IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

A modified fluctuation-test framework characterizes the population dynamics and mutation rate of colorectal cancer persister cells

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1869137 since 2022-07-12T10:24:28Z

Published version:

DOI:10.1038/s41588-022-01105-z

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

04 February 2023

A modified fluctuation-test framework characterizes the population dynamics and mutation rate of colorectal cancer persister cells

Mariangela Russo[1,2,*], Simone Pompei[3,*], Alberto Sogari[1,2,*], Mattia Corigliano[3,4,*], Giovanni Crisafulli[1,2], Alberto Puliafito[1,2], Simona Lamba[2], Jessica Erriquez[2], Andrea Bertotti[1,2], Marco Gherardi[3,4], Federica Di Nicolantonio[1,2], Alberto Bardelli[1,2,†], and Marco Cosentino Lagomarsino[3,4,†]

[1]Department of Oncology, University of Torino, Candiolo (TO), Italy.

[2]Candiolo Cancer Institute, FPO–IRCCS, Candiolo (TO), Italy.

[3]IFOM Foundation, FIRC Institute of Molecular Oncology, Milan, Italy.

[4]Dipartimento di Fisica, Università degli Studi di Milano, and I.N.F.N, via Celoria 16 Milano, Italy

[*] These authors contributed equally.

† These authors jointly supervised this work.

Correspondence:

Alberto Bardelli (alberto.bardelli@unito.it)

Marco Cosentino Lagomarsino (marco.cosentino-lagomarsino@ifom.eu)

Abstract

Compelling evidence shows that cancer persister cells represent a major limit to the long-term efficacy of targeted therapies. However, the phenotype and population dynamics of cancer persister cells remain unclear. We developed a quantitative framework to study persisters by combining experimental characterization and mathematical modeling. We found that, in colorectal cancer, a fraction of persisters slowly replicate. Clinically approved targeted therapies induce a switch to drug-tolerant persisters and a temporary 7- to 50-fold increase of their mutation rate, thus increasing the number of persisters-derived resistant cells. These findings reveal that treatment may influence persistence and mutability in cancer cells and pinpoint inhibition of error-prone DNA polymerases as a strategy to restrict tumor recurrence.

Main Text

When cancer patients are treated with targeted agents, tumor relapse is often observed after an initial response [1,2]. Emergence of resistance after prolonged response and disease stabilization is also frequent [3,4]. Indeed, when cancer cells are exposed to lethal doses of targeted therapies, the emergence of a sub-population of drug-tolerant "persister" cells prevents tumor eradication[5-10]. Unlike genetically resistant cells, persisters tolerate drug pressure through reversible, non-genetical, non-inheritable mechanisms of resistance[5,10,11]. However, it is unclear if persisters enter a quiescent state or slowly progress through the cell cycle. It is also unknown if the persister phenotype is drug-induced, or is pre-existing. Additionally, the population dynamics governing persisters evolution to resistance are only partially elucidated [8].

23 Exposure of colorectal cancer (CRC) cells to targeted therapies induces DNA damage and impairment
24 of DNA repair proficiency, a phenotype recently confirmed in other studies [12,13]. Drug treatment leads
25 to error-prone DNA replication in cancer cells, suggesting that mutability of persisters could increase
26 during therapy-induced stress [14].
27
28 Measuring mutational processes by DNA sequencing is challenging, owing to tumor heterogeneity
29 and the difficulties of tracking lineages [12]. A complementary strategy is represented by the
30 "fluctuation test" developed by Luria and Delbrück to characterize the onset of resistance in bacterial
31 populations [15]. This assay exploits multiple replicates of clonal populations to bypass lineage-tracking
32 issues and provides an elegant strategy to estimate mutation rates.
33
34 The fluctuation test has been previously modified to infer the acquisition of resistance to therapy in
35 tumors[16-19], particularly for the evolution of pre-existing resistant cells before treatment initiation, and
36 estimation of cancer cells' mutation rate in basal conditions[20]. However, it is not designed to quantify
37 mutation rates in cancer persisters during drug treatment.
38
39 Here, we present a general quantitative methodology to characterize the transition of cancer cells to
40 persistence and measure their population dynamics during drug treatment. We deployed a two-step
41 fluctuation test to quantify phenotypic mutation rates of CRC cells. Importantly, our assay
42 discriminates pre-existing resistant clones from persister-derived ones, allowing a quantification of
43 spontaneous (i.e., in untreated conditions) and drug-induced mutation rates.
44

45 Results

46

47 Growth dynamics of CRC cells.

48 We first aimed to quantify how CRC growth dynamics parameters were affected by drug treatment.
49 Our experiments included: (i) growth rates in standard conditions, (ii) population dynamics under
50 treatment and (iii) population dynamics of persisters (Supplementary note).
51
52 We used two microsatellite-stable (MSS) CRC cell models, RAS/RAF wild type DiFi and BRAF
53 V600E-mutated WiDr, which are respectively sensitive to the anti-EGFR antibody cetuximab alone
54 [13,21] or in combination with BRAF inhibitor[13], two clinically approved regimens for CRC[22-24]. To
55 reduce the possibility that pre-existing resistant cells were present in the populations at the beginning
56 of the assay, we isolated individual clones for each cell model, with growth and drug sensitivity
57 profiles comparable to those of the parental population (Supplementary figure 1a and b).
58
59 We measured birth and death rates of clones in standard cell-culture conditions (Supplementary figure
60 2, and Supplementary table 1¸ Supplementary note). We collected data from two sets of drug-response
61 growth assays (Fig. 1a). The first, the *doses-response assay,* in which CRC clones were exposed to
62 increasing concentrations of targeted therapies (Fig. 1b, Extended data fig. 1a), was used to analyze
63 growth curves defined as number of live cells *vs* time and drug concentration (Fig. 1c, Extended data
64 fig. 1b), and to quantify the intertwined processes of growth and transition to persister state upon drug
65 treatment. Supplementary figure 3 shows the normalization process of the *doses-response assay* data
66 used to obtain growth curves (Supplementary note). The second, the *single-dose assay,* consisting of
67 3 weeks of exposure to a constant drug concentration, highlighted a biphasic (two time-scale) killing
68 curve (Fig. 1d, Extended data fig. 1c), characterized by a rapid decline of sensitive cells followed by
69 a slower decline [25,26], a hallmark of the emergence of persisters in bacteria [25]. The fraction of surviving
70 persisters displayed a slow but measurable decay in cell number (Extended data fig. 2), suggesting a
71 tendency of persisters to slowly die over time.
72

CRC persister cells slowly replicate during drug treatment.

We next aimed to elucidate dynamics of persisters proliferation and death under drug treatment. Staining of CRC persisters with Carboxy Fluorescein Succinimidyl Ester (CFSE), a cell-permeable fluorescent dye allowing flowcytometric monitoring of cell divisions[27], and EdU (5-ethynyl-2'-deoxyuridine), a modified thymidine analogue that is efficiently incorporated into DNA during active DNA replication, revealed that a fraction between 0.2 and 2.5 % of persisters slowly replicates during treatment (Extended data fig. 3a-c; Supplementary note), in line with recent data [28]. We next used a live cell microscopy imaging assay (Supplementary note). While the majority of CRC persisters were non-replicating, cell division events were visible in all the CRC clones analyzed (Extended data fig. 3d and Supplementary movies 1-4). The cell division was occasionally successful and viable (Extended data fig. 3d). We also detected cell death events after cell division and in non-dividing cells (Extended data fig. 3e and f).

The persister phenotype is induced by targeted therapies in CRC.

To quantify cell population dynamics during drug treatment, we developed a mathematical model of the transition of CRC cells to the persister state, which we term "*transition to persisters*" (*TP*) model. This model incorporates birth-death parameters and phenotypic switching in the deterministic limit (i.e., neglecting fluctuations due to stochastic demographic effects, see Supplementary note) [29,30]. Fig. 2a summarizes the TP model dynamics. We exploited the model to quantify the transition rate and assess whether a sub-population of persisters predated drug administration or if the persister phenotype emerged upon drug treatment. The TP model considers three possible fates for drug-treated cells: i) death; ii) replication; iii) switching to persister state at a rate λ in the presence of the drug; it further considers the pre-existence of an arbitrary steady fraction f_0 of persisters (Fig. 2a).

The following equations define the dynamics of sensitive ($X(t)$) and persister cells ($Z(t)$) according to the TP model (under drug treatment):

$$\begin{cases} \frac{d}{dt} X(t) = & (B - D([M]))\, X(t) - \lambda([M])\, X(t) \\ \frac{d}{dt} Z(t) = & -D_p\, Z(t) + \lambda([M])\, X(t) \end{cases}, \quad (1)$$

99 where B and $D([M])$ are, respectively, the birth and drug-dependent death rates of sensitive cells,
100 while [M] is the drug concentration. Persisters emerge with a drug-dependent transition rate $\lambda([M])$.
101 Under drug treatment, persisters die with rate $D_p > 0$ (Extended data figs.2 and 3). The model assumes
102 that persisters that attempt to divide before acquiring drug-resistance mutations die; therefore a
103 possible back-switching from persister to sensitive in the presence of drug would effectively
104 contribute to the death rate (Supplementary note).
105
106 The initial condition that specifies the solution to Equation (1) is key for quantifying to what extent
107 the transition to persister state is induced by the drug treatment. Specifically, if the sensitive-to-
108 persister transition is fully drug-induced, then untreated populations would not contain any persisters,
109 i.e., $f_0 = \frac{Z(t_0)}{N_{(t_0)}} = 0$.
110
111 Conversely, if some persisters pre-exist, then the initial fraction of persister cells has a finite positive
112 value ($f_0 > 0$). If f_0 is very small, some persisters may pre-exist, but the transition is mainly drug-
113 induced. If, f_0 is actually comparable to the fraction of residual persisters after weeks of treatment,
114 then the transition to persistence is not drug-induced.
115
116 To determine the most likely scenario, we used experimental data collected from drug-response
117 assays (Fig. 1). Using results from the *doses-response assay*, we defined parameters governing the
118 dynamics of the model over a short timeframe, such as the initial fraction of persisters f_0 and the
119 effective growth rate of treated cells. Similarly, the *single-dose assay* was used to quantify model
120 parameters that affect long-term dynamics, such as the transition rate of sensitive to persister cells (λ)
121 and the effective death rate of persisters (D_p). By constraining model parameters from experimental

122 data, we established which scenario would best describe the cell-based results. The inferred
123 parameters are compatible with the values obtained by live cell microscopy assay, supporting a
124 balance between proliferation and cell death skewed slightly towards the latter (Extended data fig.
125 3g, Supplementary note).

126

127 Upon treatment, the number of cells started to decline within 1 to 3 days (t_0), depending on the initial
128 seeding density (Supplementary figure 4). The observed cell dynamics were coherent in experiments
129 with different seeding densities once the growth curves were scaled (both in time and measured
130 viability) to the maximum value reached at $t = t_0$ (Supplementary figure 4).

131

132 The parameters of the TP model were inferred with a standard Bayesian inference framework for both
133 cell lines (Supplementary table 2 and Supplementary note). DiFi displayed slower 'dying' dynamics
134 compared to WiDr. In light of this, in WiDr we performed a joint fit of both the *doses-response* and
135 *single-dose* datasets, while in DiFi we assessed growth curves in response to multiple doses of
136 targeted therapies for up to 19 days, which allowed performing a model fit based on the *doses-*
137 *response* dataset only (Supplementary note).

138

139 We identified the best-fit TP model parameters given the experimental data, considering different
140 values of the initial number of persisters (f_0). The best fit between the inferred TP model and
141 experimental data occurs when f_0=0, while the concordance decreases when f_0 increases; we note that
142 a value of f_0 of 10% already leads to significantly deviations from the data (Fig. 2b). Therefore, the
143 TP model is consistent with the persister phenotype being predominantly drug-induced. In addition,
144 the model properly describes the dynamics of *the single-dose assay* (Fig. 2c).

145

146 To further confirm the validity of the TP model, we next focused on the Bayesian statistics of the two
147 model parameters describing the dynamics of persisters: the transition rate λ and initial fraction of

148 persisters f_0. The joint posterior distribution of the Bayesian inference of these two parameters is
149 shown in Fig. 2d and Supplementary figure 5. We found that the transition rate to persistence λ
150 estimated by the model fit does not vary when considering different values of the initial fraction of
151 persisters f_0 (Fig. 2d). The marginalized posterior probability of f_0 is peaked at zero (Fig. 2d, bottom
152 panel), and its upper bound is much smaller than the ratio between the persister population size (after
153 all persister cells have emerged) and the total population size at the beginning of treatment. This
154 implies that the inferred value of transition rate λ is independent from f_0, and that the best concordance
155 of the TP model to the experimental data is obtained for f_0=0.
156

157 Finally, to compare the scenarios f_0=0 and f_0>0, we used the Bayesian Information Criterion (BIC)
158 and the Akaike Information Criterion (AIC), which are standard Bayesian criteria used for model
159 selection. According to both, the TP model with f_0=0 is preferred over $f_0 > 0$ (Extended data fig. 4
160 and Supplementary table 3, Supplementary note). We summarize the best-fit TP model parameters in
161 Supplementary table 4. Interestingly, we found the transition rate of WiDr and DiFi cells to
162 persistence to be drug-dependent (Extended data fig. 4 and Supplementary figure 6). These results
163 indicate that even if few persisters exist in the population before drug treatment, the majority of them
164 must have transitioned to the persister phenotype after drug exposure. Our finding that WiDr cells
165 show a transition rate to persistence that increases with drug concentration could be applied to design
166 innovative strategies to restrict persisters evolution. Notably, our analysis predicts that a linear
167 increase of drug concentration, compared to a constant dosage, might reduce the number of persisters
168 (Extended data figure 5).
169

170 **Persisters distribution supports a drug-induced scenario.**

171 We then measured how the number of persisters varied across multiple wells, since the distribution
172 of this parameter is expected to be different between a drug-induced and a pre-existing scenario[31].
173 We seeded DiFi cl.B6 and WiDr cl.B7 in multiple 96-well plates and quantified the distribution of

174 persisters (residual cell viability) among >400 independent wells after 3 weeks of drug treatment
175 (Extended data fig. 6a). The observed abundance distribution across wells was consistent with a
176 Poisson distribution (Extended data fig. 6b), supporting a drug-induced scenario, as confirmed by
177 computational simulations (see [31], where a similar method was used for mutational processes, and
178 Extended data fig. 6c). These numerical simulations show that, provided that persisters do not pre-
179 exist drug treatment, the number of persisters emerging from sensitive cells under treatment is
180 Poisson distributed. Conversely, pre-existing persisters would be generated with a constant rate from
181 an exponentially expanding population before treatment administration. Hence the number of pre-
182 existing cells is not Poisson distributed, but is described by a Luria-Delbruck[15] distribution (with
183 Variance ≫ Mean). We found that the final distribution of persisters across wells is Poisson, in line
184 with emergence after drug treatment.
185
186
187 **A fluctuation assay quantifies persisters' mutation rates.**
188 Measurement of mutation rates in the absence or presence of anticancer drugs required the
189 development of a second model, hereafter the "Mammalian Cells–Luria-Delbrück" or "MC-LD"
190 model. The MC-LD model is a fully stochastic birth-death branching process, describing the growth
191 of resistant cells before and during drug treatment (Fig. 3a). We designate with μ the effective rate at
192 which one individual (cell) develops resistance while μ_s and μ_p indicate mutation rates of sensitive
193 (untreated) and persister cells, respectively (Fig. 3b, Supplementary note).
194
195 WiDr and DiFi cells were seeded in twenty 96-multiwell plates each and allowed to grow for a fixed
196 number of cell divisions in drug-free standard culture conditions (Fig. 3c); afterwards, a constant
197 clinically relevant drug concentration was applied (Fig. 3d). The number of wells, the initial
198 population size in each well and the time of cell replication in the absence of drug treatment were set

199 by theoretical considerations incorporating the population dynamics parameters we previously
200 measured (Supplementary tables 1 and 4; Supplementary note).
201
202 In accordance with our previous work [6], a small number of early-emerging resistant colonies was
203 detected after 3-4 weeks of treatment (Fig. 3d, e). Conversely, in the vast majority of the wells
204 sensitive cells died, while drug-tolerant persisters survived, as detected by measurement of residual
205 cell viability (Extended data fig. 6) [6]. After several weeks of constant treatment of the residual
206 persister cells (Fig. 3d), late-emerging resistant colonies appeared in a subset of wells in which
207 persisters had previously been detected (Fig. 3d, e).
208
209 We ran multiple MC-LD model simulations, with input parameters inferred with the TP model, and
210 found that resistant clones emerging at late time points (>4 weeks of treatment) are unlikely to
211 originate from pre-existing resistant cells (Fig. 4a and Extended data fig. 7). In accordance with
212 previous work [6] we considered the resistant colonies that became microscopically visible within the
213 first 4 weeks of drug treatment (early-emerging resistant) as those representing pre-existing resistant
214 cells, i.e., mutant cells that emerged during the expansion phase by spontaneous mutation. We also
215 reasoned that resistant colonies that slowly emerged after ≥10 weeks of drug treatment (late-emerging
216 resistance) in persisters-containing wells could have developed drug resistance mutations through the
217 adaptive mutability process which we and others have observed [12,13] (Fig. 4a and Extended data fig.
218 7).
219
220 As in a standard fluctuation test, the mutation rate can be inferred from the observed fraction of wells
221 containing resistant cells. In the model, this fraction corresponds to the expected probability of
222 observing a resistant clone in a well in a given time interval $[0,T]$. To compute this probability in the
223 MC-LD model, we assumed that resistant cells divide with rate b and die with rate d, just like
224 untreated cells. Supplementary figure 7 supports the stability of the inferred values of the mutation

225 rates against variation of the division rate of resistant cells. Because of reproductive fluctuations
226 (genetic drift), cells carrying drug-resistance mutations can still go extinct, and only a fraction of the
227 mutants, which we refer to as "established mutants", survive stochastic drift. The probability of
228 surviving stochastic drift in a time interval Δt, denoted here as $\psi(\Delta t)$, is a well-known result of the
229 birth-death process [32,33] (Supplementary note).
230
231 We derived analytically an approximate solution of the model, by considering that the number of
232 mutant cells established in the time interval $[0,T]$ follows a Poisson distribution with expected value
233 $\mathcal{M}(T)$. Consequently, the probability of having at least one mutant is given by

$$P(T) = 1 - e^{-\mathcal{M}(T)} \,. \tag{2}$$

234
235 In order to quantify the spontaneous mutation rate of cancer cells before drug administration, we
236 focused on the resistant cells established by the time T_{treat} before treatment administration. The
237 expected number of resistant cells that emerged from sensitive cells in this time interval reads:

$$\mathcal{M}_{sensitive}(T_{treat}) = \mu_s \int_0^{T_{treat}} X(t)\psi(T_{treat} - t)dt \,, \tag{3}$$

238
239 where μ_s is the mutation rate of sensitive cells.
240
241 To quantify the mutation rate of persister cells μ_p, we consider resistant cells that emerged by the time
242 T since the beginning of the drug treatment. The expected number of resistant cells emerged from
243 persister cells reads:

$$\mathcal{M}_{persisters}(T) = \mu_p \int_0^T Z(t)\psi(T - t)dt \,. \tag{4}$$

244

245 We emphasize that Equation (2-4) are connected to the solution of the TP model, Equation (1). Hence,
246 the solution of the MC-LD model is defined in terms of the same parameters that were estimated with
247 the TP model (Supplementary note).
248
249 We used this solution of the MC-LD model to derive estimators of mutation rates of sensitive cells μ_s
250 (encompassing the fraction of wells with early-emerging resistant cells) and of persister cells μ_p
251 (corresponding to the fraction of wells with late-emerging resistant clones).
252
253 **Persisters show increased mutation rate under treatment.**
254 Data collected with two-step MC-LD fluctuation tests for each clone allowed inferring mutation rates
255 of sensitive (μ_s) and persisters (μ_p). We conservatively evaluated the mutational processes as
256 chronological (measured in mutations per day) rather than replicative (mutations per generation). This
257 choice is safe, as the ratio between replicative mutation rates of cells displaying the two phenotypes
258 must always be higher than for chronological rates, since (beyond any uncertainty) measured cell
259 division in persister cells was very low compared to that of untreated cells.
260
261 We found that mutation rates were increased by a factor of 7- to 50-fold in cells that survived and
262 tolerated for several weeks doses of targeted therapies that were lethal for the majority of the parental
263 population (Fig. 4b and Supplementary table 5). This result was consistent across multiple biological
264 replicates, both in DiFi and WiDr cells and in response to clinically relevant concentrations of either
265 EGFR blockade or EGFR/BRAF concomitant inhibition, respectively (Fig. 4b and Supplementary
266 table 5).
267
268 To further validate the consistency of the mutation rate inference based on the MC-LD model, we ran
269 multiple simulated replicates of the experiment, using a set of sensitive (μ_s) and persister (μ_p) mutation
270 rates, and we then used the MC-LD estimators on the synthetic data. Fig. 4c compares boxplots of

271 the estimated mutation rates across replicates of simulated experiments with the actual values of
272 mutation rates used as inputs to the simulations. The agreement between these values validates our
273 estimates.
274
275 We next assessed whether and to what extent the inferred value of the mutation rate is affected by the
276 presence of different numbers of pre-existing persister cells f_0 using our estimators within a Bayesian
277 framework (Fig. 4d). This approach returns the mutation rate, considering a range of realistic values
278 of f_0, and their probability. We obtained the fold increase of the mutation rate of persister cells as a
279 function of f_0, in the entire range of values that are compatible with the dynamics observed in the
280 growth curve assays experimentally assessed in Figs. 1 and 2. Fig. 4d summarizes the results of this
281 inference. We found that, considering all representative values of f_0 that are compatible with our
282 experimental data, the increase of mutation rate in persister cells remains strongly supported.
283
284 To corroborate these results we replicated the full set of experiments and ran the analysis pipeline for
285 two additional clones, one for each cell model (WiDr cl. B5 and DiFi cl. B3), thereby confirming our
286 findings and excluding a clonal bias effect (Extended data figs. 8 and 9). Molecular profiling of
287 persisters-derived resistant clones isolated from the fluctuation assays revealed acquisition of single-
288 nucleotide variations (SNV) or copy-number alterations (CNA) in genes involved in the RAS-MEK
289 pathway, which are known drivers of resistance to anti-EGFR/anti-BRAF inhibitors in CRC[21,34,35]
290 (Supplementary figure 8 and Supplementary table 6).
291
292 We propose a quantitative model for the evolutionary dynamics of CRC cells exposed to targeted
293 therapies (Fig. 5a). Untreated cancer cells replicate and spontaneously acquire mutations that can
294 confer resistance to targeted therapies (pre-existing resistant mutations) at a replicative mutation rate
295 μ_s. However, when cells are exposed to targeted therapies, the majority quickly die while a subset of
296 parental cells switch to a long-lasting surviving persister state at a rate λ and in a drug-induced

297 fashion. Previous and current findings indicate that persister cells, under constant exposure to lethal
298 doses of drugs, initiate a stress response that affects DNA replication fidelity [12,13], thus leading to a
299 measurable increase of their mutation rate (μ_p), therefore raising the probability that alterations
300 conferring drug resistance could occur.

301

302 **Inhibition of mutagenic REV1 extends the efficacy of targeted therapy.**

303 We previously reported that, in response to drug treatment, cancer cells switch from high- to low-
304 fidelity DNA replication through downregulation of DNA repair genes and upregulation of
305 specialized error-prone DNA polymerases [13]. This in turn could foster the temporary increase of
306 mutation rate observed in persister cells as quantitively measured in the current study. Among the
307 DNA polymerases that are upregulated in cancer cells upon targeted therapy[13], REV1 carries out
308 translesion synthesis (TLS), a mutagenic process that allows cells to tolerate DNA damage by
309 bypassing lesions that block normal DNA replication, resulting in the introduction of mutations[36,37].
310 Interfering with TLS with a REV1 inhibitor has been shown to enhance chemotherapy efficacy and
311 suppress tumor growth both *in vitro* and *in vivo*[38,39]. Based on the above, we hypothesized that
312 inhibition of mutagenic TLS would likely increase the cytotoxic effects of targeted therapies-induced
313 DNA damage, therefore delaying the acquisition of resistance during adaptive mutability.

314

315 To assess this possibility, we performed a Time-To-Progression (TTP) assay, an approach we
316 previously established [40] to monitor the development of secondary resistance in cancer cells. MSS
317 DiFi and WiDr CRC cells, as well as the MSS BRAF V600E-mutated cell line (JVE207), were treated
318 with either a MAPK pathway inhibitor, the REV1 inhibitor or their combination. Pharmacological
319 blockade of REV1 significantly delayed or prevented the development of secondary resistance to
320 EGFR/BRAF inhibitors (Figure 5b, Extended data fig. 10).

321

322

Discussion

We present an experimental framework that integrates biological assays and mathematical modelling to investigate population dynamics of cell lines exposed to environmental perturbations. The MC-LD assay allows quantitative comparisons of spontaneous and drug-induced mutation rates and could in principle be applied to measure whether and how a wide range of environmental conditions affect persister phenotype and mammalian cells' mutation rates in a considerable number of biological systems.

An important caveat of our technique is that fluctuation tests measure "phenotypic" mutation rates, i.e., rates of conversion to a phenotype (here, resistance to targeted therapies) that could result from different mutational routes, including SNV and CNA, both of which we found to drive resistance in our persister-derived resistant clones. Nevertheless, our approach has the advantage to bypass several hurdles associated to sequencing-based measurement of mutation rates.

Although we and others have recently shown that adaptive mutability fosters the acquisition of secondary resistance by increasing genomic instability in surviving persister cells [12,13], the lack of models to quantitatively characterize the behavior of persisters under treatment has so far prevented reliable quantification of persisters' mutation rate. While limited to cell lines, our controlled two-step fluctuation assay overcomes these issues. The results could be used to infer features of more complex systems (such as patient samples), where mathematical models are postulated and cannot be analogously validated.

Our results indicate that drug-induced sensitive-to-persister transition is a predominant path to the development of this phenotype. This is in line with recent evidence of a chemotherapy-induced

persister state in CRC[41]. Although our results are coherently explained by the existence of a phenotypic switch of sensitive to persister cells, no direct observation of the switching is yet available. Alternative models whereby slower-proliferating tolerant cells generate faster-proliferating non-tolerant phenotypes would also give rise to a biphasic killing curve[42,43]. In our framework, this would correspond to the case where $f_0 \gg 0$, which is ruled out by our analysis. Hence, the interpretation linking the phenotypic switch to persistence with treatment appears to be the most likely scenario.

Importantly, even a small subset of pre-existing persisters does not affect our findings that mutation rates of cancer cells significantly increase under treatment. Persister-derived resistant clones keep emerging after several weeks of continuous drug treatment. In the absence of an increased mutation rate, it would typically take (in a conservative estimate) >100-1,000 weeks for the cells in a single well to develop resistance based on the mutation rate of sensitive (untreated) cells. Equivalently a fluctuation assay performed on $2\text{x}10^5$-$2\text{x}10^6$ wells would be required to observe few resistant clones after 10 weeks. Additionally, the contribution of sensitive cells to resistance is exhausted after a few weeks of treatment, since we show that they become extinct within a few days (Fig. 1d). The evidence of active cell cycle progression alongside an increase of mutagenic rate in persister cells further supports previous findings of ongoing adaptive mutagenesis fostering acquisition of resistance [13].

A recent study concluded that the impact of adaptive mutability on the mutational load in the protein-coding genome is small [44]. This in line with previous findings showing that tumor mutational burden is not significantly increased in cells that acquired therapy resistance [12,13]. There are many confounding factors in these not precisely controlled systems. Indeed, the adaptive mutability phenotype is likely restricted in time (i.e., when the cells are maladapted to the new environment[13]), confined to a small sub-population of cells and masked by the outgrowth of pre-existing resistant cells. Bulk analysis on tumor samples at relapse cannot disentangle these factors. Our characterization of persisters' dynamics and increased mutation rate have potential clinical relevance. First, the finding

374 that higher drug concentrations induce an increased death rate of sensitive cells, and an increased
375 transition to persistence, a reservoir for the emergence of resistance, provides a rationale for
376 therapeutic strategies to impair the emergence of persistence. Second, the involvement of error-prone
377 DNA polymerases during adaptive mutability[13] offer opportunities for non-obvious combinatorial
378 strategies to restrict drug resistance. Indeed, we show that inhibition of mutagenic TLS significantly
379 delays the acquisition of secondary resistance.
380
381 Our methodology infers that clinically approved anticancer therapies can induce a temporary increase
382 in the mutation rate of CRC cells. Our framework can be used to systematically measure mutation
383 rates in mammalian cells exposed to a wide range of environmental stressors and to define drug
384 combinations to restrict the emergence of therapeutic resistance.
385
386
387 **Acknowledgments**
388 We are grateful to A. Ciliberto, M. Osella, L. Trusolino and A. Amir for useful discussions and critical
389 reading of the manuscript. This work was supported by FONDAZIONE AIRC under 5 per Mille 2018
390 - ID. 21091 program – P.I. Alberto Bardelli , Group leaders Di Nicolantonio Federica and Bertotti
391 Andrea; AIRC, Associazione Italiana per la Ricerca sul Cancro AIRC-IG (REF: 23258) – P.I. Marco
392 Cosentino Lagomarsino; AIRC IG 2017 N° 20697 – P.I. Andrea Bertotti; European Research Council
393 Consolidator Grant 724748, BEAT – PI Andrea Bertotti; AIRC under IG 2018 - ID. 21923 project –
394 P.I. Alberto Bardelli; AIRC under MFAG 2021-ID 26439 project-P.I. Mariangela Russo; MFAG
395 2020-ID 25040 project – P.I Alberto Puliafito; International Accelerator Award, ACRCelerate, jointly
396 funded by Cancer Research UK (A26825 and A28223), FC AECC (GEACC18004TAB) and AIRC
397 (22795) to Alberto Bardelli; Ministero Salute, RC 2020 to A.Bardelli; European Research Council
398 (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement
399 n. 101020342) – P.I. Alberto Bardelli; IMI contract n. 101007937 PERSIST-SEQ to Alberto Bardelli

400 and Mariangela Russo; BiLiGeCT - Progetto PON ARS01_00492 to Alberto Bardelli. Simone

401 Pompei was supported by Fondazione Umberto Veronesi.

402

403 **Author contributions**

404 M.R., A.Bardelli and M.C.L. conceived the study and contributed with key ideas at different stages.

405 M.R. and A.S. performed the biological experiments. S.P., M.C., M.G. and M.C.L. conceived the

406 modeling framework. S.P., M.C. and M.G. performed data analysis, model simulations, and analytical

407 calculations. J.E. performed time-lapse live microscopy experiments. A.P. performed time-lapse live

408 microscopy data analysis. G.C. performed WGS data analysis. S.L. performed ddPCR analysis.

409 A.Bertotti and F.D.N. contributed to data discussion. M.R., S.P., A.Bardelli and M.C.L. wrote the

410 paper. All authors read and approved the final version of the paper. M.R., S.P., A.S. and M.C. equally

411 contributed to the study. A.Bardelli and M.C.L. jointly supervised the study.

412

Competing interests

Alberto Bardelli reports receiving commercial research grants from Neophore, AstraZeneca, Boehringer; is an advisory board member/unpaid consultant for Inivata, Neophore; holds ownership interest in Neophore; and is an advisory board member/consultant for Illumina, Guardant Health, Inivata, Roche/Genentech Global CRC. All the other authors declare no competing interests.

Figure legends

Figure 1. Population dynamics of CRC clones in response to targeted therapies. **a**, Schematic representation of the drug screening growth curve assays performed on CRC clones. **b**, In the *doses-response assay*, WiDr cells were treated with increasing concentrations of dabrafenib (Dab) + 50µg/ml cetuximab (CTX), while DiFi received increasing concentrations of cetuximab. Cell viability was measured by the ATP assay at the indicated time points. Results represent the average ± SD (n=3

biologically independent experiments for WiDr; n=5 biologically independent experiments for DiFi). **c**, Growth curves of the indicated cells under treatment, reported as fold-change of viable cells (log scale) vs time of drug exposure, were calculated from *doses-response assay* data, by normalizing cell viability at the indicated time points by the viability measured at day 0. Growth curves for three different drug concentrations for each clone are shown as average ± SD (n=3 for WiDr; n=5 for DiFi). **d**, Fold-change of viable cells (log scale, assessed by ATP assay) vs time of drug exposure for indicated cells in the *single-dose assay*. The total number of viable cells is compatible with an exponential decay with two-time scales, supporting the outgrowth of persisters (the dashed line indicates the initial slope). Symbols indicate means (n=2 biologically independent experiments).

Figure 2. Emergence of CRC persister cells is predominantly drug-dependent. **a**, Schematic representation of cell population dynamics under constant drug treatment. When cancer cells are exposed to targeted therapies, the number of viable cells starts to decline. A homogeneous population of sensitive cells (grey cells) would shrink exponentially to extinction (grey dashed line). Some cells survive drug treatment due to the transition to a persister phenotype (dark yellow cells, dark yellow lines) at a rate λ, and residual cells (solid black line) show a bi-phasic decay. A finite fraction of persister cells (f_0) might be present in the population prior to drug treatment (f_0>0, dashed dark yellow line) or not (f_0=0, solid dark yellow line). Persister cells show a reduced death rate during treatment, which results in a slow exponential decline of the cell population (dark yellow dashed line). **b**, Growth curves of CRC clones under treatment, calculated from the *doses-response assay*. Black symbols and bars represent averages ± SD of the *doses-response* dataset (n=3 biologically independent experiments for WiDr, n=5 biologically independent experiments for DiFi). Continuous lines indicate the TP model fit to the experimental data for different values of the initial fraction of persister cells (f_0, color coded). **c**, Fold-change of viable cells *vs* time of drug exposure for the indicated cells assessed based on the *single-dose* dataset. Black symbols represent averages of the experimental data

(n=2 biologically independent experiments). The black dashed line indicates the fit of the TP model to the data, while the expected fraction of persisters cell is shown with the dark yellow solid line. **d**, Joint posterior distribution (contour plot, color coded with the normalized likelihood function) and marginalized posterior distributions (left and bottom panel, grey area indicates the Probability Density Function) of TP parameters describing the dynamics of persister cells: (i) the initial fraction of persister cells (f_0, bottom panel) and (ii) transition rate of sensitive to persister cells (λ) induced by the drug treatment. The likelihood function measures the agreement of the model to the experimental data as a function of the value of the parameters considered.

Figure 3. A modified Luria-Delbrück fluctuation test to measure mutation rates in mammalian cells. a, The modified fluctuation test, based on the inferred population dynamics and the MC-LD model, allows estimation of spontaneous (μ_s) and persisters (μ_p) mutation rates. **b**, Schematic representation of cell population dynamics of CRC cells during the fluctuation test. During the initial expansion in the absence of drug treatment CRC cells mutate with the spontaneous mutation rate (μ_s). When cells are exposed to targeted therapies, pre-existing resistant cells are selected by the drug and give rise to early emerging resistant colonies (red dashed line), while sensitive cells start to decline in number (black solid line) and switch to the persister state (dark yellow solid line). Resistant cells derive from persister with a mutation rate μ_p and give rise to late-emerging resistant colonies (blue dashed line). **c**, Schematic representation of the experimental design underlying the fluctuation assay. WiDr and DiFi cells were seeded in twenty 96-multiwell plates, for a total of 1,920 wells, and allowed to expand in the absence of drug for about 8 generations (reaching ~20,000 cells/well). After the expansion, all the wells were treated with targeted therapy (100 µg/ml cetuximab for DiFi and 1µM dabrafenib + 50 µg/ml cetuximab for WiDr). **d**, Two sets of resistant clones were identified during the MC-LD experimental assay: the early-emerging resistant clones grown after 3-4 weeks (Stage 1), and the late-emerging resistant clones arising after >10 weeks (Stage 2) of constant drug treatment. Scale bar 100µm. **e**, Each bar graph lists the number of resistant clones counted at the indicated

timepoints during MC-LD experiment for each CRC clone. Red bars indicate early-emerging resistant clones (appearing in the first 4 weeks of drug treatment); blue bars indicate late-emerging resistant clones (appearing after ≥10 weeks of drug treatment). Results of two independent biological replicates for each clone are shown.

Figure 4. Quantification of mutation rates in persister cells. a, Simulated data for the assay described in Fig. 3. The experimentally measured MC-LD model parameters and the model-derived estimators of mutation rate, for sensitive and persister cells (dark yellow solid lines), were used to simulate the time of appearance of pre-existing (red) and persisters-derived (blue) resistant cells. **b**, Quantification of mutation rates for sensitive (red) and persister (blue) cells in the MC-LD experiment. The indicated cell models were seeded and treated as described in Fig. 3. Mutation rates were calculated from the experimental data based on population parameters and the number of pre-existing (early-emerging) and persisters-derived (late-emerging) resistant clones as described in Fig. 3. Results represent inferred mutation rates (for each clone of sensitive and persister cells) with bar plots showing mean of the posterior distributions of the mutation rates (n=2 biologically independent experiments). Here, the bar chart is used as graphical representation of inferred mutation rates (Supplementary table 5 for the corresponding numerical values). **c**, Validation of mutation rates estimator with model simulations. The box plots represent the distribution of the estimated mutation rates for n=100 independent simulations of the entire experiment using the parameters reported in Supplementary tables 1 and 4. Red and blue boxes indicate the interquartile ranges (25% and 75% percentiles) of the estimated mutation rates of sensitive and persisters, respectively, while the upper and lower whiskers represent the maximum and minimum value of the distribution. The median of the distribution, reported as a black line in the center, is shown together with its 95% confidence interval (nuanced area), indicated by the notches on both sides of the box. The mean of the distribution and the input value of the mutation rate used in the simulation are reported as a dashed white line and

a black dashed line, respectively. **d**, Joint posterior distribution (contour plot, color coded with the normalized likelihood function) and marginalized posterior distributions (left and bottom panel, grey area shows the Probability Density Function) of (i) the initial fraction of persister cells (f_0, bottom panel) and (ii) fold increase of the mutation rate of persister cells compared to mutation rate of sensitive cells (μ_p/μ_s). The likelihood function measures the agreement of the model to the experimental data as a function of the value of the parameters considered.

413

414 **Figure 5. Inhibition of error-prone DNA polymerases delays the onset of acquired resistance to**
415 **targeted therapies. a**, Schematic representation of CRC cells mutational dynamics during drug
416 treatment. Untreated cells spontaneously acquire resistant mutations at a replicative spontaneous
417 mutation rate μ_s. When cancer cells are exposed to targeted agents, a surviving persister phenotype is
418 induced in a drug-dependent manner. Persister cells under constant drug exposure reduce DNA
419 replication fidelity and increase their mutation rate at rate μ_p. This in turn boosts genetic diversity and
420 favors the emergence of resistant clones driving tumor recurrence and treatment failure. **b**, The
421 indicated CRC cells were treated with the anti-EGFR inhibitor cetuximab (CTX) alone, or in
422 combination with anti-BRAF inhibitor dabrafenib (DAB), the REV1 inhibitor was added where
423 indicated. The number of cells was monitored during the treatment, until the emergence of resistance
424 (n=1 biological experiment for each cell line).

425

426

427

References

428

429 1. Wagle, N. *et al.* Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor
430 genomic profiling. *J Clin Oncol* **29**, 3085-96 (2011).
431 2. Russo, M. *et al.* Tumor Heterogeneity and Lesion-Specific Response to Targeted Therapy in
432 Colorectal Cancer. *Cancer Discov* **6**, 147-53 (2016).

433 3. Parikh, A.R. *et al.* Liquid versus tissue biopsy for detecting acquired resistance and tumor
434 heterogeneity in gastrointestinal cancers. *Nat Med* **25**, 1415-1421 (2019).
435 4. Siravegna, G. *et al.* Clonal evolution and resistance to EGFR blockade in the blood of
436 colorectal cancer patients. *Nat Med* (2015).
437 5. Sharma, S.V. *et al.* A chromatin-mediated reversible drug-tolerant state in cancer cell
438 subpopulations. *Cell* **141**, 69-80 (2010).
439 6. Hata, A.N. *et al.* Tumor cells can follow distinct evolutionary paths to become resistant to
440 epidermal growth factor receptor inhibition. *Nat Med* **22**, 262-9 (2016).
441 7. Salgia, R. & Kulkarni, P. The Genetic/Non-genetic Duality of Drug 'Resistance' in Cancer.
442 *Trends Cancer* **4**, 110-118 (2018).
443 8. Vallette, F.M. *et al.* Dormant, quiescent, tolerant and persister cells: Four synonyms for the
444 same target in cancer. *Biochem Pharmacol* **162**, 169-176 (2019).
445 9. Lupo, B. *et al.* Colorectal cancer residual disease at maximal response to EGFR blockade
446 displays a druggable Paneth cell-like phenotype. *Sci Transl Med* **12**(2020).
447 10. Shen, S., Vagner, S. & Robert, C. Persistent Cancer Cells: The Deadly Survivors. *Cell* **183**,
448 860-874 (2020).
449 11. De Conti, G., Dias, M.H. & Bernards, R. Fighting Drug Resistance through the Targeting of
450 Drug-Tolerant Persister Cells. *Cancers* **13**(2021).
451 12. Cipponi, A. *et al.* MTOR signaling orchestrates stress-induced mutagenesis, facilitating
452 adaptive evolution in cancer. *Science* **368**, 1127-1131 (2020).
453 13. Russo, M. *et al.* Adaptive mutability of colorectal cancers in response to targeted therapies.
454 *Science* **366**, 1473-1480 (2019).
455 14. Rosenberg, S.M., Shee, C., Frisch, R.L. & Hastings, P.J. Stress-induced mutation via DNA
456 breaks in Escherichia coli: a molecular mechanism with implications for evolution and
457 medicine. *Bioessays* **34**, 885-92 (2012).
458 15. Luria, S.E. & Delbrück, M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance.
459 *Genetics* **28**, 491-511 (1943).
460 16. Kendal, W.S. & Frost, P. Pitfalls and practice of Luria-Delbrück fluctuation analysis: a
461 review. *Cancer Res* **48**, 1060-5 (1988).
462 17. Attolini, C.S. & Michor, F. Evolutionary theory of cancer. *Ann N Y Acad Sci* **1168**, 23-51
463 (2009).
464 18. Komarova, N.L. & Wodarz, D. Drug resistance in cancer: principles of emergence and
465 prevention. *Proc Natl Acad Sci U S A* **102**, 9714-9 (2005).
466 19. Iwasa, Y., Nowak, M.A. & Michor, F. Evolution of resistance during clonal expansion.
467 *Genetics* **172**, 2557-66 (2006).
468 20. Bozic, I. *et al.* Evolutionary dynamics of cancer in response to targeted combination therapy.
469 *Elife* **2**, e00747 (2013).
470 21. Misale, S. *et al.* Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy
471 in colorectal cancer. *Nature* **486**, 532-6 (2012).
472 22. Prahallad, A. *et al.* Unresponsiveness of colon cancer to BRAF(V600E) inhibition through
473 feedback activation of EGFR. *Nature* **483**, 100-3 (2012).
474 23. Kopetz, S. *et al.* Encorafenib, Binimetinib, and Cetuximab in *BRAF* V600E–Mutated
475 Colorectal Cancer. *N Engl J Med* (2019).
476 24. Misale, S., Di Nicolantonio, F., Sartore-Bianchi, A., Siena, S. & Bardelli, A. Resistance to
477 anti-EGFR therapy in colorectal cancer: from heterogeneity to convergent evolution. *Cancer*
478 *Discov* **4**, 1269-80 (2014).
479 25. Brauner, A., Fridman, O., Gefen, O. & Balaban, N.Q. Distinguishing between resistance,
480 tolerance and persistence to antibiotic treatment. *Nat Rev Microbiol* **14**, 320-30 (2016).
481 26. Balaban, N.Q. *et al.* Definitions and guidelines for research on antibiotic persistence. *Nat Rev*
482 *Microbiol* **17**, 441-448 (2019).

483 27. Di Nicolantonio, F. *et al.* Replacement of normal with mutant alleles in the genome of normal
484 human cells unveils mutation-specific drug responses. *Proc Natl Acad Sci U S A* **105**, 20864-
485 9 (2008).
486 28. Oren, Y. *et al.* Cycling cancer persister cells arise from lineages with distinct programs.
487 *Nature* **596**, 576-582 (2021).
488 29. Gardner, S.N. A mechanistic, predictive model of dose-response curves for cell cycle phase-
489 specific and -nonspecific drugs. *Cancer Res* **60**, 1417-25 (2000).
490 30. Kussell, E., Kishony, R., Balaban, N.Q. & Leibler, S. Bacterial persistence: a model of
491 survival in changing environments. *Genetics* **169**, 1807-14 (2005).
492 31. Lang, G.I. & Murray, A.W. Estimating the per-base-pair mutation rate in the yeast
493 Saccharomyces cerevisiae. *Genetics* **178**, 67-82 (2008).
494 32. Bailey, N.T.J. *The Elements of Stochastic Processes with Applications to the Natural*
495 *Sciences*, (Wiley, 1990).
496 33. Van Kampen, N.G. *Stochastic Processes in Physics and Chemistry*, 480 (Elsevier Science,
497 1992).
498 34. Misale, S. *et al.* Blockade of EGFR and MEK intercepts heterogeneous mechanisms of
499 acquired resistance to anti-EGFR therapies in colorectal cancer. *Sci Transl Med* **6**, 224ra26
500 (2014).
501 35. Oddo, D. *et al.* Molecular Landscape of Acquired Resistance to Targeted Therapy
502 Combinations in BRAF-Mutant Colorectal Cancer. *Cancer Res* **76**, 4504-15 (2016).
503 36. Rattray, A.J. & Strathern, J.N. Error-prone DNA polymerases: when making a mistake is the
504 only way to get ahead. *Annu Rev Genet* **37**, 31-66 (2003).
505 37. Goodman, M.F. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu*
506 *Rev Biochem* **71**, 17-50 (2002).
507 38. Wojtaszek, J.L. *et al.* A Small Molecule Targeting Mutagenic Translesion Synthesis Improves
508 Chemotherapy. *Cell* **178**, 152-159.e11 (2019).
509 39. Chatterjee, N. *et al.* REV1 inhibitor JH-RE-06 enhances tumor cell response to chemotherapy
510 by triggering senescence hallmarks. *Proc Natl Acad Sci U S A* **117**, 28918-28921 (2020).
511 40. Russo, M. *et al.* Reliance upon ancestral mutations is maintained in colorectal cancers that
512 heterogeneously evolve during targeted therapies. *Nat Commun* **9**, 2287 (2018).
513 41. Rehman, S.K. *et al.* Colorectal Cancer Cells Enter a Diapause-like DTP State to Survive
514 Chemotherapy. *Cell* **184**, 226-242.e21 (2021).
515 42. Werner, B. *et al.* The Cancer Stem Cell Fraction in Hierarchically Organized Tumors Can Be
516 Estimated Using Mathematical Modeling and Patient-Specific Treatment Trajectories.
517 *Cancer Res* **76**, 1705-13 (2016).
518 43. Michor, F. *et al.* Dynamics of chronic myeloid leukaemia. *Nature* **435**, 1267-70 (2005).
519 44. Woolston, A. *et al.* Mutational signatures impact the evolution of anti-EGFR antibody
520 resistance in colorectal cancer. *Nat Ecol Evol* **5**, 1024-1032 (2021).
521

522

523

524 **METHODS**

525 **Experimental setup and data collection**

526 ***Cell cultures.*** Cells were routinely supplemented with FBS 10%, 2mM L-glutamine, antibiotics
527 (100U/mL penicillin and 100 mg/mL streptomycin) and grown in a 37°C and 5% CO_2 air incubator.

528 Cells were routinely screened for absence of Mycoplasma contamination using the Venor® GeM
529 Classic kit (Minerva biolabs). All the cell lines used were confirmed negative for Mycoplasma
530 contamination in all the tests performed. The identity of each cell line was checked no more than
531 three months before performing the experiments using the PowerPlex® 16 HS System (Promega),
532 through Short Tandem Repeats (STR) tests at 16 different loci (D5S818, D13S317, D7S820, D16S539,
533 D21S11, vWA, TH01, TPOX, CSF1PO, D18S51, D3S1358, D8S1179, FGA, Penta D, Penta E, and
534 amelogenin). Amplicons from multiplex PCRs were separated by capillary electrophoresis (3730
535 DNA Analyzer, Applied Biosystems) and analyzed using GeneMapper v.3.7 software (Life
536 Technologies). STR results for all the cell lines and corresponding clones matched the profiles
537 previously published[45]. WiDr and DiFi CRC cell populations were obtained by Prof. Bernards and
538 Prof. Baselga, respectively, as we previously reported[45]. JVE207 CRC cells were obtained by Dr.
539 Wezel, Department of Pathology, University Medical Center, Leiden (The Netherlands).
540
541 ***Isolation of CRC-derived clones.*** CRC clones were obtained by seeding WiDr and DiFi CRC
542 populations at limiting dilution of 1 cell/well in 96-multiwell plates in complete medium. Clones
543 were then selected for having growth kinetics and drug sensitivity comparable to the parental
544 counterparts. For growth testing, WiDr and DiFi populations and derived clones were seeded in 96-
545 multiwell plates ($2x10^3$ cells/well and $3x10^3$ for WiDr and DiFi respectively) in complete medium.
546 Plates were incubated at 37°C in 5% CO_2. Cell viability, assessed every day for 4 days by measuring
547 ATP content through Cell Titer-Glo® Luminescent Cell Viability assay (Promega) using the Tecan
548 Spark™ 10M plate reader with the Tecan SparkControl Magellan™ software (v. 2.2), was compared
549 to cell viability assayed at day 1. For drug sensitivity testing, cells were seeded at different densities
550 ($2x10^3$ cells/well and $3x10^3$ for WiDr and DiFi respectively) in medium containing 10% FBS in 96-
551 multiwell plates at day 0. The following day, serial dilutions of the indicated drugs in serum-free

552 medium were added to the cells (ratio 1:1) in technical triplicates, while DMSO-only treated cells
553 were included as controls. Plates were incubated at 37°C in 5% CO_2 for the indicated time. Cell
554 viability was assessed by measuring ATP content with the Cell Titer-Glo® Luminescent Cell Viability
555 assay (Promega), using the Tecan Spark™ 10M plate reader with the Tecan SparkControl Magellan™
556 software (v. 2.2). Dabrafenib was obtained from Selleckchem. Cetuximab was kindly provided by
557 MERCK.
558
559 ***Growth rates of CRC cell clones before drug treatment.*** Clonal spontaneous growth is defined by
560 the following parameters: the rate at which cells are born (birth rate, *b*), the rate at which cells die
561 (death rate, *d*) and the net growth rate *b*−*d*. To estimate the *b*−*d* rate, CRC cell clones were seeded
562 at 3.5-4x10^5 cells/well in 6-multiwell plates. Plates were incubated at 37°C in 5% CO_2. Starting from
563 the following day, the number of viable cells was assessed by manual count in trypan blue 0.4%
564 (Gibco™) by two operators independently at the indicated time points, in order to obtain the clones'
565 net growth rate. To estimate *d/b*, cells were seeded at different densities (3.5-4x10^5 cells/well) in
566 multiple 6-multiwell plates. Plates were incubated at 37°C in 5% CO_2. At each time point, cells were
567 collected and stained with Propidium Iodide (Sigma Aldrich) following manufacturer's instructions.
568 Cells were then analyzed by flow cytometry. Data were acquired with the Beckman Coulter CyAn™
569 ADP instrument using the Summit v. 4.3 software and analyzed with the FlowJo software (v. 7.6).
570 The cells in sub-G1 phase were considered dead and used to estimate *d/b*. The values of birth (*b*)
571 and death rate (*d*) were then obtained by combining *b*−*d* and *d/b* estimates (Supplementary note).
572
573 ***Doses-response growth curve assay.*** CRC cell clones were seeded at different densities (2x10^3
574 cells/well and 3x10^3 for WiDr and DiFi respectively) in medium containing 10% FBS in multiple 96-
575 multiwell plates at day 0. The following day, serial dilutions of the indicated drugs in serum-free

576 medium were added to the cells (ratio 1:1) in technical triplicates, while DMSO-only treated cells
577 were included as controls. Cell viability of WiDr and DiFi clones was assessed at indicated time points
578 over 5 and 19 days of constant treatment, respectively, by measuring ATP content through Cell Titer-
579 Glo® Luminescent Cell Viability assay (Promega), using the Tecan Spark™ 10M plate reader with the
580 Tecan SparkControl Magellan™ software (v. 2.2).

581

582 ***Single-dose growth curve assay.*** DiFi and WiDr CRC cell clones were seeded in multiple 96-multiwell
583 plates at 1,000 or 500 cells/well, respectively. Cells were allowed to expand for a fixed number of
584 generations until a population size of 10,000-20,000 cells/well was reached. At that point, treatment
585 was added (100 µg/ml cetuximab for DiFi and 1µM dabrafenib + 50µg/ml cetuximab for WiDr).
586 Plates were then incubated at 37°C in 5% CO_2 and cell viability was assessed at the indicated time
587 points by measuring ATP content with the Cell Titer-Glo® Luminescent Cell Viability assay (Promega),
588 using the Tecan Spark™ 10M plate reader with the Tecan SparkControl Magellan™ software (v. 2.2),
589 over 22 and 32 days of constant treatment (for WiDr and DiFi respectively). Medium and treatment
590 were renewed once a week. To test the effect of different seeding densities on the residual viability
591 assayed, each clone was seeded at different densities (3-20x10^3 cells/well) in complete medium.
592 The following day, treatment was added (100 µg/ml cetuximab for DiFi and 1µM dabrafenib +
593 50µg/ml cetuximab for WiDr) and viability was assessed at the indicated time points by measuring
594 ATP content.

595

596 ***Staining with Carboxy fluorescein succinimidyl ester (CFSE).*** CRC clones were seeded at 2.5x10^5
597 (WiDr) and 6.5x10^5 (DiFi) cells in multiple 10cm dishes. The following day, untreated cells were
598 stained with CellTrace™ CFSE Cell Proliferation Kit (Invitrogen™) according to manufacturer's
599 instructions. At the indicated timepoints, starting from the day after staining (T0), cells were

600 collected and resuspended in 1mL PBS with Zombie Violet™ 1,000x (BioLegend®) to exclude dead
601 cells. Cells were then analyzed by flow cytometry. For persisters proliferation analysis, CRC clones
602 were seeded at $2x10^4$ cells/well in several 24-multiwell plates. The following day, cells were treated
603 with 100 µg/ml cetuximab (for DiFi) or 1µM dabrafenib + 50µg/ml cetuximab (for WiDr) and
604 incubated at 37°C in 5% CO_2 for 14 days (renewing treatment after 1 week) until a population of
605 persister cells emerged in each well. Then, cells were stained with CellTrace™ CFSE Cell Proliferation
606 Kit (Invitrogen™) according to manufacturer instructions. At the indicated timepoints, starting from
607 the day after staining (T0), plates were checked to exclude from the analysis wells containing
608 resistant clones, while cells from the remaining wells were collected, resuspended in 1mL PBS with
609 Zombie Violet™ 1,000x (BioLegend®) to exclude dead cells, and analyzed by flow cytometry.
610 Medium and treatment were renewed once every week throughout the experiment. Flow
611 cytometry was performed using the Beckman Coulter CyAn™ ADP analyzer with the Summit v. 4.3
612 software (Beckman Coulter) and analyzed with a Python 3 script based on standard libraries
613 (FlowCal, FlowKit). The following gating strategy was used. First, cells were selected with a light
614 scattering gate (FSLin vs SS), excluding cell doublets with a single cell gate (FSArea vs SSArea). The
615 following cutoffs were used (Supplementary figure 9): (i) *FS Lin:* lower 5000 and upper 60000; (ii) *SS*
616 *Lin:* lower 3000 and upper 63000; (iii) *FS Area*: lower 3000 and upper 60000; *SS Area*: lower 2000
617 and upper 63000. We then evaluated the bi-dimensional distribution of the remaining data points
618 in the space of the coordinates FS Area and SS Area, and retained all the data-points that were
619 included in the 99nth percentile of the distribution. Viable cells were selected by excluding Zombie
620 Violet™-positive cells and CFSE signal was detected by measuring Fitc signal.
621
622 ***Staining with 5-ethynyl-2'-deoxyuridine (EdU).*** DiFi and WiDr clones were plated on several glass
623 coverslips at $5x10^4$ and $4x10^4$ cells/coverslip, respectively. The following day, cells were treated with

624 100 µg/ml cetuximab (for DiFi) or 1µM dabrafenib + 50µg/ml cetuximab (for WiDr) and incubated
625 at 37°C in 5% CO_2 for 14 days (renewing treatment after 1 week) until a population of persister cells
626 emerged on each coverslip. Then, at indicated timepoints (renewing medium and treatment once
627 every week throughout the experiment), residual cells were stained with the Click-iT™ EdU Cell
628 Proliferation Kit for Imaging (Invitrogen™) according to manufacturer's instruction. Briefly, cells
629 were incubated with 10µM EdU for 4 hours. After that, cells with fixed in 4% paraformaldehyde for
630 20 minutes at room temperature and permeabilized with 0.5% Triton® X-100 in PBS for 20 minutes
631 at room temperature. Coverslips were then incubated in Click-iT® reaction cocktail for 30 minutes,
632 followed by nuclei staining with DAPI and F-Actin staining with Alexa Fluor® 555 Phalloidin (50
633 µg/ml). Slides were then mounted using the fluorescence mounting medium (Dako, Glostrup, DK).
634 For quantification of EdU-positive persister cells, DAPI- and EdU-stained nuclei were detected with
635 a Leica DMI6000B fluorescence microscope (Leica Microsystems, Wetzlar, Germany) under a 40X
636 dry objective using the Leica Application Suite Advanced Fluorescence software (v. 2.6.3.8173).
637 Images were analyzed with "Analyze particles" function in ImageJ (v. 1.53a) in order to calculate the
638 percentage of EdU-positive cells out of the total number of cells in each slide (based on DAPI
639 staining). Two separate technical replicates, with a minimum of 200 cells each, were analyzed for
640 each timepoint for each biological replicate. Resistant colonies that had grown on each slide were
641 manually identified in each image and excluded from the analysis. Representative images shown for
642 each cell clone were acquired with a confocal laser scanning microscope (TCS SPE II; Leica, Wetzlar,
643 Germany) using the Leica Application Suite Advanced Fluorescence software (v. 2.6.3.8173) and
644 processed with Adobe Photoshop CS5.
645
646 ***Time-lapse microscopy assay.*** For live cell imaging experiments, DiFi and WiDr clones were seeded
647 in 24-multiwell plates suitable for microscopy (µ-Plate 24 Well Black, ibidi®) at $5x10^4$ and $4x10^4$

648 cells/well, respectively. The following day, cells were treated with 100 μg/ml cetuximab (for DiFi) or
649 1μM dabrafenib + 50μg/ml cetuximab (for WiDr) and incubated at 37°C in 5% CO_2 for 14 days
650 (renewing treatment after 1 week) until a population of persister cells emerged. Then, surviving
651 persister cells were labeled with a fluorescent stain for nuclei (Nucblue ®, Invitrogen™, at 0.5
652 drops/well) to track cell number and detect cell divisions, and a live fluorescent marker for the
653 activation of Caspase3/7 (CellEvent™ Caspase-3/7 Green Detection Reagent, Invitrogen™, at 1
654 drop/well) to detect cell death. Both dyes were used as recommended by the manufacturer. After
655 labeling, cells were monitored for 5 days under an inverted widefield microscope (Nikon Lipsi, 20X
656 Plan Apo objective with 0.75 NA) acquiring images every 45 minutes with the Nis-Element AR
657 software (v. 5.21.03 64 bit; Nikon). For each clone, two separate wells with 16 fields of view each
658 were monitored, for a total of more than 1300 cells for each cell model. By digital image
659 segmentation carried out with Ilastik (v. 1.3.3 opensource) [46,47], the number of cells as a function of
660 time for each frame was quantified and then fitted to an exponential function in order to extract
661 the net growth rate ($b - d$). Data analysis and manipulations were performed by means of custom
662 written Matlab R2121a (the Mathworks) scripts. Representative snapshots of cell division events
663 and apoptotic events (obtained with Fiji v. 1.53 opensource) are reported in Extended data fig. 3,
664 while Supplementary Movies report whole time-lapse experiments for selected fields of view for
665 each clone. Scale bar in supplementary movies: 200μm.
666
667 ***Characterization of distribution of persister cells.*** DiFi and WiDr cell clones were seeded in multiple
668 96-multiwell plates at 1,000 or 500 cells/well, respectively. Subsequently, cells were allowed to
669 expand until they reached 10,000-20,000 cells/well. Cell viability was then assessed by measuring
670 ATP content to normalize for cell number prior to treatment initiation. The remaining plates were
671 treated with targeted therapies (100 μg/ml cetuximab for DiFi and 1μM dabrafenib + 50 μg/ml

672 cetuximab for WiDr). Medium and treatment were renewed once a week. After 3 weeks of constant
673 drug treatment, residual viability was assessed by measuring ATP content with the Cell Titer-Glo®
674 Luminescent Cell Viability assay (Promega), using the Tecan Spark™ 10M plate reader with the
675 Tecan SparkControl Magellan™ software (v. 2.2).

676

677 ***Two-steps fluctuation assay.*** DiFi and WiDr clones were seeded at 1,000 or 500 cells/well,
678 respectively, in twenty 96-multiwell plates each, for a total of 1,920 independent replicates. Cells
679 were allowed to expand for a fixed number of generations until they reached 10,000-20,000
680 cells/well. Next, treatment was administered (100 µg/ml cetuximab for DiFi and 1µM dabrafenib +
681 50 µg/ml cetuximab for WiDr). Plates were incubated at 37°C in 5% CO_2 for the indicated time.
682 Media and drug treatment were renewed once a week. After 3-4 weeks of treatment, pre-existing
683 resistant colonies were clearly distinguishable at the microscope and counted by two independent
684 observers. The number of pre-existing resistant clones was used to estimate the spontaneous
685 mutation rate of CRC clones (Supplementary note section *Estimator of mutation rate for sensitive*
686 *cells*). After 10-11 weeks, resistant colonies started to emerge in wells where only persisters were
687 previously present. The number of persister-derived resistant clones was used to estimate the
688 mutation rate of persister cells under constant treatment (Supplementary note section *Estimator of*
689 *mutation rate for persister cells*). Pictures of the resistant colonies were acquired using a ZEISS Axio
690 Vert. A1 microscope equipped with a True Chrome HD II camera.

691

692 ***Droplet digital PCR analysis.*** Genomic DNA (gDNA) was extracted using ReliaPrep® gDNA Tissue
693 Miniprep system System (Promega). Purified gDNA was amplified using ddPCR Supermix for Probes
694 (Bio-Rad) using *RAS* (PrimePCR ddPCR Mutation Assay, Bio-Rad or custom designed) ddPCR assay
695 for point mutations detection. ddPCR was performed according to manufacturer's protocol. Briefly,

696 5 µl of DNA template was added to 10µL of ddPCR Supermix for Probes (Bio-Rad), 1 µL of the primer
697 and probe mixture. Droplets were generated using the Automated Droplet Generator (Auto-DG,
698 Bio-Rad) and transferred to a 96 well plate and then thermal cycled with the following conditions:
699 10 min at 95 °C, 40 cycles of 94 °C for 30 s, 55 °C for 1 min followed by 98 °C for 10 min (Ramp Rate
700 2.5 °C/s). Droplets were analyzed with the QX200 Droplet Reader (Bio-Rad) with the QuantaSoft
701 software (v. 1.7.4.0917; Bio-Rad) for fluorescent measurement of FAM and HEX probes. Gating was
702 performed based on positive and negative controls, and mutant populations were identified. The
703 ddPCR data were analyzed with QuantaSoft analysis software (v. 1.7.4.0917; Bio-Rad) and results
704 were reported as the percentage (fractional abundance) of mutant DNA alleles to total (mutant plus
705 wild-type) DNA alleles. Fractional Abundance is calculated as follows: F.A.% = (Nmut/(Nmut + Nwt))
706 × 100), where Nmut is the number of mutant events and Nwt is the number of wild-type events per
707 reaction. ddPCR analysis of normal control DNA (from cell lines) and no DNA template controls were
708 always included.
709
710 ***Library preparation and genetic analysis of Whole Genome Sequencing.*** Genomic DNA (gDNA) was
711 extracted using ReliaPrep® gDNA Tissue Miniprep system System (Promega) and sent to IntegraGen
712 SA (Evry, France) that performed library preparation. DNA libraries were paired-end sequenced on
713 Illumina HiSeq4000 and FASTQ files produced by IntegraGen were analysed at Candiolo Cancer
714 Institute. BWA-mem algorithm was performed to align sequences on the reference human genome
715 version 19. Resulting files were cleaned of PCR duplicates by "rmdup" sam tools command. For each
716 cell line, somatic mutation analysis was performed subtracting variations found in parental
717 (sensitive) samples to resistant counterparts according to what has been previously published [48].
718 For each cell line pre- and post-treatment, gene copy number (GCN) was computed as follows: first
719 the median read depth of all genomic regions was calculated; next, for each gene the median read

720 depth was obtained and then divided by the former value. For each gene, its GCN in the pre- and

721 post-treatment samples and the corresponding copy number variation (CNV, ratio between

722 matched GCNs) were reported. DNAcopy R module was performed to cluster CNV using circular

723 binary segmentation (CBS) algorithm.

724

725 ***Time-to-progression assay.*** Time-to-progression assays were conducted as previously described[40].

726 Briefly, 5 million cells (for WiDr and DiFi cells) and 4.5 million cells (for JVE207 cells) were plated for

727 each treatment condition. Then, cells were treated with either MAPK pathway inhibitors

728 (dabrafenib 1µM + cetuximab 30µg/mL for WiDr and JVE207; cetuximab 50µg/mL for DiFi), or REV1

729 inhibitor (2µM) or their combination, in parallel. Media and treatment(s) were renewed weekly.

730 Cells were counted each week; counts as 0 represent time points in which cells were too few and

731 only medium and drug refreshments were done.

732

Materials availability

The CRC cell clones generated in this study are available through Alberto Bardelli (Department of

Oncology, University of Torino) under a Material Transfer Agreement.

733 **Theoretical modeling and computational analyses.** In this study we have developed and used two

734 distinct mathematical models to investigate the dynamics of cell populations. The first model

735 describes the transition to persister state ("TP model"), and is a birth-death model with phenotypic

736 switching, which we explored in the deterministic limit. We considered four different model variants

737 and we compared them to experimental data in order to infer the most likely scenario for the sensitive-

738 to-persister transition. The second model, which we named *Mammalian Cells-Luria Delbrück* or

739 "MC-LD" model, is a fully stochastic birth-death branching process that includes the mutational

740 processes of sensitive (untreated) and persister cells (under treatment). In order to measure the

741 mutation rate, stochastic fluctuations cannot be neglected. We simulated individual trajectories of the
742 Markov process underlying the evolution of the MC-LD model by a coarse-grained version of the
743 Gillespie algorithm[49], which groups together all stochastic events happening in discrete time intervals
744 of fixed duration Δt.
745 For the inference of the birth-death rates b and d, we used the data on growth rates of CRC clones
746 before drug treatment. Our inference scheme is summarized in Supplementary figure 2a. The
747 parameters of the TP model were inferred using a Bayesian framework using data from the *single-*
748 *dose* and *dose-response assays*. Posterior distributions of the model parameters were sampled using
749 a Hamiltonian Monte Carlo (HMC) algorithm (*Python 3*, package *pymc3*, NUTS sampler) [50]. TP
750 model variants were compared by means of the standard Bayesian Information Criterion (BIC) and
751 the Akaike Information Criterion (AIC). To infer mutation rates for the MC-LD model, we computed
752 an approximate analytical expression for the probability of the emergence of one mutant in an
753 expanding population of cells in a given time interval $[0,T]$, and we used it to derive estimators for
754 the emergence of mutations before and during treatment administration (from persisters). The
755 mutation rate of persister cells was inferred with a Bayesian framework, in order to account for the
756 uncertainty of the value of the initial fraction of persister cells, f_0.
757 All the details on the theoretical/computational protocols are provided in the **Supplementary Note**
758 of the Supplementary Information document.
759

Data availability

760 Data used for the analysis, source data images of Edu staining and live microscopy assay are available
761 as a repository on Mendeley Data (doi:10.17632/mvfm7hs9kw.1)[51]. Sequencing data are available at
762 PRJEB49483 (ENA; https://www.ebi.ac.uk/ena/browser/home). The CRC cell clones generated in
763 this study are available through Alberto Bardelli (Department of Oncology, University of Torino)
764 under a Material Transfer Agreement.
765

766 **Code availability**

767 All the custom code used in our analyses is available as a repository on Mendeley Data
768 (doi:10.17632/mvfm7hs9kw.1)[51]. Bioinformatics code for sequencing data are available at
769 https://bitbucket.org/irccit/idea/src/master/. The code have been written in C++ (C++ 14 and g++
770 10.3.0); Mathematica (Mathematica 10); Python (Python 3.9.7); Matlab (Matlab R2121a, The
771 mathworks); Microsoft Excel (version 16.48).

772

773 # Methods-only References

774 45. Medico, E. *et al.* The molecular landscape of colorectal cancer cell lines unveils clinically
775 actionable kinase targets. *Nature Communications* **6**(2015).
776 46. Sommer, C., Straehle, C., Köthe, U. & Hamprecht, F.A. Ilastik: Interactive learning and
777 segmentation toolkit. in *2011 IEEE International Symposium on Biomedical Imaging: From*
778 *Nano to Macro* 230-233 (2011).
779 47. Berg, S. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nat Methods* **16**,
780 1226-1232 (2019).
781 48. Crisafulli, G. *et al.* Whole exome sequencing analysis of urine trans-renal tumour DNA in
782 metastatic colorectal cancer patients. *ESMO Open* **4**(2019).
783 49. Gillespie, D.T. Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem* **58**, 35-55
784 (2007).
785 50. Hoffman, M.D. & Gelman, A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in
786 Hamiltonian Monte Carlo. *Journal of Machine Learning Research* **15**, 1593-1623 (2014).
787 51. Russo, M. *et al.* Data and code from Russo et al. "A modified fluctuation-test framework
788 characterizes population dynamics and mutation rate of colorectal cancer persister cells".
789 DOI: 10.17632/mvfm7hs9kw.1 (Mendeley Data, 2022).

790

791