

ARTICLE

Received 21 Apr 2016 | Accepted 28 Jul 2016 | Published 3 Oct 2016

DOI: 10.1038/ncomms12760

OPEN

# Excess of mutational jackpot events in expanding populations revealed by spatial Luria–Delbrück experiments

Diana Fusco<sup>1,2</sup>, Matti Gralka<sup>1</sup>, Jona Kayser<sup>1,2</sup>, Alex Anderson<sup>1</sup> & Oskar Hallatschek<sup>1,2</sup>

The genetic diversity of growing cellular populations, such as biofilms, solid tumours or developing embryos, is thought to be dominated by rare, exceptionally large mutant clones. Yet, the emergence of these mutational jackpot events is only understood in well-mixed populations, where they stem from mutations that arise during the first few cell divisions. To study jackpot events in spatially structured populations, we track mutant clones in microbial populations using fluorescence microscopy and population sequencing. High-frequency mutations are found to be massively enriched in microbial colonies compared with well-shaken liquid cultures, as a result of late-occurring mutations surfing at the edge of range expansions. Thus, jackpot events can be generated not only when mutations arise early but also when they occur at favourable locations, which exacerbates their role in adaptation and disease. In particular, because spatial competition with the wild type keeps most mutant clones in a quiescent state, strong selection pressures that kill the wild type promote drug resistance.

<sup>1</sup>Department of Physics, University of California, Berkeley, California 94720, USA. <sup>2</sup>Department of Integrative Biology, University of California, Berkeley, California 94720, USA. Correspondence and requests for materials should be addressed to O.H. (email: ohallats@berkeley.edu).

One of the hallmarks of spontaneous mutations in growing populations is the emergence of mutational jackpot events—large mutant clones arising from mutations that by chance occur early in the development of a cellular population so that their progenitors benefit from prolonged growth. Due to their sheer size, these jackpot events, first discovered by Luria and Delbrück<sup>1</sup>, are thought to have momentous roles in short-term evolutionary processes, including the adaptation from standing variation<sup>2–4</sup>, evolutionary rescue<sup>5</sup>, drug resistance evolution<sup>6–10</sup> and the somatic evolution of genetic diseases<sup>11,12</sup>. However, because the emergence of jackpot events has been understood only in uniformly growing populations<sup>1,10,13</sup>, it is currently impossible to predict their impact on the evolution of many naturally structured populations.

The original Luria–Delbrück experiment studied mutant clones arising in well-mixed microbial populations and detected the size of resistant clones by counting single colonies on selective plates. Here, we generalize the assay in two ways: (i) by studying mutant clones arising in spatially structured populations and (ii) by using a combination of next-generation sequencing and fluorescence microscopy techniques to accurately detect size and structure of high-frequency clones.

We find that high-frequency clones are massively enriched in microbial colonies compared with well-shaken liquid cultures, as a result of late-occurring mutations surfing at the edge of range expansions<sup>14–16</sup>. We provide a mathematical theory that explains the observed excess of jackpot events and predicts their role in promoting rare evolutionary outcomes. In particular, we show that resistant clones generated by surfing can become unleashed under high selection pressures, and thus represent a drug resistance hazard for high-dose drug treatments. In this context, our theory offers an innovative explanation for the phenomenon of ‘competitive release’, initially observed in ecology<sup>17–19</sup>, and more recently in tumour evolution<sup>20</sup>, where the craved resource is space rather than nutrients.

## Results

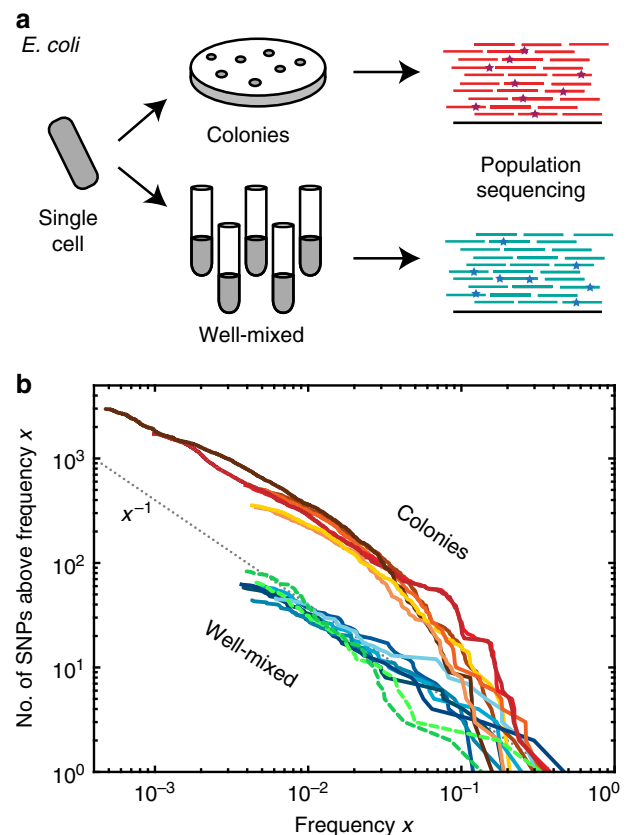
**Generalized Luria–Delbrück experiments.** To measure the size of mutational jackpot events, we employed population sequencing with low error rates (see Methods section), which returned frequencies of new mutations at many genomic sites simultaneously and independently of their phenotypic effect. Specifically, we sequenced populations of a mutator strain of *Escherichia coli* cultured in well-mixed liquid medium, where growth is uniform, and as colonies on solid agar medium, where most growth occurs at the colony edge (see Methods section)<sup>21,22</sup>. By growing from a small number of initial cells to a similar final size, all populations went through a comparable number of cell divisions (between  $1 \times 10^9$  and  $7 \times 10^9$ , Supplementary Table 1). Counting the observed frequencies of Single Nucleotide Polymorphisms (SNPs) in the populations, we obtained the number of sites in the genomes where the clonal sub-population carrying the derived mutation had a frequency larger than a given frequency value  $x$ , shown in Fig. 1b. Our deep sequencing procedure allowed us to detect all clones that have frequencies larger than about  $10^{-3}$  (see Methods section), yielding around 600 such high-frequency events in each colony, which characterize the statistics of jackpot events.

Populations of the same size and mutation rate are expected to experience the same total number of mutational events, on average, independently of the mode of growth (liquid versus solid medium). Yet, Fig. 1 shows that colonies had approximately ten times more mutant clones above frequencies of 1%—corresponding to clones of at least  $10^7$  cells—than well-mixed populations. This difference cannot be explained by the variation in final population size, since some well-mixed populations had a larger final size than some colonies (Supplementary Table 1, Supplementary Fig. 1).

To test whether the discrepancy was caused by different mutation rates in liquid culture and on agar plates, we also sequenced ‘plated well-mixed’ populations whose growth was kept uniform by regularly spreading the cells across the plate (see Methods section). The resulting distribution of SNPs is consistent with the well-mixed populations in liquid culture (green dashed and blue solid lines in Fig. 1b), confirming that the mutation rate is not significantly affected by the mode of growth (see also Supplementary Methods and Supplementary Table 2). Thus, we conclude that the observed difference in the clone size distribution must be a consequence of the non-uniform growth in colonies, which results in a surprising number of mutations at high frequency.

## Fluorescence microscopy reveals spatial structure of clones.

To study the nature of the high-frequency clones in colonies, we monitored the spatial distribution of mutant clones using separate fluorescent marker experiments. We employed a genetically engineered budding yeast strain capable of switching from a red-fluorescing state to a green-fluorescing state at a rate of about  $1.6 \times 10^{-3}$  per replication<sup>23</sup>. This heritable, non-reversible switch is mediated by the stochastic expression of Cre recombinase



**Figure 1 | Population sequencing reveals an excess of jackpot events in spatially growing populations.** (a) Starting from few cells of a mutator strain of the bacterium *E. coli* (*mutT* deletion, see Methods section), we grew 6 colonies and 6 liquid cultures up to an average population size of  $3 \times 10^9$  cells and sequenced each population at a coverage of at least  $1000 \times$  (Supplementary Table 1). The number of SNPs that occurred at a frequency higher than  $x$  is displayed in panel (b) (solid lines: colonies in warm colours, well-mixed in cold colours; dashed lines: well-mixed on plates). We found about 10 times more mutants above a frequency of 1% in colonies than in well-mixed population, even when the latter were grown on plates (Methods). The dotted black line is the fit to the well-mixed data with a mutation rate of  $\mu = 0.4$  per genome per replication.

(see Fig. 2a and Methods section for details). As shown in Fig. 2c, the resulting colonies exhibited both elongated speckles (dark arrows), which we termed ‘bubbles’, as well as previously described spoke-like sectors (white arrow)<sup>21</sup>. Importantly, image analysis of the clone area obtained from 343 colonies yielded a histogram consistent with the shoulder-like distribution obtained from our sequencing approach (Fig. 2b and Methods section). Thus, both the fluorescent data from just one ‘engineered’ site and the sequencing data covering many genomic sites seem to reflect the same mechanism shaping the clone size distribution.

The fluorescence data, moreover, revealed where clones emerge and how they grow. Time-lapse movies showed that most high-frequency clones first arise near the front of the growing colony (Fig. 2e, Supplementary Movie 1). The resulting clonal patches grow with the advancing frontier until they lose contact to the front, whereupon they become trapped as bubbles in the non-growing bulk of the population. Rarely, clones are able to ‘surf’ at the front until the end of the experiment and give rise to sectors. Such allele surfing is a characteristic feature of range expansions<sup>14,16,24</sup> and has been demonstrated to be pervasive in microbial communities<sup>21,25–27</sup>.

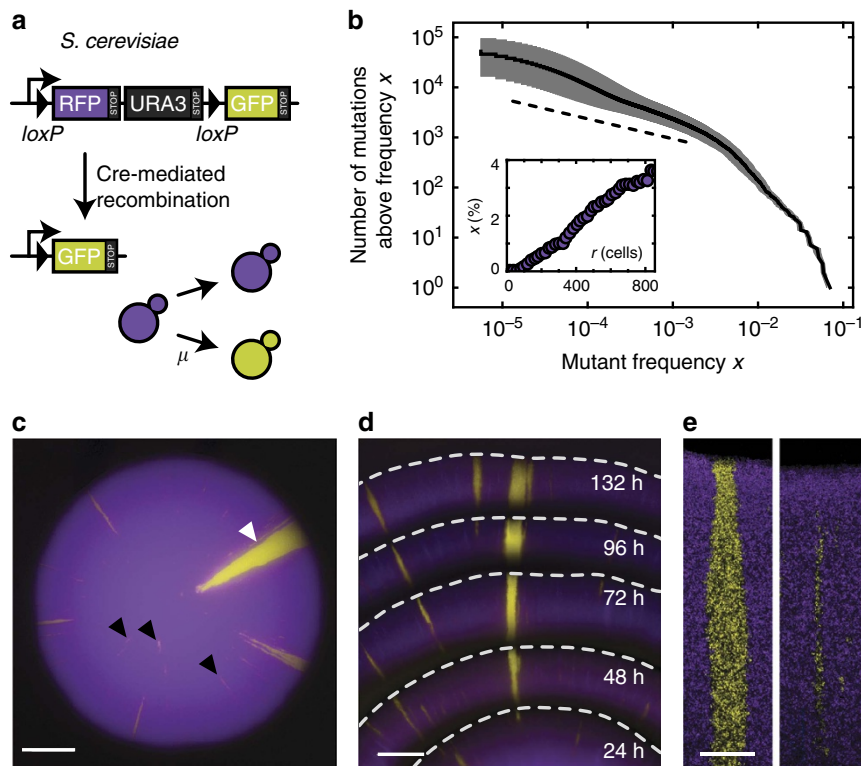
### Gene surfing theory explains mutational jackpot events.

To understand how gene surfing generates clones of different sizes, we studied their emergence in range expansion simulations. Specifically, colony growth was implemented by the random addition of new demes to the advancing frontier. The newly

added demes inherited their ancestral genotype unless they mutated, which occurred at a fixed rate (see Methods section; Supplementary Movie 2). Interestingly, this simple meta-population model generated a clone size distribution that accurately reproduced the measured one, as can be seen in Fig. 3. Our simulation results, covering over four orders of magnitude, also reveal that the distribution crosses over between two power-law distributions with distinct exponents. Analysis of the clone shapes produced in our simulations shows that the power-law regimes of low and high frequencies characterize bubbles and sectors, respectively.

The effect of gene surfing is not limited to two-dimensional (2D) growth: simulations of spherically growing meta-populations (Fig. 4b), such as those used to model solid tumours<sup>28,29</sup>, still result in mutant spectra with two distinct power-law regimes corresponding to bubbles and sectors.

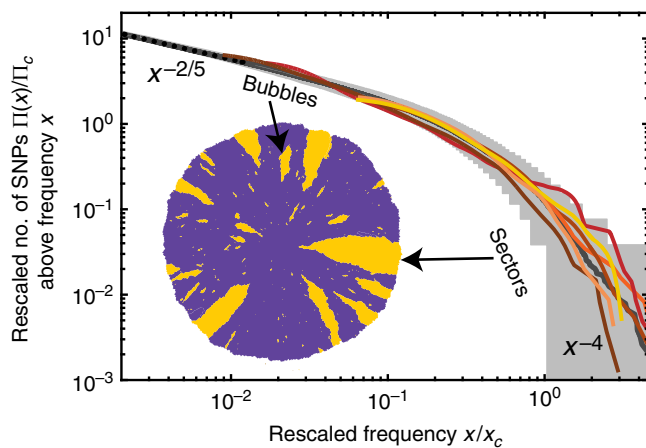
The power-law exponents can be derived analytically by treating the boundaries of the mutant clones as annihilating random walks. The statistical properties of these random walks determine the relationship between the length  $L_{\parallel}$  of clones parallel to the growth direction and the corresponding perpendicular length  $L_{\perp}$  (see Fig. 4a). For instance, purely diffusive clone boundaries, which occur when the population front is completely flat, result in  $L_{\parallel} \sim L_{\perp}^2$ ; more generally,  $L_{\parallel} \sim L_{\perp}^z$ , where the dynamical exponent  $z$  depends on the details of how the population grows (dimensionality and roughness of the growing front). The Kardar–Parisi–Zhang (KPZ) interface growth model<sup>30</sup>, which has been found consistent with bacterial growth patterns<sup>21</sup>, predicts  $z_{2D} = 3/2$  (exact) and  $z_{3D} = 1.63$



**Figure 2 | Visualizing mutational jackpot events in colonies.** (a) An engineered strain of *S. cerevisiae* that stochastically switches from RFP (purple) to GFP (yellow) at a rate of  $1.6 \times 10^{-3}$  per cell division enables us to visualize high frequency jackpot events as they arise during colony growth (c, scale bar, 1 mm). (b) Image analysis (see Methods section) of mutant clones in 343 colonies reveals a shoulder-like distribution of clone sizes, roughly consistent with our predictions for idealized 2D populations (dashed line, see also Methods section). The inset shows the median mutant frequency as a function of radius  $r$ . Error bars are smaller than symbols. (d, scale bar, 0.5 mm) Monitoring the spatial distribution of mutants using fluorescence time-lapse microscopy (see also Supplementary Movie 1) reveals that mutant clones come either as sectors<sup>21</sup> with actively growing front regions (e, left, scale bar, 0.2 mm; c, white arrow) or as ‘bubbles’ (e, right; c, black arrows), which are non-growing mutant clones that have lost contact with the expanding edge.

(numerical)<sup>31</sup>. By computing the area enclosed between two annihilating random walks, which are unbiased in the neutral case, one can determine the power-law exponent for bubbles (see Methods section). The power-law exponent for sectors follows from the one for bubbles via an exact scaling relationship (see Methods section). The results for different scenarios are summarized in Table 1.

Importantly, our scaling arguments predict that the clone size distributions obtained for different population sizes and mutation rates should collapse onto one master curve when the clones frequencies are measured in terms of the characteristic frequency of the largest occurring bubble  $x_c$ , and the clone number in terms of the average probability for a new mutation to establish a sector,  $\Pi_c$ . Indeed, after rescaling, our sequencing data show remarkable agreement with the master curve obtained from our simulations (Fig. 3, Supplementary Methods, Supplementary Table 3, and Supplementary Fig. 2).



**Figure 3 | Simulations of range expansions reproduce the measured distribution of clone sizes.** The clone size distribution obtained from our meta-population simulations (grey solid line) exhibits two predictable power-law tails that characterize bubbles (dotted line) and sectors, respectively (see Methods section). The experimental mutant spectra (coloured lines) collapse onto this theory line upon scaling the two axes by the frequency of the largest bubble  $x_c$  and the probability  $\Pi_c$  of observing a clone larger than  $x_c$ , respectively (the shaded area represents the 95% confidence interval, see Methods section for details).

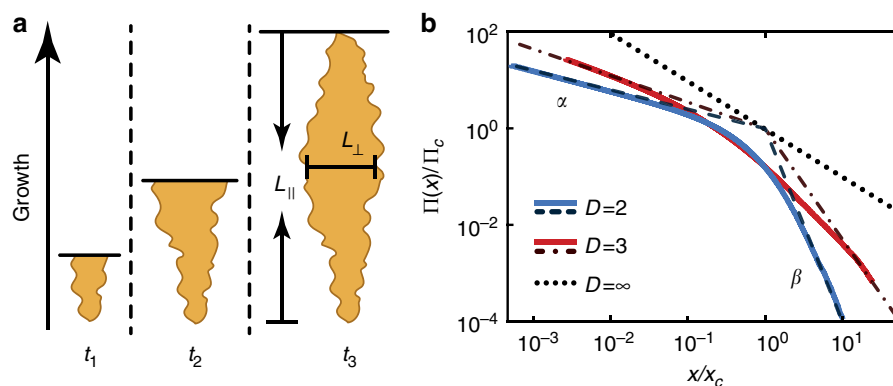
**Role of jackpot events in the evolution of drug resistance.** One of the striking features of the clone size distribution under range expansion is the excess of high frequency mutations over the well-mixed expectation. The primary consequence of more jackpot events is that, typically, the total number of mutants will be much larger in a spatially growing population compared with an equally large well-mixed population. This can be understood from the following simple mathematical argument.

If mutations arise at a low rate  $\mu$  per cell division as the population is growing to a final size  $N$ , one expects  $\mu N$  mutational events to occur. In each generation, the frequency of mutants in the population increases by  $\mu$ , on average. In a well-mixed uniformly grown population, the number of generation is  $\log_2 N$ , and hence the expected total number of mutants in this case is proportional to  $\mu N \log_2 N$ . In a range expansion, only cells near the edge of the colony have access to sufficient nutrients, leading to the formation of a layer of growing cells of width  $\lambda$  (in units of cell diameters). Since a length  $\lambda$  is added per generation to the radius of the colony, we can estimate that  $R/\lambda \propto N^{1/2}/\lambda$  generations elapse at the frontier during the growth process. Hence, the final total number of mutants created during a range expansion is proportional to  $\mu N^{3/2}/\lambda$ , which for large  $N$  is much larger than in a uniformly grown population of the same size. Note that, as in the classic Luria–Delbrück case, the mean is usually not a useful quantity because it is dominated by very rare, large events. Nevertheless, both the typical and the mean number of mutations exceed the well-mixed expectations as our stochastic analysis shows (Fig. 5, Supplementary Note 1 and Supplementary Fig. 3).

Jackpot events can be key in acquiring complex phenotypes, such as drug resistance or the onset of cancer, which often require the accumulation of multiple mutations<sup>11,32</sup>. Range expansions may promote the acquisition of secondary mutations because the pool of individuals carrying the first (driver) mutation is, both on average and typically, larger than in uniformly grown populations (Fig. 5a). As shown in Fig. 5b, the probability of secondary mutations can be almost an order of magnitude larger in spatial populations compared with well-mixed ones, especially when mutation rates are low (Supplementary Note 2).

**Drugs can trigger competitive release of mutant clones.**

Evolutionary dynamics is influenced by mutational jackpot events not only because of their size but also because of their particular spatial structure. The emergence of sectors, which sporadically arise from neutral mutations, is strongly suppressed when mutants carry a cost (Fig. 6a, Supplementary Note 3 and

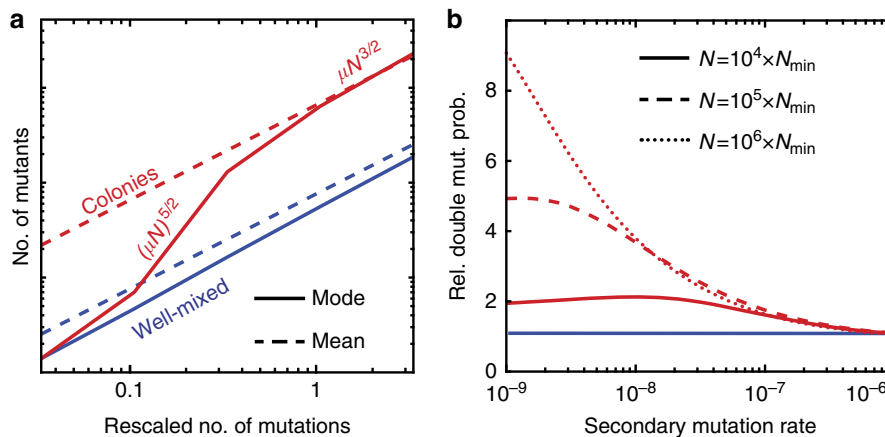


**Figure 4 | Fractal properties of clone boundaries control the size distribution of bubbles.** (a) Sketch of the stochastic growth of a bubble, which is bounded by two annihilating random walks. (b) Rescaled clone size distribution from simulations in two and three dimensions approach the predicted asymptotic power-law behaviour for bubbles,  $x^{-\alpha}$ , and sectors,  $x^{-\beta}$ , in the limit of small and large frequencies, respectively (dashed lines: asymptotic theory predictions). The values of the exponents are reported in Table 1.

**Table 1 | Predictions for the asymptotic behaviour of the clone size distribution.**

Scenario	$z$	Low frequency	High frequency
Well-mixed	0	$x^{-1}$	$x^{-1}$
2D flat front	2	$x^{-1/3}$	$x^{-3}$
3D flat front	2	$x^{-1/2}$	$x^{-2}$
2D rough front*	3/2	$x^{-2/5}$	$x^{-4}$
3D rough front	1.61	$x^{-0.55}$	$x^{-2.3}$

The dynamical exponent  $z$  summarizes the statistical properties of the sector and bubble boundaries. The low-frequency regime corresponds to bubbles, the high-frequency regime to sectors. The well-mixed scenario, where bubbles and sectors are not distinct, corresponds to  $D = \infty$  and is characterized by a single power-law regime. The experimental scenario presented here is well described by the case of 2D rough fronts (\*), where  $z = 3/2$  is given by KPZ (Kardar-Parisi-Zhang) interface growth dynamics<sup>30</sup>.



**Figure 5 | Gene surfing theory predicts an excess of jackpot events and its consequences.** (a) Sampling the experimental clone size distribution  $\Pi$ , we found that the typical (mode, solid line) and the mean (dashed line) number of mutants as a function of the rescaled number of mutations  $\mu N$ , which corresponds to the number of mutations establishing into sectors, is higher in range expansions (red) than in equally large uniformly grown populations (blue). (b) The excess of large-frequency clones promotes multi-step evolutionary processes, such as the emergence of double mutants. The relative probability (compared with the well-mixed expectation) of producing a double mutant is always higher in a colony, especially when the secondary mutation rate is low. Here,  $N_{\min} = x_{\min} N$  (see Methods section), which in our *E. coli* colonies corresponds to  $N_{\min} \approx 10^5$ .

Supplementary Fig. 4), commonly observed for drug resistance in the absence of antibiotics<sup>33</sup>: deleterious mutants can surf only briefly before they are overtaken by faster-growing wild-type cells and fall behind the growing frontier.

In the absence of antibiotics, then, costly resistant clones are expected to reside in bubbles, encased by an expanding wild-type population. Upon a sudden environmental change, for example, by a strong antibiotic attack killing the susceptible wild type, the trapped mutants may become unleashed, regrowing and thus rescuing the population from extinction. This evolutionary rescue is brought about by a particular kind of competitive release, in which the indispensable resource is space. Competition in this case is extreme: trapped clones are not only at a disadvantage, but they have no chance of escaping unless the surrounding wild type vanishes. Consistently with this idea, we expect minimal net population growth and successful containment of resistant cells at intermediate drug concentrations, sufficiently strong to slow down proliferation of the wild type without eradicating it. Indeed, when we implemented drug treatment as a tunable death rate for wild-type cells (Fig. 6b, Supplementary Movie 3, see Methods section), our simulations showed the smallest net population growth for intermediate death rates. In contrast, high death rates not only failed to eradicate the population, but promoted the spread of resistance by generating an entirely resistant population.

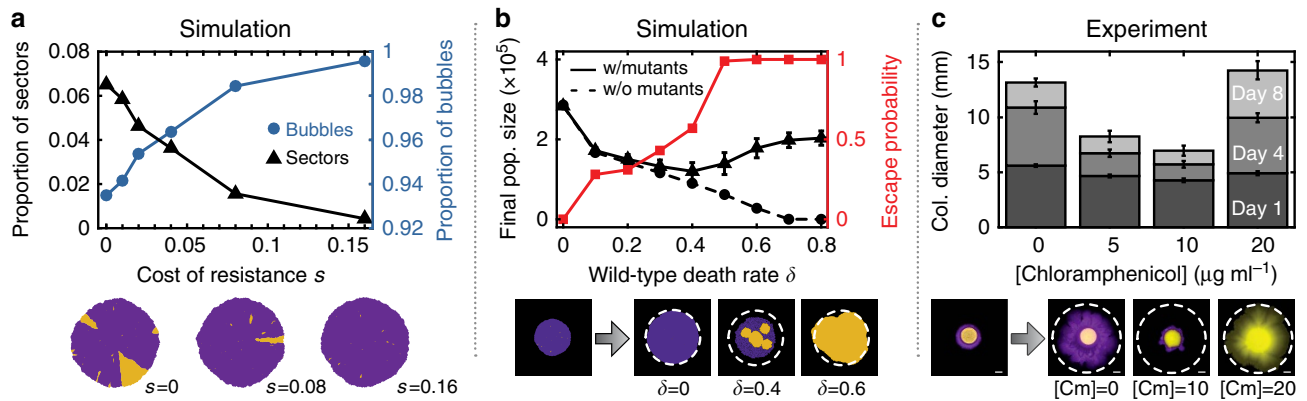
Our simulations provide a uniquely spatial mechanism of how excessively high drug concentrations can promote the spreading of drug resistance. This effect can be reproduced in conceptual

experiments in which we embedded resistant cells within a colony of susceptible cells. We found that the resistant cells stayed trapped even at intermediate drug concentrations that severely limit the wild-type growth. For higher drug concentrations, the mutants were released and grew rapidly (Fig. 6c). Hence, to optimally curtail microbial growth in our particular set-up, the drug concentration should indeed be set at an intermediate sweet spot.

## Discussion

In combination with a generalization of the Luria-Delbrück theory, our experiments show that the process of allele surfing generates an excess of high-frequency clones: mutations have a much higher chance of being carried by a high proportion of the population. These high-frequency clones come as growing sectors and non-growing bubbles, which are spatially encased by wild-type cells. Our theory suggests that the excess of jackpot events is not limited to microbial colonies but arises generally in populations that exhibit non-uniform growth rates in two or three dimensions.

In addition to antibiotic resistance evolution in pathogenic biofilms, an excess of jackpot events could thus be relevant also during the somatic evolution of some types of cancer<sup>12,34,35</sup>, as growing solid tumours often exhibit less growth in necrotic core regions<sup>36</sup> and sectoring has been recently documented<sup>29</sup>. Moreover, it has been argued that jackpot events play a crucial role in the predisposition to cancer and other genetic diseases<sup>11</sup>.



**Figure 6 | High drug concentration treatments can promote the spreading of resistance by unleashing mutants that are otherwise trapped within the colony bulk by competitive release.** (a) When the mutants carry a cost  $s$  in the absence of antibiotics, the fraction of mutants trapped in bubbles is increased and sectors become increasingly rare. (b) Starting with (i) colonies pre-grown with  $s = 0.4$  (black triangles, solid line) or (ii) resistance-free control colonies (black circles, dashed line) as initial condition, we simulated the effects of treatment with antibiotics at varying concentrations by tuning the death rate  $\delta$  of the wild type (see Methods section). After 140 generations, resistance-free colonies showed decreasing population size up to complete eradication for  $\delta > 0.7$ , as expected. In contrast, colonies with resistant bubbles exhibited an initial decrease in population size followed by a steady increase for  $\delta > 0.4$  (error bars represent standard deviations over 100 simulations). Visualization of the simulated colonies showed that in this range of death rates, the previously encapsulated mutants escape the surrounding wild type (red squares) and can then grow indefinitely (Supplementary Movie 3). A high dose of antibiotics may thus not only fail to eradicate the population, but even promote the spreading of resistance. This effect can be reproduced in conceptual experiments with *E. coli*, shown in (c), which demonstrate that high antibiotic concentrations can release trapped mutant clones. A droplet of resistant cells (yellow) embedded in a larger droplet of susceptible cells (purple) was inoculated at different antibiotic concentrations (see Methods section). After 8 days of growth intermediate antibiotic concentrations exhibited the least amount of total population growth. The highest drug concentrations eradicated the wild type and thus allowed the resistant mutants to spread freely. Error bars represent standard deviations over 16 replicates and scale bars correspond to 2 mm.

Based on the classic Luria–Delbrück theory, it was predicted that a large fraction of cancers may arise from predisposed stem-cell lineages<sup>32</sup>. Our results suggest that if growth is non-uniform during the development of the stem cell pool, predisposed stem-cell lines may occur even more frequently than previously hypothesized.

Our sequencing study design can be applied to test this prediction by measuring the site frequency spectrum in tumours as well as in multi-cellular organisms at different stages of their development. Some clues about large-frequency clones could also be collected from visible phenotypes resulting from somatic mosaicism, such as Blaschko’s lines<sup>11</sup>. Plants or animals with pigmentations caused by mobile elements are particularly amenable to such pattern analyses.

Our experiments, simulations and theory not only found an excess of jackpot events but also elucidated their spatial patterns. Since drug resistance often comes at a selective cost in the absence of the drug, most resistance mutant clones are expected to be small and hidden in the bulk of the susceptible wild-type population. A high-dose treatment removes the wild type as the natural competitor of the resistant mutants and, consequently, triggers a competitive release of the dormant mutants. This form of competitive release is extreme in the sense that mutants stop growing entirely before the release, due to lack of space and nutrients. Because a broad spectrum of dormant clones is also generated in three-dimensional (3D) growth, this spatial form of competitive release can be particularly relevant in solid tumours. Indeed, the hypothesis that growth control may, under certain conditions, be more effective than attempts of complete eradication has been recently proposed in the context of cancer<sup>37,38</sup> based on mathematical modelling of exponentially growing tumours treated with a varying dose of chemotherapeutic drugs over time. A recent study<sup>20</sup> tested this hypothesis experimentally in a mouse model and found that prolonged chemotherapy with low doses was the most effective at keeping the tumour in check. In contrast, tumours that had regrown after

an initial shrinkage following high-dose treatment did not respond to a second round of treatment, possibly indicating the emergence of resistance. Our results provide a novel mechanism, based on competition for space, that can explain these observations.

As sequencing costs continue to decrease, a growing number of studies utilizes population sequencing to draw conclusions on how cellular populations evolve. In this context, our results urge for caution when employing the classic Luria–Delbrück theory as a general theory for neutral evolution, as has been proposed in a recent meta-study of intra-tumour heterogeneity<sup>12</sup>. For one thing, spatial effects complicate the estimation of mutation rates and the amount of resistant tumour cells. More importantly, deviations from the classic Luria–Delbrück theory may simply indicate non-uniform growth rather than non-neutral evolution.

## Methods

**Scaling of clone size distribution in expanding populations.** We consider a population expanding from one to  $N$  individuals without death. Our goal is to characterize the probability  $\Pi(x)$  that a mutation randomly introduced at the birth of one of the  $N$  cells generates a clone of frequency equal or larger than  $x$  (Fig. 1a). In other words,  $\Pi(x)$  is the reverse cumulative distribution of the size of clone introduced by a random mutation. The derivative  $-\partial_x \Pi(x)$  represents the population site frequency spectrum that can be obtained directly from population sequencing. Two extreme values of  $\Pi(x)$  are known *a priori*:  $\Pi(1/N) = 1$  because any mutation will certainly reach at least frequency  $1/N$  if we ignore death. On the other hand,  $\Pi(1) = 1/N$  because this requires the mutation to be introduced in the very first birth, that is, at the root of the genealogical tree.

The behaviour of  $\Pi(x)$  in between these boundary points is known in the well-mixed or infinite-dimensional case, studied by Luria and Delbrück; it is a single power law  $(Nx)^{-1}$  (ref. 13). Our experiments and simulations suggest that the finite-dimensional case is characterized by two asymptotic regimes: a low-frequency regime  $(Nx)^{-\alpha}$  with exponent  $\alpha < 1$ , corresponding to bubbles, and a high-frequency regime  $N^{-1}x^{-\beta}$  with exponent  $\beta > 1$ , corresponding to sectors. The  $N$ -dependent pre-factors of both regimes are fixed by the boundary conditions, and the crossover point  $x_c$  follows from continuity (Supplementary Fig. 5). The described crossover behaviour can be captured mathematically by the scaling form

$$\Pi(x) = \Pi_c \chi\left(\frac{x}{x_c}\right), \quad (1)$$

in terms of the crossover frequency  $x_c = N^{-\frac{1-\alpha}{\beta-2}}$  and a crossover probability scale  $\Pi_c = N^{-\frac{\alpha(\beta-2)}{\beta-2}}$ . The scaling function  $\chi(\xi)$  exhibits two power-law regimes

$$\chi(\xi) \sim \begin{cases} \xi^{-\alpha}, & \xi \ll 1, \\ \xi^{-\beta}, & \xi \gg 1. \end{cases} \quad (2)$$

Simulated clone size distributions, displayed in Fig. 4b, confirm the predicted scaling form. In real systems, the scaling form has a range of validity, say between some minimal and maximal frequency  $x_{\min}$  and  $x_{\max}$  that are set by details of the growth processes, which are beyond the scope of our model. For instance, the maximal frequency  $x_{\max} = O(1)$  accounts for the discrete nature of growth during the first few cell divisions. Similarly,  $x_{\min} = O(1/N)$  may reflect mutations that are born behind the front such that our gene surfing theory does not apply.

The two power-law exponents  $\alpha$  and  $\beta$  are constrained by the scaling behaviour of the mean frequency of mutants, as follows. For a given mutation rate  $\mu$  per cell division, the expected frequency  $\langle X_{\text{tot}}(t) \rangle$  of mutants at the surface of a growing  $D$ -dimensional sphere satisfies

$$\partial_t \langle X_{\text{tot}}(t) \rangle = \mu(1 - \langle X_{\text{tot}} \rangle), \quad (3)$$

where time is measured in units of generations at the front. Hence, as long as  $\langle X_{\text{tot}}(t) \rangle \ll 1$ , we have  $\langle X_{\text{tot}}(t) \rangle = \mu t$ . Every generation, the radius grows by  $\lambda$ , such that it takes the population  $t = R/\lambda$  generations to grow to size  $N \propto R^D$ . The length  $\lambda$  can be interpreted as a measure for the thickness of the growth layer at the edge of the colony in units of the linear dimension of the cells. The mean frequency at radius  $r$  then is  $\langle X_{\text{tot}}(r) \rangle = \mu r/\lambda$ , and the population mean follows as  $\overline{\langle X_{\text{tot}} \rangle} \propto (\mu/\lambda)N^{1/D}$ . This implies an expected total number of mutants of

$$\overline{\langle X_{\text{tot}} \rangle} N \propto \frac{\mu}{\lambda} N^{1+1/D}. \quad (4)$$

As we ignore cell death, there are on average  $\mu N$  mutations occurring during the growth of the population. Hence, the frequency of a single clone is, on average,  $\langle X \rangle \propto N^{\beta-1}/\lambda$ , which constrains the integral over  $\Pi(x)$  since  $\langle X \rangle = \int_0^1 \Pi(x)$ . Inserting the scaling relation reported in equation (1), we obtain for large population  $N \gg 1$

$$\langle X \rangle = \int_0^1 \Pi(x) \approx \Pi_c x_c \int_0^{x_c^{-1}} \chi(\xi) = N^{-\frac{1+\beta-2\alpha}{\beta-2}}. \quad (5)$$

Here, we used that  $\alpha < 1$  and  $\beta > 1$  such that integral over the scaling function  $\chi(\xi)$  is finite and dominated by  $\xi$ -values in the crossover region,  $\xi = O(1)$ . Equation (5) can only hold for all  $N$  if we have the scaling relation

$$\frac{(\alpha-1)(\beta-1)}{\alpha-\beta} = \frac{1}{D} \quad (6)$$

Thus, computing the exponent for either the bubble or the sector regime uniquely determines the exponent of the other regime. Note that in the well-mixed case,  $D \rightarrow \infty$ , equation (6) correctly reproduces a single power law with  $\alpha = \beta = 1$ . Here, we derive the exponent for the bubble regime in two dimensions; the sector regime and higher dimensions are described in Supplementary Notes 4 and 5.

**Bubble distribution in two dimensions.** The area  $A$ , length  $L_{\parallel}$  and width  $L_{\perp}$  of a bubble are determined by the statistical properties of its boundaries, which result in the general relationship  $L_{\parallel} \sim L_{\perp}^z$  (see Fig. 4a). We can relate the distribution of the area  $A \sim L_{\perp} L_{\parallel} \sim L_{\perp}^{1+z}$  of a bubble to the distribution of the lateral bubble size  $L_{\perp}$  through

$$\Pr(A > a) = \Pr(L_{\perp} > L_{\perp}(a) \sim a^{\frac{1}{1+z}}). \quad (7)$$

For neutral mutations, we must also have the conditional probability  $\Pr(L_{\perp} > w_{\perp} | \ell_{\perp}) \sim \ell_{\perp}/w_{\perp}$  of reaching transverse size  $w_{\perp}$  given an initial size  $\ell_{\perp}$ , because, asymptotically, each front segment of size  $\ell_{\perp}$  has the same chance of expanding up to size  $w_{\perp}$ . This implies  $\Pr(L_{\perp} > w_{\perp})w_{\perp} \sim \Pr(L_{\perp} > \ell_{\perp})\ell_{\perp}$ . Since the choice of  $w_{\perp}$  is arbitrary, it must hold that  $\Pr(L_{\perp} > \ell_{\perp}) \sim \ell_{\perp}^{-1}$  for large enough  $\ell_{\perp}$ . Combining this with equation (7) yields the distribution of clone frequencies

$$\Pr(A > a) \sim a^{-\frac{1}{1+z}} \quad (8)$$

for asymptotically large bubbles, that is,  $\alpha = 1/(1+z)$  in two dimensions. In  $D$  dimensions,  $\alpha = [1 + z/(D-1)]^{-1}$  (Supplementary Note 5). The scaling relation in equation (6) can be used to find the exponent in the sector regime  $\beta = [1 + D(z-1)]/[D-1)(z-1)]$ . Numerical values for the exponents in a variety of scenarios are given in Table 1.

**Meta-population model simulations.** We simulated range expansions using a meta-population model based on the Eden model<sup>32</sup>: space is divided into a 2D or 3D square grid, whose voxels can be empty, wild type or mutant type. In general, each voxel represents a sub-population, also called deme, consisting of multiple individuals, and the lattice size  $a$  is a model parameter that characterizes the spatial extent of a deme.

The grid is initialized by filling the central voxel with wild type. In each step, we choose a filled voxel  $i$  that has at least one empty neighbour and one of its empty neighbouring voxels  $j$  at random and copy the state of  $i$  into  $j$ . If  $i$  is wild type, we switch ('mutate') voxel  $j$  to the mutant type with probability  $\mu$ . A generation

corresponds to a number of steps equal to the number of voxels that have at least one empty neighbour at the beginning of the generation.

To accommodate the limited number of observable SNPs in the experimental data when comparing with simulations, we sample  $M$  mutations with frequency above the minimum experimentally detectable frequency from the simulated distribution. Here,  $M$  is the number of observed mutations in one experiment (Supplementary Table 1). The sampling is repeated 10,000 times, and the resulting distributions binned across frequencies. The result is shown as the grey area in Fig. 3 containing 95% of the sampled distributions.

To simulate the effect of intermediate antibiotic concentrations on colony growth and spreading of resistant individuals observed in experiments (Fig. 6c), we extended the meta-population model simulation to accommodate sudden changes in the environment and death of the wild type. Environmental changes are modelled by changing the relative growth rate of mutants,  $g_{\text{mut}}$ , and wild type,  $g_{\text{wt}}$ , at time  $T$  during the simulations, which corresponds to when the antibiotic is administered. We define the selective advantage/disadvantage  $s$  of the mutant as  $s = g_{\text{mut}}/g_{\text{wt}} - 1$ . Without death of the wild type, the environmental change would only be felt by the individuals at the front, in contradiction to experiment, where we frequently observe the escape of mutants from the bulk of the population. Therefore, we introduce the possibility of death: after time  $T$ , any of the  $n_{\text{wt}}$  wild type cells has a chance  $\delta$  of dying each generation. In addition to the standard algorithm above, each generation then incurs  $\delta n_{\text{wt}}$  additional iterations, in which one wild-type voxel is deleted. More details on the algorithm are reported in Supplementary Methods. To simulate the experimental scenario in Fig. 6, we first grow a population with deleterious mutations until time  $T$  ( $\mu = 2 \times 10^{-4}$ ,  $s = -0.4$ ), whereafter new mutations are not allowed, mutants switch to having a selective advantage ( $s = 0.4$ ), and wild-type death is switched on. Supplementary Movie 3 shows the qualitative agreement with the experimental results in identifying an intermediate death rate that minimizes colony growth.

**E. coli experiments.** For the sequencing experiments, we used a K12 *E. coli* strain MG1655 where the *mutT* gene was replaced by *Cm<sub>R</sub>*, conferring resistance to chloramphenicol. This strain has an elevated mutation rate for A/T → C/G transversions over wild-type MG1655 by a factor of about 150 on average<sup>40</sup>. For the trapping experiments in Fig. 6, we used a pair of MG1655 strains, one expressing CFP from a plasmid, the other expressing YFP and a resistance gene to chloramphenicol.

In liquid culture, *E. coli* were grown continuously shaken at 37 °C in Luria-Bertani (LB) broth (10 g l<sup>-1</sup> tryptone, 5 g l<sup>-1</sup> yeast extract and 10 g l<sup>-1</sup> NaCl). For plates, 2% w/v bacto agar was added to the media before autoclaving. Antibiotics were added after autoclaving to cooled media.

To prepare cells of *mutT E. coli* for sequencing, we grew them in liquid culture up to a density of 10<sup>8</sup> cells per ml and then (i) plated single cells for colonies and (ii) inoculated well-mixed cultures from single cells following a 10<sup>7</sup> dilution. Parallel tubes (86) were inoculated with 100 μl of the dilution and incubated well shaken for about 10 h (around 30 generations). Cells were harvested in log phase, judged by OD600 (Supplementary Table 1). To grow well-mixed populations on plates, 16 LB plates were inoculated with 100 μl of the dilutions and incubated. Every 90 min, 100 μl PBS was pipetted onto each plate and vigorously spread using glass beads. After ~20 h, seven plates displayed a uniform bacterial lawn. Two lawns were resuspended via vortexing. Colonies were grown for 3 to 5 days up to a diameter between 1 and 1.5 cm (Supplementary Fig. 2). For five colonies (colonies 1–5 in Supplementary Table 1 and Supplementary Fig. 6), cells were resuspended by vortexing and the genomic DNA extracted for each population. For the last colony (colony 6), the colony was cut in four parts via a glass pipette for the centre portion (IN) and a razor blade for the remaining outer ring. The outer ring was divided into three parts: 1/8, 1/4 and the remainder of the ring (around 1/2 ring), as shown in Supplementary Fig. 6. The DNA of each portion was extracted separately (details are in Supplementary Table 1). Genomic DNA was extracted from all population following the Epicentre MasterPure DNA Purification Kit Protocol.

Illumina library preparation was performed on each sample and final libraries were used at similar concentrations to achieve similar coverage across samples. The average insert size of the library was ~200 bps. The sample libraries were then sequenced on the HiSeq 2500 at the QB3 Vincent J. Coates genomic facility at UC Berkeley using 150 paired-end reads. Because the library insert size is on average smaller than 300 bps, the two paired-end reads overlap, providing two independent calls for each base in the overlapping region<sup>41,42</sup>. Each sequencing lane accommodated six distinct samples providing, on average, a coverage of 1,000 × per sample. One colony (colony 5 in Supplementary Table 1) was also sequenced separately in one lane, generating a 6,000 × coverage for this sample.

**Processing of sequencing data.** The reads of each sample were processed according to the following pipeline. Read quality was assessed via FastQC to check for base sequence quality, GC content bias, length distribution and adaptor contamination. Because some adaptor contamination was determined, reads were filtered for library adaptors and trimmed when necessary.

Paired-end reads were then merged using BBmerge (available at <http://sourceforge.net/projects/bbmap>), which identifies the optimal relative position of corresponding read pairs and generates a unique consensus read with combined quality scores (on average, 90% of the reads were uniquely merged and

used for the subsequent analysis). The resulting sequencing error was on average lower than  $10^{-6}$  for all samples.

BreSeq<sup>43</sup> was then used to map the merged reads to the reference genome of *E. coli* strain MG1655 (NCBI id: NZ\_CP009685.1, ref. 44). SNP-calling and frequency calculation was also performed via BreSeq conditioning on at least four independent read calls.

Genomic regions with unusually high density of SNPs were omitted. To determine whether SNPs were to be disregarded, a sliding window of 5 kbps was passed across the genome counting the number of SNPs with frequency lower than a varying threshold. For each frequency threshold  $i$ , the average number of SNPs per window  $n_i$  was recorded. If a window showed a number of SNPs with  $P$ -value lower than 0.001 (assuming a Poisson distribution with mean  $n_i$ ), all the SNPs in that window were removed. This procedure takes into account that SNPs with lower frequency are more numerous and thus more densely populate the genome, while providing a more conservative approach than simply removing regions with high SNP density regardless of their frequency. The flagged regions were often shared among samples and annotated as repetitive regions in the reference genome.

Among the remaining putative mutations, SNPs present in more than 40% of the population were deemed mutations carried by the seeding cell(s) and removed. Three of the liquid cultures appeared to have been seeded by multiple cells, since they contained SNPs at frequency equal to 50%. The remaining SNPs with corresponding frequency were then used to generate the clone size distribution in Fig. 1 for all samples, with the exception of colony 6. The details on how we combined the SNPs from the different portions of colony 6 are reported in Supplementary Methods.

**S. cerevisiae experiments.** For the fluorescent marker experiments in budding yeast (Fig. 2), we used the W303 *S. cerevisiae* strain JRY10643 derived from JRY9628 (ref. 23). This strain employs the Cre-loxP recombination system to switch stochastically from a red to a green fluorescent state at a rate  $\mu = 1.6 \times 10^{-3}$  per cell division. *S. cerevisiae* were grown at 30°C in yeast extract peptone dextrose (YPD) (20 g l<sup>-1</sup> peptone, 10 g l<sup>-1</sup> yeast extract and 20 g l<sup>-1</sup> dextrose). For plates, 2% w/v bacto agar was added to the media before autoclaving. To grow colonies from single cells, saturated overnight culture was diluted 1:10 in fresh media and grown for another 4.5 h. The resulting culture was diluted in PBS to give about 50 cells per ml. A volume of 100 µl of this dilution was spread on YPD plates (containing roughly 20 ml of YPD with 2% agar) that had been dried at room temperature for at least 24 h. After drying, the plates were wrapped with parafilm and incubated at 30°C for 5 days.

**Image analysis.** For the timelapse movie (Supplementary Movie 1), single cells of JRY10643 were inoculated on YPD plates, incubated in a stage-top incubator fitted to a Zeiss AxioZoom microscope and grown overnight. One colony was selected and imaged every 30 min in both the red and the green fluorescent channel. For Fig. 2d, a colony was imaged every 24 h and images overlaid in Adobe Photoshop. To image sectors and bubbles on the single-cell scale, a colony was cut out from the agar plates and imaged on a Zeiss LSM700 confocal microscope, using 488 and 555 nm lasers. A  $z$ -stack was recorded and later combined by maximum intensity projection.

Colonies were imaged on a Zeiss AxioZoom v16 upright microscope. The red and green fluorescent channel were recorded separately, and exposure times were set automatically by the software for each colony and channel. To measure the mutant clone size distribution in the converting budding yeast strain, we used an automated thresholding algorithm with a locally adaptive threshold. To detect large clones, we removed small object by computing the geodesic opening of the green channel image before binarizing with a locally (50 pixel radius) adaptive threshold. For the detection of small bubbles, we computed the top hat transform of the green channel, using a 15 pixel radius disk as the structuring element, which effectively removes large elements from the image. The resulting image was then segmented using an adaptive threshold in a 15 pixel radius neighbourhood. Finally, the two segmented images were overlaid and eroded by 1 pixel to obtain the final segmentation. In Fig. 2b, we also show the result for no erosion and 2 pixel erosion (grey area).

The imaged budding yeast colonies are not strictly 2D but have a roughly conical shape. Small clones thus occupy stretched 3D volumes  $V \sim L_{||} L_{\perp}^2$ , of which only the projection can be observed under the microscope. To take this projection error into account, we consider the volume  $V$  of bubbles growing in 3D, and, assuming isotropic growth, we compute the projected area  $A_{\text{proj}} \sim V/L_{\perp}$ , where  $L_{\perp}$  is the size of the bubble section transverse to the growth direction. Using  $z_{3D} = 1.61$ , this leads to

$$\text{Pr}(A_{\text{proj}} > A) \sim A^{-0.77}, \quad (9)$$

which serves as an upper bound for the case where bubbles are equally extended horizontally as they are vertically. Experimentally, the best fit to the low-frequency regime of the clone size distribution in Fig. 2b gives an exponent of roughly 0.61, which is consistent with bubbles that have a small degree of three-dimensionality but grow mostly in the  $x$ - $y$  plane.

**Data availability.** The alignment files obtained from sequencing the *E. coli* populations are available in the Sequence Read Archive (SRA) with access code

SRP078606. The rest of the data that support the findings of this study are available from the corresponding author upon request.

## References

- Luria, S. E. & Delbrück, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**, 491–511 (1943).
- Hermisson, J. & Pennings, P. S. Soft sweeps molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335–2352 (2005).
- Barrett, R. D. & Schluter, D. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44 (2008).
- Messer, P. W. & Petrov, D. A. Population genomics of rapid adaptation by soft selective sweeps. *Trends Ecol. Evol.* **28**, 659–669 (2013).
- Bell, G. & Gonzalez, A. Evolutionary rescue can prevent extinction following environmental change. *Ecol. Lett.* **12**, 942–948 (2009).
- Coldman, A. & Goldie, J. A stochastic model for the origin and treatment of tumors containing drug-resistant cells. *Bull. Math. Biol.* **48**, 279–292 (1986).
- Michor, F., Iwasa, Y. & Nowak, M. A. Dynamics of cancer progression. *Nat. Rev. Cancer* **4**, 197–205 (2004).
- Haeno, H., Iwasa, Y. & Michor, F. The evolution of two mutations during clonal expansion. *Genetics* **177**, 2209–2221 (2007).
- Durrett, R. & Moseley, S. Evolution of resistance and progression to disease during clonal expansion of cancer. *Theor. Popul. Biol.* **77**, 42–48 (2010).
- Kessler, D. A., Austin, R. H. & Levine, H. Resistance to chemotherapy: patient variability and cellular heterogeneity. *Cancer Res.* **74**, 4663–4670 (2014).
- Frank, S. A. Somatic mosaicism and disease. *Curr. Biol.* **24**, R577–R581 (2014).
- Williams, M. J., Werner, B., Barnes, C. P., Graham, T. A. & Sottoriva, A. Identification of neutral tumor evolution across cancer types. *Nat. Genet.* **48**, 238–244 (2016).
- Lea, D. & Coulson, C. A. The distribution of the numbers of mutants in bacterial populations. *J. Genet.* **49**, 264–285 (1949).
- Edmonds, C. A., Lillie, A. S. & Cavalli-Sforza, L. L. Mutations arising in the wave front of an expanding population. *Proc. Natl Acad. Sci. USA* **101**, 975–979 (2004).
- Klopfstein, S., Currat, M. & Excoffier, L. The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol.* **23**, 482–490 (2006).
- Hallatschek, O. & Nelson, D. R. Gene surfing in expanding populations. *Theor. Popul. Biol.* **73**, 158–170 (2008).
- Connell, J. H. The influence of interspecific competition and other factors on the distribution of the barnacle *Chthamalus stellatus*. *Ecology* **42**, 710–723 (1961).
- Keane, R. M. & Crawley, M. J. Exotic plant invasions and the enemy release hypothesis. *Trends Ecol. Evol.* **17**, 164–170 (2002).
- De Roode, J. C., Culleton, R., Bell, A. S. & Read, A. F. Competitive release of drug resistance following drug treatment of mixed plasmodium chabaudi infections. *Malar. J.* **3**, 1 (2004).
- Enriquez-Navas, P. M. *et al.* Exploiting evolutionary principles to prolong tumor control in preclinical models of breast cancer. *Sci. Transl. Med.* **8**, 327ra24–327ra24 (2016).
- Hallatschek, O., Hersen, P., Ramanathan, S. & Nelson, D. R. Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl Acad. Sci.* **104**, 19926–19930 (2007).
- Mitri, S., Clarke, E. & Foster, K. R. Resource limitation drives spatial organization in microbial groups. *ISME J.* **10**, 1471–1482 (2015).
- Dodson, A. E. & Rine, J. Heritable capture of heterochromatin dynamics in *Saccharomyces cerevisiae*. *Elife* **4**, e05007 (2015).
- Excoffier, L. & Ray, N. Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol. Evol.* **23**, 347–351 (2008).
- Buttery, N. J. *et al.* Structured growth and genetic drift raise relatedness in the social amoeba *Dictyostelium discoideum*. *Biol. Lett.* **8**, 794–797 (2012).
- Nadell, C. D. *et al.* Cutting through the complexity of cell collectives. *Proc. R. Soc. Lond. B: Biol. Sci.* **280**, 20122770 (2013).
- van Gestel, J., Weissing, F. J., Kuipers, O. P. & Kovács, A. T. Density of founder cells affects spatial pattern formation and cooperation in *Bacillus subtilis* biofilms. *ISME J.* **8**, 2069–2079 (2014).
- Waclaw, B. *et al.* A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature* **525**, 261–264 (2015).
- Ling, S. *et al.* Extremely high genetic diversity in a single tumor points to prevalence of non-darwinian cell evolution. *Proc. Natl Acad. Sci.* **112**, E6496–E6505 (2015).
- Kardar, M., Parisi, G. & Zhang, Y.-C. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56**, 889 (1986).
- Halpin-Healy, T. & Zhang, Y.-C. Kinetic roughening phenomena, stochastic growth, directed polymers and all that. Aspects of multidisciplinary statistical mechanics. *Phys. Rep.* **254**, 215–414 (1995).
- Frank, S. A. & Nowak, M. A. Cell biology: developmental predisposition to cancer. *Nature* **422**, 494–494 (2003).



33. Levin, B. R., Perrot, V. & Walker, N. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics* **154**, 985–997 (2000).
34. Sottoriva, A. *et al.* A big bang model of human colorectal tumor growth. *Nat. Genet.* **47**, 209–216 (2015).
35. Lavrentovich, M. O. & Nelson, D. R. Survival probabilities at spherical frontiers. *Theor. Popul. Biol.* **102**, 26–39 (2015).
36. Vaupel, P., Kallinowski, F. & Okunieff, P. Blood flow, oxygen and nutrient supply, and metabolic microenvironment of human tumors: a review. *Cancer Res.* **49**, 6449–6465 (1989).
37. Gatenby, R. A. A change of strategy in the war on cancer. *Nature* **459**, 508–509 (2009).
38. Gatenby, R. A., Silva, A. S., Gillies, R. J. & Frieden, B. R. Adaptive therapy. *Cancer Res.* **69**, 4894–4903 (2009).
39. Eden, M. A two-dimensional growth process. *Dyn. Fractal Surf.* **4**, 223–239 (1961).
40. Fowler, R. & Schaaper, R. The role of the *mutT* gene of *Escherichia coli* in maintaining replication fidelity. *FEMS Microbiol. Rev.* **21**, 43–54 (1997).
41. Chen-Harris, H., Borucki, M. K., Torres, C., Slezak, T. R. & Allen, J. E. Ultra-deep mutant spectrum profiling: improving sequencing accuracy using overlapping read pairs. *BMC Genom.* **14**, 1 (2013).
42. Preston, J. L. *et al.* High-specificity detection of rare alleles with paired-end low error sequencing (pe-seq). *BMC Genom.* **17**, 1 (2016).
43. Deatherage, D. E. & Barrick, J. E. *Engineering and Analyzing Multicellular Systems: Methods and Protocols* 165–188 (Springer, 2014).
44. Berlin, K. *et al.* Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nat. Biotechnol.* **33**, 623–630 (2015).

### Acknowledgements

We thank Anne Dodson and Jasper Rine for helpful discussion and the construction of the switching *S. cerevisiae* strain. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM115851, by a National Science Foundation Career Award and by a Simons Investigator award from the Simons Foundation (O.H.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The library preparation and sequencing was

performed by the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303. The simulations were run on the Savio computational cluster resource provided by the Berkeley Research Computing programme.

### Author contributions

O.H., D.F. and M.G. designed the study. D.F. performed *E. coli* experiments and analysed the sequencing data. M.G. performed *S. cerevisiae* experiment and image analysis. M.G. and J.K. recorded the time-lapse movie and the confocal images. D.F. wrote and ran the simulations and performed the competitive release experiments. D.F., M.G., A.A. and O.H. performed the theoretical calculations. D.F., M.G., J.K. and O.H. interpreted the data and wrote the manuscript.

### Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Fusco, D. *et al.* Excess of mutational jackpot events in expanding populations revealed by spatial Luria–Delbrück experiments. *Nat. Commun.* **7**:12760 doi: 10.1038/ncomms12760 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016