



Duplex sequencing provides detailed characterization of mutation frequencies and spectra in the bone marrow of MutaMouse males exposed to procarbazine hydrochloride

Annette E. Dodge^{1,2} · Danielle P. M. LeBlanc¹ · Gu Zhou¹ · Andrew Williams¹ · Matthew J. Meier¹ · Phu Van³ · Fang Yin Lo³ · Charles C. Valentine III³ · Jesse J. Salk³ · Carole L. Yauk^{2,4} · Francesco Marchetti^{1,4}

Received: 23 February 2023 / Accepted: 17 May 2023 / Published online: 21 June 2023
© Crown 2023

Abstract

Mutagenicity testing is an essential component of health safety assessment. Duplex Sequencing (DS), an emerging high-accuracy DNA sequencing technology, may provide substantial advantages over conventional mutagenicity assays. DS could be used to eliminate reliance on standalone reporter assays and provide mechanistic information alongside mutation frequency (MF) data. However, the performance of DS must be thoroughly assessed before it can be routinely implemented for standard testing. We used DS to study spontaneous and procarbazine (PRC)-induced mutations in the bone marrow (BM) of MutaMouse males across a panel of 20 diverse genomic targets. Mice were exposed to 0, 6.25, 12.5, or 25 mg/kg-bw/day for 28 days by oral gavage and BM sampled 42 days post-exposure. Results were compared with those obtained using the conventional *lacZ* viral plaque assay on the same samples. DS detected significant increases in mutation frequencies and changes to mutation spectra at all PRC doses. Low intra-group variability within DS samples allowed for detection of increases at lower doses than the *lacZ* assay. While the *lacZ* assay initially yielded a higher fold-change in mutant frequency than DS, inclusion of clonal mutations in DS mutation frequencies reduced this discrepancy. Power analyses suggested that three animals per dose group and 500 million duplex base pairs per sample is sufficient to detect a 1.5-fold increase in mutations with > 80% power. Overall, we demonstrate several advantages of DS over classical mutagenicity assays and provide data to support efforts to identify optimal study designs for the application of DS as a regulatory test.

Keywords Error-corrected sequencing · eNGS · Genetic toxicology · Duplex sequencing · Mutagenesis · Transgenic rodent assay

Introduction

Somatic mutations are a hallmark of cancer and exposure to mutagenic substances is associated with an increased risk of carcinogenesis (Jackson et al. 2018). Thus, testing the ability of chemicals to induce genotoxic effects is required prior to regulatory approval and a battery of internationally standardized assays is available for this purpose. There is also growing recognition that mutations in themselves represent an adverse effect that should be considered in regulatory decision-making (Heflich et al. 2020). Assays with available Organisation for Economic Co-operation Development (OECD) test guidelines are routinely employed as the gold standards to assess in vitro and in vivo mutagenicity (OECD 2020, 2022a, b).

✉ Carole L. Yauk
carole.yauk@uottawa.ca

✉ Francesco Marchetti
francesco.marchetti@hc-sc.gc.ca

¹ Environmental Health Science and Research Bureau, Health Canada, Ottawa, Ontario, Canada

² Department of Biology, University of Ottawa, Ottawa, Ontario, Canada

³ TwinStrand Biosciences Inc., Seattle, Washington, USA

⁴ Department of Biology, Carleton University, Ottawa, Ontario, Canada

In vivo mutagenesis assays, such as the transgenic rodent (TGR) gene mutation assay (OECD 2022b), provide an important early indication of oncogenic risk; these assays are used to screen for potential hazard prior to the more time-intensive and costly rodent cancer bioassays. However, these assays have limitations. First, the TGR assay uses unique strains of animals that prevents integration with other standard toxicity tests. Second, the TGR plaque-based assay measures mutation induction in a single exogenous reporter gene that does not accurately represent the diverse structure and function of the mammalian genome. Many genomic features are thought to modulate differential sensitivity to chemical mutagens (Hodgkinson and Eyre-Walker 2011; Makova and Hardison 2015); these cannot be assessed when relying on a single reporter gene (Salk and Kennedy 2020). Third, the TGR assay, like many other existing mutagenicity tests, yields little mechanistic data. The pattern of mutation types induced by a chemical can provide clues to the mechanism behind the chemical's mutagenic effect; with the TGR assay this is impractical given the need to laboriously pick and independently sequence many viral plaques (Beal et al. 2020). Novel methodologies that can interrogate a broad selection of genomic targets and concurrently enumerate the spectrum of mutations across a diversity of genomic regions would provide mechanistic understanding of a chemical's mutagenic effect for better prediction of human cancer risk.

Duplex Sequencing (DS) is an error-corrected next-generation sequencing (ecNGS) technology that detects ultra-rare mutations using a customizable panel targeting a portion of the endogenous genome (Kennedy et al. 2014). Recent work with established environmental mutagens, including benzo[a]pyrene (BaP) (LeBlanc et al. 2022), aflatoxin (Chawanthayatham et al. 2017), N-ethyl-N-nitrosourea (ENU) and urethane (Valentine et al. 2020), has demonstrated that DS is effective at measuring increases in mutation frequencies (MF) in mouse somatic tissues. In addition to being able to precisely quantify spontaneous and induced MF across diverse genomic contexts, chromatin states and functional domains, DS provides complementary mechanistic information in the form of detailed mutation spectral data (Salk and Kennedy 2020) that can be linked to mechanism-associated mutational signatures extracted from genome sequences of human cancers (Alexandrov et al. 2020) or cloned ex vivo treated cells (Kucab et al. 2019). For instance, BaP, the main carcinogenic agent found in cigarette smoke, produces a mutation spectrum that is highly similar to mutational signatures observed in human lung cancers (Nik-Zainal et al. 2016; Kucab et al. 2019; LeBlanc et al. 2022). These signature analyses provide empirical evidence for the contribution of environmental mutagens to the mutation spectrum observed in human cancers. Given these features,

it is conceivable that ecNGS may contribute to reducing reliance on, or even replacing the need for, TGR and other mutation reporter assays.

Before DS can be used as a regulatory test for chemical mutagenicity assessment, more work is needed to compare its performance to standard assays across a range of chemicals with differing modes of action and potency. Here, we test the performance of DS compared to the TGR assay following exposure of mice to procarbazine hydrochloride (PRC), an anti-neoplastic agent that is classified as a Class 2A probable carcinogen by the International Agency for Research on Cancer (IARC Working Group 1981). PRC is a SN1-type methylating agent that is well characterized as a potent in vivo clastogen; although studies have shown that PRC can induce mutations in mouse somatic tissues (Maurice et al. 2019; Revollo et al. 2017; Suzuki et al. 1999), its mutagenic activity is still not fully understood. We measured mutations in the bone marrow (BM) of MutaMouse males exposed to PRC using the DS mouse mutagenesis panel comprised of twenty 2.4 kb targeted regions (Mouse Mutagenesis Kit, TwinStrand Biosciences Inc., Seattle WA, USA), chosen as a balanced representation of the entire genome with respect to location, GC content, trinucleotide abundances and genic status (LeBlanc et al. 2022). Our overall aim is to demonstrate the potential of DS as a robust in vivo mutagenicity test that could replace the TGR assay as the next “gold standard” in research and regulatory applications.

The objectives of this study are as follows: (1) use DS to study the dose–response of mutation induction by PRC in MutaMouse BM; (2) characterize regional variability in spontaneous and PRC-induced MF across the panel of 20 DS loci; (3) assess the concordance of DS with the TGR *lacZ* assay; (4) compare the mutation spectrum measured by DS to that measured by sequencing mutant *lacZ* plaques; and (5) explore optimal DS study designs for assessing mutagenicity by modeling sample size and sequencing data yield.

Methods

Animal exposure and tissue collection

All animal handling, exposures and tissue collection were approved by the Health Canada Ottawa Animal Care Committee and followed the Canadian Council on Animal Care guidelines. All MutaMouse animals used in this study were a random subset of those included in a previous TGR study investigating the effects of sampling time on mutant frequencies (Marchetti et al. 2021). Adult MutaMouse males, 8–10 weeks of age at the beginning of the exposure, were exposed by oral gavage to 6.25, 12.5, or 25 mg/kg-bw/day PRC (CAS 366-70-1; Sigma-Aldrich, Oakville, ON Canada)

or phosphate-buffered saline without calcium and magnesium (PBS; Corning cellgro, Manassas, VA, USA) as vehicle control (VC) for 28 consecutive days. Six mice for each treatment group, and three mice for VC were euthanized 42 days after the final dose. Three control mice, sampled 70 days after the final dosing, were added to achieve a matched control group of six animals. At euthanasia, BM was collected from the femurs, flash frozen, and stored at -80°C . BM was selected as the tissue to study because it is commonly used in regulatory mutagenicity testing and is a known target for PRC (Marchetti et al. 2021; Souliotis et al. 1994).

Transgenic rodent gene mutation assay (*lacZ* assay)

The *lacZ* assay was performed as per the OECD TG 488 (2022b). Genomic DNA was isolated from one femur using phenol/chloroform extraction (Gingerich et al. 2014). The quantity and purity of DNA were assessed using a NanoDrop spectrophotometer (Thermo Scientific Canada, Ottawa, Canada). λ gt10 phage vectors in genomic DNA were excised using commercial packaging extract kits (Agilent Technology, Santa Clara, CA, USA), according to the manufacturer's instructions. Transgene mutant frequency was determined using the phenyl β -d-galactopyranoside (P-gal)-positive selection assay (Lambert et al. 2005). A minimum number of 125,000 plaque-forming units (pfu) were counted per sample. Mutant frequency (phenotypically selectable mutants per *lacZ* locus) was expressed as the ratio of mutant plaques to total plaque forming units (pfu).

DNA extraction and library preparation for duplex sequencing

DNA from the second femur was extracted using the Qiagen DNeasy blood and tissue kit, as described in the Qiagen User Manual (Cat. # 69504, Qiagen, Hilden, Germany). DNA quality was assessed using an Agilent TapeStation to ensure a DNA Integrity Number greater than 7.0. DNA was prepared for sequencing as described previously (Valentine et al. 2020) by TwinStrand Biosciences (Seattle, WA, USA). Briefly, 500 ng of DNA was ultrasonically sheared to a mean fragment size of approximately 300 base pairs (bp). The DNA was then end-polished, A-tailed and ligated to DuplexSeq™ Adapters (TwinStrand Biosciences, Seattle WA, USA).

The DS Mouse Mutagenesis panel has one target on each autosome, except for chromosome 1 that has two targets. Nine of the targets are located within genic regions and 11 within intergenic regions (Suppl. Table 1). DS targets were classified as genic or intergenic based on gene intersection information obtained from the gene database GENCODE, the Mouse GenCode gene set, version M25 (Frankish et al.

2019). The chosen targets are presumed to be selectively neutral, with none of the targeted regions having known roles in cancer either in mouse or human orthologs. The panel was chosen for optimal performance in hybrid capture and contained no significant pseudogenes or repetitive elements that could potentially confound alignment or variant calling. Panel nucleotide composition shows a strong correlation with that of the whole mm10 genome (Supp. Fig. 1). Target chromatin status and %GC content were determined using the University of California Santa Cruz (UCSC) Mouse Genome Browser, reference genome mm10. Descriptions of the 20 targets are provided in Supp. Table 1. The panel was captured using 120mer biotinylated oligo probes (Integrated DNA Technologies, Coalville, IA, USA).

Prepared libraries were sequenced on a NovaSeq 6000 (Illumina, San Diego CA, USA) to a mean on-target duplex depth of $\sim 16,000\times$. An average of 400 million raw reads were sequenced per sample yielding an average of 16,000-fold coverage of the targeted loci. All samples met a target of 500 million duplex bases and 10,000X sequencing depth. The sequencing FASTQ files used in the current study are available in the Sequence Read Archive (SRA) BioProject ID PRIJNA909196. Per-library FASTQ files were analyzed using the TwinStrand DuplexSeq Mutagenesis App v.3.4.2 (TwinStrand Biosciences, Seattle, WA, USA) as functionally described in Valentine et al. (2020). This App produced sequencing summary metrics, mutagenesis metrics including mutation frequencies (MF), base substitution spectra, and a variant call file (VCF) per library. Briefly, the DuplexSeq Mutagenesis App involves the following: extracting unique molecular identifier (UMI) sequences, correcting UMI sequences, raw read alignment, grouping reads by their UMI and strand defining element, quality trimming, error-correction of read groups by duplex consensus calling, consensus post-processing, re-alignment, and variant calling. Importantly, the DuplexSeq Mutagenesis App does not include variants with a variant allele fraction > 0.01 in mutagenesis metrics because highly clonal variants were considered germline polymorphisms and unlikely to be somatic in origin. Post processing handled outside of the DuplexSeq Mutagenesis App involved calculating mutation frequencies per-target instead of just panel-wide.

Mutations that occurred more than once in a single sample were considered under the following two assumptions: (1) all identical mutations represent the clonal expansion of a single mutational event; or (2) all identical mutations occurred as independent mutational events, possibly due to local genomic features that render it more susceptible to mutation. Making the former assumption and counting multi-read mutations as a single mutation in the MF calculation guards against over-estimating mutant frequency that could otherwise be artificially increased by clonal expansion. This calculation is the

minimum mutation frequency (MF_{Min}) that could explain the data. Alternatively, counting all identical mutations independently guards against the possibility of undercounting true independent mutations. This calculation is the maximum mutation frequency (MF_{Max}) that could explain the observed data. Since it is rarely possible to be certain which of these two assumptions apply to any given mutation (and it could be a mixture of both), here, we consider both alternatives. It is also worth noting that clonal expansion of adverse mutations can lead to genetic disease and thus such events warrant consideration.

Mutation data interpretation and statistical analysis

First, mutation induction was measured as mutant frequencies (phenotypically scorable mutants per *lacZ* locus) for the *lacZ* assay and as MF_{Min} or MF_{Max} (represented as mutations per bp) for DS using the entire mouse mutagenesis panel. Estimated frequencies by dose were obtained for both assays using generalized linear models (GLM) in the R software with the quasibinomial distribution. Pairwise comparisons were conducted using the doBy R library (Højsgaard and Halekoh 2018) and these estimates were then back-transformed. Here, the delta method was employed to approximate the back-transformed standard errors (standard error of the mean; SEM) of the estimated MF. A Holm-Sidak correction was applied to adjust *p*-values for multiple comparisons for both MF_{Min} and MF_{Max} . Next, we estimated MFs for each 2.4 kb target using generalized linear mixed models (GLMM) with a binomial error distribution in the R software. Pairwise comparisons based on dose and location relative to genes were conducted using the doBY R library as described above. The MF_{Min} and MF_{Max} response for each target was then compared individually against the *lacZ* assay.

The mutation spectra data generated by DS was evaluated by considering first only the mutated base and then by extending the analyses to include the flanking bases. To determine which mutation substitution types differed between control and treated groups, a modified contingency table approach was used as previously described (Piegorisch and Bailer 1994). Pairwise comparisons as described above were used to determine which mutation types were significantly increased from VC. Subsequently, we generated a 96-trinucleotide mutation spectrum and used cosine similarity to determine the Catalogue of Somatic Mutations in Cancer (COSMIC) single base substitutions (SBS) signatures (Alexandrov et al. 2020) that most closely resembled the PRC mutation spectra, using the MutationalPatterns package in R (Manders et al. 2022).

Benchmark dose comparison of *lacZ* and DS results

Benchmark dose (BMD) modeling was conducted using the PROAST web application (version 70.1, <https://rivm.nl/en/proast>) to mathematically model the PRC dose–response curves for both assays. The BMD estimated in our study represents a dose that produces a 50% increase (i.e., the benchmark response; BMR) in *lacZ* mutant frequency, MF_{Min} or MF_{Max} ; this BMR was selected based on previous recommendations for genotoxicity assessment (White et al. 2020). The BMD is a measure of the potency of a chemical, with a lower BMD indicating a more potent mutagen. To estimate the BMD, different mathematical models (Hill, exponential, reverse exponential, and lognormal) were applied to fit the dose–response data. The Akaike Information Criterion was used to select the model with the best fit. 95% confidence intervals based on model averaging were generated from 200 bootstrap runs and were used to statistically evaluate the differences between BMDs derived with DS versus the *lacZ* assay, and across genomic loci.

Power analysis on DS study design

We performed a power analysis to determine the minimum sample size and total number of duplex base pairs needed to observe a significant increase in MF using DS. First, we used a binomial distribution to randomly generate control samples based on parameters from our experiment. Hypothetical mutation counts were generated based on the observed average MF_{Min} from our VC group (1.31×10^{-7} mutations per bp) and the average total number of bases sequenced per animal (depth; 7.98×10^8 bp). Over-dispersion was incorporated assuming a normal distribution with mean zero and standard deviation taken from the sample variance. A sample variance of 0.062 was estimated from the standard-deviation obtained from the GLMM used for per target dose–response analysis. We randomly generated a treatment group using the experimental parameters and a selected fold-increase. Next, we used a GLM to calculate the significance of the fold-change in MF between control and treated samples, as described above, and calculated the power of that analysis under the set conditions. Finally, the bisection method was used to estimate the minimum detectable effect size (MDES) of the simulated experiment yielding a power of 80% and a $p < 0.05$. We varied the sample size to calculate the power and MDES of the various conditions. Finally, we investigated the effect of sample variance and total duplex bases on our power results. We increased sample variance from 0.06 to 0.1 to roughly double our observed value and varied the total number of duplex bases. We then calculated the total number of duplex bases required to detect a 1.5-fold increase in MF with 80% power under each of the various simulated conditions.

Results

LacZ mutant frequency

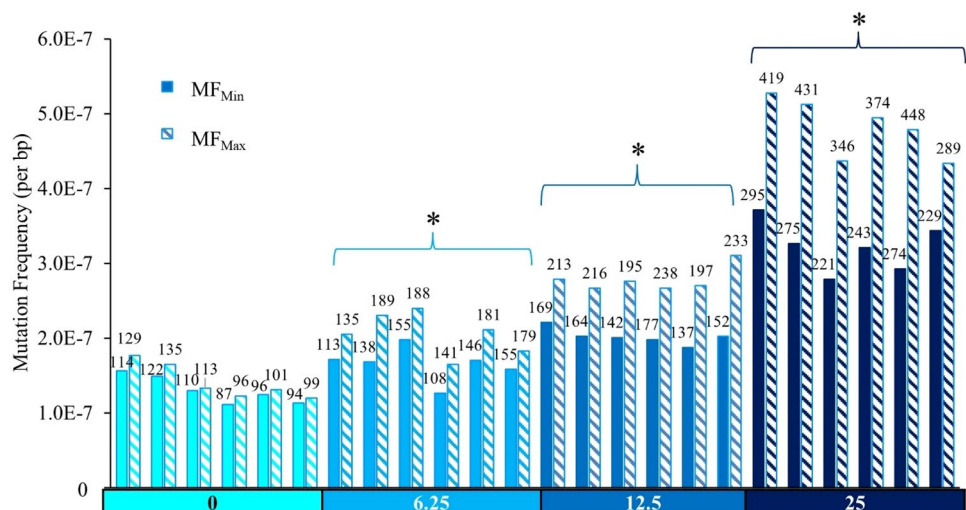
LacZ assay mutant frequency was measured by dividing the number of identified mutant plaques by the total number of plaques counted. On average, 14, 21, 51, and 100 mutants were counted for VC, 6.25, 12.5, and 25 mg/kg-bw/day PRC, respectively (Supp. Fig. 2). PRC induced a significant increase in mutant frequency ($\times 10^{-5}$) relative to VC in the middle and high-dose groups from 7.33 ± 1.76 mutations per *lacZ* locus (\pm SEM) in the VC to 24.9 ± 3.07 and 46.7 ± 4.14 at the middle and high PRC doses, respectively. No statistically significant increase was observed at the low PRC dose (Table 1). There was substantial inter-individual variation within the *lacZ* data, as indicated by the large SEM for the average mutant frequency of each dose group. These results are highly comparable to those reported by Marchetti et al. (2021) for the same animals.

Table 1 Mean mutant frequency estimates and corresponding pairwise comparisons between dose groups for the *lacZ* assay obtained using a GLM

Dose (mg/kg-bw/day)	Mutant frequency \pm SEM ($\times 10^{-5}$)	Fold changes (\pm SEM)	<i>p</i> -value*
0	7.3 ± 1.76	–	
6.25	12.7 ± 2.48	1.74 ± 0.54	8.84×10^{-2}
12.5	24.9 ± 3.07	3.39 ± 0.91	4.05×10^{-4}
25	46.7 ± 4.14	6.37 ± 1.63	1.54×10^{-6}

*Bold font indicates *p*-value < 0.05

Fig. 1 Duplex Sequencing mutation frequencies (MF) in the bone marrow of MutaMouse males at various doses of PRC. Bars represent the MF (mutations per bp) for each animal using MF_{Min} or MF_{Max}. Data labels indicate total mutation count per animal. X-axis indicates PRC dose group (mg/kg-bw/day). Asterisks indicate a significant difference (*p* < 0.05 relative to the control in the average MF across animals)



DS mutation frequency

DS reads were distributed relatively evenly across the 20 targets in the 24 samples (Supp. Fig. 3). Sequencing yielded an average of approximately 800 million duplex base pairs per sample for a total of about 20 billion duplex bases across the cohort. All samples met a minimum target of 500 million duplex bases per sample.

MF was calculated by dividing the number of identified mutants by the total number of bases sequenced per sample. Mutations include single nucleotide variants (SNVs), insertions and deletions (indels), and multi-nucleotide variants (MNVs). Under the conservative MF_{Min} assumption, 104, 136, 157, and 256 unique mutations on average were identified in mice treated with VC, 6.25, 12.5, and 25 mg/kg-bw/day PRC, respectively, with minimal inter-individual variability within each dose group (Fig. 1, Suppl Table 2). There was no significant difference in MF_{Min}, nor in the proportion of mutation subtypes, between VC animals sampled 42 or 70 days after the last dosing. PRC induced a significant dose-dependent increase in MF_{Min} ($\times 10^{-7}$) relative to VC in all PRC dose groups, from 1.31 ± 0.07 mutations per bp in the control to 3.22 ± 0.12 in the high dose group. A 1.3-, 1.6-, and a 2.5- fold increase in MF_{Min} relative to VC was induced in 6.25, 12.5, and 25 mg/kg-bw/day dose groups, respectively (Table 2).

To assess the impact of counting all mutations independently, we performed the same analyses using the MF_{Max} assumption. We identified an additional 50, 198, 351, and 770 mutations for VC, 6.25, 12.5, and 25 mg/kg-bw/day PRC, respectively. PRC induced a significant dose-dependent increase in MF_{Max} ($\times 10^{-7}$; Table 3, *p* < 0.01) in all groups relative to VC from 1.41 ± 0.08 mutations per bp (\pm SEM) in the control to 4.82 ± 0.16 in the high dose group (Fig. 1). A 1.5-, 2.0-, and a 3.4- fold increase in MF_{Max} relative to VC was induced in 6.25, 12.5, and 25 mg/kg-bw/day PRC

Table 2 Mean mutation frequency (MF_{Min}) estimates and corresponding pairwise comparisons between dose groups for DS obtained using a GLM

Dose (mg/kg-bw/day)	MF_{Min} assumption		
	$MF \pm SEM (\times 10^{-7})$	Fold Changes ($\pm SEM$)	<i>p</i> value*
0	1.31 ± 0.07	–	–
6.25	1.67 ± 0.08	1.26 ± 0.10	6.03×10^{-3}
12.5	2.03 ± 0.09	1.55 ± 0.11	1.65×10^{-5}
25	3.22 ± 0.12	2.46 ± 0.16	6.57×10^{-11}

*Bold font indicates *p*-value < 0.05

dose groups, respectively. The MF_{Max} for the middle and high dose groups were significantly higher than the MF_{Min} for those dose groups (*p* < 0.01). Inter-individual variability did not meaningfully differ between the MF_{Min} and MF_{Max} assumptions.

DS mutation frequency by target

Next, we analyzed mutation induction across the 20 genomic targets using the MF_{Min} assumption. The background MF_{Min} differed by 2.8-fold across the targets, with the highest MF_{Min} observed for the second target on chromosome 1 (chr1.2: 2.41×10^{-7}) and the lowest MF_{Min} on chromosome 10 (0.858×10^{-7}). When ranking MF across targets within the VC, we observed that the highest and lowest responding targets are similar to those of a previous study (LeBlanc et al. 2022). This reproducibility suggests that we

Table 3 Mean mutation frequency (MF_{Max}) estimates and corresponding pairwise comparisons between dose groups for DS obtained using a GLM

Dose (mg/kg-bw/day)	MF_{Max} assumption		
	$MF \pm SEM (\times 10^{-7})$	Fold changes ($\pm SEM$)	<i>p</i> value*
0	1.41 ± 0.08	–	–
6.25	2.05 ± 0.10	1.45 ± 0.11	9.51×10^{-5}
12.5	2.78 ± 0.12	1.97 ± 0.14	1.16×10^{-8}
25	4.82 ± 0.16	3.42 ± 0.23	6.48×10^{-14}

*Bold font indicates *p*-value < 0.05

have sequenced a sufficient number of mutations to have reproducible MF per target within the control group. PRC induced a significant increase in MF_{Min} relative to VC in 14 of the 20 targets at the high PRC dose (Fig. 2). At the middle PRC dose, only the target on chromosome 17 had a significant increase in MF_{Min} . No target had a significant increase in MF_{Min} compared to VC at the low dose. We observed a maximum 2.5-fold difference between the target with the highest MF_{Min} at the high PRC dose (chromosome 17; 5.21×10^{-7}) and the target with the lowest MF_{Min} at the high dose (chromosome 13; 2.13×10^{-7}). Similarly, we observed a 2.7- and 2.6-fold change between the targets with the highest and lowest MF_{Min} for the middle and low PRC doses, respectively. The targets with the highest MF_{Min} were on chromosome 17 for the high and middle PRC doses, and on the second target on chromosome 1 (i.e., chr1.2) for

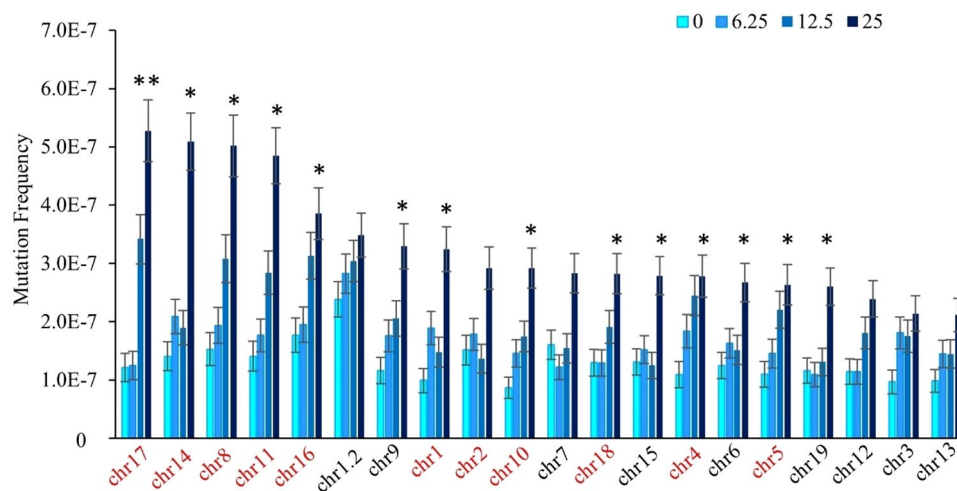


Fig. 2 Spontaneous and PRC-induced MF by Duplex Sequencing target ordered from highest MF_{Min} in the high PRC dose group to the lowest. Data are mean $MF \pm SEM$ mutations per bp, for each target (*n* = 6), separated by dose (mg/kg-bw/day). Intergenic targets are labelled in red and genic targets are in black. Asterisks indicate a sig-

nificant increase in the high PRC dose relative to the controls. The double asterisk indicates a significant increase in the high and middle PRC doses relative to controls. (Generalized linear mixed model, *p* < 0.05)

the VC and low PRC dose. Three of the five targets with the lowest background MF_{Min}, those on chromosomes 3, 5 and 13, were also among the five targets with the lowest MF_{Min} at the high PRC dose, suggesting these targets may be less sensitive to mutation induction. Finally, responsiveness to PRC varied between targets. The fold-increase from VC at the high-dose group for chr17 was 3 × higher than the target with the lowest fold-increase, chr1.2 (Fig. 3). Thus susceptibility to mutations induction varies by genomic locus.

We observed that genomic targets in intergenic regions tended to have a higher susceptibility to PRC than targets in genic regions, which, if expressed, can be subjected to transcription-coupled repair (TCR). The five targets with the highest MF_{Min} at the high PRC dose were all located in intergenic regions, while 4/5 targets with the lowest MF_{Min} were genic (Fig. 2). Overall, a combined analysis of intergenic targets showed a significantly higher mean MF_{Min} than the combined genic targets in the middle and high PRC dose groups by 1.3- and 1.4-fold, respectively ($p < 0.001$) (Supp. Fig. 4). These results did not change when using MF_{Max}. Furthermore, intergenic targets showed a significant increase in MF_{Min} and MF_{Max} at all PRC doses compared to the VC, whereas the genic targets showed a significant increase only at the high dose for MF_{Min} and the middle and high dose for MF_{Max}. Conversely, the mean background MF_{Min} observed in intergenic targets (1.26×10^{-7}) was not significantly different from the mean background MF_{Min} in genic targets (1.28×10^{-7}). These results were not changed when using MF_{Max}.

DS and TGR concordance

To assess DS concordance with the TGR assay, we compared the DS MF_{Min} to the *lacZ* mutant frequency in the

same samples. There was a significant positive correlation between the induced DS MF_{Min} and *lacZ* mutant frequencies (Fig. 4; Pearson's correlation; $r = 0.85$, $p = 6.19 \times 10^{-7}$), with *lacZ* displaying more inter-individual variability. The correlation increased further when we used MF_{Max} (Pearson's correlation; $r = 0.87$, $p = 2.88 \times 10^{-8}$), although the magnitude of the response to PRC was still higher in the *lacZ* assay (Tables 1, 2 and 3). The *lacZ* fold changes relative to VC for each dose group were greater by an average of twofold than DS MF_{Min} and 1.5-fold compared with DS MF_{Max}.

We then assessed the concordance of MF of individual DS targets with the *lacZ* assay for the two PRC doses that significantly increased *lacZ* mutations. The fold-change in MF between VC and the PRC dose groups was estimated for each target, using either the MF_{Min} or MF_{Max} DS data (Fig. 3). Using MF_{Min}, the target on chromosome 17 (4.34 ± 0.98 -fold-increase) was the only one within range of the fold-change observed using the *lacZ* assay (6.37 ± 1.63). Conversely, using MF_{Max}, we observed six targets whose fold-increase in MF falls within the *lacZ* assay range, namely the targets on chromosome 17, 14, 1, 11, 8, and 10. These six targets were all intergenic. Similar trends were observed at the middle PRC dose, where four and ten targets were within range of the *lacZ* fold-change (3.39 ± 0.91) when MF_{Min} and MF_{Max} values were used, respectively. Of these, all targets except for the one on chromosome 3 were intergenic (Supp. Fig. 5). Thus, the data show intergenic targets generated fold changes in MF that more closely approximated *lacZ* assay results than genic targets.

Finally, we examined DS and *lacZ* concordance using BMD modelling to identify the dose at which a 50% increase in MF occurred above background. The BMD for the *lacZ* assay was 3.7 mg/kg-bw/day (95% CI of 1.6–6.6 mg/kg-bw/day), which was significantly lower than the BMD

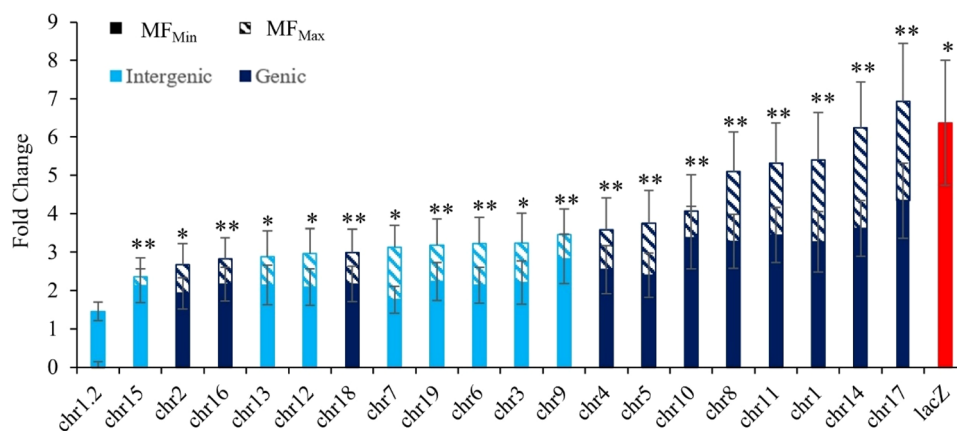


Fig. 3 Fold change in MF between PRC high-dose group and VC for the 20 Duplex Sequencing targets as well as for the *lacZ* gene (red bar). Estimated using a general linear mixed model. DS targets are listed along the X-axis. Errors bars are SEM. Asterisks indicate the

fold change was a significant increase in MF from VC for MF_{Max}. Double asterisks indicate significance for both MF_{Max} and MF_{Min}. The targets chr17, chr14, chr1, chr11, chr8, and chr10 were within range of the *lacZ* fold change

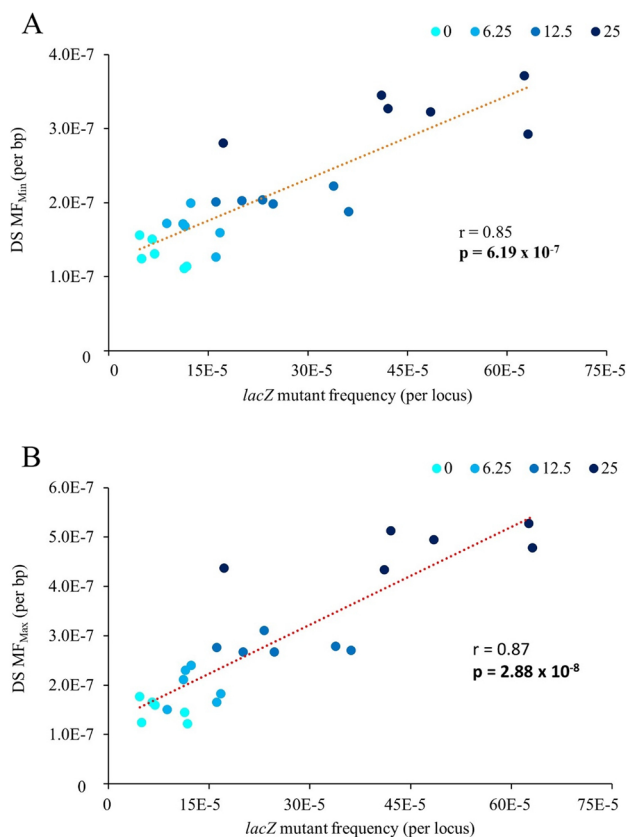


Fig. 4 Pearson's correlation analysis between *lacZ* mutant frequency and Duplex Sequencing using MF_{Min} (A) or MF_{Max} (B) for Mutamouse males exposed to PRC. Data are presented as the individual MFs for each animal

of 11.1 mg/kg-bw/day (95% CI of 8.7–13.8 mg/kg-bw/day) obtained using DS MF_{Min} (Fig. 5A). However, using MF_{Max} produced a BMD of 8.4 mg/kg-bw/day (95% CI of 6.5–10.4 mg/kg-bw/day), which overlaps with the CI of the *lacZ* assay. Thus, quantitative concordance of DS with the *lacZ* assay improves when MF_{Max} are used. We also produced BMDs for each of the 20 DS targets individually calculated as MF_{Max} (Fig. 5B) and compared these to the *lacZ* BMD. Every target had CIs that overlapped with the *lacZ* CI, except for targets on chromosomes 7, 15, and 19, which were significantly higher. Chr 1.2 failed to produce a BMD due to its weak dose–response and high mutation background frequency. Overall, BMD concordance between the assays improved when MF_{Max} was used.

Mutation spectrum

Both the spontaneous and PRC-induced mutation spectra (MF_{Min}) consisted primarily of SNVs, followed by small deletions, then MNVs, and last, small insertions (Fig. 6). Pairwise comparison of DS mutation data revealed that the overall mutation spectra were significantly different from

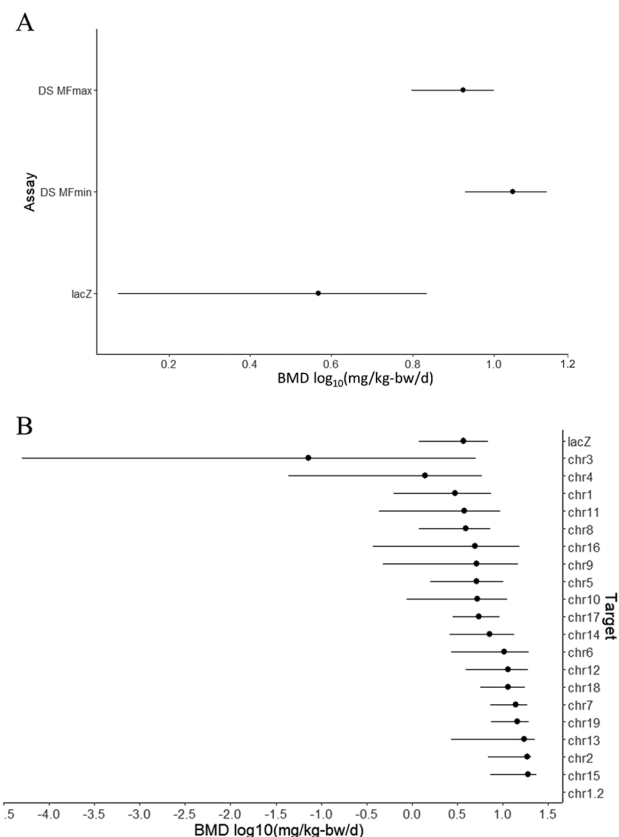
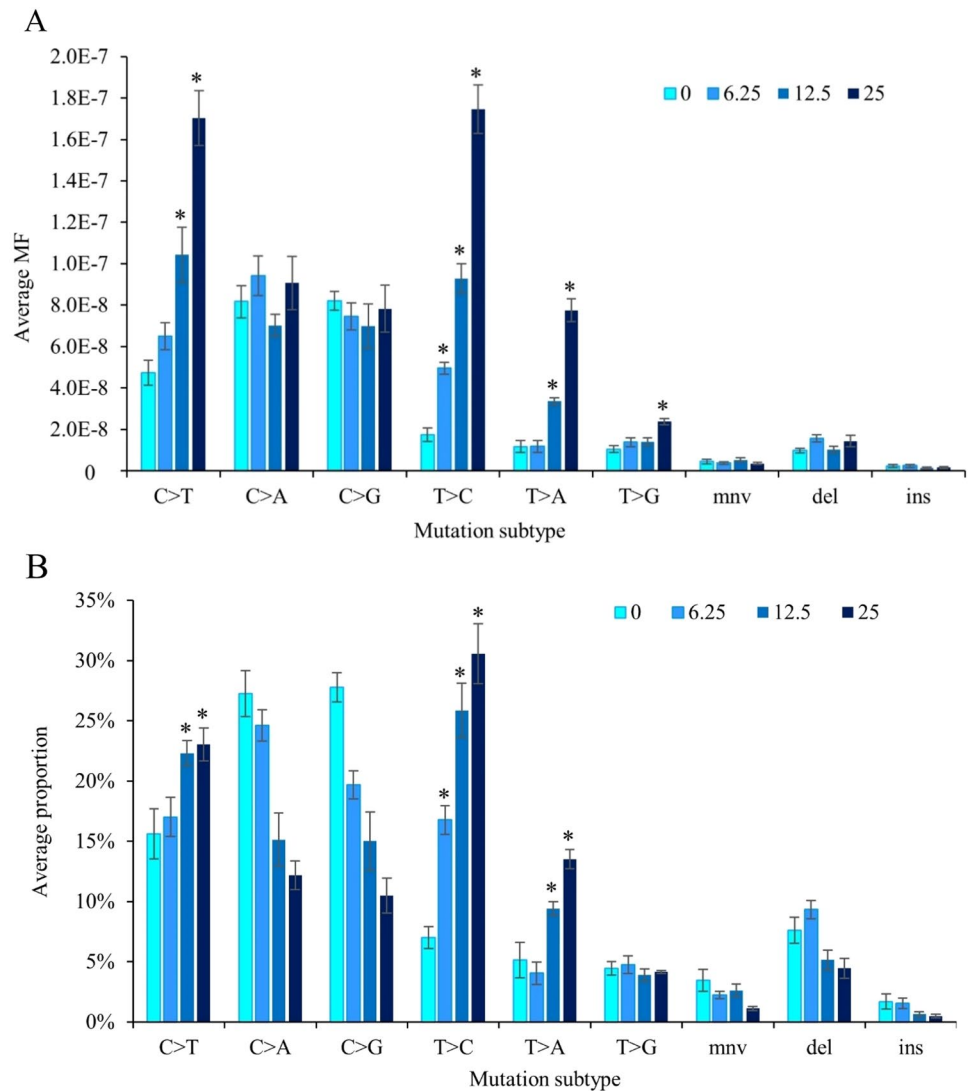


Fig. 5 95% confidence intervals (CI) based on BMD model averaging analyses. A log₁₀ BMD for a 50% increase of *lacZ* mutant frequency and the Duplex Sequencing MF, using either MF_{Min} or MF_{Max}. B log₁₀ BMD of *lacZ* mutant frequency and Duplex Sequencing MF_{Max} for each genomic target. DS targets are organized from smallest BMD to highest. Target chr1.2 failed to produce a BMD due to poor dose response

each other at all PRC doses (modified contingency table; $p < 0.05$). PRC primarily induced T:A > C:G and C:G > T:A transitions accounting for 29% and 28% of total mutations at the high dose, followed by T:A > A:T transversions (13%) and T:A > G:C transversions (4%). This differed from the background mutation spectrum, which consisted primarily of C:G > G:C transversions (33%), C:G > A:T transversions (33%), and C:G > T:A transitions (17%). We observed significant dose-dependent increases in T:A > C:G transitions in all three dose groups compared with VC; T:A > A:T transversions were significantly increased at the middle and high PRC dose compared with VC; and T:A > G:C transversions and C:G > T:A transitions were increased at the high dose only (Fig. 6).

We further analyzed the sequence context of induced mutations by considering the flanking base on either side of the mutated base to produce trinucleotide spectra and examine mutational signatures (Alexandrov et al. 2013). This revealed a clear enrichment in T:A > C:T transitions and

Fig. 6 Mutation spectrum of controls and PRC dose groups measured by Duplex Sequencing in the bone marrow of MutaMouse males. **A** Mutation subtypes are presented as the average $MF_{Min} \pm SEM$. **B** Mutation subtypes are represented by the proportion of total mutations. Values are mean \pm SEM. Plots represent MF_{Min} . Mutation subtypes include the six single nucleotide variants using a pyrimidine reference, multi-nucleotide variants (mnv), small insertion (ins), and small deletion (del) mutations. Asterisks indicate a significant increase in MF from controls



T:A>A:T transversions, together with a shift from mutations at CpG sites to non CpG sites for C:G>T:A transitions (Suppl. Fig. 6). We compared our trinucleotide frequencies to single-base substitution (SBS) signatures of the COSMIC Mutational Signatures database (mm10, version 3.3) using cosine similarity. We observed a dose-dependent increase in cosine similarity to several cancer signatures. The signature that displayed the greatest difference between its similarity to the control spectrum and its similarity to the high PRC spectrum was SBS 21, followed by SBS 26, and SBS 12 (Suppl. Fig. 6).

Power analysis for sample size and total duplex basepairs

We performed a power analysis on datasets simulated from the observed conditions of our DS experiment. Using our true sample size of six animals per dose, we had >95%

power to detect a 1.5-fold increase in MF_{Min} (Table 4). Reducing the number to three animals per dose group retained 92% power. The MDES that could be achieved after reducing the sample size to three animals was a 1.4-fold increase in MF from background (Table 4). This is higher than the 1.26-fold increase in MF_{Min} that was observed for the lowest PRC dose, but is lower than the 1.55-fold and 2.46-fold increase observed at the middle and high PRC doses. Thus, under the observed conditions of our experiment, a reduction in sample size to three animals per group would result in sufficient power to classify PRC as mutagenic and detect the changes we observed at the middle (95% power) and high (>95% power) PRC dose but would be under-powered to detect the smaller increase observed at the low PRC dose (49% power). The use of MF_{Max} did not alter the results of the power analysis (data not shown). Finally, we assessed whether these results were robust to simulated increases in sample variance. We performed the

Table 4 Power analysis for duplex sequencing dose–response studies

	Experimental conditions				Simulated condition	
	MF _{Min}		MF _{Max}		Increased variance	
Background MF	1.31×10^{-7}		1.41×10^{-7}		1.31×10^{-7}	
Total DS base pairs	7.98×10^8		7.98×10^8		7.98×10^8	
Sample variance	0.062		0.063		0.10	
Sample Size	Power (%) 1.5-FC ^a	MDES ^b	Power (%) 1.5-FC ^a	MDES ^b	Power (%) 1.5-FC ^a	MDES ^b
6	100	1.23	100	1.22	100	1.27
5	100	1.25	100	1.25	98	1.31
4	100	1.30	99	1.29	95	1.38
3	92	1.40	91	1.38	81	1.51

^aPower (%) to detect a 1.5-fold-change of MF

^bMDES: minimum detectable effect size with 80% power

same analyses as above but with twice the observed sample variance (0.1). We observed an 81% power to detect a 1.5-fold increase in MF_{Min} above background with a sample size of three and the increased sample variance.

Next, we investigated the number of total duplex bases required for analysis (Fig. 7). When using a sample size of three animals, and our observed sample variance (0.062), a minimum of 500 million DS bases per animal is required to achieve sufficient power (81%) to detect a 1.5-fold increase in MF_{Min}. However, this reduction in data yield causes analyses to become more susceptible to increases in sample variance. When the sample variance is 0.1, and the sample size is three animals, the power to detect a 1.5-fold increase in MF_{Min} is reduced to 71% and a minimum data yield of ~800 million duplex bases is required to achieve sufficient power. Increasing the number of animals per group to four, with a minimum of 300 million duplex bases, would be sufficient to detect a 50% increase in MF_{Min} for datasets with a sample variance of 0.1 (78.2% power).

Discussion

We applied DS to study spontaneous and PRC-induced mutations in the BM of MutaMouse males across a panel of diverse genomic targets and compared the results with those obtained with the OECD TGR assay on the same samples. Low intra-group variability within DS samples resulted in statistically significant increases at all PCR doses, whereas the *lacZ* assay failed to detect a significant increase at the low dose. The *lacZ* assay initially showed a stronger fold-response to PRC than DS based on MF_{Min}; however, when using MF_{Max}, several DS targets had fold-increases consistent with the *lacZ* assay. DS revealed a PRC mutation spectrum that was distinct from the background spectrum, with PRC primarily inducing T:A > C:G transitions, followed by

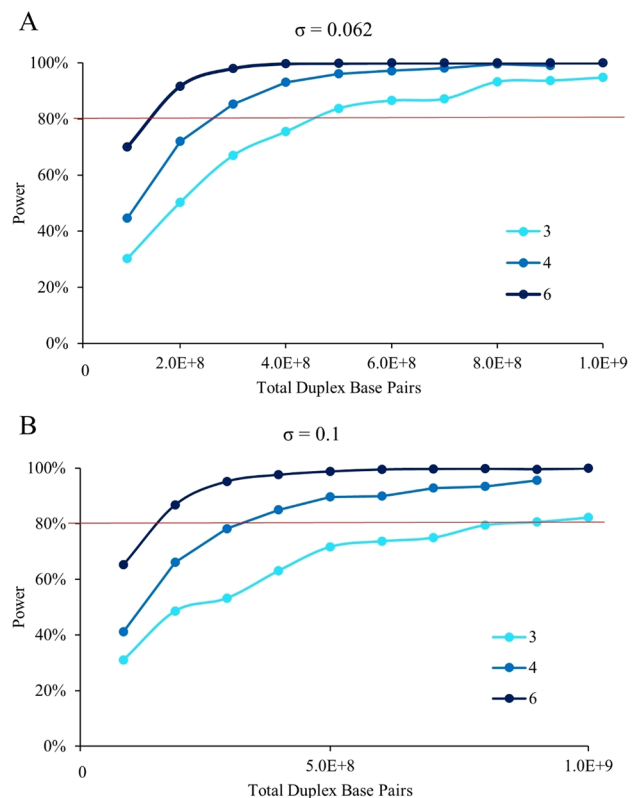


Fig. 7 The power to detect a 50% increase in MF by Total Duplex Base Pairs. Results are displayed for a sample size of 6 animals per group (dark blue), 4 animals per group (blue) and 3 animals per group (light blue). **A** Sample variance of 0.062 (based on current study). **B** Sample variance of 0.1. Red line indicates the threshold for >80% power and $p < 0.05$ (color figure online)

C:G > T:A transitions, and T:A > A:T transversions. Finally, power analyses suggested that reducing the sample size to three animals per dose group and per-sample data yield of 500 million bases retains sufficient power to detect a 1.5-fold

increase in MF in our model. Thus, fewer animals per dose-group can be used with respect to the TGR models without losing power to detect an effect. Overall, our study provides strong evidence to support the use of DS in place of the plaque-based TGR assay for qualitative and quantitative chemical mutagenicity assessment with greater sensitivity, richer data and fewer animals.

DS loci show patterns of mutation sensitivity consistent with TCR

There are many genomic and epigenomic features that have been linked to susceptibility to mutation induction including transcription (Monroe et al. 2022; Xia et al. 2020), sequence context (Hodgkinson and Eyre-Walker 2011), and chromatin state (Hodgkinson et al. 2012; Makova and Hardison 2015). The use of the Mouse Mutagenesis panel allowed us to observe differences in spontaneous and PRC-induced MF among the 20 targets. Specifically, we found that DS target loci in intergenic regions had a higher response to PRC mutagenesis than loci in genic regions, which is consistent with the protective effect of TCR. Transcribed regions of the genome are subject to increased rates of nucleotide excision repair (Nadkarni et al. 2016) that may contribute to the decreased MF in genic targets. Similar results were observed following BaP exposure in MutaMouse BM (LeBlanc et al. 2022). Remarkably, there was also a consistent pattern of relative sensitivity to mutation induction across targets between the two studies. In fact, the five targets with the highest MF_{Min} at the high PRC dose (chromosome 17, 14, 11, 8, and 16; all intergenic) were the same five targets with the highest MF_{Min} at the high BaP dose. These chemicals have different modes of action, which implies that there is an inherent feature of these loci that renders them more susceptible to mutagenesis than the others. Our results provide further support for the role of TCR in the variation of mutation among loci.

Clonal mutations and target responsiveness underlie discrepancy between DS and *lacZ*

One of the main objectives of this study was to assess the concordance of DS with the gold standard TGR assay for in vivo mutagenesis testing. Initial analyses comparing the results of the *lacZ* assay with those of DS using MF_{Min} showed a significant positive correlation, but lower fold-increases in MF and a higher BMD for DS compared to *lacZ*. Similarly, an attenuated DS response compared to the Big Blue[®] plaque-based assay was observed in mouse BM analyzed 3 days following exposure to BaP (Valentine et al. 2020). These authors suggested that the difference might be attributed to unrepaired DNA adducts that were fixed into mutations during the in vitro portion of the TGR assay (i.e.,

ex vivo mutations). However, in this study, we had a 42-day waiting period between the last daily exposure and tissue collection. It is highly unlikely that DNA adducts could remain for such an extended period. A similar study using BaP also found the same difference following a 28-day sampling time (LeBlanc et al. 2022). Thus, ex vivo mutations are unlikely to contribute to the observed difference between DS and TGR assay. It is possible that applying the MF_{Min} approach to the DS data is too conservative and contributes to the observed difference between the two assay results.

The higher response measured by the *lacZ* assay may be the result of inherent differences between the two assays in how they quantify mutations. The *lacZ* assay scores mutants based on positive selection of plaques carrying mutant *lacZ* genes; thus, it cannot identify synonymous mutations, which do not alter the function of the *lacZ* gene, nor can it discount clonal expansion events unless paired with NGS sequencing. DS, on the other hand, identifies all types of mutations without bias and can easily filter out clonally expanded mutations using the MF_{Min} approach. A single mutation may produce large mutant frequencies following clonal expansion of the mutated cell (Heddle 1999). Due to the long interval between treatment and sampling time in our experiment, it is possible for the highly proliferative BM to accumulate small clonal populations of mutated cells. Indeed, we found that clonal mutations increased with PRC dose. When we included clonally expanded mutations in our calculations (using the MF_{Max} assumption), we observed that the concordance between DS and the *lacZ* assay improved significantly. DS fold-changes in several endogenous target loci became equivalent to fold changes observed with the *lacZ* gene. Furthermore, BMD confidence intervals generated using MF_{Max} overlapped with those for the *lacZ* BMD. Thus, it is possible that some of the discrepancies between DS and *lacZ* can be attributed to the exclusion of clonally expanded mutations when analyzing DS data. Whether clonal mutations should be included in the DS mutation counts (MF_{Max}) or corrected for (MF_{Min}) will require consideration of the context with which the assay is performed and should be a subject of further investigation by the expert community.

Differences in the responsiveness of the *lacZ* and DS assays to PRC may also be the result of assessing different genomic loci with differing sensitivities. Several studies have shown that exogenous bacterial genes, such as the *lacZ* gene, tend to have higher mutant frequencies than endogenous genes. Previous work using Big Blue[®] mouse splenic tissue demonstrate that the *lacI* transgene exhibits higher spontaneous (Skopek et al. 1995) and BaP-induced (Skopek et al. 1996) mutant frequencies than the endogenous *Hprt* gene. Similarly, Valentine et al. (2020) found that the *cII* transgene has an elevated response to BaP mutagenesis compared to two expressed endogenous genes (*Cttnb1*, *Polr1c*) using DS. We posit that the main reason for the seemingly

attenuated response of DS compared to the *lacZ* assay is the result of the biologically higher susceptibility of the *lacZ* transgene to mutagenesis relative to many of the endogenous DS loci. The *lacZ* gene may exhibit increased susceptibility to mutation induction because it is a non-transcribed transgene, thus unprotected by TCR. Of the DS loci that had fold-increases in MF_{Max} rivalling that of the *lacZ* transgene, most were in intergenic regions. Future use of DS to directly sequence the *lacZ* gene would aid in understanding how the differences between the two assays influence MF.

DS identifies induction of mutation subtypes not feasible with TGR

A major benefit of DS is its ability to generate high-quality mutation spectra alongside its measurement of MF. Here, we showed that PRC induces mainly T:A > C:G, C:G > T:A, and T:A > A:T mutations. This is consistent with the known PRC mode of action as a monofunctional SN₁-methylating agent (Fong et al. 1990; Fu et al. 2012; Gerson et al. 2018). The main mutagenic lesions produced by PRC include O₄ and O₂-methylthymine leading to T:A > C:G and T:A > A:T mutations, respectively (Jenkins et al. 2005), and O₆-methylguanine—a highly mutagenic adduct that results in C:G > T:A transitions (Fu et al. 2012). Previous work on the *lacZ* and *Pig-A* loci did not show significant increases in the proportions of C:G > T:A mutations following PRC exposure (Beal et al. 2020; Pletsa et al. 1997; Revollo et al. 2017). We believe our study is the first to report a significant increase in these mutations. This was likely possible because of the substantially larger amount of spectral data generated by DS than what was produced using laborious manual clone picking and sequencing with the other approaches. Studies relying on TGR approaches characterized PRC mutation spectrum from only 20–120 mutations. In contrast, our study utilized a total of 3926 mutations, substantially improving the robustness of our analyses. Furthermore, spontaneous C:G > T:A mutations are extremely common at CpG sites as a result of spontaneous deamination of methylated cytosine to form thymine. We observed a dose-dependent proportional reduction in these mutations occurring at CpG sites, alongside an increase in the context of non-CpG cytosine bases, consistent with the formation of O₆-methylguanine adducts by PRC. Previous studies with low mutation counts likely were underpowered to observe a significant increase in C:G > T:A mutations due to their abundance in the VC group. As a mutagenicity test, DS enables the detection of lower frequency changes in specific mutation types to facilitate the discovery of potential secondary mutagenic mechanisms or less prominent DNA lesions induced by the mutagen that would not be observed using traditional assays.

The trinucleotide mutation spectrum generated by DS following PRC exposure was linked to several COSMIC

mutational signatures. The PRC mutation spectrum was most similar to SBS 21 and 26, which are both associated with defects in mismatch repair. Interestingly, the mutation spectra of similar acting alkylating agents (N-methyl-N-nitrosourea and N-ethyl-N-nitrosourea) within human pluripotent stem cells (Kucab et al. 2019) also showed high similarity with SBS 26 and SBS 12, suggesting that these SBS may be signatures of alkylation-induced chemicals. Additionally, we observed similarity with SBS 25, a signature observed in several Hodgkin's cell line samples where the patients were treated with chemotherapy (Alexandrov et al. 2020). PRC was historically used to treat Hodgkin's lymphoma; thus, it is possible that this cancer signature is related to the exposure of these patients to PRC, or the similar acting methylating agent dacarbazine now used more commonly. These results support the utility of DS mutational signature analyses derived from mutagen-exposed rodents in identifying exposure-associated cancers in humans.

DS study design accommodates reducing animal use and the cost of sequencing

The analyses of Piegorsch et al. (1995) revealed that the main source of statistical variability observed in a TGR assay was significant variation between animals within a treatment group. These analyses provided the basis for the number of animals and plaques necessary to define a negative result with acceptable power (80%) for the TGR assay. OECD TG 488 requires a minimum of 5 animals per dose group for in vivo mutagenicity testing. In the current study, DS displayed much lower inter-animal variation than is typically observed using the *lacZ* assay. We note that the low sample-to-sample variability observed with DS is consistent with what other studies have found (LeBlanc et al. 2022). This prompted us to investigate whether the DS study design could be revised to accommodate a reduced number of animals. Our computations suggested that good power (80%) to detect a biologically significant (50%) increase in MF is attainable with as few as three animals and a minimum total depth of 500 million bases per animal. However, to observe more subtle increases in MF, such as that observed in the low PRC dose group, five animals per group is still required. Nevertheless, it appears that DS reduces intragroup variability, thus requiring fewer animals per dose group to detect an effect.

Studies with different variances would require different study designs. To ensure the robustness of our analyses to increased variance, we performed the same analyses with a hypothetical sample variance that was about twice as large as our observed variance (0.1 vs 0.06). We found that a minimum of three animals per group is still sufficient to achieve good power for a 50% increase in MF; however, the minimum number of bases sequenced per sample needs

to be increased to 800 million to achieve this power. Overall, our analysis suggests that DS requires fewer animals per group than the TGR assay even when the variance is increased.

In the context of regulatory decision making, there remain many open questions on how to make a hazard call when targets within a panel vary in sensitivity to mutagenesis. For example, should a significant increase in mutations be observed in a single or a small subset of targets, but not with the overall panel, would this be sufficient for a positive result? Further work using DS to test weak and non-mutagenic chemicals across a broad panel of loci is needed to answer this question. Overall, the use of a specific custom panel of targets is an effective methodology to enable an assessment of the various genomic features that influence a chemical's mutagenic effects.

Conclusion

In summary, this study demonstrates the strong potential of DS for assessing *in vivo* mutagenicity and contributes key data towards efforts to understand optimal study design for its application as a regulatory test. With the inclusion of potentially clonally amplified mutations and the use of highly mutable targets, DS displays comparable fold-induction measurements to the gold-standard *lacZ* assay. In addition, DS provides high resolution mutation spectrum data that enable detailed investigations of mutagenic mechanisms and potential contributions to human carcinogenesis. Further advantages of DS with respect to the TGR assay include: (1) the potential to reduce the number of animals per group necessary to see a significant effect mitigating, in part, the costs associated with DNA sequencing; (2) possible integration with standard toxicity tests, thus negating the need for standalone mutagenicity testing with TGR animals, and further improving the efficiency and reducing costs and animal use for regulatory testing; and (3) application to any tissue from any species, including humans. Future work to establish DS as a mutagenicity test will require testing DS with different chemicals, including weak mutagenic and non-mutagenic chemicals, and examining response across different loci. Further studies are required to quantify relative mutational sensitivity associated with transcription status, chromatin state, and relationship to 3D genomic architecture across different cell types, tissues, and species to increase our understanding of the underlying biology of mutagenesis. Although there is still much work to be done to standardize DS, it shows great promise for regulatory mutagenicity assessment because it overcomes many of the current limitations faced by traditional assays and provides novel insights into chemical mechanisms that underlie cancer.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00204-023-03527-y>.

Acknowledgements The authors thank the extended team at Twin-Strand Biosciences for their work and continuous bioinformatics and administrative support. Additionally, we thank the two Health Canada internal reviewers for helpful comments on the manuscript.

Author contributions Conceptualization: FM, CLY; Methodology: AED, DPML; GZ; Formal Analysis and Investigation: AED, GZ, DPML, AW, MJM, PV, FYL, CCV, JJS; Writing—Original Draft Preparation: AED, CLY, FM; Writing—Review and Editing: All authors; Funding acquisition: FM, CLY; Supervision: FM, CLY.

Funding Open access funding provided by Health Canada.

Data availability Individual animal data available upon request.

Code availability PROAST in R (version 70.1, <https://rivm.nl/en/proast>).

Declarations

Conflict of interest P.V., F. Y. L., C.C.V., and J.J.S. are employees and equity holders at TwinStrand Biosciences, Inc. and are authors on one or more Duplex Sequencing-related patents.

Ethical approval Health Canada Ottawa Animal Care Committee under protocols 2012-017 and 2014-016.

Consent to participate Not applicable.

Consent for publication Not applicable.

Disclaimer Health Canada does not endorse or recommend the products or services of any commercial entity, including TwinStrand Biosciences, Inc.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR (2013) Deciphering signatures of mutational processes operative in human cancer. *Cell Rep* 3(1):246–259. <https://doi.org/10.1016/j.celrep.2012.12.008>
- Alexandrov LB, Kim J, Haradhvala NJ, Huang MN et al (2020) The repertoire of mutational signatures in human cancer. *Nature* 578(7793):94–101. <https://doi.org/10.1038/s41586-020-1943-3>
- Beal MA, Meier MJ, LeBlanc DP, Maurice C, O'Brien JM, Yauk CL, Marchetti F (2020) Chemically induced mutations in a Muta-Mouse reporter gene inform mechanisms underlying human

- cancer mutational signatures. *Comms Biol* 3:438. <https://doi.org/10.1038/S42003-020-01174-Y>
- Chawanthayatham S, Valentine CC, Fedeles BI et al (2017) Mutational spectra of aflatoxin B1 in vivo establish biomarkers of exposure for human hepatocellular carcinoma. *Proc Natl Acad Sci USA* 114(15):E3101–E3109. <https://doi.org/10.1073/PNAS.1700759114>
- Fong LYY, Jensen DE, Magee PN (1990) DNA methyl-adduct dosimetry and O6-alkylguanine-DNA alkyl transferase activity determinations in rat mammary carcinogenesis by procarbazine and N-methylnitrosourea. *Carcinogenesis* 11(3):411–417. <https://doi.org/10.1093/CARCIN/11.3.411>
- Frankish A, Diekhans M, Ferreira AM, Johnson R et al (2019) GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 47(D1):D766–D773. <https://doi.org/10.1093/NAR/GKY955>
- Fu D, Calvo JA, Samson LD (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat Rev Cancer* 12(2):104–121. <https://doi.org/10.1038/NRC3185>
- Gerson SL, Caimi PF, William BM, Creger RJ (2018) Pharmacology and molecular mechanisms of antineoplastic agents for hematologic malignancies. In: Hoffman R, Benz EJ Jr, Silberstein LE, Heslop HE, Weitz JI, Anastasi J, Salama ME, Abutalib SA (eds) *Hematology: basic principles and practice*. Elsevier, Amsterdam, pp 849–912
- Gingerich JD, Soper L, Lemieux CL, Marchetti F, Douglas GR (2014) Transgenic rodent gene mutation assay in somatic tissues. In: Sierra LM, Gaivão I (Eds.) *Genotoxicity and DNA repair*. Springer, New York, pp 305–321. https://doi.org/10.1007/978-1-4939-1068-7_18
- Heddle JA (1999) On clonal expansion and its effects on mutant frequencies, mutation spectra and statistics for somatic mutations in vivo. *Mutagenesis* 14(3):257–260. <https://doi.org/10.1093/MUTAGE/14.3.257>
- Heflich RH, Johnson GE, Zeller A, Marchetti F, Douglas GR, Witt KL, Gollapudi BB, White PA (2020) Mutation as a toxicological endpoint for regulatory decision-making. *Environ Mol Mutagen* 61(1):34–41. <https://doi.org/10.1002/em.22338>. (Epub 2019 Oct 23 PMID: 31600846)
- Hodgkinson A, Eyre-Walker A (2011) Variation in the mutation rate across mammalian genomes. *Nat Rev Genet* 12(11):756–766. <https://doi.org/10.1038/NRG3098>
- Hodgkinson A, Chen Y, Eyre-Walker A (2012) The large-scale distribution of somatic mutations in cancer genomes. *Human Mutat* 33(1):136–143. <https://doi.org/10.1002/HUMU.21616>
- Højsgaard S, Halekoh U (2018) doBy: Groupwise Statistics, LSmeans, Linear Estimates, Utilities (R package 4.6.10).
- IARC Working Group (1981) Procarbazine Hydrochloride. *IARC Monogr* 26:311–340
- Jackson M, Marks L, May GHW, Wilson JB (2018) The genetic basis of disease. *Essays Biochem* 62:643–723. <https://doi.org/10.1042/EBC20170053>
- Jenkins GJS, Doak SH, Johnson GE, Quick E, Waters EM, Parry JM (2005) Do dose response thresholds exist for genotoxic alkylating agents? *Mutagenesis* 20(6):389–398. <https://doi.org/10.1093/MUTAGE/GEI054>
- Kennedy SR, Schmitt MW, Fox EJ, Kohn BF, Salk JJ, Ahn EH, Prindle MJ, Kuong KJ, Shen JC, Risques RA, Loeb LA (2014) Detecting ultralow-frequency mutations by duplex sequencing. *Nat Protoc* 9(11):2586–2606. <https://doi.org/10.1038/NPROT.2014.170>
- Kucab JE, Zou X, Morganello S et al (2019) A compendium of mutational signatures of environmental agents. *Cell* 177(4):21–836.e16. <https://doi.org/10.1016/J.CELL.2019.03.001>
- Lambert IB, Singer TM, Boucher SE, Douglas GR (2005) Detailed review of transgenic rodent mutation assays. *Mutation Res* 590(1–3):1–280. <https://doi.org/10.1016/J.MRREV.2005.04.002>
- LeBlanc DPM, Meier M, Lo FY, Schmidt E, Valentine C, Williams A, Salk JJ, Yauk CL, Marchetti F (2022) Duplex sequencing identifies genomic features that determine susceptibility to benzo(a)pyrene-induced in vivo mutations. *BMC Genom* 23:542. <https://doi.org/10.1186/S12864-022-08752-W>
- Makova KD, Hardison RC (2015) The effects of chromatin organization on variation in mutation rates in the genome. *Nat Rev Genet* 16(4):213. <https://doi.org/10.1038/NRG3890>
- Manders F, Brandsma AM, de Kanter J et al (2022) Mutational Patterns: the one stop shop for the analysis of mutational processes. *BMC Genom* 23:134. <https://doi.org/10.1186/S12864-022-08357-3>
- Marchetti F, Zhou G, LeBlanc D, White PA, Williams A, Yauk CL, Douglas GR (2021) The 28 + 28 day design is an effective sampling time for analyzing mutant frequencies in rapidly proliferating tissues of MutaMouse animals. *Arch Toxicol* 95(3):1103–1116. <https://doi.org/10.1007/S00204-021-02977-6>
- Maurice C, Dertinger SD, Yauk CL, Marchetti F (2019) Integrated in vivo genotoxicity assessment of procarbazine hydrochloride demonstrates induction of Pig-a and lacZ mutations, and micronuclei, in MutaMouse Hematopoietic Hells. *Environ Mol Mutagen* 60(6):505–512. <https://doi.org/10.1002/EM.22271>
- Monroe JG, Srikant T, Carbonell-Bejerano P et al (2022) Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature* 602(7895):101–105. <https://doi.org/10.1038/s41586-021-04269-6>
- Nadkarni A, Burns JA, Gandolfi A, Chowdhury MA, Cartularo L, Berens C, Geacintov NE, Scicchitano DA (2016) Nucleotide excision repair and transcription-coupled DNA repair abrogate the impact of DNA damage on transcription. *J Biol Chem* 291(2):848–861. <https://doi.org/10.1074/JBC.M115.685271>
- Nik-Zainal S, Davies H, Staaf J et al (2016) Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 534:47–54. <https://doi.org/10.1038/nature17676>
- OECD (2020) Test No. 471: Bacterial reverse mutation test. OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. <https://doi.org/10.1787/9789264071247-en>
- OECD (2022a) Test No. 470: Mammalian erythrocyte pig-a gene mutation assay. OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. <https://doi.org/10.1787/4faea90e-en>
- OECD (2022b) Test No. 488. Transgenic rodent somatic and germ cell gene mutation assay. OECD guidelines for the testing of chemicals, section 4. OECD Publishing, Paris. <https://doi.org/10.1787/9789264203907-EN>
- Piegorsch WW, Bailer AJ (1994) Statistical approaches for analyzing mutational spectra: some recommendations for categorical data. *Genetics* 136(1):403–416. <https://doi.org/10.1093/GENETICS/136.1.403>
- Piegorsch WW, Margolin BH, Shelby MD, Johnson A, French JE, Tennant RW, Tindall KR (1995) Study design and sample sizes for a lacI transgenic mouse mutation assay. *Environ Mol Mutagen* 25(3):231–245. <https://doi.org/10.1002/EM.2850250310>
- Pletsa V, Valavanis C, van Delft JHM, Steenwinkel MJST, Kyrtopoulos SA (1997) DNA damage and mutagenesis induced by procarbazine in lambda lacZ transgenic mice: evidence that bone marrow mutations do not arise primarily through miscoding by O6-methylguanine. *Carcinogenesis* 18(11):2191–2196. <https://doi.org/10.1093/CARCIN/18.11.2191>
- Revollo J, Bhalli JA, Tebbe C, Noteboom J, Thomas D, McKinzie P, Felton N, Pearce MG, Dobrovolsky VN (2017) Spectrum of Pig-a mutations in T lymphocytes of rats treated with procarbazine. *Mutagenesis* 32(6):571–579. <https://doi.org/10.1093/MUTAGE/GEX032>

- Salk JJ, Kennedy SR (2020) Next-Generation genotoxicology: using modern sequencing Technologies to assess somatic mutagenesis and cancer risk. *Environ Mol Mutagen* 61(1):35–151. <https://doi.org/10.1002/EM.22342>
- Skopek TR, Kort KL, Marino DR (1995) Relative sensitivity of the endogenous hprt gene and lacI transgene in ENU-treated big blue™ B6C3F1 mice. *Environ Mol Mutagen* 26(1):9–15. <https://doi.org/10.1002/EM.2850260103>
- Skopek TR, Kort L, Marino DR, Mittal LV, Umbenhauer DR, Laws GM, Adams SP (1996) Mutagenic response of the endogenous hprt gene and lacI transgene in benzo[a]pyrene-treated Big Blue™ B6C3F1 mice. *Environ Mol Mutagen* 28(4):376–384. [https://doi.org/10.1002/\(sici\)1098-2280\(1996\)28:4%3c376::aid-em11%3e3.0.co;2-c](https://doi.org/10.1002/(sici)1098-2280(1996)28:4%3c376::aid-em11%3e3.0.co;2-c)
- Souliotis VL, Valavanis C, Boussiotis VA, Pangalis GA, Kyrtopoulos SA (1994) Comparative dosimetry of O6-methylguanine in humans and rodents treated with procarbazine. *Carcinogenesis* 15(8):1675–1680. <https://doi.org/10.1093/CARCIN/15.8.1675>
- Suzuki T, Uno Y, Idehara K, Baba T et al (1999) Procarbazine genotoxicity in the MutaMouse; strong clastogenicity and organ-specific induction of lacZ mutations. *Mutat Res* 444(2):269–281. [https://doi.org/10.1016/S1383-5718\(99\)00060-1](https://doi.org/10.1016/S1383-5718(99)00060-1)
- Valentine CC, Young RR, Fielden MR, Kulkarni R, Williams LN, Li T, Minocherhomji S, Salk JJ (2020) Direct quantification of in vivo mutagenesis and carcinogenesis using duplex sequencing. *Proc Natl Acad Sci USA* 117(52):33414–33425. <https://doi.org/10.1073/PNAS.2013724117/-/DCSUPPLEMENTAL>
- White PA, Long AS, Johnson GE (2020) Quantitative interpretation of genetic toxicity dose-response data for risk assessment and regulatory decision-making: current status and emerging priorities. *Environ Mol Mutag* 61(1):66–83. <https://doi.org/10.1002/EM.22351>
- Xia B, Yan Y, Baron M et al (2020) Widespread transcriptional scanning in the testis modulates gene evolution rates. *Cell* 180(2):248–262.e21. <https://doi.org/10.1016/J.CELL.2019.12.015>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.