A random mutation capture assay to detect genomic point mutations in mouse tissue

Jocelyn H. Wright[1],*, Kristina L. Modjeski[1], Jason H. Bielas[2], Bradley D. Preston[1], Nelson Fausto[1], Lawrence A. Loeb[1] and Jean S. Campbell[1]

[1]Department of Pathology, University of Washington and [2]Department of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

ABSTRACT

Herein, a detailed protocol for a random mutation capture (RMC) assay to measure nuclear point mutation frequency in mouse tissue is described. This protocol is a simplified version of the original method developed for human tissue that is easier to perform, yet retains a high sensitivity of detection. In contrast to assays relying on phenotypic selection of reporter genes in transgenic mice, the RMC assay allows direct detection of mutations in endogenous genes in any mouse strain. Measuring mutation frequency within an intron of a transcribed gene, we show this assay to be highly reproducible. We analyzed mutation frequencies from the liver tissue of animals with a mutation within the intrinsic exonuclease domains of the two major DNA polymerases, δ and ε. These mice exhibited significantly higher mutation frequencies than did wild-type animals. A comparison with a previous analysis of these genotypes in Big Blue mice revealed the RMC assay to be more sensitive than the Big Blue assay for this application. As RMC does not require analysis of a particular gene, simultaneous analysis of mutation frequency at multiple genetic loci is feasible. This assay provides a versatile alternative to transgenic mouse models for the study of mutagenesis *in vivo*.

INTRODUCTION

DNA mutations have been shown to be the underlying cause of many diseases including cancer (1,2). Mutations can be caused by endogenous oxidative DNA damage, damage from genotoxic agents or by error-prone DNA replication or repair. Measuring the mutation frequency in tissues is essential to understanding the mechanisms of mutagenesis *in vivo*, as well as for measuring the genotoxicity of drugs and compounds. Cellular mutation rates can be measured in immortalized fibroblasts (3) or lymphocytes (4) by fluctuation analysis. However, this method is only amenable to cell types that can be explanted from the animal, and the mutations being measured are generated *in vitro* in immortalized cells grown on plastic, which may behave differently than cells within tissues *in vivo*.

A sensitive method called random mutation capture (RMC) was developed and used to study the frequency of random point mutations in tumor and non-tumor tissues (5,6). RMC is a genotypic selection technique in which genomic DNA is digested with a restriction enzyme, followed by single-copy template PCR designed to specifically amplify molecules with mutations that eliminate a particular restriction site. This method allows one to capture and selectively sequence rare mutations. The RMC concept is based on earlier techniques called Restriction Site Mutation (RSM) and restriction fragment length polymorphism–polymerase chain reaction (RFLP–PCR) assays (7,8).

The standard methods in use for measuring the frequency of random mutations *in vivo* use transgenic mouse models developed for this purpose. The Big Blue (9,10) and MutaMouse (11) transgenic models contain lambda phage vectors inserted into the mouse genome in a tandem array. Mutations in reporter genes within these vectors are measured by phenotypic selection after they are shuttled out of the mouse genomic DNA, packaged into phage and used to infect bacteria. As these assays use engineered mouse strains, the transgene must be genetically crossed into the mouse genotype(s) of interest before experimentation. In contrast, the RMC assay is a direct mutation detection method that can be done in any mouse model without genetic cross or purchase of proprietary mice and associated reagents, making it potentially faster and less expensive than using transgenic mouse models.

---

*To whom correspondence should be addressed. Tel: +1 206 616 4796; Fax: +1 206 543 3967; Email: jhw5@uw.edu

Mutations in intron 6 of the human Trp53 locus were previously captured using the RMC assay (5,6). Cultured human cells had a very low background mutation frequency, and demonstrated an increase in mutation frequency with mutagen treatment of the cells that was linear with dose (5). The target site demonstrated genetic neutrality in that there was no selection for or against mutations over time (5). RMC analysis of five human tumor and normal tissues indicated a large increase in point mutation frequency in tumor as compared to normal tissue (6).

Here, we present a detailed protocol for using RMC to measure nuclear mutation frequency in mouse tissue. In addition to adapting the method for use in mice, we have significantly simplified the original method (5). Removal of a lengthy and technically difficult enrichment step allows this protocol to be faster than the original method, while retaining a high sensitivity. In general, no mutations were detectable in young wild-type mice. In one mouse where a single mutation was detected, 55 million base pairs of DNA were screened giving a mutation frequency of less than 1×10^{-8} mutations/bp. The frequency of mutations detected in normal tissues was similar to that detected in other endogenous genes, and less than that measured by transgenic mouse assays (12).

DNA polymerase δ (Pol δ) and DNA polymerase ε (Pol ε) are the two polymerases in the cell that are responsible for replication of the bulk of nuclear genomic DNA during S phase. These polymerases contain an intrinsic 3′ to 5′ exonuclease domain (13). Experiments where this domain was mutated have shown that this domain functions to ensure the fidelity of the replication process in yeast (14,15). Homozygous knock-in mice that express Pol δ or Pol ε with a point mutation within this exonuclease domain were created and shown to have higher incidence of cancer (12,16). *In vitro* fluctuation analysis performed on immortalized mouse embryo fibroblasts from these mice demonstrated that mutation of the exonuclease domains in either of these DNA polymerases gave much higher mutations rates (12). In addition, these mutant mouse stains were crossed into a Big Blue mouse background and the mutation frequency within the *cII* transgene was analyzed. In this *in vivo* mutagenesis assay, these mice demonstrated measurable increases in mutation frequency in four different tissues (12). Herein, we have analyzed these same strains of mice to determine how mutation frequency was affected as measured using our new RMC assay. Data in the results section illustrate that, by RMC, the mice with the polymerase mutant genotypes showed a greater increase in mutation frequency over wild type than was seen in the Big Blue assay. Additional findings based on analysis of the mutation spectra are also presented.

MATERIALS AND METHODS

Genomic DNA isolation from mouse tissue

Obtain fresh tissue, or store tissue in vials submerged in liquid nitrogen prior to DNA isolation (>100 mg best). DNA damage is minimized during isolation by not using phenol, and by the addition of anti-oxidants and iron chelators to the lysis buffer (17). Use Roche kit for DNA isolation (Cat # 11814770001), which uses the high salt method for DNA purification. Our procedure follows the kit procedure precisely, except for addition of the following anti-oxidants to the isolation buffer: 4 mM histidine, pH 7.0 (Sigma catalog# 151688; make 100 mM stock in water and store at 4°C); 3 mM reduced glutathione (Sigma catalog# G4705; make 300 mM stock in water and store at −20°C in aliquots) and 1 mM desferrioxamine (Sigma catalog# D9533; make fresh 50 mM stock in water). Place tissue (fresh or frozen in $N_{(l)}$ and quick thawed) and lysis buffer in 50 ml screw cap centrifuge tube at room temperature. Grind tissue using a tissue homogenizer in 5–10 s pulses until all macroscopic pieces are gone. Treat with Proteinase K and RNAse as described in the kit. Transfer to 40 ml polypropylene, high speed centrifuge tubes prior to the addition of protein precipitation solution and spin to remove insoluble protein as described.

When transferring the DNA containing-supernatant to clean 30 ml Corex tube, do not transfer the last bit of volume in order to minimize the transfer of any particulate material as it will interfere with TaqI digestion later. Precipitate DNA with isopropanol and wash pellet with 70% EtOH as described in kit. Resuspend DNA in 1.4 ml TE per 400 mg original wet tissue weight (this will give a [DNA] = ∼1 mg/ml for liver). Spin samples in a microfuge to pellet any remaining insoluble matter and quantify using a spectrophotometer. A260/280 should be ≥1.8.

Generation of genomic DNA standard curve for DNA copy number calculations

The number of copies per well in the standard curve dilutions are first approximated by spectra photometric absorbance at 260 nm (OD_{260}). Measure the OD_{260} of an uncut total genomic DNA sample to obtain a DNA concentration. Genomic DNA rather than plasmid DNA should be used for the standard curve. Divide the concentration of your DNA expressed in pg/μl by the molecular weight of the mouse genome (2.74 pg/copy) to determine the approximate number of copies per microliter in your DNA sample. Make your first dilution well above the concentration you will use per well in your quantitative PCR (qPCR) with experimental primers. We recommend starting with 200 000 copies/5 μl. Make a 4-fold dilution series for 11 dilutions. Mix sample thoroughly before making the next dilution. (Note: Make a large volume of each dilution—∼1 ml per dilution final volume—and freeze in single use aliquots.) Store your dilutions at 4°C until validated as freeze–thaw cycles alter the effective DNA concentration.

Test your dilution series for linearity in a real-time PCR machine with a fluorescence detector (Bio-Rad Opticon 2, catalog# CFB-3120 or comparable machine) using the control primers (Forward: 5′-CTCACCAAAAACAAAA ACAGCC-3′, Reverse: 5′-CTTTTGTCCCTCCCACTT TGG-3′). For each well in a 96-well plate (Bio-Rad catalog# HSP9655), add 12.5 μl SYBR green master mix (Applied Biosystems catalog# 4309155), 0.5 μl

Uracil–DNA glycosylase (UDG, New England Biolabs catalog# M0280L), 0.5 µl of 15 µM stock of control primers (Fwd and Rev), 6.5 µl irradiated water and 5 µl of diluted DNA. Seal the plate with a plastic seal (Bio-Rad catalog# MSB1001), and use the following PCR protocol:

(i) $37°C$ 10 min
(ii) $95°C$ 10 min
(iii) $95°C$ 30 s
(iv) $60°C$ 30 s
(v) $74°C$ 10 s
(vi)   Read plate.
(vii) Repeat steps (iii)–(vi) for 47 cycles.
(viii) Run melting temperature curve to determine relative $T(m)$ of products.

You likely will have to discard a few points from both ends of the series, but six or seven dilutions should fall on a line in a graph plotting $C_{(T)}$ versus log (copy number) as shown in Figure 1. If the series is linear, proceed with validation by limiting dilution.

In order for the absolute copy number to be accurate, the approximate DNA copy number values determined by optical density are then corrected by limiting dilution analysis and estimation of copy number by Poisson distribution. For limiting dilution analysis, dilutions that approximate 1–10 copies per well are tested in multiple wells. The number of wells without amplification (those that have 0 copies) are used to calculate the average copy number per well by Poisson analysis. Select two dilutions from your genomic DNA standards (the last one the amplified, and the first one that did not amplify—these two should be 1–10 copies/5 µl) and run qPCR reactions on multiple wells (at least 36 for each of the two dilutions). Calculate resultant average copy number per well using the Poisson equation: average copy number/well = −ln [negative wells/total wells]. Multiply average copy number/well by a factor of 4 in series to determine absolute copy number per well for each DNA standard in the curve. See Figure 1 for an example of absolute copy number values calculated for a standard curve. Freeze your standard curve samples in single use aliquots at $−80°C$ for use on all subsequent experimental samples that will be analyzed together. Test one of your frozen aliquots before proceeding with control primers and mutant-specific primers side-by-side to ensure that both primer sets amplify with equal efficiency. The $C_{(T)}$ values for a given dilution should be similar for both sets of primers with these uncut genomic DNA samples (Ideally, within $0.5\,C_{(T)}$ values of each other) for the mutation frequency calculation to be accurate.

TaqI digestion of genomic DNA

High concentration $Taq^{\alpha}I$ (100 000 U/ml; New England Biolabs catalog# R0149M) is used to maximize enzyme concentration in the reaction. $Taq^{\alpha}I$ is a mutant form of the enzyme that has a two amino acid replacement at its amino terminus. This allows for a higher level of expression without interfering with its catalytic properties. Achieving sufficient digestion of genomic DNA with TaqI is a critical step in this protocol. The quality control testing provided by the supplier is designed to ensure sufficient digestion efficiency for analysis of digested DNA in bulk. For our application, where DNA molecules are analyzed individually, we have found that the quality of $Taq^{\alpha}I$ can vary between lots. Small quantities of several lots of $Taq^{\alpha}I$ enzyme should be tested first with one DNA sample and then the best lot then acquired in an adequate amount to complete the experiment (Figure 2).
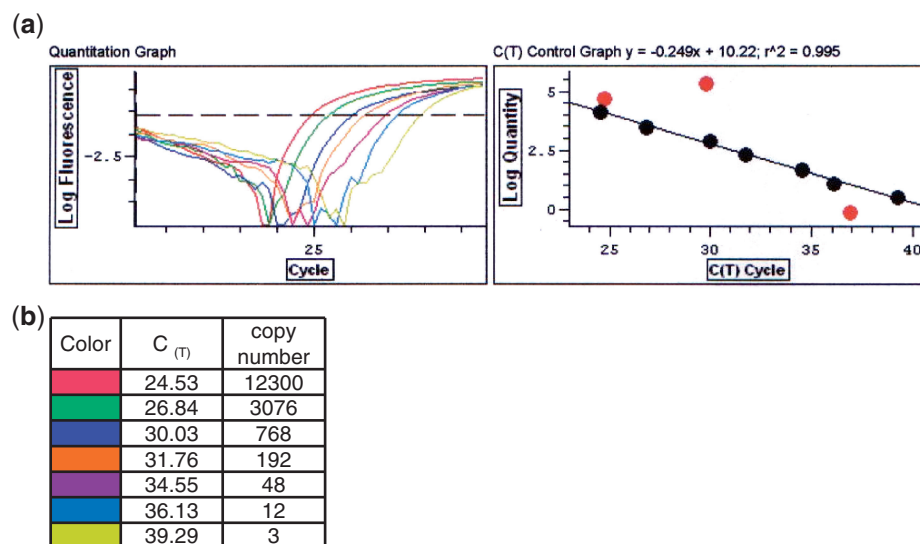
Color	$C_{(T)}$	copy number
	24.53	12300
	26.84	3076
	30.03	768
	31.76	192
	34.55	48
	36.13	12
	39.29	3

Figure 1. DNA standard curve for normalization of total genome copy number between tissue samples. (**a**) A typical display for a qPCR performed on standard curve DNA samples run on the Opticon real-time PCR system. In the graph relating $C_{(T)}$ to log quantity, the data points in black are the samples that behaved in a linear fashion and were used to generate the standard curve, and those is red were not in the linear range and therefore not used. (**b**) The $C_{(T)}$ values and copy number values (calculated from limiting dilution analysis) are shown in tabular form, color coded to correspond to the Opticon graph of amplification curves shown in (a).
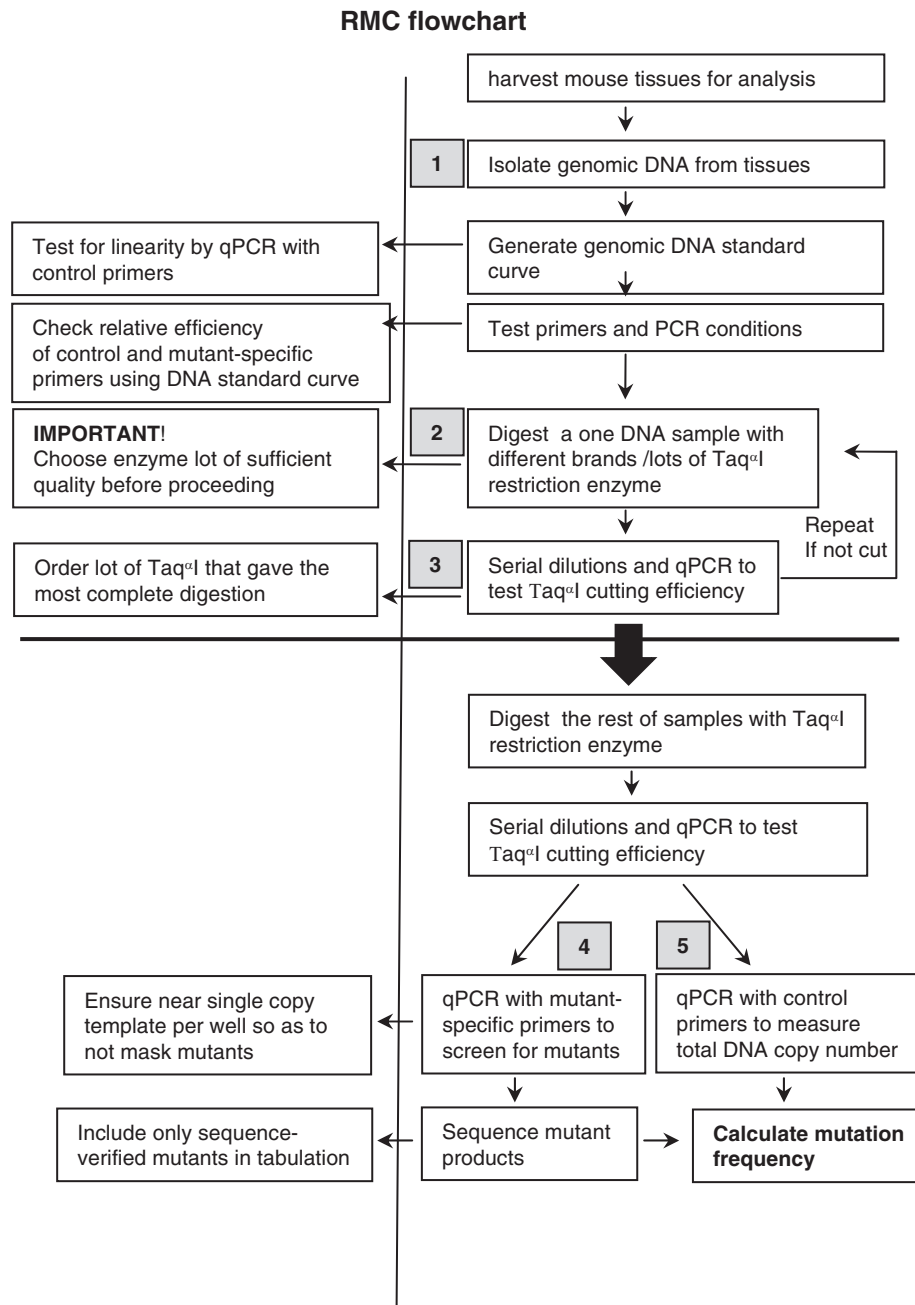
RMC flowchart

Figure 2. Flow diagram of RMC protocol. The sequence of experimental steps is shown with important considerations for each of those steps shown on the left. Boxed numbers refer to the same numbers step illustrated in Figure 3. A troubleshooting guide for this protocol is available in Supplementary Table I.

TaqI digestion can be done all at once, or over a period of several days, storing the samples at 4°C between digestions. Use one 1.7 ml Eppendorf type tube with 64 μg DNA for each sample, bring up the reaction volume to 800 μl total volume with NEB4 buffer (10X), BSA (100X) and double-deionized water. Add 8 μl of Taq^{α}1 (100 000 U/ml) to each tube and digest at 65°C for 1 h, shaking at 1000 rpm in an Eppendorf thermomixer or in a water bath. Add more Taq^{α}I every 60 min for seven more hours. Spin samples in a microfuge for 30 s prior to each Taq^{α}I addition to remove liquid that collects on the lid. A precipitate may form due to denaturation and precipitation of the protein in the reaction. With the fourth addition or if the samples were held at 4°C overnight, give samples a quick microfuge spin, transfer the samples to new tubes and add fresh BSA along with the Taq^{α}I (to replace denatured, precipitated BSA). (Note: loss of BSA carrier protein reduces the efficiency of Taq^{α}I digestion.) Next, spin samples in a microfuge and transfer to Microcon YM-50 concentrator tubes (Millepore catalog# 42416). (Note: do not transfer any pellet.) Follow the protocol located in the Microcon
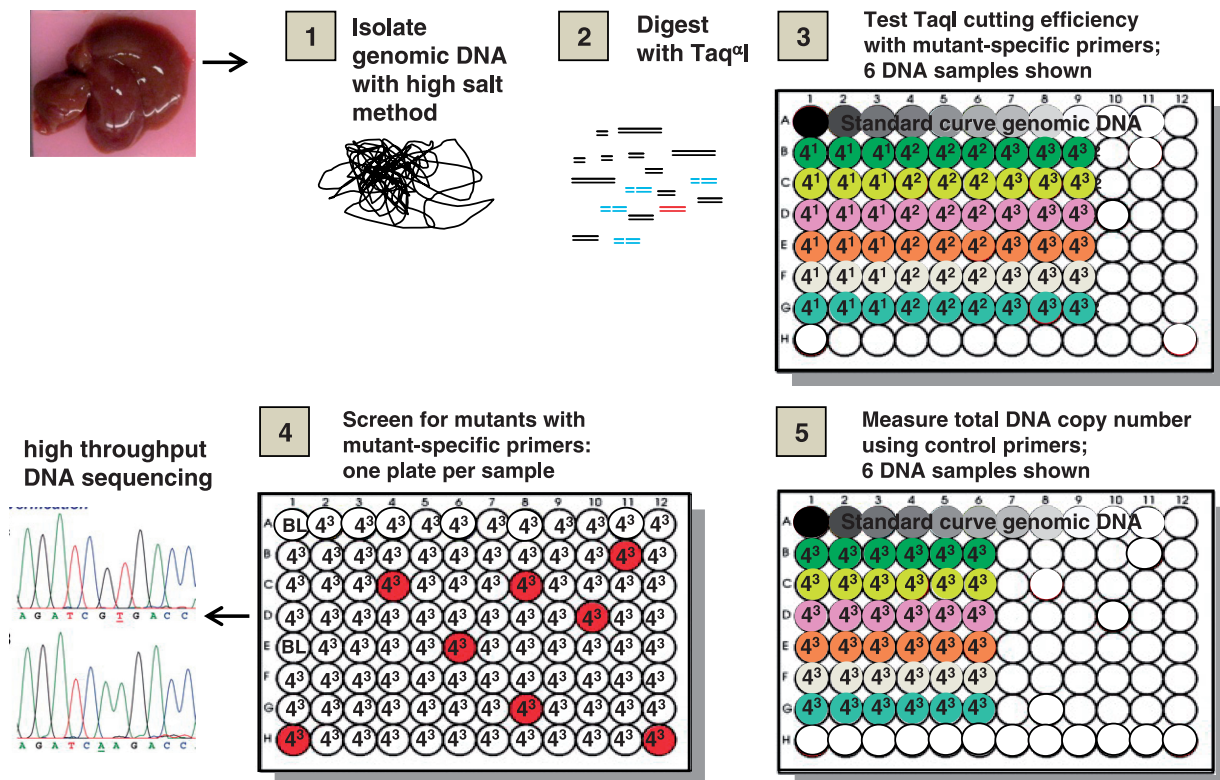
Figure 3. RMC assay: pictorial illustration of experimental steps. Photograph of a normal mouse liver is shown. Individual steps in the protocol are indicated with boxed numbers. In step 2, target sequences in the digested DNA are color coded with blue for wild type and red for mutant. Steps 3, 4 and 5 illustrate a typical experimental design using 96-well qPCR plates. The dilution factors suggested to test are shown in powers of 4 in the wells. In steps 3 and 5, the DNA standard curve is shown in gray, and analysis of six different DNA samples are shown in different colors. In step 4, wells marked 'Blk' are blanks to test for DNA contamination during the mutant screening step. Numbered experimental steps shown relate to numbers in the experimental flow chart in Figure 2.

booklet to buffer-exchange the DNA into water, and to concentrate the DNA; wash $2 \times$ with $400 \, \mu l$ double-deionized water before concentrating the DNA to $\sim 100 \, \mu l$. Bring sample up to $800 \, \mu l$ again in 1X TaqI cutting buffer (i.e. NEB4 plus BSA). Add fresh $Taq^\alpha I$ enzyme again for two more 60-min incubations at $65°C$. Perform a buffer-exchange again in the Microcon YM-50 tubes with $2 \times 400 \, \mu l$ washes with UV-irradiated double-deionized water, and then concentrate the DNA to a final volume of $\sim 100 \, \mu l$. (Note: do not over-concentrate as it leads to variable yields.) UV-irradiated, double-deionized water is prepared by irradiating $1.0 \, ml$ aliquots in open Eppendorf tubes in a hood or box equipped with short-wave UV light source for 15 min. Irradiation destroys any contaminating uncut mouse genomic DNA. Finally, dilute samples to $800 \, \mu l$, typically by adding $\sim 700 \, \mu l$ UV-irradiated TE (10 mM Tris, pH 8.0; 1 mM EDTA) and store samples at $4°C$. (Note: do not freeze DNA samples as it alters the effective DNA concentration.) This is considered the first 4-fold dilution (4^1) as described in the next section.

Testing for TaqI digestion efficiency

Quantitative PCR (qPCR) reactions designed to test the efficiency of TaqI digestion are performed prior to the initiation of screening for mutants. An illustration of a typical qPCR plate used to determine TaqI cutting efficiency is shown in Figure 3. Post-PCR restriction analysis of products with TaqI is used to differentiate between wild-type and mutant products. Using samples from the TaqI digestion protocol above make three 4-fold serial dilutions with $200 \, \mu l$ DNA and $600 \, \mu l$ 1X TE (4^2, 4^3, 4^4). Mix samples thoroughly between dilutions to make the series linear. Serial dilutions of the digested DNA are tested with mutant-specific primers (Forward: 5'-CTTCTCACCAAAA ACAAAAACAGCC-3', reverse: 5'-ATGGCCAACACACC TGACTTCTC-3') by qPCR. [Note: the target TaqI site used in our protocol in the same one used previously (18); however, we used different primers that were optimized for this new protocol.] Test triplicate wells with each dilution (4^2–4^4) for each sample. Run in a 96-well plate format, using each of the following per well: $12.5 \, \mu l$ SYBR Green master mix, $6.75 \, \mu l$ UV-irradiated water, $0.5 \, \mu l$ UDG, $0.25 \, \mu l$ of $15 \, \mu M$ stock of each primer, along with $5 \, \mu l$ of diluted DNA sample. Contamination with mouse genomic DNA is always a potential issue with single copy template PCR so always use barrier tips, and change gloves and aliquots of reagent frequently, and work, if possible in a hood equipped with a short-wave UV light source and irradiate working area prior to the start of each experiment.

QPCR products generated with ABI SYBR green mix contain some UDP in place of TDP. Addition of UDG to the reaction removes any contamination with previously generated PCR products. The qPCR protocol to be used with the mutant-specific primers is as follows:

  (i) 37°C 10 min (allows for digestion of UTP containing DNA).
 (ii) 95°C 10 min
(iii) 95°C 30 s
 (iv) 62°C 60 s
  (v) 74°C 10 s
 (vi) read plate.
(vii) repeat steps (iii)–(vi) for 47 cycles.
(viii) run a melting temperature curve to determine relative $T_{(m)}$ of products

When the reaction is done, place qPCR plate immediately at −20°C upon completion as UDG will digest products even at 4°C.

In wells where product is seen, verify that you have the correct product by examining the melt temperature curve. The correct melting temperature [$T_{(m)}$ for products with mutant-specific primers is 81.5–82°C. Wells with correct $T_{(m)}$ product contain either wild-type or mutant product. In order to distinguish between wild-type and mutant products, perform a post-PCR TaqI digestion and agarose gel electrophoresis analysis, using wild-type PCR products with and without TaqI digestion as molecular weight markers. Digest 8 μl of the each PCR reaction with 0.25 μl Taq^{α}1 (NEB, 100 000 U/ml), 0.5 μl Uracil–DNA glycoslyase inhibitor (UGI, New England Biolabs catalog# M0281L), 0.2 μl NEB4 buffer and 1.0 μl double-deionized water added per reaction. Incubate for 10 min at 65°C, 1 min at 95°C and then 1 min at 25°C (digestions are easiest to do in a PCR machine). Analyze digested sample using 2% agarose gel electrophoresis. Mutant products have one band at 146 bp, while wild-type products have two bands at 96 and 50 bp. If the products are all wild-type in this test, then your sample is not sufficiently digested for analysis. Go back and repeat the TaqI digestion step. Otherwise, you will waste a lot of reagents and time sequencing wild-type molecules that were not cut prior to the mutant screening step. We recommend that no more than 10 wild-type molecules should be detected per mutant. If mutant products are detected, you can use this qPCR analysis of the dilution series to determine which dilution will give you approximately one mutant copy per well to use for the mutant screening (usually 4^2–4^3 in our hands). A dilution that approximates one copy of template per well is often the dilution where no amplification had occurred; amplification will only occur 37% of the time based on a Poisson distribution. As a rule of thumb, choose the last dilution where some amplification occurred.

Running qPCR plates to screen for mutants

Set-up and analyze one full plate for each DNA sample using the DNA dilution determined in qPCR plate for TaqI cutting efficiency test described above. Use the same PCR protocol with the mutant-specific primers described above.

Include 94 wells, all at the same DNA dilution and two blank wells with water in place of the DNA template to test for DNA contamination. If blanks are contaminated, data obtained with this plate must be discarded and sample re-run with new reagents. Freeze plates immediately at −20°C after the run as UDG can function quite well at 4°C and will degrade your products.

Analyze the plate for potential mutants. In all wells where product is seen, verify that you have the correct product by examining the melt temperature curve. QPCR analysis with too high of a template copy number can cause the masking of mutations with either wild-type product or other spurious products if the PCR is not optimized. To avoid masking mutations, no more that 59 wells in each 96-well plate should be positive for product to insure near single copy template in every well. Determine the sequence of material from all wells with a product of the correct melting temperature to identify mutations. In our experience, given the low rate at which mutant molecules are present in the excess of wild-type sequences in genomic DNA, a significant number of the PCR products obtained will be wild-type in spite of extensive TaqI digestion.

Unpurified PCR products can be directly sequenced in most high-throughput sequencing facilities in which case, 2 μl or each positive well can be directly sequenced with either the forward or reverse PCR primer. Add 0.5 μl of UGI to each sequencing reaction to inhibit residual UDG activity in sample which will degrade the template. In the case where low-throughput or costly sequencing is the only option, screening samples by post-PCR TaqI digestion and gel analysis as described above can reduce the number of samples to be sequenced.

Determining total DNA copy number

Set up a 'total copy number plate' using the control primer pair and the same PCR protocol described above for the generation of the DNA standard curve. Run each plate with a standard curve to allow you to convert the $C_{(T)}$ values into DNA copy number per well (Figure 1). Up to 14 DNA samples can be analyzed on one plate for determining the total copy number. An example of a total copy number calculation plate is shown in Figure 3, step 5. Run six wells of each DNA sample, using the same dilution that was used in the mutant screening plate. Take the average DNA copy number calculated for the six-well replicates and use this as your total DNA copy number/well for that sample. Calculate mutation frequency: mutation frequency = (number of sequence verified mutants) / (total copies per well) × (number of wells screened) × (4 bp per copy).

Given the stochastic nature of mutagenesis as well as the very low frequency of mutation, multiple mutations need to be detected before a difference between cohorts can be considered real. If two sets of DNA samples being compared were to have, for example, similarly low frequencies as our young wild-type mice presented in the results section, more base pairs would need to be screened to

determine if there was a difference between cohorts. As a rule of thumb, if only 0 or 1 mutation were detected per mouse in both sets of samples to be compared, more base pairs must be screened until multiple mutants are obtained for at least one cohort. A Mann–Whitney statistical test can be used to assess statistical significance.

RESULTS

Description of the method

A TaqI restriction site located in intron IV of the Trp53 locus on chromosome 11 was used to develop the mouse RMC assay. The target site is illustrated in Figure 4. This site is located in an intronic region of no known function that is poorly conserved across species, suggesting that it is likely to be genetically neutral. Two pairs of primers from the Trp53 Intron IV target sequence were selected for use that have equivalent amplification efficiencies, one pair adjacent to the target TaqI site (control primers) and one pair flanking the TaqI site (mutant-specific primers). Total genomic DNA is extensively digested with TaqI restriction enzyme. The frequency of uncut wild-type TaqI sites that remain in the genomic DNA after digestion must be very low in order for rare mutations to be detectable. If the frequency of uncleaved, wild-type products captured is not in the same range as the frequency of mutant products, the vast majority of molecules captured and sequenced will be wild type making detection of mutants laborious.

TaqI digested genomic DNA is screened by qPCR with mutant-specific primers to amplify mutant molecules. In order to be able to sequence mutant products, the mutant screening must be done at high dilution of the DNA template, such that only one or two template copies are present per well. An illustration of a mutant-screening plate is shown in Figure 3. TaqI-resistant products are verified as mutants by direct sequencing of PCR products. The total number of genomes that was screened for mutants is measured in parallel for each DNA sample. QPCR analysis is performed using the control primers, which do not flank the TaqI site, on the same DNA dilution that was used for the mutant screening. An illustration of a typical total copy number plate template is shown in Figure 3. As control primers will amplify all
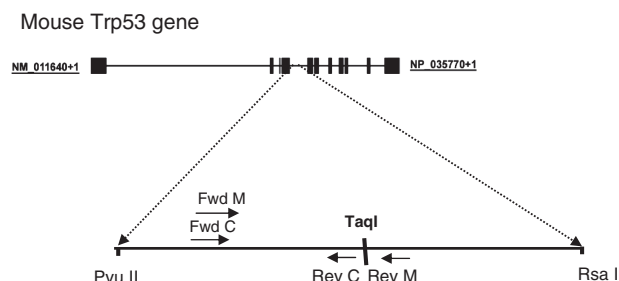
genomic copies present in each well, the number of total DNA molecules screened per well is determined. A DNA standard curve is used to convert cycle threshold [$C_{(T)}$] values to DNA copy number. An example of a standard curve used for this conversion is shown in Figure 1. As the accuracy of the final mutation frequency number is dependent on the accuracy of the total DNA copy number, it is recommended that a significant effort be made to construct a robust set of DNA standards to be used on all plates throughout the experiment.

The final result is a mutation frequency expressed as mutations/bp. This number is calculated as: (number of sequence verified mutants)/(total copy number per well) × (number of wells screened) × (4 bp)

Reproducibility of RMC analyses

We analyzed a cohort of 9-month-old, wild-type C57BL/6 mice using the RMC assay. Mice were sacrificed, and the left lobe of each mouse liver was harvested and quickly frozen. Genomic DNA was isolated from tissue pieces between 200 and 400 mg in weight. In independent experiments performed 6 months apart, the same DNA samples were digested with two different lots of Taq^αI enzyme, and analyzed with different batches of qPCR reagents. Roughly one-fourth of a liver was homogenized to generate each DNA sample, and about 1/500 of that DNA sample was analyzed in each of the two independent experiments. Figure 5 shows a comparison of the mutation frequencies calculated for the mice from the two different experiments. Each bar represents the mutation frequency calculated for the individual mice.

It has been suggested that, as the mutant screening in the RMC assay is performed at limiting dilution, significant sampling error is introduced (19). We observed that the frequency measured for a given mouse did vary between the experiments indicating some sampling error, yet the relative difference between the mice was reproducible (Figure 5).

The median mutation frequency for these mice was 5.3×10^{-7} mutations/bp, with >10-fold variation in frequency between individual mice. These 9-month-old mice

Figure 4. Target sequence in Trp53 locus used for RMC assay. A schematic drawing of the mouse Trp53 locus with the location of RMC target sequence is shown. Location of the TaqI restriction site is indicated, and control forward (Fwd C), control reverse (Rev C), mutant-specific forward (Fwd M) and mutant-specific reverse (Rev M) primers are designated with small arrows.
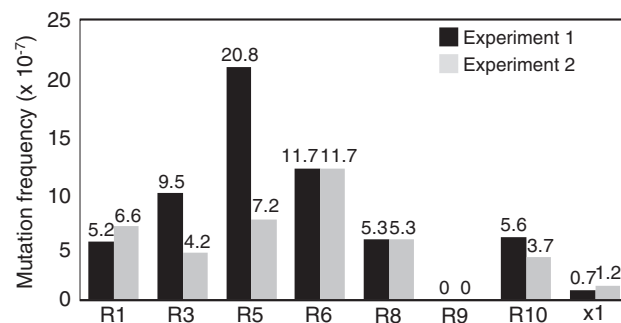
Figure 5. The reproducibility of data from two RMC analyses of 9-month-old wild-type C57Bl/6 mice. Results from two different experiments performed on the same DNA samples are presented. Results from experiment 1 are shown in black, and those from experiment 2 are shown in gray. Each bar represents mutation frequencies calculated for individual mice. Mouse 'code' names are indicated below each pair of data bars.
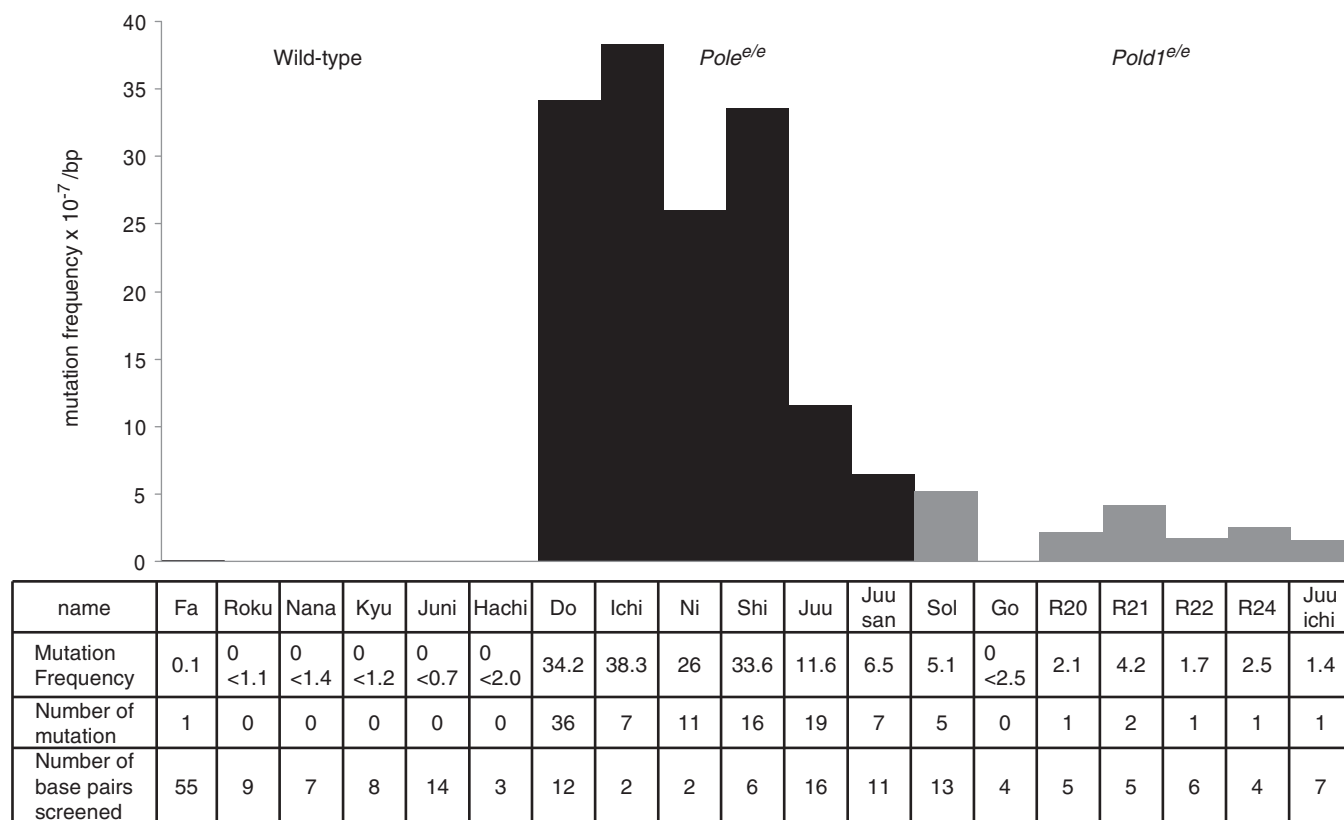
name	Fa	Roku	Nana	Kyu	Juni	Hachi	Do	Ichi	Ni	Shi	Juu	Juu san	Sol	Go	R20	R21	R22	R24	Juu ichi
Mutation Frequency	0.1	0 <1.1	0 <1.4	0 <1.2	0 <0.7	0 <2.0	34.2	38.3	26	33.6	11.6	6.5	5.1	0 <2.5	2.1	4.2	1.7	2.5	1.4
Number of mutation	1	0	0	0	0	0	36	7	11	16	19	7	5	0	1	2	1	1	1
Number of base pairs screened	55	9	7	8	14	3	12	2	2	6	16	11	13	4	5	5	6	4	7

Figure 6. Relative mutation frequency for individual mice from *Pole*[e/e], *Pold1*[e/e] and wild-type C57BL/6 genotypes. Mice with different genotypes are shown in different colors; *Pole*[e/e] are black, *Pold1*[e/e] in gray and wild type in light gray. Each bar represents data for an individual mouse, which was assigned a 'code' at necropsy to blind the experimenter to genotype during analysis. Mouse names are indicated below each datum bar, along with the mutation frequency, the number of sequence-verified mutations identified and the total number of base pairs (in millions) screened for that individual mouse.

had significantly higher mutation frequencies than did other C57BL/6 mice at 2 months of age shown in Figure 6 and described in the following section. These results suggest that mice accumulate random mutations with age as has been shown previously using other methods (20,21). Based on these results, we recommend that at least six mice be tested per cohort and that the cohorts be age matched.

Point mutation frequencies in Pol δ and Pol ε exonuclease-deficient mice

To directly compare the mouse RMC assay to an established method, we analyzed DNA from mice with proof-reading deficient DNA polymerase ε (*Pole*[e/e]) and DNA polymerase δ (*Pold1*[e/e]). Previously, these mice were crossed with C57BL/6 Big Blue mice and analyzed in a Big Blue assay (12). Elevated mutation frequencies were measured within the *cII* transgene for both *Pole*[e/e] and *Pold1*[e/e] Big Blue mice as compared with wild-type littermates (12).

Cohorts of six to seven mice, 6–8 weeks of age, were sacrificed from *Pole*[e/e] and *Pold1*[e/e] colonies along with wild-type littermates. DNA was isolated from the left lobe of the liver, digested with Taq[α]I and diluted for qPCR analysis. For each sample, greater than or equal to four million base pairs were screened with

mutant-specific primers. Results of this analysis are shown in Figure 6. Each bar represents the mutation frequency calculated for the individual mice. Multiple mutations were detected in all the *Pole*[e/e] mice and in six out of seven *Pold1*[e/e] mice. These high mutation frequencies in the mutant mice were in sharp contrast to wild-type mice where one mutation was detected in only a single individual. Thus, elimination of proofreading activity from either of the major DNA polymerases caused a significant increase in mutation frequency within the RMC target TaqI site (Figures 6 and 7). The majority of the wild-type mice had no mutations identified giving them mutation frequencies of zero. We included less-than values shown at the bottom of Figure 6 which reflect the number of total base pairs screened for each mouse. In order to obtain a background mutation frequency value for a wild-type mouse, 55 million base pairs were screened from one sample until a mutation was identified. The calculated mutation frequency for that one wild-type mouse was 9×10^{-9} mutations/bp.

The median mutation frequency in the *Pole*[e/e] mice was significantly higher than in *Pold1*[e/e] mice, and both were significantly higher than wild type (Figure 7a). Mutation frequencies do not have a normal distribution and therefore a standard *t*-test for parametric data would not be applicable. The statistical test we used to analyze these
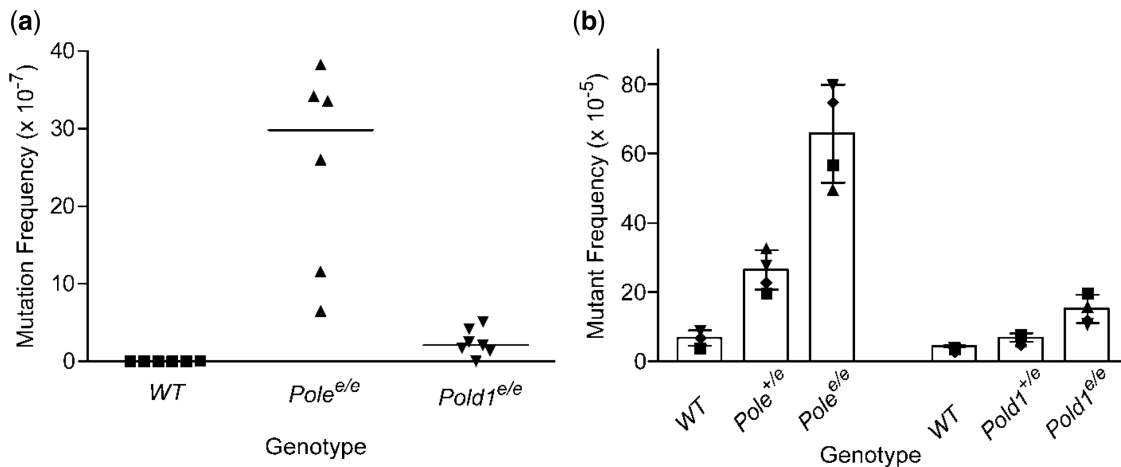
Figure 7. Comparison of RMC analysis of mutation frequency in Trp53 intron IV with analysis of frequency in the *cII* transgene using Big Blue mice. (**a**) RMC data from the 6- to 8-week-old mice with the different genotypes are shown. Frequencies for individual mice are presented as points around the median value that is indicated as a horizontal line. Median values are as follows: 0.0 for wild type, 29.8 for *Pole*$^{e/e}$ and 2.1 for *Pold1*$^{e/e}$. The median mutation frequencies were found to be statistically different between both mutator mouse genotypes and wild type: wild type versus *Pole*$^{e/e}$, P = 0.003; wild type versus *Pold1*$^{e/e}$, P = 0.009 and *Pole*$^{e/e}$ versus *Pold1*$^{e/e}$, P = 0.001. (**b**) Big Blue mouse data from a previous study (12) on 6- to 8-week-old mice of the different genotypes. Mutation frequencies measured in four different tissues are shown: small intestine (squares), lung (diamonds), thymus (triangles), bone marrow (upside down triangles). The three bars on the left show the effect of a mutation in DNA polymerase epsilon, and the three bars on the left show the effects of a mutation in DNA polymerase delta.

data was a Mann–Whitney test with a correction for ties. By performing the 'asymptotic' test rather than the 'exact' test, the calculation accounts for tied values. We chose to use this test given the large number of zero values obtained for the wild-type cohort of mice. Analysis was performed using the SPSS 18 statistical program (IBM). In Figure 7b, we present a reprinted version of Figure 2A from Albertson *et al.* (12), where they measured the mutation frequency in the phage *cII* gene from *Pole*$^{e/e}$ and *Pold1*$^{e/e}$ mice using the Big Blue mouse assay. In this side-by-side comparison, we see that the relative mutation frequencies measured in the liver by RMC analysis (Figure 7a) were very similar to those measured in four different tissues using the Big Blue mouse assay (Figure 7b). *Pole*$^{e/e}$ mice had the highest frequency, *Pold1*$^{e/e}$ mice had a lower frequency, and wild-type mice had the lowest.

When the median values for the *Pole*$^{e/e}$ and *Pold1*$^{e/e}$ mice are divided by 249 (the number of mutable base pairs in the *cII* gene) to express them as mutations/bp, the values measured by RMC and by the Big Blue assay are remarkably similar (27×10^{-7} and 26×10^{-7} for *Pole*$^{e/e}$, and 2.5×10^{-7} and 6.8×10^{-7} for *Pold1*$^{e/e}$, respectively). However, when the same analysis is done to compare the values for the wild-type mice as measured by the two methods, RMC gave a lower mutation frequency than did the Big Blue assay (0.09×10^{-7} for the one mouse where 55 million base pairs were screened versus 2.6×10^{-7}). The larger difference between wild-type and mutator mice that we measured with RMC as compared to the Big Blue assay may reflect the fact that spontaneous mutation frequencies are lower within endogenous, transcribed genes. In another study where mutation frequencies at the *HPRT1* locus and the *cII* reporter gene were compared, a lower frequency of spontaneous mutation was measured at *HPRT1* than at

cII in untreated cells, while the mutagen-treated cells gave similar values at the two loci (22).

Comparison of mutation spectra from Pol δ and Pol ε exonuclease-deficient mice and older wild-type mice

The very low mutation frequency in young wild-type mice precluded our ability to compare the spectra of spontaneous mutations in the polymerase mutant mice with wild-type mice from this same study. We therefore used the spectra from our RMC analysis of the older cohort of wild-type C57BL/6 mice that were used in our reproducibility study to compare with the polymerase mutant mice spectra. In the wild-type mice, a variety of mutational events were seen, with the predominant events being C-to-T and G-to-A transitions (Figure 8). These are the most common mutational events seen in the spectra of spontaneous mutations within the *cII* transgene from the Big Blue mouse model (23). It is also the most common event seen by deep sequence analysis of human tumor tissue (24). The fact that we obtained similar spectra to those obtained by other methods strongly suggest that we are detecting mutations that occurred *in vivo*. We detected very few T-to-A transversions, which are a common mistake made by TaqI polymerase during PCR, suggesting that a PCR artifact did not contribute significantly to our detected mutations.

C-to-T and G-to-A transitions, which could be considered as the same type of mutations occurring on opposite DNA strands within the palindromic TaqI site, were measured with similar frequency in wild-type mice. *Pold1*$^{e/e}$ mice had similar spectral profiles as that from the wild-type cohort, with the predominant mutational events being C-to-T and G-to-A transitions (Figure 8). In contrast, the *Pole*$^{e/e}$ strain had very large number of C-to-T events, but only a single G-to-A event in one
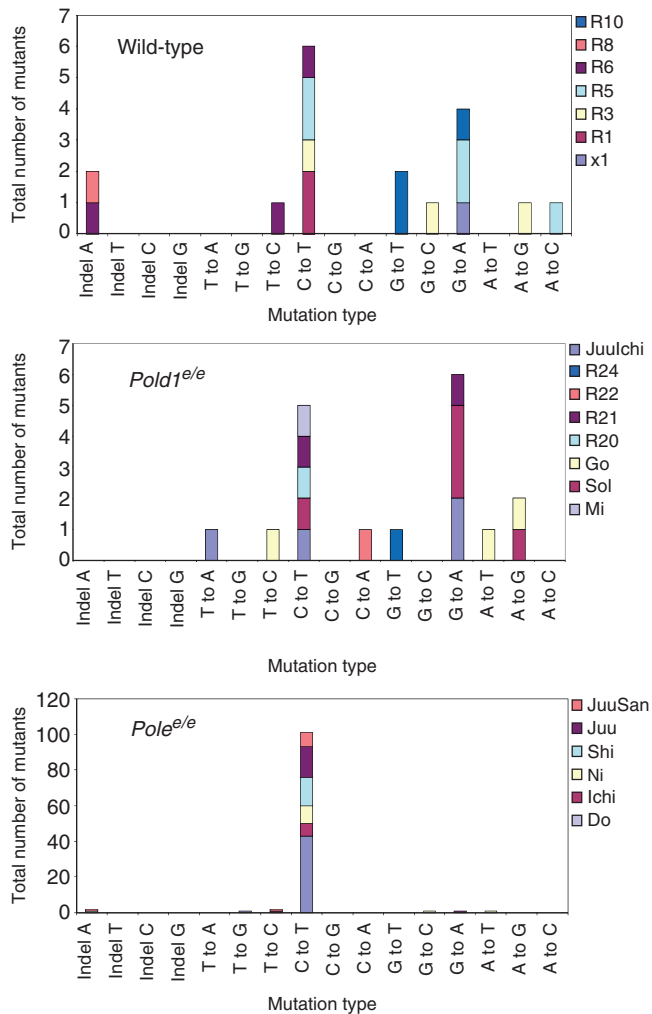
Figure 8. Comparison of the spectra of total mutations identified from mice with different genotypes. Along the *x*-axis are shown the different mutations detected within the TaqI site; each set of stacked bars corresponds to one particular type of mutation detected. The events that occurred within a particular mouse are color coded, with one color per mouse. 'Code' names for the mice in the legends correspond to those in Figures 5 and 6. The spectra shown for the wild-type mice are from a 9-month-old cohort of C57Bl/6 mice used to show the reproducibility of RMC analysis shown in Figure 5. These data are included in this figure for comparison with the mutator mouse spectra.

mouse. These results suggest that *Pole^{e/e}* mice may have a stand bias in the mutational events generated. A similar phenomenon was observed with the analogous exonuclease mutant in yeast (25). Another interpretation of these data is that the 'C' residue within our target TaqI site is a mutational hot-spot for this particular polymerase mutant-strain. These data illustrates the potential for RMC, even using only one genomic site, to shed light on mutational or DNA replication and repair mechanisms.

Sequence analysis of mutations: independent events versus total events

There are two ways that mutation spectra can be tabulated: (i) as 'total' mutations where all mutations are included in the tabulation or (ii) as 'independent'

mutations where, if the same mutation is obtained multiple times from the same mouse, it is only counted once. The spectra of the 'total' mutations detected by RMC from wild-type and polymerase proofreading-deficient mice are shown in Figure 8, and the spectra for 'independent' mutation can be seen in Supplementary Figure 1. For mutagenesis assays with larger genetic targets such as those used in transgenic mouse assays, duplicate mutations from the same mouse are often excluded (26). Given the small size of the TaqI target sequence, if only known independent events (different mutations) were counted, it is likely that the number of independent mutations would be underestimated. Using total or independent events to calculate mutation frequency both have possible issues. It is best to consider the particular experiment and determine the best choice. For example, with measuring mutations after acute mutagen exposure, one would always count total mutations as little or no cell proliferation occurred during the experiment. With experiments testing for possible mutator genotypes that could cause changes in mutation frequency throughout the lifetime of the animal, using total mutations to score for frequency runs the risk of including some clonal expansions. In our analysis of the data presented here, we chose to count the total mutations rather than risk excluding mutations that are repeated simply because they are a common event. In our data, the only events that were identified more that once from the same mouse were C-to-T and G-to-A transitions, which are the most common event overall.

DISCUSSION

Spontaneous and induced mutagenesis in mice is monitored *in vivo* using transgenic mouse models (26,27) or *in vitro* with fluctuation assays using immortalized mouse cell lines (9,11,28). The mouse RMC assay provides a useful alternative to these other assays for measuring mutation frequency, both *in vivo* and *in vitro*. RMC can be used to study changes in mutation frequency as a function of genotype, age or disease state. In addition, one could compare the mutation frequencies between different tissues.

In transgenic mouse assays, mutations are measured within exogenous genes that are transcriptionally silent in the mouse and thus would not be subject to transcription-coupled repair. Transcription coupled repair is a major pathway for cellular DNA repair, especially in quiescent cells (29). Comparisons of mutational events at the endogenous *Hprt1* locus and the *lacI* and *cII* transgenes in Big Blue mice revealed that both transgenes had a higher spontaneous mutation frequency than did *Hprt1* (22). As RMC does not rely on a reporter gene for mutation detection, mutation frequency can be measured in any locus in the genome. Using the target TaqI site in TRP53 intron IV described in this protocol, the mutation frequency within an endogenous, transcribed gene is measured and should reflect the rate of mutational events in actively transcribed chromatin. Another advantage of RMC over other random mutagenesis assays is

that there is no phenotypic selection step required for the identification of mutations. The spectrum of mutations sequenced will accurately reflect the events that occurred *in vivo*, not just those that generate a phenotypic change.

While the TaqI target sequence (TCGA) is useful for detecting all types of point mutations in that one of each the four bases is represented, the size of the genetic target in RMC is significantly smaller than that used in the transgenic mouse assays. A TaqI site has 4 bp as compared to 249 bp for the smallest target (*cII*) used in transgenic mouse assays. A smaller genetic target would be predicted to decrease the sensitivity of mutation detection. This can have the positive effect of lowering the background frequency in a case where effects of a mutagen or mutator genotype are being tested, but could also reduce the ability to detect changes at very low mutation frequency. In Figure 7, where data from RMC are directly compared with those from the *cII* reporter gene measured in a Big Blue mouse assay, a greater difference was measured between wild-type and mutator mice with RMC than in the Big Blue mouse assay. Thus for this application, RMC appeared to be the more sensitive method for detecting a mutator phenotype.

One disadvantage of this RMC protocol as we have designed it is that large deletions or insertions are not detected. The RMC assay is designed to specifically detect point mutations. Given the small size of the PCR amplified target DNA, only very small indels are detectable [Figure 8 and see reference (5)]. However, the assay can be modified to detect larger indels (30).

While we describe our protocol for measuring mutation frequency *in vivo*, mouse RMC could also be used to measure mutation rate *in vitro* using immortalized cell lines and fluctuation analysis. RMC used *in vitro* could obviate the need for the time consuming drug selection step where individual clones of cells must be expanded into individual colonies. Fluctuation analysis could be performed directly on polyclonal cultures of cells that are screened in parallel for random mutations at any endogenous genetic locus.

Given that the RMC assay does not require analysis of a particular reporter gene, multiple genetic loci could be examined in parallel to study the position effect on random mutation frequency at the different genomic sites. RMC could be used, for example, to compare the mutation frequencies between introns and exons, or between loci that are transcriptionally active or silent. One could also compare euchromatic DNA with heterochromatic DNA. We describe this protocol using a particular genomic site, yet this assay can be adapted to examine any TaqI restriction site (TCGA), or even multiple loci in parallel from the same digested DNA samples using different PCR primers. For expanding RMC analysis to new genomic sites, a description of important design parameters is presented in S1 supplementary methods.

Currently, *in vivo* genotoxic screening of novel drugs or chemicals in mice is done using indirect measurements such as unscheduled DNA synthesis in the liver (27). Tests using either selection or screening for mutations in endogenous genes such as *HPRT* or *Dbl* are also useful yet they are limited to a few tissues. Transgenic mouse assays have the advantage of mutation detection in virtually any tissue (27), yet their utility is limited due to a high spontaneous mutation rate as compared with endogenous genes (22,26). RMC could provide a faster and more sensitive screening method for both acute and chronic mutagenic effects of novel compounds in mice. This simplified RMC protocol is amenable to automation, operational in any mouse strain, and could also be adapted for use in other animal model systems such as rat, dog, or primate. Currently, DNA isolation techniques as well as qPCR amplification and sequencing have both been successfully adapted to automation. A procedure that combines both, coupled with repeated incubations with restriction enzyme could allow for an automated RMC assay, which would be, by far, the most high-throughput method for measuring mutation frequency from cells or tissues.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Salk,J., Fox,E. and Loeb,L. (2010) Mutational heterogeneity in human cancers: origin and consequences. *Annu. Rev. Pathol.*, **5**, 51–75.
2. Erickson,R. (2003) Somatic gene mutation and human disease other than cancer. *Mutat. Res.*, **543**, 125–136.
3. van Zeeland,A. and Simons,J. (1976) The use of correction factors in the determination of mutant frequencies in populations of human diploid skin fibroblasts. *Mutat. Res.*, **34**, 149–158.
4. Seshadri,R., Kutlaca,R., Trainor,K., Matthews,C. and Morley,A. (1987) Mutation rate of normal and malignant human lymphocytes. *Cancer Res.*, **47**, 407–409.
5. Bielas,J. and Loeb,L. (2005) Quantification of random genomic mutations. *Nat. Methods*, **2**, 285–290.
6. Bielas,J., Loeb,K., Rubin,B., True,L. and Loeb,L. (2006) Human cancers express a mutator phenotype. *Proc. Natl Acad. Sci. USA*, **103**, 18238–18242.

7. Parry,J., Shamsher,M. and Skibinski,D. (1990) Restriction site mutation analysis, a proposed methodology for the detection and study of DNA base changes following mutagen exposure. *Mutagenesis*, **5**, 209–212.

8. Sandy,M., Chiocca,S. and Cerutti,P. (1992) Genotypic analysis of mutations in Taq I restriction recognition sites by restriction fragment length polymorphism/polymerase chain reaction. *Proc. Natl Acad. Sci. USA*, **89**, 890–894.

9. Kohler,S., Provost,G., Fieck,A., Kretz,P., Bullock,W., Putman,D., Sorge,J. and Short,J. (1991) Analysis of spontaneous and induced mutations in transgenic mice using a lambda ZAP/lacI shuttle vector. *Environ. Mol. Mutagen.*, **18**, 316–321.

10. Kohler,S., Provost,G., Fieck,A., Kretz,P., Bullock,W., Sorge,J., Putman,D. and Short,J. (1991) Spectra of spontaneous and mutagen-induced mutations in the lacI gene in transgenic mice. *Proc. Natl Acad. Sci. USA.*, **88**, 7958–7962.

11. Gossen,J., de Leeuw,W., Tan,C., Zwarthoff,E., Berends,F., Lohman,P., Knook,D. and Vijg,J. (1989) Efficient rescue of integrated shuttle vectors from transgenic mice: a model for studying mutations *in vivo*. *Proc. Natl Acad. Sci. USA*, **86**, 7971–7975.

12. Albertson,T., Ogawa,M., Bugni,J., Hays,L., Chen,Y., Wang,Y., Treuting,P., Heddle,J., Goldsby,R. and Preston,B. (2009) DNA polymerase epsilon and delta proofreading suppress discrete mutator and cancer phenotypes in mice. *Proc. Natl Acad. Sci. USA*, **106**, 17101–17104.

13. Bernad,A., Blanco,L., Lázaro,J., Martín,G. and Salas,M. (1989) A conserved 3′-5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. *Cell*, **59**, 219–228.

14. Simon,M., Giot,L. and Faye,G. (1991) The 3′ to 5′ exonuclease activity located in the DNA polymerase delta subunit of Saccharomyces cerevisiae is required for accurate replication. *EMBO J.*, **10**, 2165–2170.

15. Morrison,A. and Sugino,A. (1994) The 3′−>5′ exonucleases of both DNA polymerases delta and epsilon participate in correcting errors of DNA replication in Saccharomyces cerevisiae. *Mol. Gen. Genet.*, **242**, 289–296.

16. Goldsby,R., Hays,L., Chen,X., Olmsted,E., Slayton,W., Spangrude,G. and Preston,B. (2002) High incidence of epithelial cancers in mice deficient for DNA polymerase delta proofreading. *Proc. Natl Acad. Sci. USA*, **99**, 15560–15565.

17. Kvam,E. and Tyrrell,R. (1997) Artificial background and induced levels of oxidative base damage in DNA from human cells. *Carcinogenesis*, **18**, 2281–2283.

18. Zheng,L., Dai,H., Zhou,M., Li,M., Singh,P., Qiu,J., Tsark,W., Huang,Q., Kernstine,K., Zhang,X. *et al.* (2007) Fen1 mutations result in autoimmunity, chronic inflammation and cancers. *Nat. Med.*, **13**, 812–819.

19. Poovathingal,S., Gruber,J., Halliwell,B. and Gunawan,R. (2009) Stochastic drift in mitochondrial DNA point mutations: a novel perspective ex silico. *PLoS Comput. Biol.*, **5**, e1000572.

20. Lee,A., DeSimone,C., Cerami,A. and Bucala,R. (1994) Comparative analysis of DNA mutations in lacI transgenic mice with age. *FASEB J.*, **8**, 545–550.

21. Ono,T., Miyamura,Y., Ikehata,H., Yamanaka,H., Kurishita,A., Yamamoto,K., Suzuki,T., Nohmi,T., Hayashi,M. and Sofuni,T. (1995) Spontaneous mutant frequency of lacZ gene in spleen of transgenic mouse increases with age. *Mutat. Res.*, **338**, 183–188.

22. Monroe,J., Kort,K., Miller,J., Marino,D. and Skopek,T. (1998) A comparative study of in vivo mutation assays: analysis of hprt, lacI, cII/cI and as mutational targets for N-nitroso-N-methylurea and benzo[a]pyrene in Big Blue mice. *Mutat. Res.*, **421**, 121–136.

23. Morgan,C. and Lewis,P. (2006) iMARS–mutation analysis reporting software: an analysis of spontaneous cII mutation spectra. *Mutat. Res.*, **603**, 15–26.

24. Sjöblom,T., Jones,S., Wood,L., Parsons,D., Lin,J., Barber,T., Mandelker,D., Leary,R., Ptak,J., Silliman,N. *et al.* (2006) The consensus coding sequences of human breast and colorectal cancers. *Science*, **314**, 268–274.

25. Shcherbakova,P. and Pavlov,Y. (1996) 3′−>5′ exonucleases of DNA polymerases epsilon and delta correct base analog induced DNA replication errors on opposite DNA strands in *Saccharomyces cerevisiae*. *Genetics*, **142**, 717–726.

26. Lambert,I., Singer,T., Boucher,S. and Douglas,G. (2005) Detailed review of transgenic rodent mutation assays. *Mutat. Res.*, **590**, 1–280.

27. Dean,S., Brooks,T., Burlinson,B., Mirsalis,J., Myhr,B., Recio,L. and Thybaud,V. (1999) Transgenic mouse mutation assay systems can play an important role in regulatory mutagenicity testing in vivo for the detection of site-of-contact mutagens. *Mutagenesis*, **14**, 141–151.

28. Foster,P. (2006) Methods for determining spontaneous mutation rates. *Methods Enzymol.*, **409**, 195–213.

29. Bielas,J. and Heddle,J. (2004) Quiescent murine cells lack global genomic repair but are proficient in transcription-coupled repair. *DNA Repair (Amst)*, **3**, 711–717.

30. Vermulst,M., Bielas,J. and Loeb,L. (2008) Quantification of random mutations in the mitochondrial genome. *Methods*, **46**, 263–268.