Published in final edited form as:

Methods Mol Biol. 2018 ; 1672: 421–438. doi:10.1007/978-1-4939-7306-4_29.

Quantitative analysis of the rates for repeat-mediated genome instability in a yeast experimental system

Elina A. Radchenko¹, Ryan J. McGinty¹, Anna Y. Aksenova², Alexander J. Neil¹, and Sergei M. Mirkin^{1,*}

¹Department of Biology, Tufts University, Medford, MA 02155, USA

²Amyloid Biology, Saint-Petersburg University, Russia

Abstract

Instability of repetitive DNA sequences causes numerous hereditary disorders in humans, the majority of which are associated with trinucleotide repeat expansions. Here we describe a unique system to study instability of triplet repeats in a yeast experimental settings. Using fluctuation assay and the novel program FluCalc we are able to accurately estimate the rates of large-scale expansions, as well as repeat-mediated mutagenesis and gross chromosomal rearrangements for different repeat sequences.

Keywords

Trinucleotide repeats; repeat expansions; repeat-induced mutagenesis; fluctuation assay; expansion rate; mutation rate; MSS-MLE; FluCalc

1. Introduction

The fluctuation test has been a cornerstone of the field of microbial genetics since the Luria-Delbrück experiment, which determined that mutations arise spontaneously during growth of a culture prior to selection [1]. The mathematical methods underlying mutation rate calculations are complicated by the fact that mutations spontaneously arise by chance at different generations in the culture, and subsequently accumulate with each cell division [1, 2]. This produces fluctuations in the frequency of mutants in each independent experiment. The mathematical methods of accounting for this natural fluctuation have improved over time, with innovations to account for variables such as the relative fitness of the mutant population, partial plating of the culture, or variation in the total population size across the cultures in the experiment [3, 4, 5].

Laboratory methods have also evolved since the days of Luria and Delbrück. An experimental strain typically contains a forward-selection reporter, which allows the detection of mutant cells on particular selective media. A fluctuation test usually begins with growth of several parallel cultures in non-selective conditions, where a small number of inoculated cells grow to saturation. In our experiments, we first spread cells on a Petri dish,

*Correspondence: sergei.mirkin@tufts.edu.

where a single cell grows to form a clonal colony. Individual colonies are then resuspended in water and various dilutions are plated onto both selective and non-selective solid media. The number of colonies grown on each plate is used to determine the number of mutations per cell per generation, also known as μ or the mutation rate.

Notably, however, different colonies on selective media do not necessarily arise from the same type of mutation. A reporter gene may be inactivated by a point mutation, a small insertion or deletion causing a frameshift, a larger deletion removing part or all of the gene, or by chromosomal events such as translocations or the loss of a non-essential arm of the chromosome. Furthermore, mutations can occur within the body of the gene, or in an important regulatory region. Finally, reporter gene inactivation can also occur indirectly, for instance via mutations or downregulation of genes responsible for the uptake of a selective drug. These different events do not necessarily have the same consequences in terms of colony formation, with a regulatory change producing a weaker selective pressure than a deletion of the entire marker. Variations on the fluctuation assay may focus on one or more of these types of mutations. A classic example is the frameshift reversion assay, which begins with a selective marker inactivated by a 1 bp insertion, and selects for those mutations that restore the frame of translation.

One unusual type of mutation is microsatellite instability. Tracts of short repetitive sequences (repetitions of 2-12 bases) can mutate in their length, with the number of repeats either expanding or contracting. Expansions of microsatellite sequences have been linked to a variety of human disorders [6]. The instability of microsatellite sequences can be attributed to the formation of various non-B-form DNA secondary structures, depending on the particular sequence. Repeats can be added a few at a time, or by a single large-scale event. Different microsatellite sequences may be more or less prone to large or small-scale events, as well as chromosomal fragility and repeat-induced mutations, all of which differ in their underlying mechanisms and rates [7]. Importantly, the rates of these various events appear to increase exponentially with the length of the starting repeat tract [8, 9]. Depending on the particular repeat, the rate of such events can be higher than the average point mutation rate of the genome by many orders of magnitude.

In the protocol below, we describe the use of our experimental system to select for large-scale expansions of H-DNA-forming (GAA)_n repeats in *S. cerevisiae* [8, 10]. The yeast carry a selectable cassette consisting of a modified *URA3* gene, which is counter-selectable on media containing the drug 5-fluoroorotic acid (5-FOA) [see Fig. 1]. This cassette is located ~1 kb downstream of the early-firing replication origin ARS306. The *URA3* gene has been modified to include 100 GAA repeats within a 974 bp artificial intron. Due to the limitations of the yeast spliceosome, an intron exceeding ~1000 bp is no longer spliced, leading to 5-FOA-resistance in strains bearing our genetic construct. It is important to note that this threshold is not absolute, but rather splicing efficiency appears to diminish in an exponential manner as the intron length grows past 1000 bp. Thus, we designed our system to begin with the overall intron length close to this threshold, such that even relatively short expansions should be sufficient to induce a level of 5-FOA-resistance distinguishable from the background.

Importantly, however, *URA3* can become inactivated not only due to a repeat expansion, but also by point mutations, deletions, translocations, downregulation, or even indirectly. In practice, each of these situations creates varying degrees of 5-FOA resistance, resulting in a range of colony sizes. The smallest 5-FOA-resistant colonies are typically formed owing to indirect resistance or slight *URA3* downregulation. The largest colonies typically result from the strongest inactivating events, such as point mutations and large deletions. Expansions typically occupy a middle ground, as they severely but not totally reduce *URA3* expression, with the strength of this effect depending on the size of the expansion. Very large expansions may be the result of secondary expansions, *i.e.* expansions that occurred following a primary expansion. To distinguish between these different events, we perform a PCR on various (ideally all) 5-FOA-resistant colonies. Using primers flanking the repetitive tract, we can measure the size of each expansion, or infer the presence of a deletion/translocation (if the PCR product is absent) or a point mutation (if there is no change in repeat length).

A major advantage of our system is that it can be adapted to favor the recovery of certain types of events, simply by altering the concentration of 5-FOA in the selective media. The weakest events, including small-scale expansions, can be observed by growing on a low concentration of 5-FOA. The concentration of 5-FOA can be optimized to reduce these background colonies, favoring the recovery of large-scale primary expansions. If the concentration of 5-FOA is increased further, this can eliminate most expansions, highlighting only the most damaging types of mutations, including deletions, translocations and point mutations.

Indeed, there are a number of possible modifications to the fluctuation test, and to our expansion system, that can be made. For instance, the starting repeat length can be increased or decreased [8], or the (GAA)_n repeats can be swapped out for other sequences such as SCA10 pentanucleotide (ATTCT)_n repeats, or telomeric repeats [11, 12]. The *URA3* promoter can be swapped for an inducible *GAL1* promoter to study the effects of transcription [10]. The entire cassette can be moved to different locations, such as the non-essential arm of chromosome V, allowing for recovery of arm-loss events (double strand breaks that were not repaired) [13]. In addition to forward selection for repeat expansions, reverse-selection to a strong Ura⁺ phenotype can be used to study contractions of the repetitive tracts [14]. Thus, our system is highly adaptable to the study of various molecular mechanisms.

As written, the protocol below describes the steps used for measuring the rate of primary expansions in our wild-type strain with the (GAA)₁₀₀ selectable cassette located ~1 kb downstream of the early-firing origin ARS306. This is perhaps the most challenging use of the assay, as it requires optimization of 5-FOA concentration and growth times, as well as careful scrutiny of the distribution of repeat lengths. We have included a number of notes in the protocol to guide with this optimization, as well as suggestions for where the assay can be modified. Variations of this protocol can be found in our previous publications [8, 9, 10, 11, 12, 14], but this document represents our most up-to-date and refined version of the assay.

Finally, we introduce here a new tool, named FluCalc, for calculating mutation rates from these fluctuation tests (see Fig. 2). Although several mathematical tools such as SALVADOR, or web-tools such as FALCOR and bz-rates were designed for accurate estimation of mutation rate, [15, 16, 17, 18] they appeared either unsuitable or inconvenient for our experiments. FluCalc uses the Ma-Sandri-Sarkar maximum likelihood estimation (MSS-MLE) equations for calculation, regarded as one of the most accurate estimators available [19]. MSS-MLE is also applicable across most mutational spectra [20]. FluCalc calculates the values for mutant frequencies, mutations per culture, mutation rates and 95% confidence intervals of mutation rates. FluCalc also automatically incorporates plating efficiency calculations to account for dilutions. FluCalc displays a full description of all equations used in the calculations, and shows all values of the intermediate variables at each step. FluCalc is presented in a web format suitable for all modern browsers, and does not rely on plugins that become outdated and broken. FluCalc allows one to download and save the data as a text file, which can be subsequently reopened by the program or imported into a spreadsheet program. This feature is especially handy when processing and comparing large amounts of data. We hope that FluCalc will be useful for many researchers wanting an accurate and efficient tool for mutation rate calculations. The tool will be available on <http://ase.tufts.edu/biology/labs/mirkin/resources/>.

2. Materials

All experiments are described on the basis of the *Saccharomyces cerevisiae* WT strain CH1585 (*MATa*, *leu2* 1, *trp1* 63, *ura3* 52, and *his3* 200) with (GAA)₁₀₀ repeats inserted in the intron-containing *URA3* gene on chromosome III. (see **Note 1**)

2.1. Non-Selective Growth Period

1. 15% Glycerol stock solution: 300 µL of autoclaved 50% glycerol solution + 700 µL of YPD.
2. YPD + Uracil plates (example makes 1 L): 1 L dH₂O, 10 g yeast extract, 20 g peptone, 20 g dextrose, 20 g agar. Autoclave, allow to cool slightly, then add 24 mL uracil from a 2 mg/mL stock solution. Pour into Petri dishes, solidify overnight, seal and store at 4°C.

2.2. Selective Growth Period

1. 5-FOA media (example makes 1 L): Autoclaved portion: 500 mL dH₂O, 20 g agar, 20 g dextrose, 2 g synthetic complete amino acid supplement minus uracil with added adenine, 6.7 g yeast nitrogen base without amino acids or carbohydrates and with ammonium sulfate. Non-autoclaved portion: 500 mL dH₂O, 1 g 5-FOA, 50 µg uracil – shake at 37°C until 5-FOA is completely dissolved and use a 0.2 µm bottle-top vacuum filter to sterilize the solution. Allow autoclaved portion to cool slightly, then combine both portions, mix on stirrer and immediately pour into Petri dishes. Let the media solidify overnight and use the following day (see **Note 2**).

2. YPD media (example makes 1 L): 1 L dH₂O, 10 g yeast extract, 20 g peptone, 20 g dextrose, 20 g agar. Autoclave, allow to cool slightly, and pour into Petri dishes. Let the media solidify overnight, seal and store at 4°C.

2.3. Colony PCR

1. Lyticase solution: 2.5 mg/mL lyticase (L2524 Sigma 2,000 units/mg protein) in 0.9M sorbitol, 0.1M EDTA (pH 7.4) (*see Note 3*). Zymolyase (e.g. 100T Zymo Research) performs equally well.
2. PCR primers (*see Note 4*).
3. PCR mastermix (example makes 24 µL):
 - H₂O 13.5 µL
 - 5× Green GoTaq buffer (Promega) 5 µL
 - 10mM dNTPs 0.25 µL
 - 10uM Forward Primer 2.5 µL
 - 10uM Reverse primer 2.5 µL
 - SibenzymeTaq 0.25 µL

2.4. Gel Electrophoresis

1. 10× TBE stock solution: 0.89M Tris, 0.89M Borate, 0.02M EDTA. Store at room temperature.
2. Ultrapure high resolution agarose.
3. NEB Quick-Load 100 bp or 50 bp DNA ladder.
4. Ethidium bromide 1% (Fisher Bioreagents).

3. Methods

3.1. Non-selective growth stage (*see Note 5*)

1. Streak a small amount of yeast from 15% glycerol frozen stock onto a YPD plate supplemented with uracil such that distinct individual colonies form (*see Note 6*).
2. Incubate 40 hours at 30°C (*see Note 7*).

3.2. Selective growth and total cell count (*see Fig. 3*)

1. Fill a sterile 96-well polypropylene, U-well micro-titer plate with 200 µL of sterile deionized water in the 1st row and 180 µL in rows 2-5.
2. Use a small plastic inoculating loop to scoop up an average-sized colony and resuspend it in the 1st row of the plate. Repeat until 8-12 similarly-sized colonies have been resuspended, one per well (*see Note 8*). Mix thoroughly with a multi-channel pipette.

3. Perform 10× serial dilutions, transferring 20 µL from the 1st to 2nd rows, and then into subsequent rows, mixing thoroughly at each step (*see Note 9*).
4. Mark corresponding YPD and 5-FOA plates with the strain name and colony number.
5. Plate 100 µL from the 1st row onto a 5-FOA plate and 100 µL from the 5th row onto a YPD plate for each sample (*see Note 10*) and spread until dry with a large disposable plastic inoculating loop.
6. Spin down micro-titer plate to collect yeast cells for PCR analysis, add 1.5 µL lyticase solution and go to step 3.3.3 (*see Note 11*).
7. Incubate plates at 30°C for 3 days (*see Note 12*).
8. Count individual colonies on all plates (*see Note 13*).
9. (Optional) Using a flatbed scanner, scan images of all plates. Load images into OpenCFU colony counting software. Adjust threshold and radius values until all 5-FOA-resistant colonies are included and all background growth is excluded from the count (*see Note 14*).

3.3. PCR analysis of repeats (*see Note 15*)

3.3.1. Option A (*see Note 16*)

1. If necessary, randomly subdivide 5-FOA plates into 1/2, 1/4 or 1/8 size wedges containing no less than 4 and no more than 16 colonies within a wedge from each plate. Count total number of colonies within random wedge of each 5-FOA plate.
2. Pick all colonies from within the selected 5-FOA plate or wedge using a pipette tip and resuspend one colony in each well of a 96-well PCR plate containing 1.5 µL lyticase solution per well.

3.3.2. Option B (*see Note 17*) and Option C (*see Note 18*)

1. For each 5-FOA plate, choose four 5-FOA-resistant colonies for PCR analysis.
2. Pick colonies using a pipette tip and resuspend one colony in each well of a 96-well PCR plate containing 1.5 µL lyticase solution per well.

3.3.3. Continue from Option A, B or C

1. Incubate cells in lyticase solution for 8-15 minutes at 37°C
2. Add 50 µL of PCR-grade water to all wells. Resuspend if necessary.
3. Incubate plate at 100°C for 5 minutes.
4. Spin plate at 2500-3000 × g for 2-5 minutes to pellet cell debris. The solution containing yeast genomic DNA can be stored for up to 12 months at -20°C and repeatedly analyzed.
5. Prepare PCR master mix for the appropriate number of samples.

6. Add 12-24 μL of PCR master mix into each well of a new 96 well PCR plate.
7. Add 1 μL of genomic DNA supernatant to each PCR reaction.
8. Run PCR program:
 1. 95°C – 15 sec
 2. 94°C – 15 sec
 3. 72°C – 2 min
 4. Go to Step 2 $\times$ 30
 5. 72°C – 2 min
 6. 12°C – hold

3.4. Gel Electrophoresis (see Fig. 4)

1. Prepare a 1.5% agarose gel using ultrapure agarose and 0.5 $\times$ TBE (see **Note 19**). Add 2 μL of ethidium bromide per 100 mL of TBE.
2. Load samples along with 100 bp DNA ladder, run gel, and image.
3. Calculate expansion sizes by subtracting the starting length of the unexpanded PCR product from expanded PCR product sizes and dividing by 3 (for triplet repeats) (see **Note 20**).
4. For each individual experimental strain, visualize the distribution of expansion sizes by graphing the percentage of expanded products that fall into 10-repeat bins (e.g. 0-10 repeats added, 11-20 repeats added, etc.) (see **Note 21**).

3.5. Expansion Rate Calculation (see Note 22)

1. Insert the following information into FluCalc to obtain an accurate expansion rate with corresponding 95% confidence interval:
 1. V_{tot} – the total volume of the culture following resuspension of colonies in step 3.2.2 (typically 200 μL)
 2. C_{sel} – the observed number of colonies with expanded repeats from each plate or sector (adjusted to reflect the results of the gel electrophoresis) (see **Note 23**)
 3. C_{com} – the observed number of colonies on non-selective media
 4. D_{sel} – the dilution factor used for plating on selective media (e.g. if liquid from the second row of a 10 $\times$ serial dilution was plated on 5-FOA, then $D_{\text{sel}} = 10$)
 5. D_{com} – the dilution factor used for plating on non-selective media (e.g. if liquid from the fifth row of a 10 $\times$ serial dilution was plated on YPD then $D_{\text{com}} = 10,000$)

6. V_{sel} – the volume of culture plated per plate or sector on selective media (see **Note 24**)
7. V_{com} – the volume of culture plated per plate on non-selective media
2. Download and save the data in FluCalc.txt format so that it can be subsequently reopened by the program.
3. Click on the “Calculate” button to get results for the expansion rate, presented as either raw values or corrected values (see **Note 25**). Frequencies for each plate will also be reported (see **Note 26**). 95% confidence intervals are also reported (see **Note 27**).

4. Notes

1. The strain used for the example in this protocol was described and thoroughly examined in Shah, et al. 2012. [18] A key feature of this strain is the “balanced” intron length just under the splicing threshold length, allowing a single expansion to inactivate the *URA3* gene. The rate of expansions for a GAA₁₀₀ repeat was determined to be $\sim 2 \times 10^{-5}$ expansions per cell per generation. Altering the starting repeat length or sequence can dramatically change this rate. Our protocol is ideal for measuring rates in the range of 10^{-6} to 10^{-3} . Measuring rates below this range is difficult due to the larger number of cells required for plating. Measuring rates above this range results in large natural fluctuations, making this type of assay inappropriate. We recommend adjusting the starting length of the repeat until the rate is close to 10^{-5} . This will allow the investigation of various mutants or experimental conditions that raise or lower the rate within the measurable range.
2. 5-FOA will remain effective when stored at 4°C for a period of weeks. However, as water slowly evaporates from the plates, this will raise the effective concentration of 5-FOA, which can shift the selection towards larger expansions, point mutations and deletions. Thus it is important, if determining the rate of expansions in a mutant strain, to concurrently re-test the wild-type strain to help control for subtle differences in the media. Plastic wrap can be used to reduce evaporation, but it is best to use the plates within days. The above recipe is for 0.1% 5-FOA media, which is ideal for the selection of primary repeat expansions. To select for deletions/point mutations/translocations, a concentration of 0.15% works well. The strength of the 5-FOA media can also be increased by lowering the pH, though we find this method of adjustment to be less precise.
3. The pH of the lyticase solution should be kept around 7.4-7.5 and is critical for the activity of the enzyme. This solution should be stored in small aliquots at -20°C.
4. Primers should be closely flanking the repeats, and should have T_m above 72°C. The T_m should be high to discourage secondary structure formation which can lead to PCR failures that create faint or smeared bands. The PCR product should

be as small as possible to facilitate accurate length determination. PCR products with longer non-repetitive flanks can sometimes help to reduce PCR failures, but these longer products will not be as accurate for measuring small changes in length.

5. If assessing any experimental variables (such as temperature sensitivity, chemical perturbation, or gene induction), it is important to introduce them during this stage of growth.
6. Extra uracil is added to the media to prevent cells from becoming dependent on the *URA3* gene, which will be inactivated by an expansion. If petites are particularly frequent following unfreezing, the strains can first be patched onto YPEG plate for 1 day. The YPEG patches can then be used to streak for single colonies on YPD+Uracil. For strains in which the *URA3* promoter is replaced by the galactose-inducible *GALI* promoter, start with a 5 mL overnight culture in YPD + raffinose to remove glucose inhibition of the *GALI* promoter. Dilute the overnight culture in deionized water and plate on YPD or YPGal, supplemented with uracil.
7. This amount of time can be extended for any slow-growing strains or alternate growth conditions. However, it is important not to grow the colonies for too many generations, which can result in a higher probability of secondary expansions. This is because a primary expansion will increase the next generation's starting repeat length, which will in turn have an exponentially higher rate of expansion.
8. It is important to pick colonies that are of the same size. Most methods to calculate mutation rates (including MSS-MLE) rely on an assumption that culture sizes are the same (within 20% of the coefficient of variation) across the experiment [5, 20]. To decrease the inter-culture variation, do not choose the largest or smallest colonies on the plate – it is better to aim for average sized colonies. When testing our wild-type GAA₁₀₀ strain, we choose single colonies that are about 1mm in diameter when uncrowded. The starting colony size can be adjusted in anticipation of higher or lower expansion rates. After colonies have been counted (step 3.2.8), if there is a YPD plate that deviates significantly from the median, both the YPD and corresponding 5-FOA plate should be excluded from further analysis.
9. Accurate pipetting is critical throughout the serial dilution, as small random errors can be magnified after each step. Accuracy can be improved through the use of low-retention tips, by discarding tips between each suspension, and by visually checking for air bubbles, droplets stuck in the tips, or other pipetting errors.
10. This recommendation is applicable to our wild-type GAA₁₀₀ strain, and should result in uncrowded single colonies. Both the volume and choice of dilution can be altered for different mutant backgrounds or growth conditions.

11. PCR analysis of the starting colonies should confirm that expansions did not take place prior to the non-selective growth period. All future analysis depends on these cells beginning from the same initial repeat length. Encountering expanded or contracted clones at this stage should be rare, but strains with very high rates of instability may require careful monitoring.
12. The growth period may be altered if necessary. Alterations to the temperature or media components at this stage will not affect the rate at which mutations were generated during the non-selective growth phase, but may affect the recovery of those events for counting. Only change these conditions if you have reason to believe you are not recovering certain events. Also, if using a galactose-inducible strain, the 5-FOA media must also contain galactose in order for the *URA3* gene to be active and thus counter-selectable.
13. For YPD plates, all colonies must be counted. For 5-FOA plates, there will be a certain amount of small background colonies, which do not have expansions or other mutations of interest. A rule of thumb is that if the colony is too small to be easily picked up with a pipette tip, it cannot be analyzed by PCR and therefore should not be counted. It is possible to empirically determine for each strain the relationship between colony size and mutation status. In any case, it is important to be consistent between those colonies counted and those chosen for PCR analysis.
14. The OpenCFU program is freely available at <http://opencfu.sourceforge.net/>. An alternative is the Image J program with a cell-counter plugin (<https://imagej.nih.gov/ij/index.html>). Use a consistent image size for every plate. (A 300 dpi scan will produce 1000×1000 pixels per plate). Image the underside of the plate with the top cover removed. This will produce the most consistent and clear images, allowing proper comparison of colony sizes across different plates.
15. The goal of the PCR analysis is to determine which 5-FOA resistant colonies contain expansions, as well as the size of individual expansions. An ideal solution would involve PCR analysis of every 5-FOA-resistant colony on a plate. To achieve a manageable number of colonies on each selective plate (*e.g.* ~10 per plate) you may need to dilute cells further or even plate several dilutions on 5-FOA plates in a given experiment. While performing PCR for each colony on every 5-FOA plate may become prohibitively costly and laborious, it eliminates sampling error from confidence intervals for the mutation rates. Because of this consideration, we recommend to PCR all 5-FOA resistant colonies; if it is not practically doable, three options to make PCR analysis feasible are presented below.
16. Option A is to plate a lower density of cells on each plate, or to count and PCR every colony only from a sector of each plate. This approach eliminates any kind of bias (especially by size) that might occur while choosing colonies for PCR analysis. Of course, the subdivision of each plate must occur randomly in order to avoid unintentionally choosing a sector with larger or smaller colonies. One solution is to make a “Wheel-of-Fortune,” marking sectors on the top cover of a

Petri plate, then place the “wheel” over each plate, spin, and mark each colony that lands within the pre-chosen sector. If there are a large number of colonies in the sector, a smaller sector may be used. Note, that the size of a sector should be the same for every 5-FOA plate within a batch. If very small sectors are needed, it may be better to perform the fluctuation test again with a lower concentration of cells. Option A also has the advantage of being mathematically simpler when performing mutation rate calculations. The caveat is that by plating fewer cells, the fluctuations in rate may be higher due to the greater number of dilutions and pipetting steps. Counting from sectors of a plate also assumes that the cells were spread evenly.

17. Option B is to PCR a set number of colonies on each plate, and to apply that percentage to the total number of 5-FOA-resistant colonies. Option B has several advantages. This method allows a higher plating density, provided that the colonies are not overcrowded to the point where they affect the growth of their neighbors. It also has the advantage of producing a predictable number of PCR reactions per fluctuation test, allowing more high-throughput processing. A caveat is that the user must choose a few colonies from each plate to represent the entire population, potentially introducing bias. Some of this bias can be eliminated, since we know roughly the relationship between colony size and mutation type. To best account for the various events that cause 5-FOA resistance, choose colonies by size in proportion to their approximate representation in the population, i.e. if one out of four colonies are larger than average, then choose 3 average sized colonies and one larger colony for analysis. Although this tactic is not completely void of observer's bias, it might provide a comprehensive analysis of the population. It is important not to choose any small background colonies if they were not included in the total colony count. Another caveat of Option B is that the additional step of the percentage calculation can affect the mathematics of the mutation rate calculation. Rate estimations can be skewed, particularly if expansions (or other events analyzed) do not represent the majority of the population. Because FluCalc applies plating efficiency adjustments, the calculated mutation rate in Option A is expected to be lower than in Option B. It is best, therefore, to be consistent in choosing either Option A, B or C for all experiments.
18. In Option C, as in Option B, a fixed number of colonies per plate are analyzed. However, this step is treated as the mathematical equivalent of a dilution, with a different dilution factor for each plate. Like Option A, the actual number of observed events can then be used for each calculation, though this requires applying an additional plating efficiency correction to each plate. The calculations accounting for this “variable dilution” are not currently integrated into FluCalc. See Note 25 and references 12 and 14 for more details.
19. In our current system, the repetitive tract starting length PCR is 350 bp, and expansions will typically be in the range of 375-1000 bp. A 1.5% agarose gel is appropriate for this size range. Switching to a 2-2.5% agarose gel in 1×TAE buffer can help provide better resolution when analyzing small-scale changes.

(Aksenova et al. 2015). Voltage and running time should be adjusted for clear separation of bands within this size range, and size estimation to the nearest 25 bp. Bands can also be analyzed electronically, for example via the BioRad QuantityOne program, for greater accuracy. We also recommend large format gels with multichannel-compatible wells, due to the high volume of PCR samples that must be analyzed per assay.

20. In some cases, a gel lane may contain two or more bands, indicating the presence of a secondary expansion or contraction that occurred following plating on 5-FOA (as opposed to secondary expansions that occurred during the non-selective growth stage, which will only produce a single band). We typically attribute the pre-plating event to the darkest band on the gel (representing the greatest number of cells within the colony), with the caveat that shorter repeat lengths typically result in a more successful PCR reaction. On the other hand, an extremely long expansion may push the limits of the PCR reaction, creating a smeared band. We typically use the darkest part of the smear as the length of the pre-plating expansion. If one of the two bands has the original size of the repeat, the expansion most likely occurred post-plating and the colony should be counted as non-expanded.
21. We have found that primary expansions occur in a normal distribution with a median of 50-60 repeats added [9]. Only experiments in which the expansion size histogram follows a unimodal distribution will produce an accurate estimate of mutation rate. A bimodal distribution may indicate the presence of secondary expansion events. Secondary expansions can be selected for if the 5-FOA concentration is too high. Secondary expansions become more and more prevalent as starting repeat tract length increases, in mutant strains with high expansion rates, or if the non-selective growth period is too long (because primary expansions will increase the likelihood of further expansions in the next generation). It is also not uncommon for the distribution to reveal that smaller-scale expansions were not allowed to grow due to a higher 5-FOA concentration. Expansion rates calculated in non-normally distributed assays will thus reflect an uncertain mixture of primary, secondary and missed events.
22. This example protocol is used to determine the rate of repeat expansions. However, FluCalc can be used to determine rates for any type of observed mutation, depending on which set of C_{sel} values are entered. For example, the user can instead enter the number of colonies with an observed deletion, or the total number of 5-FOA-resistant colonies.
23. For Options A and C, enter the number of colonies with expansions after running gel electrophoresis. For Option B, the numbers entered for C_{sel} should be the number of observed 5-FOA-resistant colonies multiplied by the overall percent of expansions (i.e. total number of colonies with expansions divided by total number of analyzed colonies for all cultures in the experiment).

24. If counting sectors (Option A), assume an even distribution of liquid sample during plating (e.g. if 100 μ L of solution was plated on 5-FOA and all colonies from a 1/8 wedge were analyzed for expansions then $V_{\text{sel}} = 12.5 \mu\text{L}$)
25. Corrected values are raw values corrected for plating efficiency, resulting in a more accurate representation of the mutation rate per cell per generation. If using Option C for PCR analysis, plating efficiency calculations must be performed manually. This is because a fixed amount of PCRs represent a variable amount of colonies, and thus the plating efficiency correction is not the same for each plate. We denote the following calculation method as VZ-MLE, or variable z maximum likelihood estimation. This method also relies on the assumption that the cultures were of roughly equal size (*see Note 8*). We hope to fully integrate VZ-MLE into FluCalc in the near future. At present, additional steps must be performed to employ FluCalc: Enter C_{sel} values as the actual number of observed expansions (e.g. 2 out of 4) obtained for each individual plate. Enter all other values (V_{tot} , D_{sel} , V_{sel} , C_{com} , D_{com} and V_{com}) as defined in Section 3.5.1 and press “Calculate” to obtain the raw mutation rate and confidence intervals. Next, calculate the z-factor for each culture by dividing the number of colonies analyzed by PCR by total number of 5-FOA resistant colonies in the culture, which is $C_{\text{sel}}D_{\text{sel}}V_{\text{tot}}/V_{\text{sel}}$. From there, determine the plating efficiency coefficient ($P=1-z/\ln(z)$) for each plate, and then the average value of P (AVP). Then multiply the rate and confidential limits obtained with FluCalc by AVP.
26. Step 3.2.5 describes resuspending and plating dilutions of each colony on corresponding 5-FOA and YPD plates. Thus the number of entries for C_{sel} and C_{com} should be the same, and FluCalc will determine individual frequencies for each corresponding set of plates, making the assumption that the first 5-FOA plate corresponds to the first YPD plate, and so forth. If the number of entries for D_{com} and C_{com} are not the same, FluCalc assumes that individual 5-FOA and YPD plates do not correspond, and instead uses the mean of C_{com} for each calculation.
27. The reported confidence intervals are accurate for any given experiment, but they may not fully reflect the inherent nature of the fluctuation assay. Due to the relatively high baseline rate of expansion events, the median expansion rate of our wild-type GAA100 strain does indeed fluctuate from trial to trial. This natural fluctuation is usually within +/- 50% of the typical rate, however we do caution at making conclusions for mutant strains with any less than a 3-fold difference from the wild-type. Despite all of the complications and caveats, we find this assay to be a reliable and consistent method to determine expansion rates.

Acknowledgments

We thank Alexander A. Shishkin and Kartik A. Shah for their invaluable contributions in developing cassettes to study repeat instability, and for developing experimental protocols for the selection and PCR procedures, Timofei S. Bondarev for developing FluCalc program and Durwood Marshall for statistical consulting. This study was funded by NIH grants GM105473 and GM60987 to S.M.M and RFBR grant #15-04-08658 and research project in the Centre for Molecular and Cell Technologies (Research Park, SPBU) for A.Y.A.

References

1. Luria SE, Delbruck M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*. 1943; 28(6):491–511. [PubMed: 17247100]
2. Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci USA*. 1991; 88:7160–7164. DOI: 10.1073/pnas.88.16.7160 [PubMed: 1831267]
3. Foster PL. Methods for determining spontaneous mutation rates. *Methods in enzymology*. 2006; 409:195–213. DOI: 10.1016/S0076-6879(05)09012-9 [PubMed: 16793403]
4. Zheng Q. A new practical guide to the Luria–Delbrück protocol. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. 2015; 781:7–13. DOI: 10.1016/j.mrfmmm.2015.08.005 [PubMed: 26366669]
5. Zheng Q. A second look at the final number of cells in a fluctuation experiment. *Journal of Theoretical Biology*. 2016; 401:54–63. DOI: 10.1016/j.jtbi.2016.04.027 [PubMed: 27113784]
6. Mirkin SM. Expandable DNA repeats and human disease. *Nature*. 2007; 447(7147):932–940. DOI: 10.1038/nature05977 [PubMed: 17581576]
7. Kim JC, Mirkin SM. The balancing act of DNA repeat expansions. *Curr Opin Genet Dev*. 2013; 23(3):280–288. DOI: 10.1016/j.gde.2013.04.009 [PubMed: 23725800]
8. Shishkin AA, Voineagu I, Matera R, et al. Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. *Molecular cell*. 2009; 35(1):82–92. DOI: 10.1016/j.molcel.2009.06.017 [PubMed: 19595718]
9. Shah KA, Shishkin AA, Voineagu I, et al. Role of DNA polymerases in repeat-mediated genome instability. *Cell reports*. 2012; 2(5):1088–1095. DOI: 10.1016/j.celrep.2012.10.006 [PubMed: 23142667]
10. Shah KA, McGinty RJ, Egorova VI, et al. Coupling transcriptional state to large-scale repeat expansions in yeast. *Cell reports*. 2014; 9(5):1594–1602. DOI: 10.1016/j.celrep.2014.10.048 [PubMed: 25464841]
11. Cherng N, Shishkin AA, Schlager LI, et al. Expansions, contractions, and fragility of the spinocerebellar ataxia type 10 pentanucleotide repeat in yeast. *Proc Natl Acad Sci U S A*. 2011; 108(7):2843–2848. DOI: 10.1073/pnas.1009409108 [PubMed: 21282659]
12. Aksenova AY, Greenwell PW, Dominska M, et al. Genome rearrangements caused by interstitial telomeric sequences in yeast. *Proc Natl Acad Sci USA*. 2013; 110(49):19866–19871. DOI: 10.1073/pnas.1319313110 [PubMed: 24191060]
13. Schmidt KH, Pennaneach V, Putnam CD, et al. Analysis of gross-chromosomal rearrangements in *Saccharomyces cerevisiae*. *Methods Enzymol*. 2006; 409:462–76. DOI: 10.1016/S0076-6879(05)09027-0 [PubMed: 16793418]
14. Aksenova AY, Han G, Shishkin AA, et al. Expansion of Interstitial Telomeric Sequences in Yeast. *Cell Reports*. 2015; 13(8):1545–1551. DOI: 10.1016/j.celrep.2015.10.023 [PubMed: 26586439]
15. Zheng Q. Statistical and algorithmic methods for fluctuation analysis with SALVADOR as an implementation. *Mathematical Biosciences*. 2002; 176(2):237–252. DOI: 10.1016/S0025-5564(02)00087-1 [PubMed: 11916511]
16. Zheng, Q. [Accessed July 7, 2016] rSalvador 1.5: an R tool for the Luria–Delbrück fluctuation assay. 2016. (<http://eeeeeric.github.io/rSalvador>)
17. Hall BM, Ma CX, Liang P, et al. Fluctuation AnaLysis CalculatOR: a web tool for the determination of mutation rate using Luria–Delbrück fluctuation analysis. *Bioinformatics*. 2009; 25(12):1564–1565. DOI: 10.1093/bioinformatics/btp253 [PubMed: 19369502]
18. Gillet-Markowska A, Louvel G, Fischer G. bz-rates: A Web Tool to Estimate Mutation Rates from Fluctuation Analysis. *G3 (Bethesda, Md)*. 2015; 5(11):2323–2327. DOI: 10.1534/g3.115.019836
19. Sarkar S, Ma WT, Sandri GH. On fluctuation analysis: a new, simple and efficient method for computing the expected number of mutants. *Genetica*. 1992; 85(2):173–179. DOI: 10.1007/BF00120324 [PubMed: 1624139]
20. Rosche WA, Foster PL. Determining mutation rates in bacterial populations. *Methods*. 2000; 20(1): 4–17. DOI: 10.1006/meth.1999.0901 [PubMed: 10610800]

Fig. 1.

Experimental system to study GAA repeats expansions in yeast. The selectable cassette contains flanking sequences from chromosome III (depicted in black and grey), *URA3* flanking and coding sequences from chromosome V (green), intron sequence inside *URA3* gene derived from *ACT1* gene from chromosome VI (blue) and *TRP1* flanking and coding sequences from chromosome IV (red).

FluCalc
Version: 2016.7.2

Enter data below or select a file with data: No file chosen

Volume of a culture (μL), V_{tot}

Selective media

Observed numbers of clones, C_{sel}

27
53
47
73
21
40
22
30
45
19
32
38

Dilution factor, D_{sel}

Volume plated (μL), V_{sel}

Complete media

Observed numbers of clones, C_{com}

34
60
55
43
34
59
42
32
44
56
31
58

Dilution factor, D_{com}

Volume plated (μL), V_{com}

Results

Raw values

m	11.0726
μ	12.1233e-06
CI_{lower}	8.4580e-06
CI_{upper}	16.2564e-06
Frequency mean	8.1569e-05

Corrected values

m	15.9743
μ	1.7490e-05
CI_{lower}	1.2736e-05
CI_{upper}	2.2781e-05

Fig. 2.

Screenshot of the FluCalc interface. A user needs to enter the number of colonies on both selective (C_{sel}) and complete media (C_{com}), a dilution factor (D_{sel} and D_{com}) and a volume of cultures (V_{sel} and V_{com}) plated on both media. A volume of the initial culture (V_{tot}) can be changed from its default value of 200 μL . Raw values and corrected values for m - number of mutations; μ - mutation rate; CI – confidence intervals for μ .

Fig.3.
Scheme of the fluctuation test.

Fig. 4. Representative gel electrophoresis of the PCR analysis of 5-FOA-resistant colonies, originated in a strain containing our selectable cassette with the $(GAA)_{100}$ repeat. L – ladder, e – expansion, d – deletion, m – mutation, c – contraction with a simultaneous mutation in the *URA3* body, de – double expansion. The initial length of $(GAA)_{100}$ is ~350 bp that is marked by arrow.