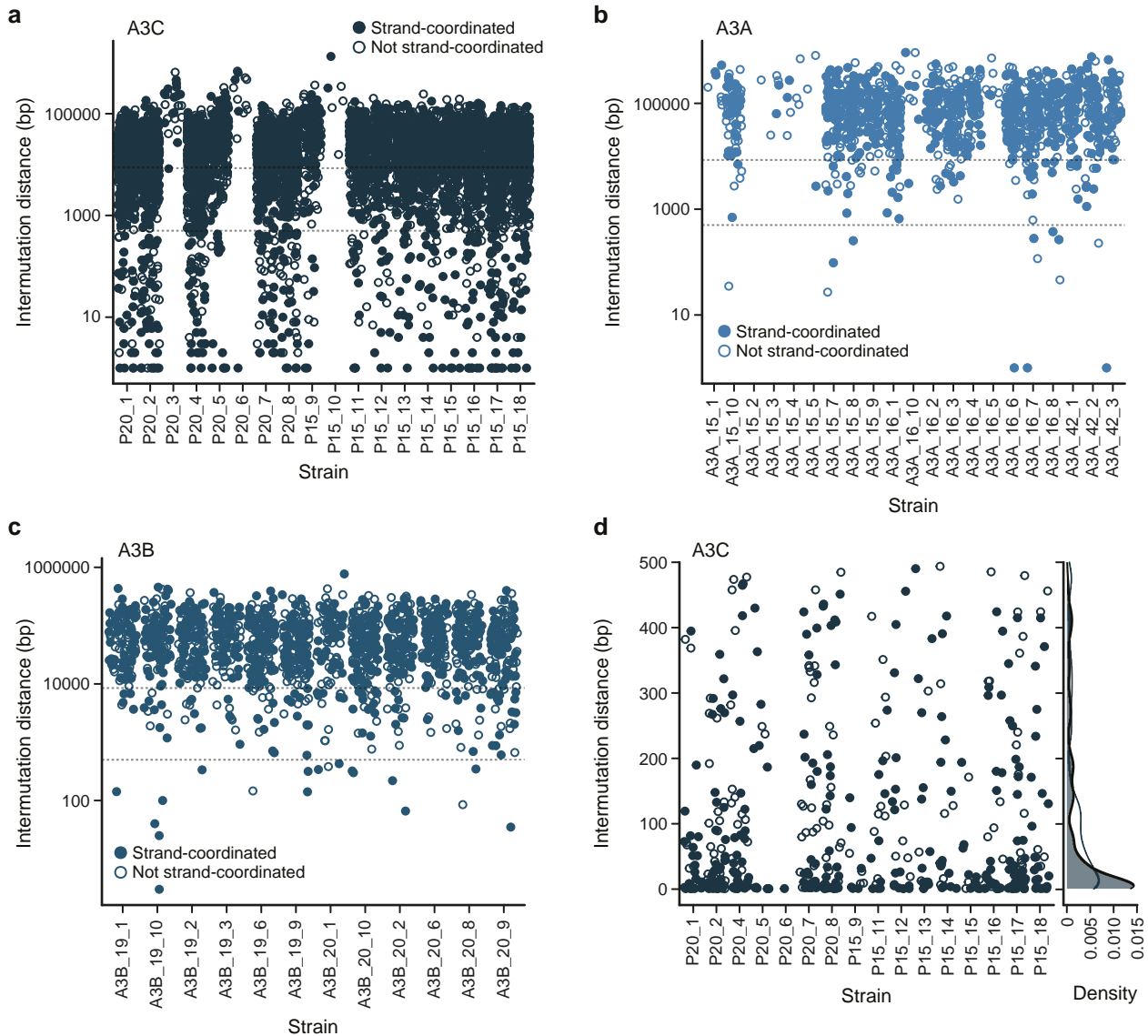


Fig. 4. APOBEC3C makes clustered strand-coordinated mutations. a) The intermutation distance is plotted for each pair of APOBEC3C SNVs for each mutation accumulation line. Mutation pairs on the same DNA strand are indicated by closed circles. The horizontal dotted lines indicate intermutation distances of 8,500 and 500 bp as cutoffs for proximal and clustered mutations, respectively. b) The intermutation distance is plotted for each pair of APOBEC3A SNVs for each mutation accumulation line. Details as in (a). c) The intermutation distance is plotted for each pair of APOBEC3B SNVs for each mutation accumulation line. Details as in (a). d) The intermutation distance is plotted for each pair of APOBEC3C SNVs for each mutation accumulation line, for intermutation distances less than 500 bp. Mutation pairs on the same DNA strand are indicated by closed circles. The density distributions of the strand-coordinated intermutation distances (shaded histogram) and the non-strand-coordinated intermutation distances (open histogram) are plotted on the right.

origins will be the earliest to replicate. Both APOBEC3C and APOBEC3A showed a modest enrichment of 1.2-fold in regions ± 7 kb of an early-firing origin, relative to potential APOBEC3 targets (Fig. 6c). APOBEC3B did not show a statistically supported enrichment proximal to early-firing origins (Fig. 6c).

Data from cancer genomes (Haradhvala et al. 2016; Seplyarskiy et al. 2016) and from humanized yeast systems (Hoopes et al. 2016; Sui et al. 2020) indicate that APOBEC3A and APOBEC3B deaminations are biased toward the lagging strand DNA replication template. I tested whether APOBEC3C showed a similar bias (Fig. 6, d–f and Supplementary Fig. 3b). Genome regions located ± 50 kb from early-firing origins were considered, as the direction of replication fork movement, and therefore the assignment of leading vs

lagging strand, is most certain in proximity to early origins. Since SNVs are annotated relative to the genome reference strand, replication strand bias can be visualized as a change in the proportion of C->T transitions as the direction of fork movement changes from right-to-left to left-to-right at replication origins (Fig. 6d). Lagging strand template bias results in a higher proportion of C->T to the right of origins and a lower proportion to the left. As previously described, both APOBEC3A (Fig. 6e) and APOBEC3B (Supplementary Fig. 3b) display a clear preference to deaminate the lagging strand template, and the bias is stronger closer to early-firing replication origins. Surprisingly, APOBEC3C does not have a substantial bias and deaminates the leading and lagging strand templates approximately equally (Fig. 6f).

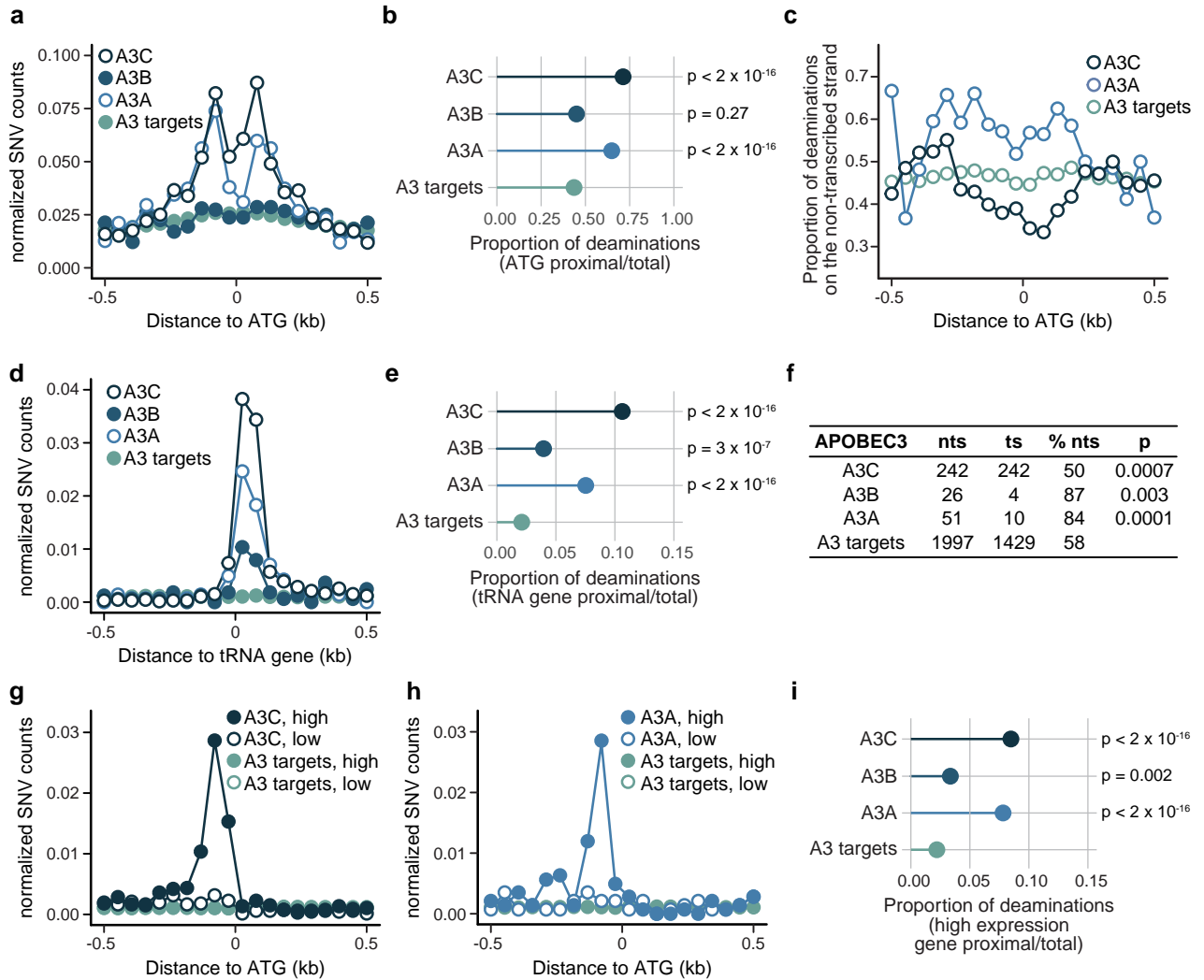


Fig. 5. APOBEC3C targets the 5' ends of pol II and tRNA genes. **a**) The proportion of SNVs in 50-bp bins surrounding pol II gene start codons is plotted for each APOBEC3. Putative APOBEC3 targets (TC) are plotted as a control. **b**) The proportion of deaminations ± 500 bp of pol II gene start codons relative to total deaminations is plotted for APOBEC3C, 3B, and 3A. Putative APOBEC3 targets (TC) are plotted as a control. Statistical support for enrichments proximal to ATGs relative to APOBEC3 targets was assessed with chi-squared tests. **c**) The proportion of deaminations on the nontranscribed strand in 50-bp bins surrounding pol II gene start codons is plotted for each APOBEC3. Putative APOBEC3 targets (TC) are plotted as a control. **d**) The proportion of SNVs in 50-bp bins surrounding tRNA genes is plotted for each APOBEC3. Putative APOBEC3 targets (TC) are plotted as a control. **e**) The proportion of deaminations ± 500 bp of tRNA genes relative to total deaminations is plotted for APOBEC3C, 3B, and 3A. Putative APOBEC3 targets (TC) are plotted as a control. Statistical support for enrichments proximal to tRNA genes relative to APOBEC3 targets was assessed with chi-squared tests. **f**) Deaminations by APOBEC3C, 3A, and 3B on the nontranscribed strand (nts) and the transcribed strand (ts) ± 500 bp of tRNA genes are tabulated, as is the percent of deaminations on the nts (% nts). Statistical support for enrichments of deaminations on the nts relative to APOBEC3 targets was assessed with chi-squared tests. **g**) The proportion of SNVs in 50-bp bins surrounding the highest 5% (closed circles) and the lowest 5% of expressed genes is plotted for APOBEC3C. Putative APOBEC3 targets (TC) are plotted as a control. **h**) The proportion of SNVs in 50-bp bins surrounding the highest 5% (closed circles) and the lowest 5% of expressed genes is plotted for APOBEC3A. Putative APOBEC3 targets (TC) are plotted as a control. **i**) The proportion of deaminations ± 500 bp of the most highly expressed genes relative to total deaminations is plotted for APOBEC3C, 3B, and 3A. Putative APOBEC3 targets (TC) are plotted as a control. Statistical support for enrichments proximal to highly expressed genes relative to APOBEC3 targets was assessed with chi-squared tests.

Discussion

Through analysis of APOBEC3C deamination mutations in a yeast model, several interesting patterns of APOBEC3C activity have emerged that clearly distinguish APOBEC3C from APOBEC3A and APOBEC3B. The APOBEC3C mutation signature is distinct, with an over-representation of T at -2 and decreased bias against C at -1 , relative to the APOBEC3A or APOBEC3B signatures. APOBEC3C produces strand-coordinated mutation clusters that are consistent with processive deamination. The absence of replication and transcription strand biases seen with other APOBEC3s further distinguishes the properties of APOBEC3C.

The unique mutation signature of A3C has some interesting implications. Explorations of the role of APOBEC3 proteins in mutating cancer genomes have largely focused on APOBEC3A and 3B (Alexandrov et al. 2013; Burns et al. 2013a, 2013b; Taylor et al. 2013; Chan et al. 2015; Petljak et al. 2021, 2022; Caswell et al. 2023; Isozaki et al. 2023), with some attention paid to APOBEC3H haplotype I (Starrett et al. 2016), APOBEC3G (Liu et al. 2023), and APOBEC3C (Qian et al. 2022). Indeed, recent analyses indicate that APOBEC3A is the predominant mutator of cancer genomes, with a lesser role for APOBEC3B (Petljak et al. 2022), although the relative contributions likely vary in different contexts (Carpenter et al. 2023). Removing both APOBEC3A and APOBEC3B does not eliminate mutagenesis

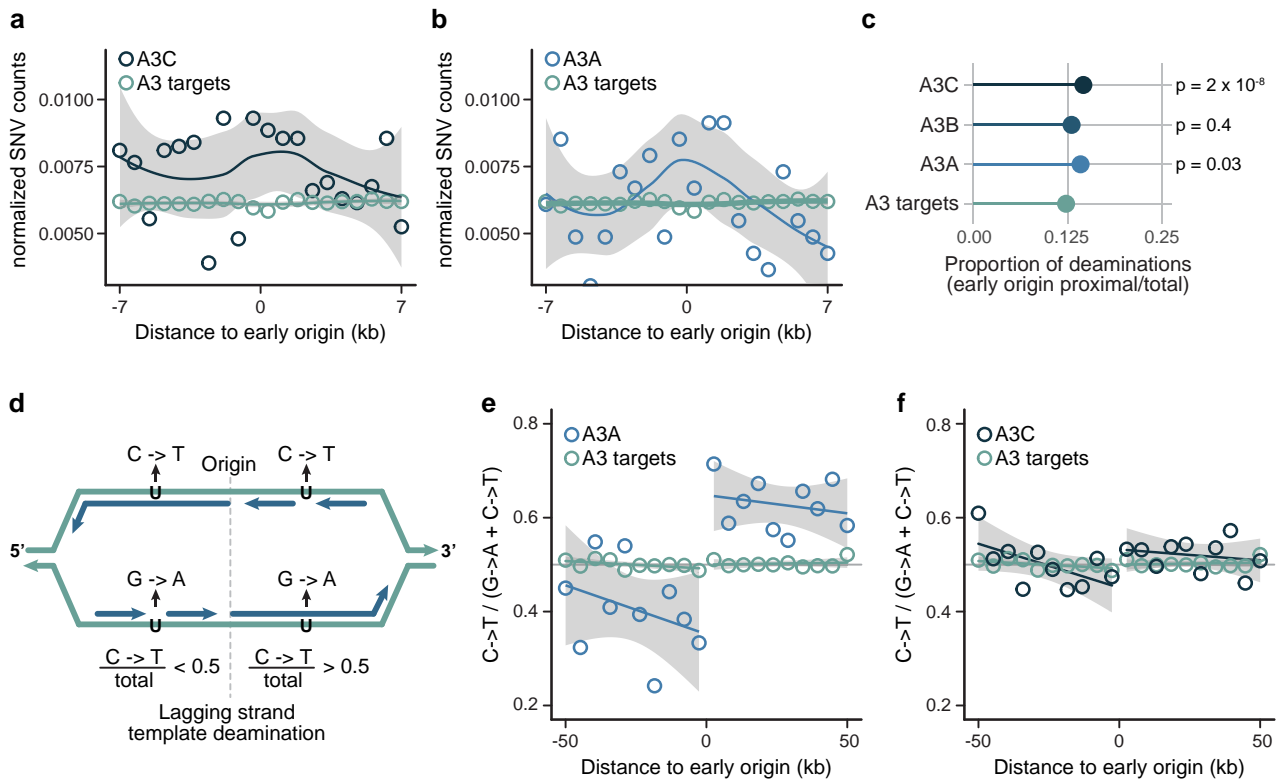


Fig. 6. APOBEC3C lacks a DNA replication strand preference. a) The proportion of SNVs in 70-bp bins surrounding early-firing DNA replication origins is plotted for APOBEC3C. Putative APOBEC3 targets (TC) are plotted as a control. The solid line is the loess smoothed conditional mean with the shaded regions showing the standard error with a 0.95 confidence interval. b) The proportion of SNVs in 70-bp bins surrounding early-firing DNA replication origins is plotted for APOBEC3A. Details as in (a). c) The proportion of deaminations in regions proximal to early-firing origins is plotted for APOBEC3C, 3B, and 3A. Putative APOBEC3 targets (TC) are plotted as a control. Statistical support for enrichments in origin-proximal regions relative to APOBEC3 targets was assessed with chi-squared tests. d) Schematic diagram depicting the change in replication strand bias on opposite sides of a replication origin. To the left of the origin, a deamination on the top strand results in a C->T transition in the genome sequencing data, whereas a deamination on the bottom strand results in a G->A transition since the genome sequences are mapped onto the top strand. To the left of the origin, the bottom strand is the lagging strand template and so a preference for lagging strand template deamination manifests as a reduced proportion of C->T transitions. To the right of the origin, the opposite is true because now the top strand is the lagging strand template. e) The proportion of C->T SNVs in 5-kbp bins surrounding early-firing DNA replication origins is plotted for APOBEC3A. Putative APOBEC3 targets (TC) are plotted as a control. The solid line is the loess smoothed conditional mean with the shaded regions showing the standard error with a 0.95 confidence interval. f) The proportion of C->T SNVs in 5-kbp bins surrounding early-firing DNA replication origins is plotted for APOBEC3C. Details as in (e).

(Petljak et al. 2022), indicating that additional APOBEC3 proteins could also cause mutations in tumor cells. APOBEC3C mRNA is more highly expressed in tumor samples, on average, than either APOBEC3A or APOBEC3B (Aaltonen et al. 2020; Goldman et al. 2020), and so could also play a role in cancer genome mutagenesis. Interestingly, APOBEC3C is widely expressed in normal tissues whereas APOBEC3A and APOBEC3B are expressed at lower basal levels, if at all (THE GTEX CONSORTIUM 2020). Therefore, APOBEC3C might also impact genome mutagenesis outside of the cancer context.

A common feature of cancer genome analyses with respect to APOBEC3 is the use of the TCN or TCW context, which manifests as mutation signatures SBS2 and SBS13, to identify mutations with a high likelihood of being caused by APOBEC3 activity (Green and Weitzman 2019; Petljak and Maciejowski 2020). However, while the TCN context captures 86% of APOBEC3A and 91% of APOBEC3B mutations analyzed here, TCN captures only 50% of the APOBEC3C mutations. Thus, it is possible that the mutagenic influence of APOBEC3C has been underestimated in current cancer genome analyses. It will be of great interest to mine cancer genome data for evidence of APOBEC3C mutation, using the mutation signature derived in the humanized yeast model.

The features that dictate mutagenesis by APOBEC3 deaminases are subjects of great interest that has focused on two

realms. The first includes explorations of contexts that could generate or reveal ssDNA substrates for APOBEC3 mutagenesis and the second examines intrinsic features of the APOBEC3 enzymes. One context that is likely to be critically important in revealing ssDNA substrates for APOBEC3 deaminases is DNA replication stress. The biases to early-replicating regions (Fig. 6, a-c), tRNA genes (Fig. 5, d and e), and highly expressed pol II genes (Fig. 5, g and h) that are evident in the humanized yeast models are typically of a modest effect size, consistent with the presence of ssDNA being at native levels in these contexts. Introducing mutations known to produce abnormal levels of ssDNA, such as deleting RAD5 (Gallo et al. 2019), increases mutagenesis by APOBEC3 proteins (Fig. 1d and Hoopes et al. 2016), as do mutations that cause DNA replication stress (Mertz et al. 2023). Similarly, DNA replication stress induced in human cells by gemcitabine results in a dramatic increase in cytidine deamination (Ubhi et al. 2024). The biases toward early-replicating regions and the lagging strand that are evident in cancer genomes (Green et al. 2016; Haradhvala et al. 2016; Kanu et al. 2016; Morganella et al. 2016; Seplyarskiy et al. 2016; Nikkilä et al. 2017) almost certainly have DNA replication stress as a contributor, and it seems likely that the specific contexts driving replication stress will be important determinants of APOBEC3 mutation patterns.

Perhaps the most surprising finding yielded by analysis of APOBEC3C mutations is the absence of DNA strand biases that are evident in mutations caused by APOBEC3A and APOBEC3B. Although the importance of transcription bubbles as APOBEC3 targets in human cells remains unresolved (Nordentoft *et al.* 2014; Kazanov *et al.* 2015; Chervova *et al.* 2021; McCann *et al.* 2023), it is clear that in the humanized yeast models for APOBEC3A, APOBEC3B, and APOBEC3C there are different biases to the 5' regions of tRNA and pol II genes. APOBEC3A and 3C show similar biases to tRNA and pol II genes, whereas APOBEC3B lacks a bias to pol II genes and causes a much smaller proportion of deaminations at tRNA genes (Fig. 5, c and f). Both APOBEC3A and APOBEC3B prefer to deaminate the nontranscribed strand, yet APOBEC3C shows no bias toward the nontranscribed strand (Fig. 5, d and g). For DNA replication strands, APOBEC3A and APOBEC3B show asymmetry with preferences for the lagging strand template which APOBEC3C lacks (Fig. 6, e and f). In the humanized yeast models, ssDNA at DNA replication forks or transcription bubbles is expected to be present in an identical fashion and any human cell specific modulators of APOBEC3 function are expected to be absent. Thus, the data indicate that the simple presence of a ssDNA target is not sufficient for deamination by APOBEC3 and suggest that intrinsic features of each APOBEC3 enzyme will continue to be important keys to understanding the specificities of APOBEC3 enzymes in genome mutagenesis, as is the case for the DNA sequence and structural properties that drive the selectivity differences between APOBEC3A and APOBEC3B (Buisson *et al.* 2019; Sanchez *et al.* 2024).

Data availability

Raw sequencing reads from the APOBEC3C mutation accumulation experiment are available from the NCBI Sequencing Read Archive, BioProject ID PRJNA938955. The author affirms that all data necessary for confirming the conclusions of the article are present within the article, figures, tables, and supplementary materials.

Supplemental material available at GENETICS online.

Acknowledgments

I thank Elçin Ünal for the pUB1306 plasmid, Damiano Fantini for assistance with the MutSignatures package, Tajinder Ubhi, Brandon Ho, Raphael Loll-Kripplleber, and Atina Coté for coding advice, Linda Chelico and Kin Chan for helpful discussions, Jiongwen Ou for performing the immunoblots, the Donnelly Sequencing Centre at the University of Toronto for performing Illumina sequencing, and Tajinder Ubhi and members of the Brown lab for careful reading of the manuscript and helpful discussions. I am grateful to work on the lands of the Mississaugas of the Credit, the Anishnaabeg, the Haudenosaunee, and the Wendat peoples, land that is now home to many diverse First Nations, Inuit, and Métis peoples.

Funding

This work was supported by the Canadian Institutes of Health Research (FDN-159913 to GWB). GWB holds a Canada Research Chair (Tier 1).

Conflicts of interest

The author(s) declare no conflicts of interest.

Literature cited

- Aaltonen LA, Abascal F, Abeshouse A, Aburatani H, Adams DJ, Agrawal N, Ahn KS, Ahn S-M, Aikata H, Akbani R, *et al.* 2020. Pan-cancer analysis of whole genomes. *Nature*. 578(7793): 82–93. doi:10.1038/s41586-020-1969-6.
- Adolph MB, Ara A, Feng Y, Wittkopp CJ, Emerman M, Fraser JS, Chelico L. 2017a. Cytidine deaminase efficiency of the lentiviral restriction factor APOBEC3C correlates with dimerization. *Nucleic Acids Res.* 45(6):3378–3394. doi:10.1093/nar/gkx066.
- Adolph MB, Love RP, Feng Y, Chelico L. 2017b. Enzyme cycling contributes to efficient induction of genome mutagenesis by the cytidine deaminase APOBEC3B. *Nucleic Acids Res.* 45(20): 11925–11940. doi:10.1093/nar/gkx832.
- Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, *et al.* 2020. The repertoire of mutational signatures in human cancer. *Nature*. 578(7793):94–101. doi:10.1038/s41586-020-1943-3.
- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale A-L, *et al.* 2013. Signatures of mutational processes in human cancer. *Nature*. 500(7463):415–421. doi:10.1038/nature12477.
- Balakrishnan R, Park J, Karra K, Hitz BC, Binkley G, Hong EL, Sullivan J, Micklem G, Cherry JM. 2012. YeastMine—an integrated data warehouse for *Saccharomyces cerevisiae* data as a multipurpose tool-kit. Database (Oxford). 2012:bar062. doi:10.1093/database/bar062.
- Balint A, Kim T, Gallo D, Cussiol JR, Bastos de Oliveira FM, Yimit A, Ou J, Nakato R, Gurevich A, Shirahige K, *et al.* 2015. Assembly of Slx4 signaling complexes behind DNA replication forks. *EMBO J.* 34(16):2182–2197. doi:10.1525/embj.201591190.
- Blokzijl F, Janssen R, van Boxtel R, Cuppen E. 2018. Mutational Patterns: comprehensive genome-wide analysis of mutational processes. *Genome Med.* 10(1):33. doi:10.1186/s13073-018-0539-0.
- Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD. 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*. 14(2):115–132. doi:10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2.
- Buisson R, Langenbucher A, Bowen D, Kwan EE, Benes CH, Zou L, Lawrence MS. 2019. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. *Science*. 364(6447):eaaw2872. doi:10.1126/science.aaw2872.
- Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB, *et al.* 2013a. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature*. 494(7437):366–370. doi:10.1038/nature11881.
- Burns MB, Temiz NA, Harris RS. 2013b. Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat Genet.* 45(9):977–983. doi:10.1038/ng.2701.
- Carpenter MA, Temiz NA, Ibrahim MA, Jarvis MC, Brown MR, Argyris PP, Brown WL, Starrett GJ, Yee D, Harris RS. 2023. Mutational impact of APOBEC3A and APOBEC3B in a human cell line and comparisons to breast cancer. *PLoS Genet.* 19(11):e1011043. doi:10.1371/journal.pgen.1011043.
- Caswell DR, Gui P, Mayekar MK, Law EK, Pich O, Bailey C, Boumelha J, Kerr DL, Blakely CM, Manabe T, *et al.* 2023. The role of APOBEC3B in lung tumor evolution and targeted cancer therapy resistance. *Nat Genet.* 56(1):60–73. doi:10.1038/s41588-023-01592-8.
- Chan K, Roberts SA, Klimczak LJ, Sterling JF, Saini N, Malc EP, Kim J, Kwiatkowski DJ, Fargo DC, Mieczkowski PA, *et al.* 2015. An

- APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. *Nat Genet.* 47(9):1067–1072. doi:[10.1038/ng.3378](https://doi.org/10.1038/ng.3378).
- Chen Z, Eggerman TL, Bocharov AV, Baranova IN, Vishnyakova TG, Patterson AP. 2021. APOBEC3-induced mutation of the hepatitis virus B DNA genome occurs during its viral RNA reverse transcription into (–)-DNA. *J Biol Chem.* 297(2):100889. doi:[10.1016/j.jbc.2021.100889](https://doi.org/10.1016/j.jbc.2021.100889).
- Chervova A, Fatykhov B, Koblov A, Shvarov E, Preobrazhenskaya J, Vinogradov D, Ponomarev GV, Gelfand MS, Kazanov MD. 2021. Analysis of gene expression and mutation data points on contribution of transcription to the mutagenesis by APOBEC enzymes. *NAR Cancer.* 3(3):zcab025. doi:[10.1093/narcan/zcab025](https://doi.org/10.1093/narcan/zcab025).
- Curran KA, Morse NJ, Markham KA, Wagman AM, Gupta A, Alper HS. 2015. Short synthetic terminators for improved heterologous gene expression in yeast. *ACS Synth Biol.* 4(7):824–832. doi:[10.1021/sb5003357](https://doi.org/10.1021/sb5003357).
- Degasperi A, Zou X, Amarante TD, Martinez-Martinez A, Koh GCC, Dias JML, Heskin L, Chmelova L, Rinaldi G, Wang VYW, et al. 2022. Substitution mutational signatures in whole-genome-sequenced cancers in the UK population. *Science.* 376(6591):science.abl9283. doi:[10.1126/science.abl9283](https://doi.org/10.1126/science.abl9283).
- Duggal NK, Fu W, Akey JM, Emerman M. 2013. Identification and antiviral activity of common polymorphisms in the APOBEC3 locus in human populations. *Virology.* 443(2):329–337. doi:[10.1016/j.virol.2013.05.016](https://doi.org/10.1016/j.virol.2013.05.016).
- Dunham MJ, Gartenberg MR, Brown GW. 2015. *Methods in Yeast Genetics and Genomics: A Cold Spring Harbor Laboratory Course Manual*. New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Fantini D, Vidimar V, Yu Y, Condello S, Meeks JJ. 2020. MutSignatures: an R package for extraction and analysis of cancer mutational signatures. *Sci Rep.* 10(1):18217. doi:[10.1038/s41598-020-75062-0](https://doi.org/10.1038/s41598-020-75062-0).
- Gallo D, Kim T, Szakal B, Saayman X, Narula A, Park Y, Branzei D, Zhang Z, Brown GW. 2019. Rad5 recruits error-prone DNA polymerases for mutagenic repair of ssDNA gaps on undamaged templates. *Mol Cell.* 73(5):900–914.e9. doi:[10.1016/j.molcel.2019.01.001](https://doi.org/10.1016/j.molcel.2019.01.001).
- Goldman MJ, Zhang J, Fonseca NA, Cortés-Ciriano I, Xiang Q, Craft B, Piñero-Yáñez E, O'Connor BD, Bazant W, Barrera E, et al. 2020. A user guide for the online exploration and visualization of PCAWG data. *Nat Commun.* 11(1):3400. doi:[10.1038/s41467-020-16785-6](https://doi.org/10.1038/s41467-020-16785-6).
- Green AM, Landry S, Budagyan K, Avgousti DC, Shalhout S, Bhagwat AS, Weitzman MD. 2016. APOBEC3A damages the cellular genome during DNA replication. *Cell Cycle.* 15(7):998–1008. doi:[10.1080/15384101.2016.1152426](https://doi.org/10.1080/15384101.2016.1152426).
- Green AM, Weitzman MD. 2019. The spectrum of APOBEC3 activity: from anti-viral agents to anti-cancer opportunities. *DNA Repair (Amst).* 83:102700. doi:[10.1016/j.dnarep.2019.102700](https://doi.org/10.1016/j.dnarep.2019.102700).
- Hamedirad M, Weisberg S, Chao R, Lian J, Zhao H. 2019. Highly efficient single-pot scarless golden gate assembly. *ACS Synth Biol.* 8(5):1047–1054. doi:[10.1021/acssynbio.8b00480](https://doi.org/10.1021/acssynbio.8b00480).
- Haradhvala NJ, Polak P, Stojanov P, Covington KR, Shinbrot E, et al. 2016. Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. *Cell.* 164(3):538–549. doi:[10.1016/j.cell.2015.12.050](https://doi.org/10.1016/j.cell.2015.12.050).
- Harris RS, Dudley JP. 2015. APOBECs and virus restriction. *Virology.* 479–480:131–145. doi:[10.1016/j.virol.2015.03.012](https://doi.org/10.1016/j.virol.2015.03.012).
- Harris RS, Petersen-Mahrt SK, Neuberger MS. 2002. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell.* 10(5):1247–1253. doi:[10.1016/s1097-2765\(02\)00742-6](https://doi.org/10.1016/s1097-2765(02)00742-6).
- Harvey CJB, Tang M, Schlecht U, Horecka J, Fischer CR, Lin H-C, Li J, Naughton B, Cherry J, Miranda M, et al. 2018. HEx: a heterologous expression platform for the discovery of fungal natural products. *Sci Adv.* 4(4):eaar5459. doi:[10.1126/sciadv.aar5459](https://doi.org/10.1126/sciadv.aar5459).
- Hoopes JI, Cortez LM, Mertz TM, Malc EP, Mieczkowski PA, Roberts SA. 2016. APOBEC3A and APOBEC3B preferentially deaminate the lagging strand template during DNA replication. *Cell Rep.* 14(6):1273–1282. doi:[10.1016/j.celrep.2016.01.021](https://doi.org/10.1016/j.celrep.2016.01.021).
- Horn AV, Klawitter S, Held U, Berger A, Vasudevan AAJ, Bock A, Hofmann H, Hanschmann KM, Trösemeier JH, Flory E, et al. 2014. Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity. *Nucleic Acids Res.* 42(1):396–416. doi:[10.1093/nar/gkt898](https://doi.org/10.1093/nar/gkt898).
- Hou S, Lee JM, Myint W, Matsuo H, Kurt Yilmaz N, Schiffer CA. 2021. Structural basis of substrate specificity in human cytidine deaminase family APOBEC3s. *J Biol Chem.* 297(2):100909. doi:[10.1016/j.jbc.2021.100909](https://doi.org/10.1016/j.jbc.2021.100909).
- Isozaki H, Sakhtemani R, Abbasi A, Nikpour N, Stanzione M, Oh S, Langenbucher A, Monroe S, Su W, Cabanos HF, et al. 2023. Therapy-induced APOBEC3A drives evolution of persistent cancer cells. *Nature.* 620(7973):393–401. doi:[10.1038/s41586-023-06303-1](https://doi.org/10.1038/s41586-023-06303-1).
- Ito F, Fu Y, Kao S-CA, Yang H, Chen XS. 2017. Family-wide comparative analysis of cytidine and methylcytidine deamination by eleven human APOBEC proteins. *J Mol Biol.* 429(12):1787–1799. doi:[10.1016/j.jmb.2017.04.021](https://doi.org/10.1016/j.jmb.2017.04.021).
- Jakobsdottir GM, Brewer DS, Cooper C, Green C, Wedge DC. 2022. APOBEC3 mutational signatures are associated with extensive and diverse genomic instability across multiple tumour types. *BMC Biol.* 20(1):117. doi:[10.1186/s12915-022-01316-0](https://doi.org/10.1186/s12915-022-01316-0).
- Kanu N, Cerone MA, Goh G, Zalmas L-P, Bartkova J, Dietzen M, McGranahan N, Rogers R, Law EK, Gromova I, et al. 2016. DNA replication stress mediates APOBEC3 family mutagenesis in breast cancer. *Genome Biol.* 17(1):185. doi:[10.1186/s13059-016-1042-9](https://doi.org/10.1186/s13059-016-1042-9).
- Kazanov MD, Roberts SA, Polak P, Stamatoyannopoulos J, Klimczak LJ, Gordenin DA, Sunyaev SR. 2015. APOBEC-induced cancer mutations are uniquely enriched in early-replicating, gene-dense, and active chromatin regions. *Cell Rep.* 13(6):1103–1109. doi:[10.1016/j.celrep.2015.09.077](https://doi.org/10.1016/j.celrep.2015.09.077).
- Khalif P, Suspène R, Caval V, Thiers V, Beauclair G, Marchio A, Bekondi C, Amougou Atsama M, Camengo-Police SM, Noah Noah D, et al. 2022. APOBEC3C s1881 polymorphism enhances context-specific editing of hepatitis B virus genome. *J Infect Dis* 226(5):891–895. doi:[10.1093/infdis/jiac003](https://doi.org/10.1093/infdis/jiac003).
- Koito A, Ikeda T. 2013. Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases. *Front Microbiol.* 4:28. doi:[10.3389/fmicb.2013.00028](https://doi.org/10.3389/fmicb.2013.00028).
- Kotopka BJ, Smolke CD. 2020. Model-driven generation of artificial yeast promoters. *Nat Commun.* 11(1):2113. doi:[10.1038/s41467-020-15977-4](https://doi.org/10.1038/s41467-020-15977-4).
- Kouno T, Silvas TV, Hilbert BJ, Shandilya SMD, Bohn MF, Kelch BA, Royer WE, Somasundaran M, Kurt Yilmaz N, Matsuo H, et al. 2017. Crystal structure of APOBEC3A bound to single-stranded DNA reveals structural basis for cytidine deamination and specificity. *Nat Commun.* 8(1):15024. doi:[10.1038/ncomms15024](https://doi.org/10.1038/ncomms15024).
- Lada AG, Kliver SF, Dhar A, Polev DE, Masharsky AE, Rogozin IB, Pavlov YI. 2015. Disruption of transcriptional coactivator sub1 leads to genome-wide re-distribution of clustered mutations induced by APOBEC in active yeast genes. *PLoS Genet.* 11(5):e1005217. doi:[10.1371/journal.pgen.1005217](https://doi.org/10.1371/journal.pgen.1005217).

- Lang GI. 2018. Measuring mutation rates using the Luria–Delbrück fluctuation assay. *Methods Mol Biol.* 1672:21–31. doi:[10.1007/978-1-4939-7306-4_3](https://doi.org/10.1007/978-1-4939-7306-4_3).
- Langlois M-A, Beale RCL, Conticello SG, Neuberger MS. 2005. Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. *Nucleic Acids Res.* 33(6):1913–1923. doi:[10.1093/nar/gki343](https://doi.org/10.1093/nar/gki343).
- Layton CJ, McMahon PL, Greenleaf WJ. 2019. Large-scale, quantitative protein assays on a high-throughput DNA sequencing chip. *Mol Cell.* 73(5):1075–1082.e4. doi:[10.1016/j.molcel.2019.02.019](https://doi.org/10.1016/j.molcel.2019.02.019).
- Liu W, Newhall KP, Khani F, Barlow L, Nguyen D, et al. 2023. The cytidine deaminase APOBEC3G contributes to cancer mutagenesis and clonal evolution in bladder cancer. *Cancer Res.* 83(4):506–520. doi:[10.1158/0008-5472.CAN-22-2912](https://doi.org/10.1158/0008-5472.CAN-22-2912).
- Love RP, Xu H, Chelico L. 2012. Biochemical analysis of hypermutation by the deoxycytidine deaminase APOBEC3A. *J Biol Chem.* 287(36):30812–30822. doi:[10.1074/jbc.M112.393181](https://doi.org/10.1074/jbc.M112.393181).
- Lu Z, Lin Z. 2019. Pervasive and dynamic transcription initiation in *Saccharomyces cerevisiae*. *Genome Res.* 29(7):1198–1210. doi:[10.1101/gr.245456.118](https://doi.org/10.1101/gr.245456.118).
- Mas-Ponte D, Supek F. 2020. DNA mismatch repair promotes APOBEC3-mediated diffuse hypermutation in human cancers. *Nat Genet.* 52(9):958–968. doi:[10.1038/s41588-020-0674-6](https://doi.org/10.1038/s41588-020-0674-6).
- McCann JL, Cristini A, Law EK, Lee SY, Tellier M, Carpenter MA, Beghè C, Kim JJ, Sanchez A, Jarvis MC, et al. 2023. APOBEC3B regulates R-loops and promotes transcription-associated mutagenesis in cancer. *Nat Genet.* 55(10):1721–1734. doi:[10.1038/s41588-023-01504-w](https://doi.org/10.1038/s41588-023-01504-w).
- McDaniel YZ, Wang D, Love RP, Adolph MB, Mohammadzadeh N, Chelico L, Mansky LM. 2020. Deamination hotspots among APOBEC3 family members are defined by both target site sequence context and ssDNA secondary structure. *Nucleic Acids Res.* 48(3):1353–1371. doi:[10.1093/nar/gkz1164](https://doi.org/10.1093/nar/gkz1164).
- McGuffee SR, Smith DJ, Whitehouse I. 2013. Quantitative, genome-wide analysis of eukaryotic replication initiation and termination. *Mol Cell.* 50(1):123–135. doi:[10.1016/j.molcel.2013.03.004](https://doi.org/10.1016/j.molcel.2013.03.004).
- Mertz TM, Collins CD, Dennis M, Coxon M, Roberts SA. 2022. APOBEC-induced mutagenesis in cancer. *Annu Rev Genet.* 56(1):229–252. doi:[10.1146/annurev-genet-072920-035840](https://doi.org/10.1146/annurev-genet-072920-035840).
- Mertz TM, Rice-Reynolds E, Nguyen L, Wood A, Cordero C, Bray N, Harcy V, Vyas RK, Mitchell D, Lobachev K, et al. 2023. Genetic inhibitors of APOBEC3B-induced mutagenesis. *Genome Res.* 33(9):1568–1581. doi:[10.1101/gr.277430.122](https://doi.org/10.1101/gr.277430.122).
- Morganella S, Alexandrov LB, Glodzik D, Zou X, Davies H, Staaf J, Sieuwerts AM, Brinkman AB, Martin S, Ramakrishna M, et al. 2016. The topography of mutational processes in breast cancer genomes. *Nat Commun.* 7(1):11383. doi:[10.1038/ncomms11383](https://doi.org/10.1038/ncomms11383).
- Nikkilä J, Kumar R, Campbell J, Brandsma I, Pemberton HN, Wallberg Fredrik, Nagy Kinga, Scheer I, Vertessy Beata G, Serebrenik Artur A, et al. 2017. Elevated APOBEC3B expression drives a kataegic-like mutation signature and replication stress-related therapeutic vulnerabilities in p53-defective cells. *Br J Cancer.* 117(1):113–123. doi:[10.1038/bjc.2017.133](https://doi.org/10.1038/bjc.2017.133).
- Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, et al. 2012. Mutational processes molding the genomes of 21 breast cancers. *Cell.* 149(5):979–993. doi:[10.1016/j.cell.2012.04.024](https://doi.org/10.1016/j.cell.2012.04.024).
- Nordentoft I, Lamy P, Birkenkamp-Demtröder K, Shumansky K, Vang S, Hornshøj H, Juul M, Villesen P, Hedegaard J, Roth A, et al. 2014. Mutational context and diverse clonal development in early and late bladder cancer. *Cell Rep.* 7(5):1649–1663. doi:[10.1016/j.celrep.2014.04.038](https://doi.org/10.1016/j.celrep.2014.04.038).
- O'Shea JP, Chou MF, Quader SA, Ryan JK, Church GM, Schwartz D. 2013. pLogo: a probabilistic approach to visualizing sequence motifs. *Nat Methods.* 10(12):1211–1212. doi:[10.1038/nmeth.2646](https://doi.org/10.1038/nmeth.2646).
- Pecori R, Giorgio SD, Lorenzo JP, Papavasiliou FN. 2022. Functions and consequences of AID/APOBEC-mediated DNA and RNA deamination. *Nat Rev Genet.* 23(8):505–518. doi:[10.1038/s41576-022-00459-8](https://doi.org/10.1038/s41576-022-00459-8).
- Petljak M, Alexandrov LB, Brammeld JS, Price S, Wedge DC, Grossmann S, Dawson K, Ju YS, Iorio F, Tubio JMC, et al. 2019. Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. *Cell.* 176(6):1282–1294.e20. doi:[10.1016/j.cell.2019.02.012](https://doi.org/10.1016/j.cell.2019.02.012).
- Petljak M, Chu K, Dananberg A, Bergstrom EN, von Morgen P, et al. 2021. The APOBEC3A deaminase drives episodic mutagenesis in cancer cells. *bioRxiv* 2021.02.14.431145. doi:[10.1101/2021.02.14.431145](https://doi.org/10.1101/2021.02.14.431145).
- Petljak M, Dananberg A, Chu K, Bergstrom EN, Striepen J, von Morgen P, Chen Y, Shah H, Sale JE, Alexandrov LB, et al. 2022. Mechanisms of APOBEC3 mutagenesis in human cancer cells. *Nature.* 607(7920):799–807. doi:[10.1038/s41586-022-04972-y](https://doi.org/10.1038/s41586-022-04972-y).
- Petljak M, Maciejowski J. 2020. Molecular origins of APOBEC-associated mutations in cancer. *DNA Repair (Amst).* 94:102905. doi:[10.1016/j.dnarep.2020.102905](https://doi.org/10.1016/j.dnarep.2020.102905).
- Qian Y, Gong Y, Zou X, Liu Y, Chen Y, Wang R, Dai Z, Tasiheng Y, Lin X, Wang X, et al. 2022. Aberrant APOBEC3C expression induces characteristic genomic instability in pancreatic ductal adenocarcinoma. *Oncogenesis.* 11(1):1–13. doi:[10.1038/s41389-022-00411-9](https://doi.org/10.1038/s41389-022-00411-9).
- R Core Team. 2021. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
- Roberts SA, Sterling J, Thompson C, Harris S, Mav D, Shah R, Klimczak LJ, Kryukov GV, Malc E, Mieczkowski PA, et al. 2012. Clustered mutations in yeast and in human cancers can arise from damaged long single-strand DNA regions. *Mol Cell.* 46(4):424–435. doi:[10.1016/j.molcel.2012.03.030](https://doi.org/10.1016/j.molcel.2012.03.030).
- Ronsmans A, Wery M, Szachnowski U, Gautier C, Describes M, Dubois E, Morillon A, Georis I. 2019. Transcription-dependent spreading of the Dal80 yeast GATA factor across the body of highly expressed genes. *PLoS Genet.* 15(2):e1007999. doi:[10.1371/journal.pgen.1007999](https://doi.org/10.1371/journal.pgen.1007999).
- Saini N, Roberts SA, Sterling JF, Malc EP, Mieczkowski PA, Gordenin DA. 2017. APOBEC3B cytidine deaminase targets the non-transcribed strand of tRNA genes in yeast. *DNA Repair (Amst).* 53:4–14. doi:[10.1016/j.dnarep.2017.03.003](https://doi.org/10.1016/j.dnarep.2017.03.003).
- Sanchez A, Ortega P, Sakhtemani R, Manjunath L, Oh S, Bournique E, Becker A, Kim K, Durfee C, Temiz NA, et al. 2024. Mesoscale DNA features impact APOBEC3A and APOBEC3B deaminase activity and shape tumor mutational landscapes. *Nat Commun.* 15(1):2370. doi:[10.1038/s41467-024-45909-5](https://doi.org/10.1038/s41467-024-45909-5).
- Schumacher AJ, Nissley DV, Harris RS. 2005. APOBEC3G hypermutates genomic DNA and inhibits Ty1 retrotransposition in yeast. *Proc Natl Acad Sci U S A.* 102(28):9854–9859. doi:[10.1073/pnas.0501694102](https://doi.org/10.1073/pnas.0501694102).
- Septyarskiy VB, Soldatov RA, Popadin KY, Antonarakis SE, Bazykin GA, Nikolaev SI. 2016. APOBEC-induced mutations in human cancers are strongly enriched on the lagging DNA strand during replication. *Genome Res.* 26(2):174–182. doi:[10.1101/gr.197046.115](https://doi.org/10.1101/gr.197046.115).
- Shi K, Carpenter MA, Banerjee S, Shaban NM, Kurahashi K, Salamango DJ, McCann JL, Starrett GJ, Duffy JV, Demir Ö, et al. 2017. Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nat Struct Mol Biol.* 24(2):131–139. doi:[10.1038/nsmb.3344](https://doi.org/10.1038/nsmb.3344).

- Siow CC, Nieduszyńska SR, Müller CA, Nieduszyński CA. 2012. OriDB, the DNA replication origin database updated and extended. *Nucleic Acids Res.* 40(D1):D682–D686. doi:[10.1093/nar/gkr1091](https://doi.org/10.1093/nar/gkr1091).
- Starrett GJ, Luengas EM, McCann JL, Ebrahimi D, Temiz NA, Love RP, Feng Y, Adolph MB, Chelico L, Law EK, et al. 2016. The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. *Nat Commun.* 7(1):12918. doi:[10.1038/ncomms12918](https://doi.org/10.1038/ncomms12918).
- Sui Y, Qi L, Zhang K, Saini N, Klimczak LJ, Sakofsky CJ, Gordenin DA, Petes TD, Zheng D-Q. 2020. Analysis of APOBEC-induced mutations in yeast strains with low levels of replicative DNA polymerases. *Proc Natl Acad Sci U S A.* 117(17):9440–9450. doi:[10.1073/pnas.1922472117](https://doi.org/10.1073/pnas.1922472117).
- Supek F, Lehner B. 2017. Clustered mutation signatures reveal that error-prone DNA repair targets mutations to active genes. *Cell.* 170(3):534–547.e23. doi:[10.1016/j.cell.2017.07.003](https://doi.org/10.1016/j.cell.2017.07.003).
- Taylor BJ, Nik-Zainal S, Wu YL, Stebbings LA, Raine K, Campbell PJ, Rada C, Stratton MR, Neuberger MS. 2013. DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. *eLife.* 2:e00534. doi:[10.7554/eLife.00534](https://doi.org/10.7554/eLife.00534).
- Taylor BJM, Wu YL, Rada C. 2014. Active RNAP pre-initiation sites are highly mutated by cytidine deaminases in yeast, with AID targeting small RNA genes. *Elife.* 3:e03553. doi:[10.7554/eLife.03553](https://doi.org/10.7554/eLife.03553).
- THE GTEx CONSORTIUM. 2020. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science.* 369(6509):1318–1330. doi:[10.1126/science.aaz1776](https://doi.org/10.1126/science.aaz1776).
- Ubhi T, Zaslaver O, Quaile AT, Plenker D, Cao P, Pham N-A, Békési A, Jang G-H, O’Kane GM, Notta F, et al., 2024. Cytidine deaminases APOBEC3C and APOBEC3D promote DNA replication stress resistance in pancreatic cancer cells. *Nat Cancer.* doi:[10.1038/s43018-024-00742-z](https://doi.org/10.1038/s43018-024-00742-z).
- Vonesch SC, Li S, Szu Tu C, Hennig BP, Dobrev N, Steinmetz LM. 2021. Fast and inexpensive whole-genome sequencing library preparation from intact yeast cells. *G3 (Bethesda).* 11(1):jkaa009. doi:[10.1093/g3journal/jkaa009](https://doi.org/10.1093/g3journal/jkaa009).
- Warner JR. 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci.* 24(11):437–440. doi:[10.1016/s0968-0004\(99\)01460-7](https://doi.org/10.1016/s0968-0004(99)01460-7).
- Wittkopp CJ, Adolph MB, Wu LI, Chelico L, Emerman M. 2016. A single nucleotide polymorphism in human APOBEC3C enhances restriction of lentiviruses. *PLoS Pathog.* 12(10):e1005865. doi:[10.1371/journal.ppat.1005865](https://doi.org/10.1371/journal.ppat.1005865).
- Young BP, Loewen CJ. 2013. Balony: a software package for analysis of data generated by synthetic genetic array experiments. *BMC Bioinformatics.* 14(1):354. doi:[10.1186/1471-2105-14-354](https://doi.org/10.1186/1471-2105-14-354).
- Zhang C, Lu Y, Chen B, Bai Z, Hervieu A, Licciardello MP, Wang M, Mitsopoulos C, Al-Lazikani B, Totorici M, et al., 2022. R-loop editing by DNA cytosine deaminase APOBEC3B determines the activity of estrogen receptor enhancers. *bioRxiv* 2022.10.21.513235. doi:[10.2139/ssrn.4292602](https://doi.org/10.2139/ssrn.4292602).
- Zhu YO, Siegal ML, Hall DW, Petrov DA. 2014. Precise estimates of mutation rate and spectrum in yeast. *Proc Natl Acad Sci U S A.* 111(22):E2310–E2318. doi:[10.1073/pnas.1323011111](https://doi.org/10.1073/pnas.1323011111).

Editor: H. Merrih