Published in final edited form as:

Science. 2016 March 4; 351(6277): 1094–1097. doi:10.1126/science.aac9786.

Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation

Stephan Uphoff[1,2,*], **Nathan D. Lord**[2], **Burak Okumus**[2], **Laurent Potvin-Trottier**[2,3], **David J. Sherratt**[1], and **Johan Paulsson**[2,*]

[1]Department of Biochemistry, University of Oxford, UK

[2]Department of Systems Biology, Harvard Medical School, USA

[3]Biophysics PhD program, Harvard Medical School, USA

Abstract

Cells rely on the precise action of proteins that detect and repair DNA damage. However, gene expression noise causes fluctuations in protein abundances that may compromise repair. For the Ada protein in *Escherichia coli*, which induces its own expression upon repairing DNA alkylation damage, we found that undamaged cells on average produce one Ada molecule per generation. Because production is stochastic, many cells have no Ada molecules and cannot induce the damage response until the first expression event occurs, sometimes delaying the response for generations. This creates a subpopulation of cells with increased mutation rates. Non-genetic variation in protein abundances thus leads to genetic heterogeneity in the population. Our results further suggest that cells balance reliable repair against toxic side-effects of abundant DNA repair proteins.

The integrity of the genome is constantly threatened by DNA damage. Most damage events are reversed by active repair systems, but the ones that escape repair can cause cell death or mutations. An intriguing question is what causes those failures. Specifically, the classic perspective suggests that failures to repair reflect the intrinsic error rate of the repair enzymes, for example due to the random search for lesions (1, 2). Alternatively, most

*Correspondence to: stephan.uphoff@bioch.ox.ac.uk; johan_paulsson@harvard.edu.

One Sentence Summary:
Success or failure to induce a DNA damage response is dictated by the stochastic presence or absence of a single protein molecule per cell, which causes mutation rate variation between distinct cell subpopulations.

Author contributions: S.U. conceived the study, generated cell strains, designed and performed experiments and analysis. S.U., D.J.S., J.P. interpreted the data. N.L., L.P.-T. developed the microfluidic imaging methods. B.O. developed the single-molecule counting method. S.U., D.J.S., J.P. wrote the manuscript. The authors declare competing financial interests. A United States Patent Application 20150247790 entitled “Microfluidic Assisted Cell Screening” was filed on behalf of B.O., J. P. and co-workers by the President and Fellows of Harvard College.

Supplementary Materials:
Materials and Methods
Figures S1 to S15
Table S1
Movies S1 and S2
References (27–35)

failures could occur in an error-prone subpopulation of cells (3, 4) in which repair is compromised by fluctuations in the abundances of the repair proteins (5–7).

To distinguish between these possibilities we quantitatively analyzed, with single-molecule resolution in single cells, the adaptive response that protects *E. coli* against the toxic and mutagenic effects of DNA alkylation damage (8). The Ada protein functions not only in the direct repair of alkylated DNA but also as the transcriptional activator of the adaptive response (Fig. 1A) (9, 10). Specifically, *ada* expression is induced by methylated Ada (meAda) following irreversible methyl transfer from DNA phosphotriester and O^6MeG lesions onto cysteine residues of Ada. Because Ada is present in low numbers before damage, this positive feedback gene regulation may amplify stochastic fluctuations and create cell-to-cell heterogeneity in the repair system (2, 11).

We imaged the endogenous expression of a functional Ada-mYPet fluorescent protein fusion (fig. S1) in cells treated with methyl methanesulfonate (MMS) (Fig. 1B). We observed a strong and uniform expression of Ada in most cells, but 20% of the cells did not respond at all, even at saturating doses of MMS (Fig. 1, B and C). Quantitatively similar results were obtained with a transcriptional fluorescent reporter in cells with untagged Ada (fig. S2), showing that the protein fusion did not affect the observations.

To visualize the dynamics of the process, we monitored Ada-mYPet abundance in real time in a microfluidic device that allows imaging of single cells over tens of generations during constant DNA damage treatment (fig. S3 and movies S1 and S2) (12, 13). At low to intermediate MMS concentrations (<200 μM MMS), cells showed random unsynchronized pulses of Ada expression (Fig. 1D). The pulse frequency increased proportionally to the MMS concentration (fig. S4), as expected when triggering is limited by Ada finding a lesion. At higher MMS concentrations, most cells rapidly induced a persistent and uniform response (Fig. 1E). However, 20 to 30% of cells were lagging even at saturating MMS, and triggered the response after exponentially distributed delays with an average of one generation time (Fig. 1F and fig. S5). Some cells thus failed to respond for several generations.

To identify the molecular determinants of this heterogeneity, we measured the Ada abundance prior to MMS treatment. Ada-mYPet was undetectable over the auto-fluorescence background of cells, suggesting that absolute amounts were on the order of a few molecules per cell. We therefore turned to single-molecule microscopy to directly count individual proteins in live cells (Fig. 1G and fig. S6). The abundance of Ada was extremely low: the observed population average was 1.4 ± 0.1 molecules per cell (±SEM) and 20 to 30% of the cells did not contain a single Ada molecule. Because the *ada* gene is strictly auto-regulatory, i.e., it can only be induced by the Ada protein (8–10, 14), cells with zero Ada molecules should be unable to trigger the adaptive response, despite high amounts of damage. This is supported by the quantitative agreement between the percentages of cells with a delayed response and with zero Ada molecules. Consequently, the delay before response activation should match the time until the first random expression event occurs in these cells. Indeed, the distribution of Ada copies before damage was very close to a Poisson distribution (Fig. 1G) with an average production rate of one molecule per cell cycle (fig.

S6), and the late-responding cells also activated the response with a Poisson rate of once per cell cycle (Fig. 1F).

These findings also mean that most cells reliably launch the response with just one or two Ada molecules to sense the damage and induce *ada* expression (Fig. 2A). We indeed observed distinct single-molecule signatures: the rates of Ada production displayed stair-case patterns with equidistant states during response activation and deactivation at low MMS concentrations (Fig. 2B and fig. S7), indicative of discrete production and loss events of the meAda molecules that control Ada expression. To further confirm the low numbers, we titrated meAda using promoter sites on a low copy-number plasmid (15), which markedly decreased steady-state Ada induction, as expected (fig. S8). Furthermore, the discrete production rate steps disappeared when meAda abundance was increased using high MMS concentrations (fig. S7).

Because failure to trigger the adaptive response seems due to a complete lack of Ada molecules in a fraction of cells, it should be possible to reduce this fraction with a slight increase in the average abundance of Ada. Specifically, for many distributions (including the Poisson) the probability mass in the tails depends sensitively on the average. We therefore moderately increased Ada numbers per cell either by inhibiting cell division – keeping the concentrations constant (16) – or by expressing additional unlabelled Ada from the P_{Ada} promoter on a very low copy-number plasmid (MiniF; ~2 copies/cell). In both cases we observed the predicted uniform MMS response and disappearance of the late-responding cell subpopulation (Fig. 2, C and D).

These observations raise the question of why the native *ada* gene is expressed at such low basal amounts. Following the fates of single cells over time showed that a failure to activate the adaptive response during MMS treatment lowered the viability of those cells, as expected (Fig. 3A and fig. S9). However, the moderate over-expression of Ada resulted in severe toxicity of MMS treatment (Fig. 3A and fig. S10) (14, 17), and caused spontaneous triggering of the response in the absence of MMS (Fig. 3B and fig. S10), something we never observed at native *ada* expression (Fig. 3B and fig. S3). The extremely low abundance of Ada can thus be advantageous to the population as a whole, implying that the repair system faces a trade-off to repair exogenous alkylation damage without introducing harmful effects. In fact, given the low numbers of molecules, the *ada* regulation is remarkably precise: First, the Poisson distribution before damage shows an almost complete absence of gene expression bursts or ‘extrinsic’ noise (Fig. 3C and fig. S6), in stark contrast to the regulation of most genes studied (5, 6, 7, 18). This can be explained by a short half-life and inefficient translation of *ada* mRNAs (19, 20), and the fact that Poisson noise tends to dominate at very low abundances. Second, a dual reporter assay (5) that simultaneously monitors expression of the endogenous P_{Ada} *ada-mYPet* and an ectopic P_{Ada} *cfp* insertion (Fig. 3, D and E, and fig. S11) showed that both the activation time upon MMS treatment and the subsequent expression dynamics were closely correlated between the two genes, with little uncorrelated noise that would indicate transcriptional bursting. Considering the central role of meAda in *ada* regulation (8–10, 14), these expression dynamics likely reflect fluctuations in meAda numbers. Indeed, the normalized standard deviation was inversely proportional to the square root of the expected average number of DNA damage sites,

quantitatively consistent with the simplest model where varying damage levels determine meAda abundances that then reliably control *ada* expression (fig. S12). Third, *ada* transcription activation is inhibited by unmethylated Ada. This may control response deactivation after the damage has been repaired (21). Indeed, removal of MMS caused all cells to switch off the adaptive response uniformly, and Ada was diluted due to cell growth (Fig. 3F).

The total number of Ada molecules directly determines a cell's repair capacity: each Ada molecule can only act once to remove one mutagenic O^6MeG lesion (10). Furthermore, a lack of Ada repair capacity cannot be compensated for by the DNA mismatch repair pathway because unrepaired O^6MeG lesions miscode for T instead of C. This leads to futile mismatch repair cycles, eventually causing stable mutations during the next round of replication (17). We therefore tested whether heterogeneity in Ada concentrations affects mutation rates. To directly measure genomic mutation rates in single cells we used the DNA mismatch recognition protein MutS as a marker for labelling nascent mutations (22). Specifically, photoactivated single-molecule tracking (23) allowed us to classify individual MutS-PAmCherry fusion proteins as DNA-bound or mobile (24, 25), while also imaging Ada-mYPet in the same live cells (Fig. 4). Without MMS treatment, the apparent mutation frequency was low (fig. S13) and most MutS molecules were mobile (average 6% bound) (fig. S14). MMS treatment of Ada deficient cells (Δ*ada*) increased both the mutation frequency (Fig. S13) and MutS binding (56% bound) (Fig 4 and fig. S14). MMS treatment of wild-type cells resulted in highly variable amounts of bound MutS molecules between cells. This variation could be entirely explained by the heterogeneity in Ada expression (Fig. 4 and fig. S15): MutS binding was increased only in the subpopulation of cells with low Ada expression (30% bound) whereas cells with abundant Ada retained low MutS activity (10% bound). Stochastic activation of the adaptive response therefore leads to an error-prone cell subpopulation that does not efficiently repair DNA alkylation damage and accumulates mutations.

We found that a cell's fate following DNA damage can be accurately predicted by the presence or absence of a single protein molecule. The resulting cell heterogeneity increases the chance of genetic adaptation in a hypermutagenic subpopulation of cells without jeopardizing the genetic integrity in the majority of the population during stress (3, 26). However, our observations that high Ada expression is toxic and that cells appear to minimize the heterogeneity in several ways suggest that this is not an adaptive bet-hedging strategy, but rather a side-effect of maximizing short-term fitness: Because proteins with the capacity to modify DNA can be detrimental, cells may be forced to express them in low amounts, such that random fluctuations are unavoidable. Mutations can then result from stochastic variation in the concentrations of DNA repair proteins. Just as genetic heterogeneity can cause phenotypic heterogeneity, the reverse is thus also true.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank R. Reyes-Lamothe, U. Alon, J-Y Bouet, A. Kapanidis, C. Lesterlin, A. Upton, P. Zawadzki for reagents and discussions. We thank C. Saenz and the Microfluidics Core Facility at Harvard Medical School. Microscopy at Micron Oxford was supported by a Wellcome Trust Strategic Award (091911) and MRC grant (MR/K01577X/1). S.U. is funded by a Sir Henry Wellcome Fellowship by the Wellcome Trust and a Junior Research Fellowship at St John's College, Oxford. J.P., N.D.L., L.P-T., B.O. are funded by NIH Grant GM095784. LPT acknowledges fellowship support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Fonds de recherche du Québec -- Nature et technologies. D.J.S. is funded by a Wellcome Trust Investigator Award (099204/Z/12Z). The primary data described in the manuscript is available upon request.

References and Notes

1. Kad NM, Van Houten B. Dynamics of lesion processing by bacterial nucleotide excision repair proteins. Prog. Mol. Biol. Transl. Sci. 2012; 110:1–24. [PubMed: 22749140]
2. Uphoff S, Kapanidis AN. Studying the organization of DNA repair by single-cell and single-molecule imaging. DNA Repair. 2014
3. Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a Stress Response and the Regulation of Evolvability. Crit. Rev. Biochem. Mol. Biol. 2007; 42:399–435. [PubMed: 17917874]
4. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nat. Rev. Cancer. 2012; 12:323–334. [PubMed: 22513401]
5. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. Science. 2002; 297:1183–1186. [PubMed: 12183631]
6. Golding I, Paulsson J, Zawilski SM, Cox EC. Real-Time Kinetics of Gene Activity in Individual Bacteria. Cell. 2005; 123:1025–1036. [PubMed: 16360033]
7. Choi PJ, Cai L, Frieda K, Xie XS. A Stochastic Single-Molecule Event Triggers Phenotype Switching of a Bacterial Cell. Science. 2008; 322:442–446. [PubMed: 18927393]
8. Samson L, Cairns J. A new pathway for DNA repair in Escherichia coli. Nature. 1977; 267:281–283. [PubMed: 325420]
9. Landini P, Volkert MR. Regulatory Responses of the Adaptive Response to Alkylation Damage: a Simple Regulon with Complex Regulatory Features. J. Bacteriol. 2000; 182:6543–6549. [PubMed: 11073893]
10. Sedgwick B. Repairing DNA-methylation damage. Nat. Rev Mol Cell Biol. 2004; 5:148–157. [PubMed: 15040447]
11. Alon U. Network motifs: theory and experimental approaches. Nat. Rev. Genet. 2007; 8:450–461. [PubMed: 17510665]
12. Wang P, et al. Robust growth of Escherichia coli. Curr. Biol. CB. 2010; 20:1099–1103. [PubMed: 20537537]
13. Norman TM, Lord ND, Paulsson J, Losick R. Memory and modularity in cell-fate decision making. Nature. 2013; 503:481–486. [PubMed: 24256735]
14. Shevell DE, LeMotte PK, Walker GC. Alteration of the carboxyl-terminal domain of Ada protein influences its inducibility, specificity, and strength as a transcriptional activator. J. Bacteriol. 1988; 170:5263–5271. [PubMed: 3141384]
15. Brewster RC, et al. The Transcription Factor Titration Effect Dictates Level of Gene Expression. Cell. 2014; 156:1312–1323. [PubMed: 24612990]
16. Locke JCW, Young JW, Fontes M, Jiménez MJH, Elowitz MB. Stochastic Pulse Regulation in Bacterial Stress Response. Science. 2011; 334:366–369. [PubMed: 21979936]
17. Fu D, Calvo JA, Samson LD. Balancing repair and tolerance of DNA damage caused by alkylating agents. Nat. Rev Cancer. 2012; 12:104–120. [PubMed: 22237395]
18. Maamar H, Raj A, Dubnau D. Noise in Gene Expression Determines Cell Fate in Bacillus subtilis. Science. 2007; 317:526–529. [PubMed: 17569828]
19. Bernstein JA, Khodursky AB, Lin P-H, Lin-Chao S, Cohen SN. Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. Proc. Natl. Acad. Sci. 2002; 99:9697–9702. [PubMed: 12119387]

20. Li G-W, Burkhardt D, Gross C, Weissman JS. Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. Cell. 2014; 157:624–635. [PubMed: 24766808]
21. Saget BM, Walker GC. The Ada protein acts as both a positive and a negative modulator of Escherichia coli's response to methylating agents. Proc. Natl. Acad. Sci. 1994; 91:9730–9734. [PubMed: 7937881]
22. Elez M, et al. Seeing mutations in living cells. Curr. Biol. CB. 2010; 20:1432–1437. [PubMed: 20674359]
23. Manley S, et al. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. Nat Methods. 2008; 5:155–157. [PubMed: 18193054]
24. Uphoff S, Reyes-Lamothe R, de Leon FG, Sherratt DJ, Kapanidis AN. Single-molecule DNA repair in live bacteria. Proc. Natl. Acad. Sci. U. S. A. 2013; 110:8063–8068. [PubMed: 23630273]
25. Liao Y, Schroeder JW, Gao B, Simmons LA, Biteen JS. Single-molecule motions and interactions in live cells reveal target search dynamics in mismatch repair. Proc. Natl. Acad. Sci. 2015:201507386.
26. Veening J-W, Smits WK, Kuipers OP. Bistability, Epigenetics, and Bet-Hedging in Bacteria. Annu. Rev. Microbiol. 2008; 62:193–210. [PubMed: 18537474]
27. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. 2000; 97:6640–6645. [PubMed: 10829079]
28. Reyes-Lamothe R, Sherratt DJ, Leake MC. Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. Science. 2010; 328:498–501. [PubMed: 20413500]
29. Baba T, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2006; 2:2006.0008.
30. Zaslaver A, et al. A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. Nat. Methods. 2006; 3:623–628. [PubMed: 16862137]
31. Bouet J-Y, Rech J, Egloff S, Biek DP, Lane D. Probing plasmid partition with centromere-based incompatibility. Mol. Microbiol. 2005; 55:511–525. [PubMed: 15659167]
32. Sliusarenko O, Heinritz J, Emonet T, Jacobs-Wagner C. High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. Mol. Microbiol. 2011; 80:612–627. [PubMed: 21414037]
33. Holden SJ, et al. Defining the limits of single-molecule FRET resolution in TIRF microscopy. Biophys J. 2010; 99:3102–3111. [PubMed: 21044609]
34. Chung SH, Kennedy RA. Forward-backward non-linear filtering technique for extracting small biological signals from noise. J. Neurosci. Methods. 1991; 40:71–86. [PubMed: 1795554]
35. Uphoff S, Sherratt DJ, Kapanidis AN. Visualizing protein-DNA interactions in live bacterial cells using photoactivated single-molecule tracking. J. Vis. Exp. JoVE. 2014

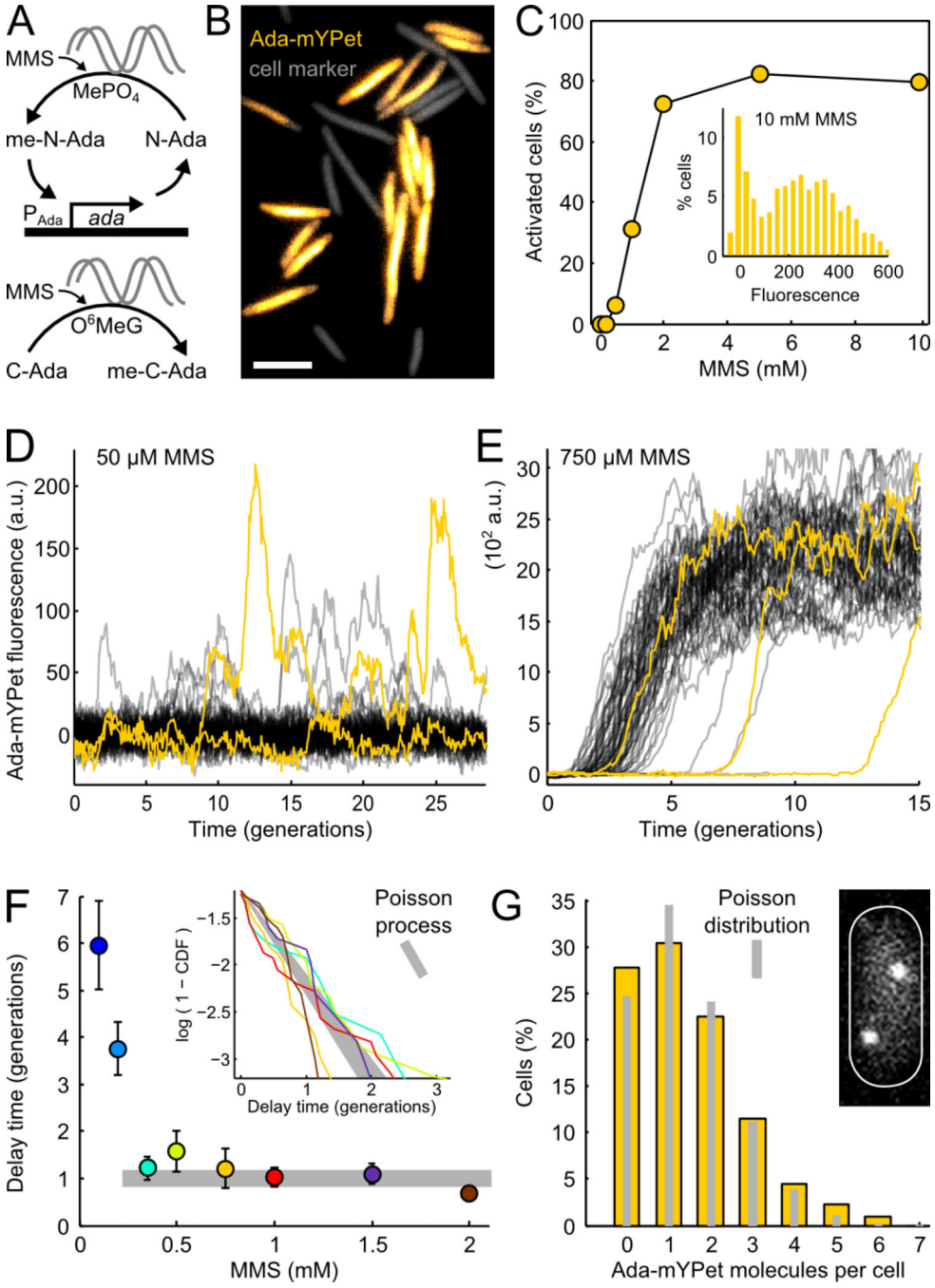

Fig. 1. Stochastic gene expression delays Ada response activation in a cell subpopulation

(A) Methylation of Ada N- and C-terminal domains functions as a damage sensor, turning Ada into an auto-regulatory activator of genes involved in DNA alkylation repair. **(B)** Ada-mYPet fluorescence (yellow) in cells treated with 10 mM MMS for 1 hour. Constitutive mKate2 serves as fluorescent cell marker (gray). Scale bar, 5 µm. **(C)** Percentage of cells that activated Ada-mYPet expression after 1 hour in MMS. (Inset) Histogram of Ada-mYPet fluorescence per cell with 10 mM MMS. **(D and E)** Time traces of Ada-mYPet fluorescence in single cells treated with 50 µM and 750 µM MMS (added at time 0). Example cells in

yellow; time in units of average generation times (42 min) throughout. **(F)** (Inset) Transformed cumulative distribution log(1-CDF) of response delay times for the last 30% of cells to activate Ada-mYPet expression upon MMS treatment in the microfluidic chip. Different MMS concentrations in colours as in main plot. Straight lines on log scale reflect exponential distributions as generated by a Poisson process; the slope corresponds to the average delay time constant. Gray area: Poisson process with a rate of 1 ± 0.1 per generation. Main plot: Average delay time constants from the inset data (±SEM). **(G)** Single-molecule counting of Ada-mYPet without MMS. Example cell shown. Poisson model was generated using measured production rate of 1 molecule per generation. Note that the actual value may be closer to 1.2 because of delayed maturation of mYPet (see supplementary materials).

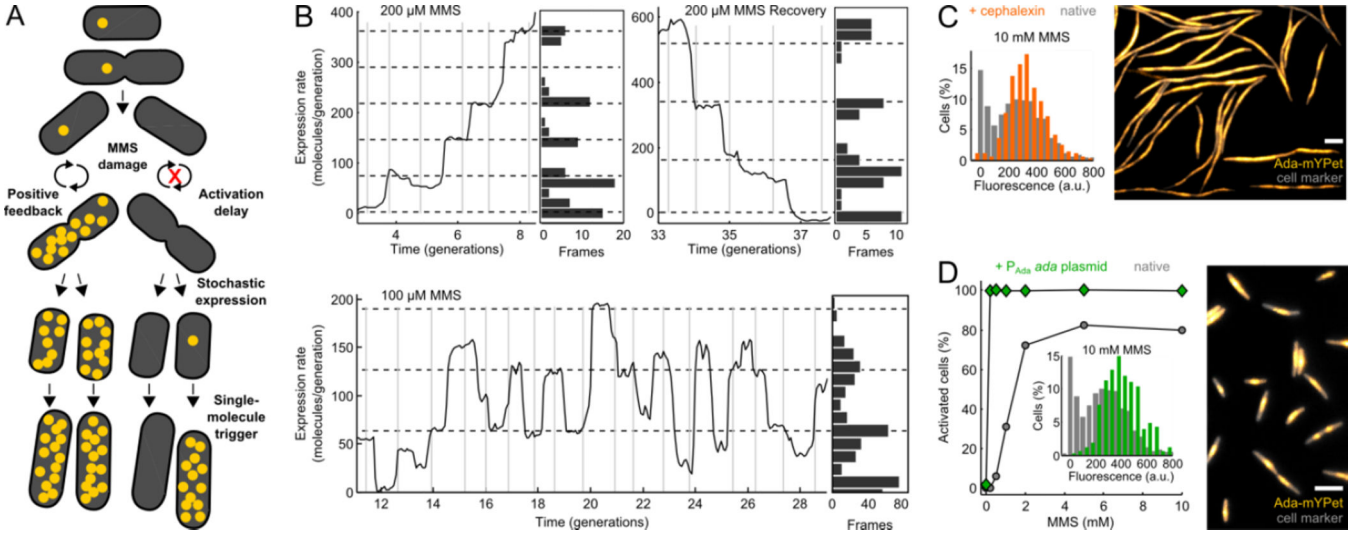

Fig. 2. Single-molecule trigger of the Ada response

(A) Stochastic expression and random segregation of molecules at cell division creates a subpopulation of cells with zero Ada molecules which therefore fails to autoinduce the adaptive response. **(B)** Sections of time traces showing distinct steps in Ada-mYPet expression rates during response activation upon 200 µM MMS treatment, deactivation after MMS removal, and stochastic activation and deactivation transitions with 100 µM MMS. Vertical lines indicate cell divisions. Histograms show number of frames spent in the expression rate states. Losses can occur due to rare meAda degradation or by segregation at cell division. At very low numbers, all meAda molecules should sometimes remain in the same cell, maintaining expression rates as observed. **(C)** Uniform Ada-mYPet induction when cell division was inhibited with cephalexin prior to MMS treatment (orange). **(D)** Uniform accumulation of endogenous Ada-mYPet with additional MiniF plasmid carrying P_{Ada} *ada* (green). Scale bars, 5 µm.

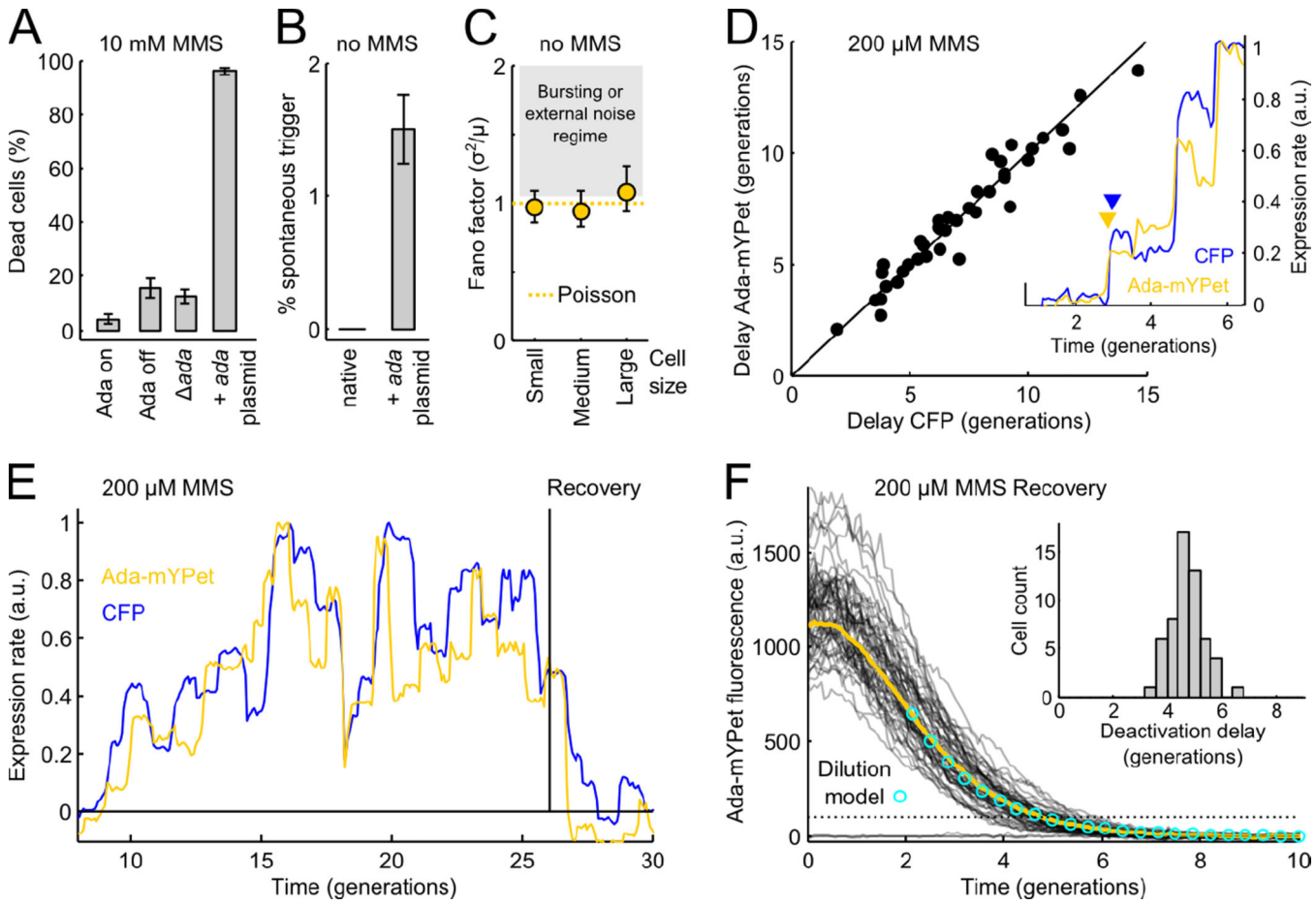

Fig. 3. High precision of the Ada response

(A) Cell fates after treatment with 10 mM MMS for 1 hour: Percentages of cells failing to recover growth during time-lapse microscopy without MMS for 3 hours (±SEM). Cells were distinguished if they had activated (Ada on) or failed the response (Ada off). **(B)** Percentages of cells spontaneously triggering Ada-mYPet expression without MMS (±SEM). **(C)** Fano factors (variance/mean) for Ada-mYPet without MMS, using single-molecule counting data (Fig. 1G, ±SEM bootstrapped). Cells grouped by size. Expression bursting would give Fano factors above Poisson limit of 1. **(D)** Dual reporter assay: Delay times between MMS addition and response activation for endogenous *ada-mYPet* and ectopic P_{Ada} *cfp* are closely correlated. Each dot represents one cell. Inset: Example expression rate time-traces with simultaneous activation of both genes. **(E)** Example time-traces showing correlated expression rate fluctuations of the dual reporter genes and simultaneous response deactivation after MMS removal. **(F)** Deterministic response deactivation: Time-traces following MMS removal at time 0 (average: yellow). The dilution model (circles) has an exponential decay constant equal to the average generation time. Inset: Narrow distribution of delay times from MMS removal until response is deactivated (dotted line threshold).

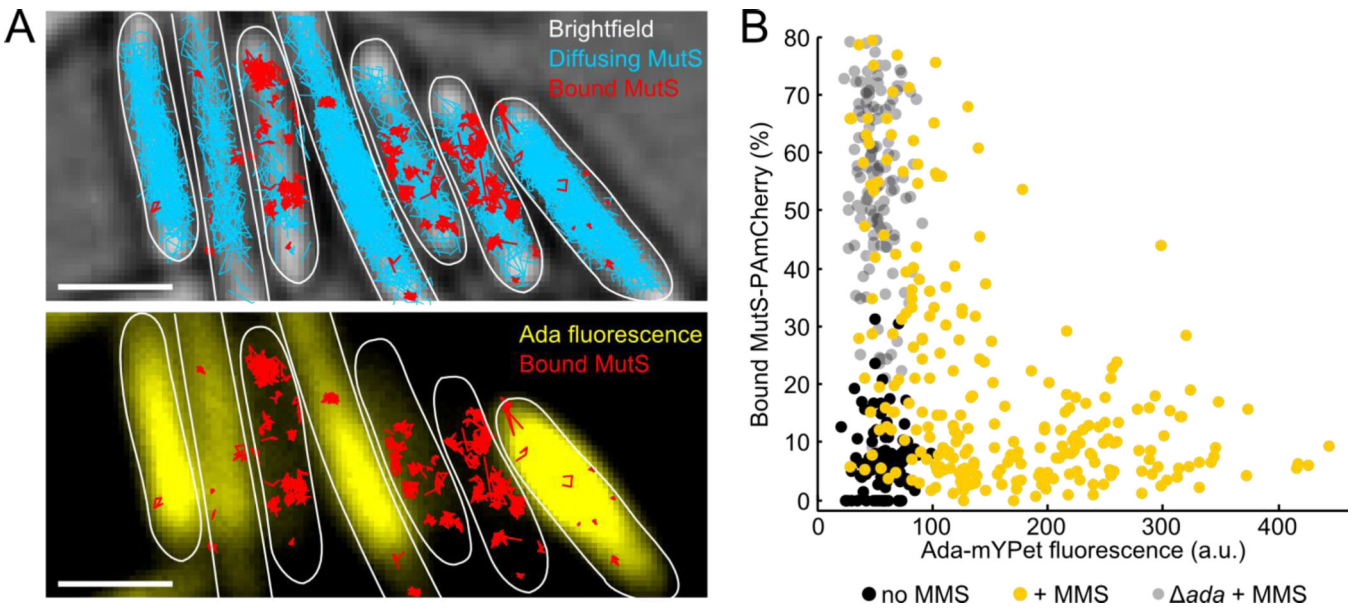

Fig. 4. Increased binding of mismatch recognition protein MutS in cells with delayed Ada response

Photoactivated single-molecule tracking of MutS-PAmCherry and Ada-mYPet fluorescence in single cells treated with 10 mM MMS for 1 hour. **(A)** Tracks of bound (red) and mobile MutS (blue). Cell outlines drawn; scale bars, 2 μm. **(B)** Percentage of bound MutS molecules vs. Ada-mYPet fluorescence per cell. Native strain with (yellow) and without MMS (black); Δ*ada* with MMS (gray).