PROSPECTS FOR THE IMPLEMENTATION OF HEURISTIC ALGORITHMS FROM THE EVOLUTIONARY COMPUTING AS NEW EXPERIMENTAL APPROACHES IN SYNTHETIC BIOLOGY: IN SILICO TESTS

Alexander Spirov[1] and Ekaterina Myasnikova[2]
[1]*The Institute of Scientific Information for Social Sciences RAS, Moscow, Russia*
[2]*Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*

ABSTRACT

In vitro evolution ("evolution in a test tube", directed evolution) is a modern technology used in bioengineering that mimics the process of natural selection to direct proteins or nucleic acids to acquire a specific function. One of the most promising directions in the development of modern technologies for the evolution of macromolecules in vitro is the use of microdroplet devices based on the principles of microfluidics. A significant increase in the efficiency and cost reduction of in vitro evolution approaches can be achieved through the development and analysis of new heuristic evolutionary algorithms. This goal is achieved by numerical tests on computer models of real molecular genetic experiments (in silico experiments). This work was inspired by our recent key observation that many heuristic evolutionary algorithms can be associated with new engineering solutions for the modules of such microfluidic devices. Here, we have proposed and simulated new approaches to in vitro evolution with the prospect of their implementation in experimental microfluidic setups. In this report, we have demonstrated in our in silico experiments that some algorithms that have proven themselves in evolutionary computing turn out to be significantly more efficient than those algorithms that model the routine approaches of modern microdroplet microfluidics. Prospects for the implementation of such promising algorithms in new experimental approaches are discussed.

KEYWORDS

Evolutionary Algorithms, Computational Biology, Selection Techniques, Hill-Climbing, Biomolecular Experiments in Silico, Algorithms Implementation in Bioengineering

1. INTRODUCTION

In vitro evolution ("evolution in a test tube", directed evolution) is a modern bioengineering technology that mimics the process of natural selection in directing proteins or nucleic acids to acquire a desired function (Darmostuk et al., 2015; Vorobyeva et al., 2018; Komarova, Kuznetsov, 2019). One of the most promising directions in the development of modern technologies for the in vitro evolution of macromolecules is the use of microdroplet devices based on the microfluidics principles (Griffiths, Tawfik, 2006; Paegel, Joyce, 2010). A significant increase in the efficiency and cost reduction of in vitro evolution approaches can be achieved through the development and analysis of new heuristic evolutionary algorithms.

Design of any new modern technology should begin with a numerical study of its promising engineering solutions (i.e. Komarova & Kuznetsov, 2019; Voigt et al., 2002), and this is the goal of our article. This approach can significantly increase the efficiency of very, very expensive and resource-intensive experiments.

1.1 Basics and Perspectives of in Vitro Evolution

Macromolecular devices (molecular devices) of modern synthetic biology are able to highly selectively identify (and specifically bind) various target molecules in order to perform certain functions (e.g. Famulok et al., 2007; Trausch, Batey, 2015). For example, they can test a certain chemical agent in the environment

and trigger the fluorescence, thereby quantitatively indicating the presence of the agent in the environment (environmental applications). Another example from biomedicine: such devices can highly-specifically recognize and bind to the specific targets on the surface of certain cancer cells specifically killing them and only them. Molecular devices are synthesized from nucleic acids (DNA and RNA) and proteins.

Molecular devices are typically synthesized experimentally by methods of directed evolution (in vitro evolution; reviewed in Spirov, Myasnikova, 2022a; Spirov, Myasnikova, 2022b). In this approach, an initial "population" is formed from macromolecules, usually of a given length, completely or partially randomized in sequence. Then molecules with a certain (desired) function are selected. The best molecules are copied (the rest are discarded) with mutations being introduced in copies, thereby restoring the population size. In the population of daughter individuals, the best ones are again selected for the desired function, and they again multiply with mutations. Such cycles are repeated until the desired fully functional molecular device is achieved.

Despite the perceived promise of in vitro evolution applications, experimenters are only at the beginning of the road to developing robust, efficient, and cost-effective methods for obtaining desired or required molecular devices.

1.2 Microfluidics for in Vitro Evolution

Experimental microdroplet microfluidics approaches to the directed evolution of biological macromolecules (nucleic acids and proteins) have been extensively developed over the past couple of decades (see, for example, Trausch and Batey, 2015; Price and Paegel, 2016). The fundamental novelty of microdroplet microfluidics methods is that evolutionary search and key experimental procedures are carried out at the level of individual sequences – each sequence in its own microdroplet (e.g. Griffiths, Tawfik, 2006; Paegel, Joyce, 2010; Ryckelynck et al., 2015).

On the other hand, microdroplet microfluidics has its limitations. It is costly and resource intensive and typically manipulates sequence libraries several orders of magnitude less than older approaches. This range of problems brings to the fore the search and testing of new approaches that have the potential of making evolutionary search more efficient (Spirov, Myasnikova, 2022a; Spirov and Myasnikova, 2022b).

1.3 Numerical Analysis of New Prospective Approaches (in Silico Experiments)

Many experimental researchers in the field of in vitro evolution systematically emphasize the importance and promise of preliminary numerical tests of the problems that experimenters have to solve (e.g. Voigt et al., 2002; Komarova & Kuznetsov, 2019; Spirov, Myasnikova, 2022a; Spirov, Myasnikova, 2022b). So, in silico tests should precede costly and resource-intensive real experiments, determining the choice of specific methods, the choice of parameters, as well as an assessment of the chances of success.

1.3.1 Transfer of Algorithms from the Evolutionary Computing

It is important that the field of evolutionary computations is a source of those methods and algorithms that could be used as numerical models of new experimental approaches in microdroplet microfluidics (Spirov, Myasnikova, 2022a; Spirov, Myasnikova, 2022b).

This approach is a clear example of the interdisciplinarity synergy, where the development of biotechnology is impossible without computational support, which, in turn, would not be feasible without the vast experience accumulated in the field of EC in computer science. Moreover, evolutionary algorithms were originally developed as a computational implementation of the evolutionary processes occurring in nature, while nowadays there is an increasing back transfer of ideas from computational theory to synthetic biology and biotechnology. It is this reverse transfer of algorithms from computer science to biology that we are going to implement (O'Hagan, Knowles, Kell, 2012; Eremeev, Spirov, 2021; Myasnikova, Spirov, 2021; Spirov, Myasnikova, 2022a; Spirov, Myasnikova, 2022b).

Our purpose is to study the effectiveness of a number of techniques and algorithms that are developed, well-researched, and well-proven in the field of EC by large-scale numerical experiments. Accordingly, these techniques and algorithms could have prospects for being implemented in experimental approaches to the in vitro evolution of biological macromolecules.

2. NEW ALGORITHMS WITH PROSPECTS FOR IMPLEMENTATION

As noted above, a clear advantage of microdroplet microfluidics is that it provides the possibility to manipulate individual molecules. This is despite the fact that today's standard microfluidics procedures in in vitro evolution are very primitive and, with rare exceptions, have not been even comparable with evolutionary computational approaches. Below we will briefly review those algorithms that may be promising for their implementation in experimental biology.

2.1 Selection in Evolutionary Computing

In EC, a number of selection techniques defining the proportion of individuals transferred from the parent population to the child one has been widely studied on standard tests and on real-life problems (Holland, 1975; Goldberg, 1989). The most well-known of these techniques are listed in Table 1.

Table 1. The best-known selection schemes in EC

Selection in evolutionary computing
Proportionate selection (roulette method). The probability for an individual to be selected is proportional to its fitness value (the value of objective function).
Tournament selection. Initially, a given number of individuals is randomly selected (usually two), and then we pick an individual with the best fitness value.
Rank selection. The selection probability depends on the place in the list of individuals, sorted by the fitness value.
Uniform rank selection is a random choice.
Sigma truncation. The fitness function is linearly transformed and scaled by its standard deviation in the parent population to prevent the premature convergence of evolutionary search.

Moreover, more general breeding strategies (we will use this term) have been widely researched and applied. Selection strategies determine whether the original individuals from the parent population are transferred to the daughter one and according to what rules. We will specifically focus on the elitism strategy.

The elitist strategy is to protect the best individuals in subsequent iterations. In a classical genetic algorithm, the fittest individuals do not always go on to the next generation. An elitist strategy is used to prevent the loss of such an individual. This individual is guaranteed to be included in the new population.

In general, the routine functioning scheme of the microfluidic device for in vitro evolution resembles the evolutionary algorithm (μ,λ) from EC. This selection method will be described in detail below. More complex procedures and algorithms (with rare exceptions) have not been considered in experimental biology.

2.2 Hill-Climbing Algorithms

Among hill-climbing algorithms, a simple random-mutation hill-climbing (RMHC) turned out to be one of the most effective for a number of benchmark problems as well as those of real-life (Mitchell, 1998).

2.2.1 The Random-Mutation Hill-Climbing (RMHC) Algorithm

This algorithm, proposed back in the early 90s, can be interpreted as a simplified version of the annealing method (namely, the Metropolis method at zero temperature) (Forrest, Mitchell, 1993; Mitchell, 1998). The procedure is aimed at maximization of objective (fitness) function and includes the following steps:

1. Choose a sequence of a given length at random.
2. Mutate each sequence position with the given probability.
3. If mutation results in an equal or higher fitness, then the current sequence is re-placed by the mutated one.
4. Go to step 2.
5. Repeat 2-4 until the optimal value of the fitness function is reached. Return the cur-rent sequence.

We note here that despite it is easy to implement a simple hill-climber on a computer, its experimental implementation using microdroplet microfluidics seems to be very cumbersome and inefficient. However, microfluidics may enable the implementation of parallel versions of hill-climbing algorithms.

2.2.2 Parallel Version of RMHC

In the parallel version, we simultaneously run N (up to several hundred) RMHC processes in parallel (with the same population size in each RMHC). At the same time, as soon as a more efficient solution is found in any of the populations, this finding is immediately copied to all the rest populations. It is in this sense that we consider each of these hundreds of RMHC procedures to be (semi-)autonomous: they are parallel and independent until one of the RMHC procedures finds a better solution.

2.3 Crossover Algorithms

In EC, crossover (recombination) has always been given great importance, so it is quite common to expect that these procedures are critical to evolutionary search and typically should enhance the role of mutations (see Holland, 1975). Therefore, in this article, we also systematically test the role of crossover in evolutionary search for our test problems.

It should also be noted that in experimental biology there are also very high expectations regarding the role of the crossover in in vitro evolution (Stemmer, 1994; Voigt et al., 2002).

3. NUMERICAL ANALYSIS OF NEW PROSPECTIVE APPROACHES (IN SILICO EXPERIMENTS)

In this section, we numerically explore those new evolutionary search schemes inspired by known EC heuristic algorithms that could be performed by new microfluidic devices.

For our numerical studies, we offer our original benchmark tests inspired by real-life problems of in vitro evolution.

3.1 New Real-Life Problems Require New Benchmark Tests

For this project, we have developed a new version of our test problems as a modification and extension of our original BioRS fitness functions (formulated in our publications Myasnikova, Spirov, 2021; Spirov, Myasnikova, 2022a).

3.1.1 Test Problem with Several Local Extrema

At first step, we generalize a single-domain (instead of multi-domain) version of BioRS (Myasnikova, Spirov, 2021; Spirov, Myasnikova, 2022a), which is based on the consensus of a specific aptamer to adenosine and some of its derivatives, known as the Sasanfar-Szostak aptamer (Sassanfar and Szostak, 1993). To make the problem realistic enough for this project (to get as close as possible to real life problems), we modify the search space in the following way. In addition to the Sasanfar-Szostak target consensus, the algorithm is able to find one (or several) other consensuses with sufficiently high levels of fitness. They represent local extrema in sequence space, but are not the goal of our numerical evolutionary experiments. The presence of several (or even many) local extrema, where the evolutionary search tends to get stuck, is a well-known problem in EC. This problem is also recognized when analyzing the results of real molecular biological experiments on the evolution of macromolecules in vitro. Specifically, in the case of aptamer sequences to adenosine and close sequences, many sequences have been obtained and characterized, differing not only in minor substitutions (and/or insertions, deletions, indels), but also qualitatively different (reviewed in Spirov, Myasnikova, 2022a).

As in the case of the Royal Road functions (Forrest, Mitchell, 1993; Mitchell, 1998), the fitness function is given by a series of sequences in the four-letter alphabet of nucleic acids. As a global extremum, we are looking for a consensus with five uniquely defined positions (out of 26 in total), nine two-valued positions, and three complementary pairs (the remaining six positions are arbitrary). At the same time, it is precisely due to the two-valued positions that the growth of the fitness level is ensured: one of the two values in each case gives an additional increase (delta) when approaching the consensus. Details are shown in Figure 1.

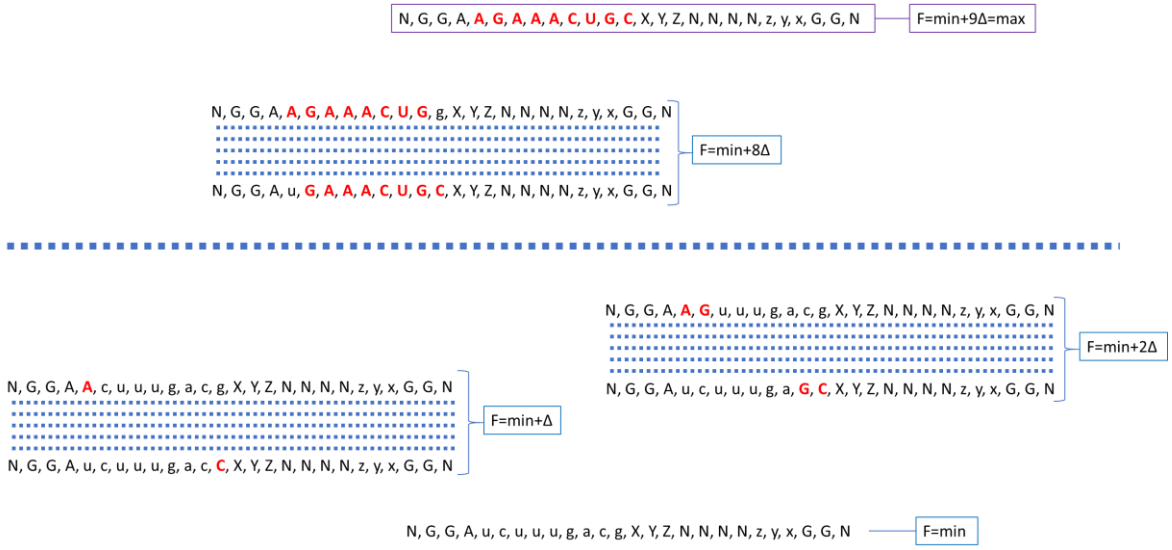

Figure 1. A family of sequences (26 characters long) in the RNA alphabet (A,T,G,U) that have a non-zero fitness level in our test problem. The family naturally breaks down into 10 groups. This is the lowest fitness sequence (Fitness=min), then eight groups with fitness levels from F=min+Δ to F=min+8Δ, and, finally, the sequence (consensus sequence) with the maximum fitness F=min+9Δ. N is arbitrary letter; X-x, Y-y, Z-z are complementary pairs. See text for further details

Further, local extrema (one or more) are characterized by only two complementary pairs, which provides a slightly higher efficiency of finding them compared to the consensus of the global extremum. Additionally, by definition, the global extremum consensus has a slightly higher level of (maximum) fitness (see Figure 1).

In arbitrary units of our tests, the finding of the minimally required Sasanfar-Szostak aptamer sequence yields 100 units (minimal fitness; Figure 1). Finding each preferred position from a set of nine two-valued positions yields an increment of delta each time (Δ =8). So, the global maximum has a fitness of 172. For local maxima, formally similar, we set Δ =4, so that one local extremum ends up with a fitness of 140 and another one of 136. All of our tests below are done on this particular version of our problem.

3.2 (μ,λ) Selection as Control in Numerical Experiments

Literature analysis in the field of in vitro evolution by means of microdroplet microfluidics leads us to the conclusion that selection above a certain fitness threshold level used as a control is sufficient formalization of a routine real experiment (compare Oh, Lee, McKay, 2011; Lee et al., 2011; Spirov, Myasnikova, 2022a). Typically, this threshold is quite low. In this case, the selected parental individuals are copied so as to achieve the required size of the child population. In copying a parent into the daughter population point mutations are introduced into the parent sequence. This is in line with experimental techniques used in in vitro evolution. Typically, this is the multiplication of nucleic acid molecules by an error-prone PCR (Lee et al., 2011; Spirov, Myasnikova, 2022a). At the same time, in the first approximation, it is reasonable to assume that all selected individuals are copied in the same proportion (they give approximately, on average, the same number of descendants to the daughter population, regardless of their fitness level). Further, the process of multiplying of a selected part of individuals by error-prone polymerase chain reaction (until a given population size is restored) can rather be interpreted as selection without elitism, because multiple copying of an original parental molecule is likely to result in mutating all copies of the original parental sequence.

Therefore, in the EC paradigm, the closest thing to real microfluidic experiments on in vitro evolution will be the (μ,λ) selection scheme without elitism. Three parameters of control tests are critical in terms of performance: mutation rate, μ/λ threshold (stringency of selection), and population size. It is essential that the efficiency depends on all three parameters, and when any of them changes, the rest must be selected for the highest search efficiency.

3.2.1 Effect of Mutation Rate in Control Tests

The level of mutagenesis is certainly one of the key factors both in EC and in vitro evolution. As it turned out, for our test problem, efficiency is the most sensitive to the mutation rate, while search speed is less dependent. As the mutation rate increases, the efficiency rapidly increases, reaching a maximum at Mmut=7, and then decreases. At the mutation rate Mmut=7, the efficiency reaches 68%. The speed of evolutionary search is gradually decreasing. The result is presented in Table 2.

Table 2. The efficiency of evolutionary search grows rapidly with the increase in the mutation rate, and then rapidly decreases; while the rate of evolutionary search decreases with increasing mutation rate (averaged over 400 runs; population size 4000, μ/λ 0.25)

Mmut	Average	StDev	Efficacy
3	197794.11	98192.81	47
5	193524.76	87847.11	58.25
7	206645.05	86566.50	68.25
9	229069.70	83337.71	60.25
11	273755.92	77253.41	32

Accordingly, based on the results of Table 2, we will take MMut=7 in subsequent tests (unless otherwise specified).

3.2.2 Effect of Parameter μ/λ Value (Proportion of Parent Population) in Control Tests

The μ/λ parameter is the key to this selection scheme and it determines the rigorousness of the selection. The search efficiency is also quite sensitive to the μ/λ parameter, comparable to the sensitivity to the mutation rate. The results are shown in Table 3.

Table 3. Effect of parameter μ/λ value on the evolutionary search (averaged over 400 runs)

Quote (μ/λ)	Average	StDev	Efficacy
5	149739.55	102762.22	70.5
10	145510.05	98499.62	66.5
25	206645.05	86566.50	68.25
40	212066	57795.55	2.5

We will typically use μ/λ = 25 in subsequent tests for this problem (unless otherwise noted). We take such a fairly high quota in order to avoid strict selection (it can be seen from the table that with a lower quota (5 - 10%), evolutionary search is better for this our control).

3.2.3 Effect of Population Size in Control Tests

We have found out that a population of a few thousand individuals is already sufficient for a successful evolutionary search for our test problems. The results of the runs are shown in Table 4. It can be seen that in the range of 1 - 6 thousand, the efficiency is kept at about 70%. while the search rate increases with the population size.

Table 4. Effect of the population size on the evolutionary search (averaged over 400 runs)

Popul	Average	StDev	Efficacy
1000	181957.71	126089.09	71.25
2000	188256.92	114757.72	67.75
4000	206645.05	86566.50	68.25
6000	265339.69	107695.56	69.5

We will usually use Popul-4000 in subsequent tests for this problem (unless otherwise noted).

Thus, in the series of runs, we eventually established the ranks of the values of the main parameters at which our tests in control are sufficiently effective to be able to perform a large series of various runs in a reasonable time on a workstation (usually several hours on a DELL Precision M6700). It is with these controls that we will compare all other numerical experiments (with different selection schemes).

3.3 Sigma Truncation vs (μ,λ) Selection (without Crossover)

Sigma truncation (Table 1) refers to selection schemes aimed at preventing premature search convergence to local extrema. In general, it can be interpreted as one of the extreme variants of proportional selection, when individuals with greater fitness leave more offspring. But unlike proportional selection, with sigma truncation, even the most successful parents leave no more than few (usually, two) offspring.

Our tests have shown that sigma selection for our test problem compared to (μ,λ) control speeds up evolutionary search. The speed almost doubles (from 207k to 117k), but the efficiency decreases slightly (from 68 to 64%), as Table 5 illustrates.

Table 5. Sigma selection speeds up the evolutionary search but does not increase its efficiency (averaged over 400 runs)

Mmut	Average	StDev	Efficacy
5	129941,727272727	100482,81351216	55
7	130137,358974359	100002,344373163	58,5
9	128855,558951965	103550,825865692	57,25
11	117341,53125	107049,314391733	64
13	109978,368	99123,6004949845	62,5

This result is somewhat discouraging, if we take into account how much attention in the theory and practice of EC is paid to the problems of premature convergence of solutions and methods to avoid this. And it is sigma-truncation that is considered effective for overcoming such search difficulties. Below we will see that for these purposes it is also necessary to add a crossover to the sigma truncation.

3.4 Tournament Selection Is Ineffective in Our Tests

Our tests have shown that for our problem (and in the parameter ranges considered in this publication), tournament selection is practically inefficient. We tested tournament teams of 2 and 4 individuals, intermediate parent population sizes (obtained as a result of tournament selection) from 10 to 35%, and mutation rates from 1 to 22.

3.5 Parallel RMHC Is the Best

The simple RMHC algorithm classified as (1+1)EA (discussed above) turns out to be one of the most efficient and fastest algorithms for a number of real life and test problems. In the tests with our benchmark problems, we also made sure of the impressive efficiency of the parallel version of this algorithm.

This version of the parallel RMHC with explicit parallelization into many (partially) autonomous (1+1)EA processes (as described above) shows both a clear evolutionary search acceleration (from 207k to 110 k) and an increase in search efficiency (from 68% to 86%), when parallelized into several hundred (semi) autonomous processes. These results are illustrated in Table 6.

Table 6. Parallel version of RMHC is faster and more efficient than standard (μ,λ) for our test problem in the series of tests with the increase of (semi)autonomous processes (averaged over 400 runs)

PopNmb	Average	StDev	Efficacy
100	112857.36	104658.83	64.5
200	113591.66	99321.62	84
400	124260.46	106594.49	86
600	110720.93	99237.41	86

Thus, the parallel RMHC turns out to be the fastest and most efficient evolutionary search procedure for our test problems. Therefore, we are ready to recommend it in the first place for implementation in real microfluidic devices (see the section 4).

3.6 Elitism Improves the (μ,λ) Selection

Elitism is applicable to all selection schemes discussed in this publication. As we will see, in several cases it demonstrates its efficiency for our test problems.

So, elitism for (μ,λ) selection increases both the speed of evolutionary search and efficiency. The speed increases by about one and a half times (162k versus 207k), while the efficiency increases from 68% to 77%, as illustrated by Table 7.

Table 7. Elitism in the (μ,λ) selection increases both the speed of evolutionary search and its efficiency (averaged over 400 runs)

QuotElit	Average	StDev	Efficacy
20	161095.66	83746.31	71.25
25	162164.10	86163.91	77
50	152675.13	74108.08	70.25
75	148118.99	63452.31	56

The fact that elitism supports a simple (μ,λ) selection scheme suggests that this technique should be implemented in in vitro evolution. In other words, it is possible to propose procedures for transferring the best part of the parent population to the child population into the schemes of real biological experiments. The implementation of elitism in the approaches of real microfluidic experiments requires some modifications of the experimental schemes (and/or microfluidic device), and we will discuss this in the section 4.

3.7 Crossover Enhances Evolutionary Search

As might be expected from the structure of our benchmark test (obvious three-modality of the search space), the crossover enhances evolutionary search.

3.7.1 Crossover in (μ,λ) Selection

For our test problem with (μ,λ) selection, we used a scheme for selecting both parents from a quota of μ best individuals for the crossover. Our tests have shown that a standard single-point crossover improves the evolutionary search efficiency of our problem from 68% to 81%. The results are shown in Table 8. At the same time, the search speed for the runs with the best efficiencies does not practically change as compared to the control.

Table 8. Single-point crossover significantly increases the efficiency of evolutionary search in (μ,λ) selection but insufficiently increases the speed (averaged over 400 runs)

Cross(%%)	Average	StDev	Efficacy
0.5	211447.36	138562.66	80.5
1.0	195252.07	131229.29	80.75
2.5	221347.47	154077.49	80.75
5.0	200784.56	125076.15	78.5

Thus, even the simplest single-point crossover with a simple selection scheme (μ,λ) reliably increases the selection efficiency. Thus, our tests confirm well-known expectations about the role of crossover in the performance of evolutionary search. We made sure that the crossover increases the chances of a “jump” from a local maximum to a global one. Therefore, we are ready to strongly recommend that the experimental procedures of in vitro evolution techniques be supplemented with real experimental crossover methods (along with the point mutation procedures routinely used there).

3.7.2 Crossover in Sigma Truncation Selection

As we have seen, the single-point crossover favors evolutionary search in the case of sigma truncation. Efficiency increases from 68% to 78%, speed also increases slightly (158k vs 201k) as shown in Table 9.

Table 9. Single-point crossover in sigma selection significantly increases the efficiency of evolutionary search for our benchmark-test and increases the speed (averaged over 400 runs)

Cross(%%)	Average	StDev	Efficacy
0.1	158158.63	149907.30	77.75
5.0	177701.15	164113.85	73
10.0	189025.01	167849.72	67

Thus, as might be expected, the crossover improves the efficiency and speed of the search in sigma selection. However, for our tests, sigma selection does not end up being better (or not much better) than (μ,λ) selection if the mutation procedure is supplemented by a crossover in both cases (compare Table 8 and Table 9).

4. CONCLUSIONS

Our tests with various selection schemes from section "Selection Procedures in Evolutionary Computing" showed that, at least for our test problem, only hill-climbing turned out to be better than control runs. Elitism can also positively influence selection. The role of the crossover is also very significant.

A microfluidic modular device with new modules with new features, briefly discussed above, would allow the implementation of a whole series of heuristic algorithms, at least some of which demonstrate in numerical experiments reproducible efficiency in comparison with routine schemes of microfluidic devices for in vitro evolution.

In addition to the promising heuristic algorithms that we have considered in this paper, there are many more algorithms that are promising for implementation in in vitro evolutionary approaches.

Among these techniques there will be various modifications of genetic algorithms, probabilistic models, methods of multi-objective optimization (e.g. Mühlenbein & Mahnig, 2001; O'Hagan, Knowles, Kell, 2012; Verma et al., 2013).

At last, it is natural to expect that the next generation algorithms which were for a long time intentionally developed for various optimization problems, may prove to be significantly more efficient than standard GAs. Accordingly, their implementation in experimental procedures of in vitro evolution can be very effective and significantly resource-saving. It is in this direction that we plan the further research.

ACKNOWLEDGEMENT

The study is supported by the Russian Science Foundation grant 22-18-00383.

REFERENCES

Darmostuk, M. et al, (2015). Current approaches in SELEX: An update to aptamer selection technology. Biotechnol Adv., Vol. 33, Issue 6, Part 2, pp. 1141-1161.

Eremeev, A.V. and Spirov, A.V., (2021). Modeling SELEX for regulatory regions using Royal Road and Royal Staircase fitness functions. Biosystems, Vol. 200, pp. 104312.

Famulok, M. et al, (2007). Functional Aptamers and Aptazymes in Biotechnology, Diagnostics, and Therapy. Chemical Reviews, Vol. 107, No. 9, pp. 3715–3743.

Forrest, S. and Mitchell, M., (1993). Relative building-block fitness and the building block hypothesis, In D. Whit-ley (ed.), Foundations of Genetic Algorithms, Vol. 2: pp. 109-126, San Mateo, CA Morgan Kaufmann.

Goldberg, D.E., (1989). Genetic Algorithms in Search, Optimization and Machine Learning. Reading, MA: Addison-Wesley.

Griffiths, A.D. and Tawfik, D.S., (2006). Miniaturising the laboratory in emulsion droplets. Trends Biotechnol., Vol. 24, No. 9, pp. 395-402.

Griffiths, A.D. and Tawfik, D.S., (2006). Miniaturising the laboratory in emulsion droplets. Trends Biotechnol. Vol. 24, No. 9, pp. 395-402.

Holland, J.H., (1975). Adaptation in Natural and Artificial Systems. Ann Arbor, MI: University of Michigan Press.

Komarova, N. and Kuznetsov, A., (2019). Inside the Black Box: What Makes SELEX Better? Molecules, Vol. 24, No 19, pp. 3598.

Lee, Y.G. et al, (2011). Investigating vesicular selection: A selection operator in in vitro evolution. Applied Soft Computing, Vol. 11, Issue 8, pp. 5528 – 5550.

Mitchell, M., (1998). An Introduction to Genetic Algorithms. Cambridge, Massachusetts: MIT Press.

Myasnikova, E. and Spirov, A., (2021). Extensions of Royal Road Functions for Use in SELEX in silico. Proceedings of 17th International Asian School-Seminar Optimization Problems of Complex Systems (OPCS), pp. 60-64.

Mühlenbein, H. and Mahnig, T., (2001). Evolutionary Algorithms: From Recombination to Search Distributions. In: Kallel L., Naudts B., Rogers A. (eds) Theoretical Aspects of Evolutionary Computing. Natural Computing Series, Springer, Berlin, Heidelberg.

O'Hagan, S. et al, (2012). Exploiting Genomic Knowledge in Optimising Molecular Breeding Programmes: Algorithms from Evolutionary Computing. PLoS ONE, 7, No. 11, pp. e48862.

Oh, I.S. et al, (2011). Simulating chemical evolution. Proceedings of 2011 IEEE Congress of Evolutionary Computation, New Orleans, LA, pp. 2717–2724.

Paegel, B.M. and Joyce G.F., (2010). Microfluidic compartmentalized directed evolution. Chem Biol., Vol. 17, No. 7, pp. 717-724.

Price, A.K. and Paegel, B.M., (2016). Discovery in Droplets. Anal Chem., Vol. 88, No. 1, pp. 339-53.

Sassanfar, M. and Szostak, J.W., (1993). An RNA motif that binds ATP. Nature, Vol. 364, pp. 550–553.

Stemmer, W.P., (1994). Rapid evolution of a protein in vitro by DNA shuffling. Nature, Vol. 370, pp. 389-391.

Spirov, A. and Myasnikova, E., (2022a). Heuristic algorithms in Evolutionary Computation and modular organization of biological macromolecules: applications to in vitro evolution. PLoS ONE, Vol. 17, No. 1, pp. e0260497.

Spirov, A and Myasnikova, E., (2022b). Problem of Domain/Building Block Preservation in the Evolution of Biological Macromolecules and Evolutionary Computation, IEEE/ACM Trans Comput Biol Bioinform., doi: 10.1109/TCBB.2022.3175908.

Trausch, J.J. and Batey, R.T., (2015). Design of modular "plug-and-play" expression platforms derived from natural riboswitches for engineering novel genetically encodable RNA regulatory devices. Methods Enzymol, Vol. 550, pp. 41-71.

Verma, D. et al, (2019). Pareto Optimization of Combinatorial Mutagenesis Libraries, IEEE/ACM Trans Comput Biol Bioinform., Vol. 16, No. 4, pp. 1143-1153.

Voigt, C.A. et al, (2002). Protein building blocks preserved by recombination. Nature Structural Biology, Vol. 9, pp. 553-558.

Vorobyeva, M.A. et al, (2018). Key Aspects of Nucleic Acid Library Design for in Vitro Selection. Int J Mol Sci., Vol. 19, No. 2, pp. 470.