

# Optimization Aspects of Carcinogenesis

B. Brutovsky<sup>a,\*</sup> and D. Horvath<sup>b,c,†</sup>

<sup>a</sup>*Institute of Physics, P. J. Safarik University, Jesenna 5, 04154 Kosice, Slovakia*

<sup>b</sup>*Centre de Biophysique Moléculaire, CNRS; Rue Charles Sadron, 45071 Orléans, France and*

<sup>c</sup>*Department of Physics, Faculty of Electrical Engineering and Informatics, Technical University, Letná 9, 042 00 Košice*

## Abstract

A process in which competing solutions replicate with errors and the numbers of their copies depend on their respective properties is the evolutionary optimization procedure. As carcinogenesis fulfills the essence of the process, we interpret it from the point of view of the theory of evolutionary optimization. Within the conceptual framework the mechanisms which are known to decrease the efficiency of any evolutionary optimization process are reviewed from the point of view of their eventual therapeutic perspective. At the end we present some counterintuitive implications stemming from the optimization facet of carcinogenesis enabling to interpret some not yet fully understood experimental findings, such as often observed more aggressive and malignant growth of therapy surviving cancer cells.

## Introduction

The term cancer refers to hundreds types of neoplasms which share specific prototypical traits, summarized by Hanahan and Weinberg [1], collectively leading to malignant growth. During the past few decades molecular biologists have produced much cancer-related data which has shown cancer as an extremely stochastic, heterogeneous and complex disease [2]. To analyze them, cancer research applies many concepts originally developed in different branches of science, such as applied mathematics, nonlinear dynamical systems, and statistical physics. At present, evolutionary nature of carcinogenesis is accepted and implications for cancer robustness (exemplified by resistance to therapy) are often emphasized [3, 4]. Darwinian view to carcinogenesis implicitly puts genetic (and epigenetic) changes into microenvironmental context [5]. Consequently, tumor microenvironment is viewed as an eventual target for chemoprevention and cancer reversion [6, 7]. On the other hand, anticancer research and therapy concentrate mainly on molecular data and tend to overlook its evolutionary nature.

Optimality model applied in experimental evolution [8] describes the evolution as simple generalized trade-offs, presuming that genomes adapt successfully and freely enough and, consequently, genetic details become irrelevant. Mathematical approaches to carcinogenesis often apply concepts of feedback and optimal control theory [9] instead of molecular or genetic data. Komarova et al. [10] have solved the optimization problem for cancerous growth and proposed optimal strategies. However, as they state, the ideal (optimal) strategy may not be realistic due to many constraints in nature which escape modeling, but can make a strategy impossible.

In the paper we concentrate on the abstract mechanisms of attaining an optimal strategy instead of the strategy itself. We view any process in which solutions replicate with errors and numbers of their copies depend on their respective qualities as an evolutionary optimization process. As carcinogenesis conforms the above definition, we identify it with an evolutionary optimization process and apply concepts and results of the long lasting research in the evolutionary optimization [11]. Keeping in mind an eventual therapeutic application, we focus on those aspects of evolutionary optimization which decrease or inhibit efficiency of the optimization process. Strict adherence to the optimization framework has led us to counterintuitive implications.

## Methods

### *Evolutionary optimization*

In the optimization theory, the quality of a solution is usually defined explicitly in the form of a *fitness function* (also *fitness landscape* or *fitness*), quantifying how well a candidate solution meets required criteria. The ultimate aim of the optimization procedure is to find a solution for which the fitness function receives optimum value. Large group of optimization algorithms, called evolutionary algorithms (EA), performs the task by mimicking biological evolution implementing the genetic-like mechanisms, such as mutation, selection and reproduction. Broad class of EA engineering applications has enabled to recognize those aspects of fitness landscapes which support efficient evolu-

---

\*Electronic address: bru@upjs.sk

†Electronic address: horvath.denis@gmail.com

tionary optimization and, at the same time, those which prevent it.

Theoretical analysis of the most popular EA variant, the genetic algorithms (GA), has been performed by J. H. Holland introducing the schema formalism [11]. It has enabled to identify the driving force behind the biological-like manipulations with binary strings representing parameters of the model. We present it as a motivation, not as a rigorous tool to analyze carcinogenesis. In the simplest engineering applications, canonical GA (CGA) applies:

1. initial population of random binary strings is generated

```
1011101011 ...
0001100101 ...

1011110010 ...
```

2. each of the bit string is projected and scaled to get the real parameter set  $\mathbf{r} = (r_1, r_2, \dots)$

```
1001...10100110...0111...
  r1          r2
```

and its fitness function value  $\phi(\mathbf{r})$  is determined

3. child population of the bit strings is constructed from the parents population applying the genetic operators - selection depending on the strings fitnesses, crossover and mutation; after it is complete it replaces the parent population
4. until some convergence criterion applies go to the step 2.

It was recognized that the information driving the above population-based optimization algorithm are the fitnesses of the correlations of bits in the binary strings. More exact analysis of GA internal workings is therefore based on the "schemas" concept [11]. The schema can be viewed as a bit pattern over the bit positions in the string. If the bit alphabet  $\{0,1\}$  is assumed, the schema can be easily constructed over the ternary alphabet  $\{0,1,*\}$ , where '\*' matches both, 0 and 1, at the respective position:

Let's have 4 binary strings

```
A 10100101
B 01011011
C 11100010
D 00010001
```

and two schemas, X and Y

```
X *1***01*
Y 0*01***1
```

The schema X is contained in the strings B and C, the schema Y in the strings B and D. It is usually said that strings B and C are instances of the schema X and the strings B and D are the instances of the schema Y.

Specificity and robustness of the schemas are quantified by the schema order,  $\mathcal{O}$ , and defining length,  $\delta$ . The schema order is the number of fixed positions in the schema. The defining length is the distance between the leftmost and the rightmost fixed positions. To predict the number of instances of a schema in generation  $t+1$ , Holland derived the schema theorem (ST) [11]

$$N^{t+1} \geq N^t \frac{\Phi^t}{\phi^t} \left[ 1 - P_c \frac{\delta}{l-1} - \mathcal{O} P_m \right], \quad (1)$$

where  $N^t$  and  $N^{t+1}$  are numbers of instances of the schema in  $t$  and  $t+1$ , respectively,  $P_c$  is the probability of strings crossover,  $P_m$  is mutation rate,  $l$  is the length of the binary string,  $\phi^t$  is the average fitness in the population, and  $\Phi^t$  is the schema fitness in  $t$  defined as the average fitness of all the instances of the schema in the population in  $t$ . ST (1) states that during the GA optimization the number of the above average schemas increases on the account of less favourable schemas. Exact convergence analysis of EA requires much better mathematical definition of the relevant fitness landscape and more obvious parametrization of a solution than one typically disposes with biological systems. Nevertheless, even without them, ST (1) can be applied to recognize the principal mechanisms beyond evolutionary optimization.

It has been demonstrated, that GA optimally allocates its trials among alternative solutions (known as  $k$ -arm bandit problem) during the search for the optimum solution as long as the schemas' fitnesses are correctly estimated [11]. As in the paper we have identified carcinogenesis with the evolutionary optimization process, any feature or mechanism which decreases efficiency of the optimization process is interesting from the point of view of its eventual therapeutic application. Recognizing ST (1) as the principal mechanism driving the evolutionary optimization, the therapy can be identified with partial or complete failure of ST. For that, the therapy must prevent reliable estimation of the schemas' fitnesses to mislead or slow down the optimization process. Below we list most often presented reasons preventing reliable estimations of the schemas' fitnesses.

*i) Too large sampling errors.* The factors influencing the reliability of statistical sampling are the number of evaluated candidate solutions and their distribution in the search space (i.e. population *heterogeneity*). They should cover as much of the search space (fitness landscape) as possible so that

the convergence to the optimum was as probable as possible. The sampling errors can be reduced by the appropriate choice of mutation rate. If mutation rate is too low, optimization sticks in a suboptimum solution (known as premature convergence). If mutation rate is too high, optimization procedure turns into so-called blind search. The existence of the critical mutation rate in evolution beyond which Darwinian selection does not operate has been predicted by Eigen's theory of quasispecies [12]. The mathematical model developed by Sole and Deisboeck [13] supports the existence of this error catastrophe also in cancer cells population.

*ii) Dynamic fitness landscape.* As the evolutionary optimization procedure converges towards optimum solution in a stationary fitness landscape, heterogeneity of the population decreases. The parts of the search space near the optimum become overpopulated, and, at the same time, other parts only sporadically populated, or even empty. The role of the observed increase of population heterogeneity in changed environment is well interpretable using the terms of evolutionary optimization, namely evolution algorithms in dynamic environments. Efficient transition from the old optimum to optimum(a) in a new fitness landscape requires i) detection of the fitness landscape change, and ii) response to that change [14]. For that, candidate solutions must be appropriately distributed in the search space so that evolutionary algorithms could perform representative statistical sampling to determine reliable schemas' fitnesses estimates which are necessary for optimal allocation of trials during optimization. If there are no (or too few) evaluations in the changed part of the fitness landscape, the change goes undetected. Therefore, mechanisms of heterogeneity maintenance have been developed in optimization theory and deeply studied [14].

*iii) Deceptiveness of fitness landscape.* To answer the question which fitness landscapes are GA-hard, Bethke [15] expresses a fitness function as a linear combination of Walsh monomials and showed the relationship between the schema's fitness and Walsh coefficients. Consequently, he applied the Walsh transform to characterize functions as easy or hard for GA optimization. It has been understood that the principal problem for GA optimization is the class of deceptive fitness functions, in which lower order (lower number of defined bits) schemas lead the search towards bad higher order schemas. Goldberg showed the possibility of constructing high-order deceptive functions using low-order Walsh coefficients in special cases [16].

## Results

### *Carcinogenesis as evolutionary optimization process*

Any process in which competing solutions replicate with errors and numbers of their copies depend on their respective qualities is an evolutionary optimization process. From this point of view carcinogenesis is an evolutionary optimization process. Regarding the above introduced schema formalism for CGA two differences should be mentioned. At first, carcinogenesis is an asexual process, therefore constant  $P_c$  in (1) equals zero. The second difference is that in (1) no spatial relation between offsprings and their parents is assumed. Neither of the two differences puts in doubts importance of reliable estimates of the schemas' fitnesses for optimal allocation of trials. Moreover, as often used in evolutionary optimization practice, we use the term optimum solution in a sense of a winning solution, i. e. the best solution obtained after reasonable (or affordable) long optimization, instead of exact, mathematically proved, solution. Below we analyze those aspects of carcinogenesis which relate to its optimization facet.

*Fitness landscape.* In biology, the fitness is usually understood in a sense of "reproduction" fitness, meaning that the more copies solution has the more fit it is. The term has no meaning without reference to specific environment and time scales. It is unexceptionable fact that solutions which are good in some environments may perform badly in others. The same is true for the time scales - solutions which are good on short scales can have very bad long term perspective (and vice versa). Without going into genetic details, we apply pleiotropic view to the genome as a whole. It relates the two units of replication: cells and organisms (multicellular bodies). Both units have evolved regarding their respective environments and time scales towards maximizing their respective fitnesses derived from the same object - genome. The genome is the physical realization of the trade-off between the two processes: i) maximization of the organism's reproduction fitness (acting during millenia), and ii) maximization of cellular reproduction fitness (individual lifespan), respectively. Thus, effectively two evolutionary processes have been sculpting the same genome, nevertheless on different time scales. The former process presumes social cooperation of cells (e. g. limited replicative potential, production of growth signals, sensitivity to antigrowth signals, cellular senescence, apoptosis, etc.) and severe prohibition of the cells' selfishness, the latter favours selfishness instead of cooperation [17]. The trade-off is mediated by the initial genomic stability, evolved to postpone short scale evolution in the respective environment beyond reproduction period of the respective organism.

Putting it into optimization context, the wild-type genome represents optimum solution in the respective environment; its further optimization in unchanged fitness landscape is, by definition, inhibited. If, however, the fitness landscape has changed, optimization of the genome becomes possible. Regarding the structure of the fitness landscape, during the optimization two fitness landscapes are independently sampled, each for the respective unit of replication - organism or cell. As there are many cellular fitness evaluations during the only organism's lifespan, only cellular fitness landscape may be sampled representatively enough to provide reliable schemas' fitnesses on the short time scale. As implied by (1), optimization process allocates its trials optimally driving the short time evolution of the genome into an optimum in the changed cellular fitness landscape. The organismal fitness landscape, selecting for intercellular cooperation, does not apply during the lifetime of the body and the optimization process is driven purely by cellular fitness landscape for which the intercellular cooperation is not desirable trait and its decay increases the genome's cellular fitness. From this point of view, any short-scale change of the fitness landscape is not only mutagenic but also carcinogenic, as it selects for destroying intercellular cooperation. Applying the quasispecies model [12], Forster and Wilke have demonstrated that competitive dynamics of finite populations of as few as two strains, adapted to the long-term and short-term environment changes, respectively, is quite complex [18].

*Increased heterogeneity.* Heterogeneous nature of neoplasms (in comparison to normal cells) reveals at different levels and number of characteristics [19]. Extreme tumor cells heterogeneity gives cancer robustness, exemplified by the resistance to therapy [3, 4], and it is the most tormenting problem in cancer research to which therapies and experimental models must face [20, 21]. Recently, remarkable genomic heterogeneity has been demonstrated in samples of colorectal cancers [2, 22]. It has been shown that sets of mutated genes in two samples of colorectal cancers overlap to only a small extent and it is anticipated to be general feature of most solid tumours [23]. Similarly, resuming studies in breast and renal cancer, Gatenby and Frieden [24] concluded that probably no prototypical cancer genotype exists and every tumor seems to possess a unique set of mutations indicating that multiple genetic pathways may lead to invasive cancer as would be expected in a stochastic non-linear dynamical system. Clonal diversity in a subset of patients with early stage haematopoietic malignancy has been demonstrated and it has been shown that such clones may arise independently [25]. It has been also observed [26], that time to disease progression and overall survival after treatment were significantly shorter in those

patients with EGFR heterogeneity. Maley et al. [27] have demonstrated that clonal diversity predicts progression to cancer and that accumulation of viable, clonal genetic variants is a greater risk for progressing to cancer than homogenizing clonal expansion. Mathematical model by Komarova et al. [10] shows that tumors thrive when cancerous cells mutate to speed up malignant transformation, and then stay that way by turning off the mutation rate. The question whether the genetic instability is a driving force or a consequence has not been satisfactorily answered yet.

On the other hand, heterogeneity is central to the evolutionary optimization. If heterogeneity is appropriate for reliable schemas' fitnesses determination, the evolutionary optimization process optimally allocates its trials among competing solutions. Most engineering applications start with heterogeneous, typically random, initial population of candidate solutions. In the case of stationary fitness landscapes, heterogeneity decreases towards some minimum level as the optimization procedure converges to the best solution (the analogy with a homogenizing clonal expansion inflicts itself), despite typically constant mutation rate. On the other hand, evolutionary optimization in changing fitness landscapes [14] shows importance of avoiding total homogenization. In experiments where mutation rate is not exempted from optimization, its increase (followed by the increase of heterogeneity) is observed after the fitness landscape has changed. Similarly, selection of mechanisms for increased mutation rate in biological systems, like RNA viruses, in unstable environments was reported [28]. Donaldson-Matasci et al. [29] have shown that optimal amount of diversity depends on environmental uncertainty which can lead to the evolution of either generalist or specialist strategy.

*Evolvability.* Heterogeneity of the genomes population comes as the combination of effective population size, genomic stability, and relation between the genome's quality and the number of its offspring. Evolutionary optimization view implies that these attributes must enable as reliable schemas' fitnesses estimates as possible. Statistically different fitness landscapes require different attributes. It has been demonstrated that rapid or extreme environmental change leads to the selection for greater evolvability [30]. As evolving clones implicitly undergo competition, the schemas' fitnesses must be, at the same time, as fast as possible determined. It means that the number of fitness evaluations per time unit, reproduction rate, implicitly enters optimization.

*Implications for therapy*

Below we present specific insights and implications for anti-cancer therapy stemming from the above presented optimization view to carcinogenesis. Some of them are intuitive and consistent with established anti-cancer therapies, some others are quite counterintuitive and, hopefully, novel and put in question some current trends in the development of anti-cancer therapies.

Accordingly to pathologists, there are three possible mechanisms by which cancer cells can outnumber the wild type genomes: i) they replicate faster; ii) a smaller proportion of them die; or iii) a greater proportion of daughter cells replicates compared with normal cells; the two first mechanisms not being observed frequently [31]. Straightforward therapies rely on distinguishing between normal and tumor cells. For instance, radiation therapy distinguishes cells as unable to repair (tumor cells) and those maintaining this capability and is more deleterious for the former. Chemotherapy is more harmful for more frequently replicating cells than those which replicate less often. Novel targeted and gene therapies go even further - they interfere directly with the specific molecules or genes participating in carcinogenesis. Nevertheless, eventual precise enough distinguishing between the tumor and wild type cells is in contradiction with the evolutionary nature of carcinogenesis which, as any other evolutionary process, crucially depends on the variability of traits observed at many levels [2, 22].

Within the frame of the above outlined identification of carcinogenesis as the evolutionary optimization process, therapy is a purposeful effort to decrease the efficiency of that optimization process or, hopefully, inhibit it completely. For that purposes, we have listed above the three most frequent obstacles to the efficient evolutionary optimization, stemming from validity of the schema theorem (1). These are: too large sampling errors, dynamic (or changing) fitness landscapes and deceptiveness of fitness landscape. In all the cases the estimation of the schemas' fitnesses is not reliable (or systematically wrong) which prevents the optimization process in optimal allocation of its trials.

*i) Too large sampling errors.* It is understood that diversification plays a central role in evolution and provides species (or clones) with a capacity to cope with environmental uncertainty. On the other hand, if genetic instability exceeds a certain threshold, the deleterious effects outweigh the above selection advantage [32]. Consequently, as defective stability pathways of tumor cells make them sensitive to stress-inducing agents, the tumor cells could be target for direct attack by instability drugs [33]. In [13] the simple mathematical model of quasispecies dynamics has been used to quantify the upper limit of affordable genetic instability (er-

ror threshold), beyond which genetic information is lost. From the point of view of the evolutionary optimization theory, to find an optimum solution, a clone (as a subset of the whole population) must optimally allocate its trials between emerging alternatives which requires reliable and prompt schemas' fitnesses estimations. Reliability of the schema estimates depends not only on genetic instability but, at the same time, on many other parameters, such as effective population (clone) size, reproduction rate, and fitness landscape. Being evolving/optimizing system, cancer cells will respond by changing other evolutionary attribute(s) to reduce sampling error, for instance, by efficiently increasing its size, reproduction rate, cellular mortality, stability (the mechanism does not matter at this point), etc.

*ii) Dynamic fitness landscape.* Changing the fitness landscape can be a double edged sword. On the one hand, cancer cells reveal increased adaptivity enabling them to respond to environmental changes to keep high (reproductive) fitness. On the other hand, higher adaptivity of cancer cells can be therapeutically exploited, as outlined by Maley et al. [34]. They proposed to create "a booster" to help a benign neoplasm to outperform malignant cells by increasing its relative fitness (increasing proliferation rate, decreasing the rate of apoptosis and other forms of cellular mortality or boosting). Alternatively, they proposed to select for the cells sensitive to cytotoxins before applying cytotoxic therapy.

*iii) Deceptiveness of fitness landscape.* As the last of the obstacles to efficient optimization we have listed deceptiveness of fitness landscape. Deceptive landscapes can be interpreted as the landscapes in which correlations of traits systematically lead away the search from the global optima. To our knowledge, there is no therapeutic approach explicitly exploiting deceptiveness of the fitness landscape. We anticipate that combining biological intuition, the results of mathematical analysis of deceptive fitness landscapes [16] and digitized evolution [35] can bring new promising insights into the evolution of cancer phenotype.

All the three above obstacles to the efficient evolutionary optimization apply in all the evolutionary optimization processes. What is, for our purposes, very important for their eventual application in anticancer treatment is their statistical nature. Evolutionary optimization view to carcinogenesis bridges the results of many papers referring connection between the statistical characteristics of the fitness landscape and the above evolutionary attributes [18, 36, 37, 38, 39, 40] with the results of long-standing research in the stochastic evolutionary algorithms, especially in dynamic fitness landscapes [14, 41]. It provides novel conceptual framework for modification of evolutionary attributes indirectly by changing *statistical* prop-

erties of the fitness landscape, such as roughness or dynamics, in purposeful cancer-inhibiting way.

### *Is therapy a penalty function?*

From the evolutionary optimization point of view therapy is a purposeful change of the fitness landscape, namely decrease of reproduction fitness in the relevant area of the search (sequence) space at reasonable time scales. All the well established traditional therapies (surgery, radiotherapy and chemotherapy) make an effort to remove all the cancer cells, or, at least, as many of them as possible. Evolutionary optimization theory implies that ultimate therapeutic success depends not only on how many cancer cells survived the therapy, but also on the distribution of the cells in the search space, i. e. statistics of the remaining population. If the population statistics is sufficient for the efficient optimization, the regrowth appears. Below we interpret the role of therapy and regrowth from the point of view of evolutionary optimization.

It has been reported that therapy-surviving tumor cells are frequently more malignant and aggressive than the initial tumor population [17, 42]. Inhibition of angiogenesis has been envisioned as promising anticancer therapeutic strategy for a long time [43]. Since then, modes of resistance to antiangiogenic therapy, such as evasive and intrinsic resistance, has been reported [44]. It has been found by Paez-Ribes et al. [45] that targeting the vascular endothelial growth factor (VEGF) induces (apart from anti-tumour effects to primary tumor) higher invasiveness and, in some cases, increased lymphatic and distant metastasis. Ebos et al. [46] have found that the VEGFR/PDGFR kinase inhibitor can accelerate metastatic tumor growth and decrease overall survival in mice receiving short-term therapy. Similarly, it has been reported that the resistance to some synergistic drug combinations evolves faster than the resistance to individual drugs [47]. In their review Kim and Tannock [48] report that repopulation of cancer cells after radiotherapy as well as chemotherapy is often accelerated in comparison to untreated cases. The mechanism of this acceleration has not yet been understood.

In the spirit of our work we attribute the above increase of invasiveness and acceleration of the evolution of resistance during repopulation to the optimization facet of carcinogenesis. In engineering applications of evolutionary optimization one often applies *ad hoc* penalty function to disadvantage some part(s) of fitness landscape to accelerate convergence of the process into the optimum in desirable parts (Figure 1). The simplification of fitness landscape enables to perform more representative schema sampling of more promising parts at the same price obtaining more reliable schemas'

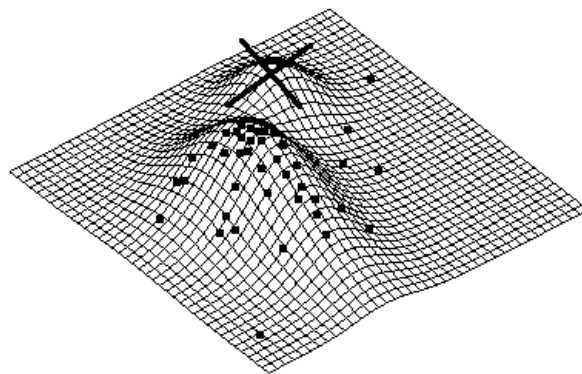


FIG. 1: 2-dimensional fitness landscape with one "hill" (crossed) prohibited by a penalty function. Dots show sampled points. Without penalty, both hills would be sampled.

fitnesses evaluations resulting, accordingly to (1), in closer-to-optimum allocation of the trials among alternative solutions. If cancer, metaphorically said, solves the optimization problem, the same mechanism applies. Therapy represents penalty function, unless it removes decisive portion of cancer cells. If not, therapy just simplifies fitness landscape and optimum can be more easily found by therapy-resistant clone(s). A paradoxical consequence of the optimization facet of carcinogenesis is that applied anti-cancer therapy can be, in fact, the causative factor of its acceleration on longer time scales.

## Discussion

Carcinogenesis is, unquestionably, driven by physical interactions. At the same time, it is, as all the evolutionary processes, an optimization process. The question stands what is the most appropriate aspect to combat cancer. Straightforward approaches study carcinogenesis and develop anticancer strategies analyzing biochemical or genetic details. In the paper we have speculated that it may be not relevant *per se*. Instead, we have proposed that cancer relates to the cells population statistics and all the therapies lead (more or less intentionally or explicitly) to its modification. Traditional therapies rely on comparison between cancerous and non-cancerous cells which may be motivated by the long lasting effort to reduce cancer cells population by some straightforward action. Evolutionary view suggests that carcinogenesis should be inhibited by a purposful modification of evolutionary attributes, such as mutation rate, effective population size or generation time of the self-renewing cells. Nevertheless, apart from the fact that we currently lack tools to measure or manipulate those attributes of neoplasms [49], evolutionary theory gives no clear answer why and

how much should be the evolutionary attributes changed (except for trivial cases). (Evolutionary) optimization view to carcinogenesis, implicitly based on interconnection between the population statistics and statistics of fitness landscape, offers a little bit more informative answer to this issue. It proposes that what is actually relevant is the allocation of trials in a sense of extracting as reliable information about correlations of traits as possible in the shortest possible time. It depends on the statistical properties of the fitness landscape itself and the above evolutionary attributes; these may intertwine in nontrivial way.

Mathematical models of physical phenomena presume well-posedness of the problem which, as defined by Hadamard a century ago [50], means that: i) a solution of the problem exists, ii) the solution is unique, and iii) the solution depends continuously on the data in some reasonable topology. On the other hand, causality in evolutionary processes is actually provided by the feedback from environment. In general, the evolutionary process is a fitting procedure, which is the method of solving (typically ill-posed) inverse problems [51]. Confronting with enormous diversity of cancers at all scales it seems that each cancer occurrence is a unique solution (fit) of the problem. It indicates that the fitting problem solved by cancer is very probably highly underdetermined, meaning that the number of effective degrees of freedom is much higher than the number of observables (at least for more advanced malignancies), which results in the arbitrariness of a fit (or model). The overfitting is consistent with the metaphoric conclusion by Witz and Levy-Nissenbaum [5], who stated "...the extreme complexity of the signaling cascades operating in the microenvironment and the interactive cross-talk between these cascades, generates the feeling that 'anything that can happen - it will'". Wild-type genome has come as one of many good enough solutions to underdetermined problem, though it probably does not represent the optimum solution in rigorous sense.

We have proposed that, as any other evolutionary optimization process, carcinogenesis progresses

by schema sampling as expressed by the schema theorem (1). Efficiency of the schema sampling depends on the number of sampled points and their distribution in the fitness landscape, as well as the fitness estimation time. These attributes adapt to statistical features of the respective fitness landscape simultaneously with the search. As, at the same time, efficiency of the sampling determines the therapeutic efficiency, we conclude that the therapeutic efficiency can be influenced by manipulation with statistical properties of the fitness landscape. The above statistical view may be relevant especially for advanced malignancies, where increased heterogeneity of the cancer cells population enables them to adapt successfully to therapeutically-changed environment.

Classifying carcinogenesis as the evolutionary optimization process does not contradict to often presented view of cancer as the result of accumulating specific mutations in the only transformed cell. It emphasizes, however, importance to combine molecular data with statistical view which may play crucial role before and during carcinogenesis. We have put some of the observed cancer features, such as increased heterogeneity in neoplasia, clonal escape, consequences of changing the fitness landscape and accelerated evolution of resistance to chemotherapy into optimization scenario. Presented view is far from complete. We conclude the paper by the question: Is the optimization facet of carcinogenesis causative enough to trigger a novel, explicitly optimization-preventing, therapeutic approach?

### Acknowledgements

The authors acknowledge financial support from VEGA, Slovak Republic (Grants 1/4021/07, 1/0300/09) and APVV (grant LPP-003006). D.H. acknowledges financial support by the LESTUDIUM fellowship of the Region Centre and the Centre National de la Recherche Scientifique.

- 
- [1] Hanahan, D. and Weinberg, R. A. (2000) The hallmark of cancer. *Cell*, **100**, 57–70.
- [2] Sjöblom, T., Jones, S., Wood, L. D., Parsons, D. W., Lin, J., Barber, T. D., Mandelker, D., Leary, R. J., Ptak, J., Silliman, N., Szabo, S., Buckhaults, P., Farrell, C., Meeh, P., Markowitz, S. D., Willis, J., Dawson, D., Willson, J. K. V., Gazdar, A. F., Hartigan, J., Wu, L., Liu, C., Parmigiani, G., Park, B. H., Bachman, K. E., Papadopoulos, N., Vogelstein, B., Kinzler, K. W., and Velculescu, V. E. (2006) The consensus coding sequences of human breast and colorectal cancers. *Science*, **314**, 268–274.
- [3] Komarova, N. L. and Wodarz, D. (2003) Evolutionary dynamics of mutator phenotypes in cancer: Implications for chemotherapy. *Cancer Res.*, **63**, 6635–6642.
- [4] Kitano, H. (2004) Cancer as a robust system: implications for anticancer therapy. *Nat. Rev. Cancer*, **4**, 227–235.
- [5] Witz, I. P. and Levy-Nissenbaum, O. (2006) The tumor microenvironment in the post-Paget era. *Cancer Lett.*, **242**, 1–10.
- [6] Anderson, A. R. A., Weaver, A. M., Cummings, P. T., and Quaranta, V. (2006) Tumor morphology and phenotypic evolution driven by selective

- pressure from microenvironment. *Cell*, **127**, 905–915.
- [7] Ingber, D. E. (2008) Can cancer be reversed by engineering the tumor microenvironment? *Semin. Cancer Biol.*, **18**, 356–364.
- [8] Heineman, R. H. and Bull, J. J. (2007) Testing optimality with experimental evolution: Lysis time in bacteriophage. *Evolution*, **61**, 1695–1709.
- [9] Chareyron, S. and Alamir, M. (2009) Mixed immunotherapy and chemotherapy of tumors: Feedback design and model updating schemes. *J. Theor. Biol.*, **258**, 444–454.
- [10] Komarova, N. L., Sadvovsky, A. V., and Wan, F. Y. (2008) Selective pressures for and against genetic instability in cancer: a time-dependent problem. *J. R. Soc. Interface*, **5**, 105–121.
- [11] Holland, J. H. (1975) *Adaptation in natural and artificial systems*. University of Michigan.
- [12] Eigen, M. (1971) Self-organization of matter and the evolution of biological macromolecules. *Naturwiss.*, **58**, 465–523.
- [13] Sole, R. V. and Deisboeck, T. S. (2004) An error catastrophe in cancer? *J. Theor. Biol.*, **228**, 47–54.
- [14] Morrison, R. W. (2004) *Designing Evolutionary Algorithms for Dynamic Environments*. Springer.
- [15] Bethke, A. D. (1980) *Genetic algorithms as function optimizers (Doctoral dissertation)*. University of Michigan.
- [16] Goldberg, D. E. (1992) Construction of higher-order deceptive functions using low-order Walsh coefficients. *Annals of Mathematics and Artificial Intelligence*, **5**, 35–48.
- [17] Greaves, M. (2007) Darwinian medicine: a case for cancer. *Nat. Rev. Cancer*, **7**, 213–221.
- [18] Forster, R. and Wilke, C. O. (2005) Tradeoff between short-term and long-term adaptation in a changing environment. *Phys. Rev. E*, **72**, 041922.
- [19] Fidler, I. J. (1978) Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer Res.*, **38**, 2651–2660.
- [20] Klein, C. A., Blankenstein, T. J. F., Schmidt-Kittler, O., Petronio, M., Polzer, B., and Stoecklein, N. H. (2002) Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. *Lancet*, **360**, 683–689.
- [21] Kitano, H. (2003) Tumour tactics. *Nature*, **426**, 125–125.
- [22] Beerenwinkel, N., Antal, T., Dingli, D., Traulsen, A., Kinzler, K. W., Velculescu, V. E., Vogelstein, B., and Nowak, M. A. (2007) Genetic progression and the waiting time to cancer. *PLoS Comput. Biol.*, **3**, e225.
- [23] Wood, L. D., Parsons, D. W., Jones, S., Lin, J., Sjöblom, T., Leary, R. J., Shen, D., Boca, S. M., Barber, T., Ptak, J., Silliman, N., Szabo, S., Dezso, Z., Ustyanksky, V., Nikolskaya, T., Nikolsky, Y., Karchin, R., Wilson, P. A., Kaminker, J. S., Zhang, Z., Croshaw, R., Willis, J., Dawson, D., Shipitsin, M., Willson, J. K. V., Sukumar, S., Polyak, K., Park, B. H., Pethiyagoda, C. L., Pant, P. V. K., Ballinger, D. G., Sparks, A. B., Hartigan, J., Smith, D. R., Suh, E., Papadopoulos, N., Buckhaults, P., Markowitz, S. D., Parmigiani, G., Kinzler, K. W., Velculescu, V. E., and Vogelstein, B. (2007) The genomic landscapes of human breast and colorectal cancers. *Science*, **318**, 1108–1113.
- [24] Gatenby, R. A. and Frieden, B. R. (2004) Information dynamics in carcinogenesis and tumor growth. *Mutation Res.*, **568**, 259–273.
- [25] Beer, P. A., Jones, A. V., Bench, A. J., Goday-Fernandez, A., Boyd, E. M., Vaghela, K. J., Erber, W. N., Odeh, B., Wright, C., McMullin, M. F., Cullis, J., Huntly, B. J. P., Harrison, C. N., Cross, N. C. P., and Green, A. R. (2009) Clonal diversity in the myeloproliferative neoplasms: independent origins of genetically distinct clones. *Brit. J. Haematol.*, **144**, 904–908.
- [26] Taniguchi, K., Okami, J., Kodama, K., Higashiyama, M., and Kato, K. (2008) Intratumor heterogeneity of epidermal growth factor receptor mutations in lung cancer and its correlation to the response to gefitinib. *Cancer Sci.*, **99**, 929–935.
- [27] Maley, C. C., Galipeau, P. C., Finley, J. C., Wongsurawat, V. J., Li, X., Sanchez, C. A., Paulson, T. G., Blount, P. L., Risques, R. A., Rabinovitch, P. S., and Reid, B. J. (2006) Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nature Genet.*, **38**, 468–473.
- [28] Tenailon, O., Toupance, B., Nagard, H. L., Taddei, F., and Godelle, B. (1999) Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. *Genetics*, **152**, 485–493.
- [29] Donaldson-Matasci, M. C., Lachmann, M., and Bergstrom, C. T. (2008) Phenotypic diversity as an adaptation to environmental uncertainty. *Evol. Ecol. Res.*, **10**, 493–515.
- [30] Earl, D. J. and Deem, M. W. (2004) Evolvability is a selectable trait. *Proc. Natl. Acad. Sci. USA*, **101**, 11531–11536.
- [31] Vineis, P. and Berwick, M. (2006) The population dynamics of cancer: a Darwinian perspective. *Int. J. Epidemiol.*, **35**, 1151–1159.
- [32] Li, C.-Y., Little, J. B., Hu, K., Zhang, W., Zhang, L., Dewhirst, M. W., and Huang, Q. (2001) Persistent genetic instability in cancer cells induced by non-dna-damaging stress exposures. *Cancer Res.*, **61**, 428–432.
- [33] Cahill, D. P., Kinzler, K. W., Vogelstein, B., and Lengauer, C. (1999) Genetic instability and darwinian selection in tumours. *Trends Biochem. Sci.*, **24**, M57–M60.
- [34] Maley, C. C., Reid, B. J., and Forrest, S. (2004) Cancer prevention strategies that address the evolutionary dynamics of neoplastic cells: Simulating benign cell boosters and selection for chemosensitivity. *Cancer Epidemiol. Biomark. Prev.*, **13**, 1375–1384.
- [35] Wilke, C. O., Wang, J. L., Ofria, C., and Lenski, R. E. (2001) Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, **412**, 331–333.
- [36] Nilsson, M. and Snoad, N. (2002) Optimal mutation rates in dynamic environments. *Bull. Math. Biol.*, **64**, 1033–1043.
- [37] Tanaka, M. M., Bergstrom, C. T., and Levin, B. R. (2003) The evolution of mutator genes in bacterial populations: The roles of environmental change and timing. *Genetics*, **164**, 843–854.
- [38] Travis, J. M. J. and Travis, E. R. (2002) Mutator dynamics in fluctuating environments. *Proc. R.*

- Soc. Lond. B*, **269**, 591–597.
- [39] Clune, J., Misevic, D., Ofria, C., Lenski, R. E., Elena, S. F., and Sanjuan, R. (2008) Natural selection fails to optimize mutation rates for long-term adaptation on rugged fitness landscapes. *PLoS Comput. Biol.*, **4**, e1000187.
- [40] Brunet, E., Rouzine, I. M., and Wilke, C. O. (2008) The stochastic edge in adaptive evolution. *Genetics*, **179**, 603–620.
- [41] Grefenstette, J. J. (1999) Evolvability in dynamic fitness landscapes: A genetic algorithm approach. *Congress on Evolutionary Computation*, pp. 2031–2038. IEEE.
- [42] Schneider, B. L. and Kulesz-Martin, M. (2004) Destructive cycles: the role of genomic instability and adaptation in carcinogenesis. *Carcinogenesis*, **25**, 2033–2044.
- [43] Folkman, J., Bach, M., Rowe, J. W., Davidoff, F., Lambert, P., Hirsch, C., Goldberg, A., Hiatt, H. H., Glass, J., and Henshaw, E. (1971) Tumor angiogenesis - therapeutic implications. *N. Engl. J. Med.*, **285**, 1182–1186.
- [44] Bergers, G. and Hanahan, D. (2008) Modes of resistance to antiangiogenic therapy. *Nat. Rev. Cancer*, **8**, 592–603.
- [45] Paez-Ribes, M., Allen, E., Hudock, J., Takeda, T., Okuyama, H., Vinals, F., Inoue, M., Bergers, G., Hanahan, D., and Casanovas, O. (2009) Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis. *Cancer Cell*, **15**, 220–231.
- [46] Ebos, J. M. L., Lee, C. R., Cruz-Munoz, W., Bjarnason, G. A., Christensen, J. G., and Kerbel, R. S. (2009) Accelerated metastasis after short-term treatment with a potent inhibitor of tumor angiogenesis. *Cancer Cell*, **15**, 232–239.
- [47] Hegreness, M., Shores, N., Damian, D., Hartl, D., and Kishony, R. (2008) Accelerated evolution of resistance in multidrug environments. *Proc. Nat. Acad. Sci. USA*, **105**, 13977–13981.
- [48] Kim, J. J. and Tannock, I. F. (2005) Repopulation of cancer cells during therapy: An important cause of treatment failure. *Nat. Rev. Cancer*, **5**, 516–525.
- [49] Pepper, J. W., Findlay, C. S., Kassen, R., Spencer, S. L., and Maley, C. C. (2009) Cancer research meets evolutionary biology. *Evolut. Appl.*, **2009**, 62–70.
- [50] Hadamard, J. (1902) Sur les problemes aux derivees partielles et leur signification physique. *Princeton University Bulletin*, **13**, 49–52.
- [51] Sabatier, P. C. (1985) Inverse problems - an introduction. *Inverse Problems*, **1**, UL1.