Genetic Algorithm Against Cancer*

F. Pappalardo[1,2], E. Mastriani[2], P.-L. Lollini[3], and S. Motta[2]

[1] Faculty of Pharmacy, University of Catania,
V.le A. Doria, 6, I-95125 Catania, Italy
Tel.: +39 095 7384071; Fax.: +39 095 7384278
[2] Department of Mathematics and Computer Science, University of Catania,
V.le A. Doria, 6, I-95125 Catania, Italy
Tel.: +39 0957383073; Fax.: +39 0957337046
francesco@dmi.unict.it, mastriani@dmi.unict.it, motta@dmi.unict.it
[3] Sezione di Cancerologia, Dipartimento di Patologia Sperimentale
and Centro Interdipartimentale di Ricerche sul Cancro "Giorgio Prodi",
University of Bologna,
Viale Filopanti 22, I-40126 Bologna, Italy
Tel.: +39 051241110; Fax.: +39 051242169
pierluigi.lollini@unibo.it

Abstract. We present an evolutionary approach to the search for effective vaccination schedules using mathematical computerized model as a fitness evaluator. Our study is based on our previous model that simulates the Cancer - Immune System competition activated by a tumor vaccine. The model reproduces pre-clinical results obtained for an immunoprevention cancer vaccine (Triplex) for mammary carcinoma on HER-2/neu mice. A complete prevention of mammary carcinoma was obtained *in vivo* using a Chronic vaccination schedule. Our genetic algorithm found complete immunoprevention with a much lighter vaccination schedule. The number of injections required is roughly one third of those used in Chronic schedule.

1 Introduction

Immunoprevention of mammary carcinoma in HER-2/neu transgenic mice was attempted using various immunological strategies, including cytokines, non-specific stimulators of the immune response, and HER-2/neu specific vaccines made of DNA, proteins, peptides, or whole cells. Most approaches achieved a delay of mammary carcinogenesis, but a complete prevention of tumor onset was not attained, particularly in the most aggressive tumor models.

A new vaccine, called Triplex, was proposed in [6]. The vaccine combined three different stimuli for the immune system. The first was p185neu, protein

* F.P. and S.M. acknowledge partial support from University of Catania research grant and MIUR (PRIN 2004: *Problemi matematici delle teorie cinetiche*). This work has been done while F.P. is research fellow of the Faculty of Pharmacy of Universiy of Catania. P.-L.L. acknowledges financial support from the University of Bologna, the Department of Experimental Pathology ("Pallotti" fund) and MIUR.

I. Bloch, A. Petrosino, and A.G.B. Tettamanzi (Eds.): WILF 2005, LNAI 3849, pp. 223–228, 2006.

product of HER-2/neu, which in this system is at same time the oncogene driving carcinogenesis and the target antigen. p185neu was combined with the two non specific adjuvants, allogeneic class I major histocompatibility complex (MHCI) glycoproteins and interleukin 12 (IL-12). Allogeneic MHCI molecules stimulate a relatively large fraction of all T cell clones, up to 10% of the available repertoire. IL-12 is a cytokine normally produced by antigen presenting cells (APC) such as dendritic cells (DC) to stimulate T helper cells and other cells of the immune system, such as natural killer cells (NK) [2].

A complete prevention of mammary carcinogenesis with the Triplex vaccine was obtained when vaccination cycles started at 6 weeks of age and continued for the entire duration of the experiment, at least one year (chronic vaccination). One vaccination cycle consisted of four intraperitoneal administrations of non-replicating (mitomycin-treated) vaccine cells over two weeks followed by two weeks of rest [4]. Other tested vaccination schedules were unable to prevent mammary carcinogenesis. Triplex is an immunoprevention vaccine: *in vivo* experiments have shown that the vaccine is no more effective after the solid tumor formation.

The question whether the chronic protocol is the minimal vaccination protocol yielding complete protection from tumor onset, or whether a lower number of vaccination cycles would provide a similar degree of protection, is still an open question.

Finding an answer to this question via a biological solution would be too expensive in time and money as it would require an enormous number of experiments, each lasting at least one year.

For this reason we developed an accurate model [9] of immune system responses to vaccination. We performed *in silico* experiments considering a large population of individual mice. Each individual mouse is characterized by a sequence of uniformly distributed random numbers which will determine the probabilistic events. As showed in [9], comparison with *in vivo* experiments shows excellent agreement.

Our evolutionary approach uses the model and its computer implementation described above as a fitness evaluator to find a schedule which controls the growth of cancer cells by a minimal number of vaccine injections.

The paper is organized as follows. In section 2, we describe in depth our algorithm. Section 3 will provide computational results and, in Section 4, we will give conclusions and final remarks.

2 The Genetic Algorithm

Evolutionary Algorithms have been applied with satisfactory results to a very long list of hard combinatorial problems. A complete description or enumeration of such results is, per se, a hard problem. The interested reader can found extensively review in [3, 7].

The approach we present in this paper differs from "standard" GA applications as we use a *simulator* to compute a parameter of the fitness function. To

our best knowledge, very few applications [1] use a complete simulator in a genetic algorithm and no applications of this type exist in cancer immunoprevention.

Each chromosome in the population represents a vaccine schedule. It is a binary string of 1200 bits, in which each gene represents a timestep in which is possible to inoculate a vaccine injection. One timestep is equal to 8 hours of mouse's life. If the i-th gene is expressed (respectively not expressed), i.e. the i-th bit is set to 1 (respectively the i-th bit is set to 0), then a vaccination has to be administered at timestep i (respectively no vaccination has to be administered at timestep i). We decide to set a population formed by 40 chromosomes.

The selection operator used is *tournament selection* [5]. Reproduction uses uniform crossover.

Mutation acts as follows: first a gene subject to mutation is chosen with probability $p = 1/g_n$, where g_n represents the number of chromosome's genes; second the chosen gene is setted to 1 with probability $p = 1/p_s$ and to 0 with probability $p = 1 - (1/p_s)$, where p_s represents the population size.

Elitism is used on two specific elements of the population: *i)* best fitness element; *ii)* second best fitness element.

Results showed in [9], fully justify the usage of the SimTriplex simulator for the definition of a good fitness function. SimTriplex simulator computes the main biological entities of the cancer - immune system competition. If the number of cancer cells overcomes 10^5, then the simulator recognizes the solid tumor formation (carcinogenesis) and simulation ends at the reached time. We will refer to this time as mice survival time. An effective vaccination must reach a mice survival time of 1200 timesteps, i.e. 400 days.

In setting up a fitness function we must take into account two fundamental and competing requirements: *i)* any schedule must be an effective one, i.e. the mice survival time must reach 400 days; *ii)* the best schedules must have a minimal cardinality, i.e. they must provide mice survival with the minimum number of vaccine injections.

Any evolutionary approach which just takes into consideration the first requirement, will produce populations of individuals *very rich* of ones, thus, not minimal. If instead, we take into consideration just the second requirement, we will have populations of individuals full of zeroes, and thus very likely we will obtain a non effective schedule. Any fitness function therefore must be, at least, a two-variables function of type $f(n, s, \ldots)$ where n is the number of injections, and s is the number of timesteps survived by the mouse. Also, f must satisfy the following two properties:

$$f(n, s, \cdots) < f(n, s', \cdots) \text{ if } s > s' \tag{1}$$

$$f(n, s, \cdots) > f(n', s, \cdots) \text{ if } n > n' \tag{2}$$

We restrict ourself the simple case of a two variable function and we chose the following function:

[1] The only one we found is reported in http://www.cs.ucl.ac.uk/staff/P.Bentley/WLBEC1.pdf

$$f(n, s) = \frac{n^2}{s} \tag{3}$$

which meets the properties (1) and (2). Obviously, the fitness function (3) has to be minimized.

Determining the quality (fitness) of each chromosome involves the use of the SimTriplex simulator which takes a non negligible amount of time. In particular, the simulator returns the timesteps survived by the mouse with the proposed therapy coded by the chromosome.

3 Results

In setting the computer experiments, we randomly chose 10 virtual mice over the 100 of the first sample set we used in [9]. Each run of the GA took about 36 hours on a 686-class PC machine. All the 10 virtual mice gave similar, but obviously not identical, results. They all got complete prevention of mammary carcinoma with 19 vaccine injections, against the 59 required by Chronic vaccination schedule. Figure 1 shows the effect of Chronic vaccination schedule for one of the chosen mice. For the same mouse the GA suggested schedule is shown in figure 2.

This shows, as in Chronic vaccination schedule, an initial burst of cancer cells. This effect is due to the time-lag of vaccine effect [9]. Chronic schedule then controls the growth of cancer cells with regular successive vaccine injections. Our GA was not required to do this, so a second burst of cancer cells appears in GA proposed schedule. This is a reasonable minimum from the point of the view of the GA. One could argue if this is also reasonable for mice safety. In this specific case we notice that both the cancer cells maximum in the GA suggested

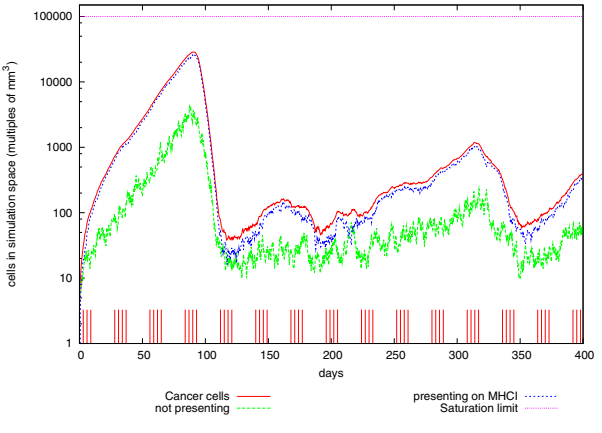

Fig. 1. Cancer cells dynamics. Cancer cells controlled by CHRONIC vaccination schedules. Red ticks above x axis represent the timing of vaccine administration. Saturation limit indicates that a solid tumor is formed, i.e. the number of cancer cells in the simulated space becomes greater than 10^5.

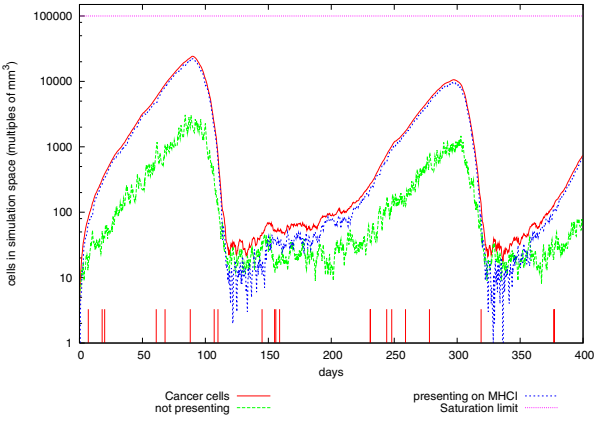

Fig. 2. Cancer cells dynamics. Cancer cells controlled by genetic algorithm proposed vaccination schedules. Red ticks above x axis represent the timing of vaccine administration. Saturation limit indicates that a solid tumor is formed, i.e. the number of cancer cells in the simulated space becomes greater than 10^5.

schedule are lower than the Chronic schedule maximum. So in this case the GA suggested schedule is safer than the Chronic one. However this could not be the case for other mice. This suggests that more requirements should be added to the GA fitness function. Those requirements must obviously be biologically driven.

4 Conclusion and Future Work

We have presented an evolutionary algorithm which turns out to be efficient in finding effetctive therapies for protecting virtual mice from mammary carcinoma.

One of the major novelties of our algorithm, is the usage of a *simulator* to compute a parameter of the fitness function. To our best knowledge, very few applications use a complete simulator in a genetic algorithm and no applications of this type exist in cancer immunoprevention.

Future work will see a GRID/parallel implementation of our GA that can be used to perform a large number of simulated experiments which suggest key *in vivo* experiments on a small set of schedules which take into account biological or clinical constraints.

References

1. Bäck, T., Eiben A., Vink, M.: A superior evolutionary algorithm for 3-SAT. Proceedings of the 7th Annual Conference on Evolutionary Programming, Lecture Notes in Computer Science, Springer (1998) 125-136
2. Colombo, MP., Trinchieri, G.: Interleukin-12 in anti-tumor immunity and immunotherapy. Cytokine Growth Factor Rev., **13:2** (2002) 155-168

3. Corne, D., Dorigo, M., Glover, F.: New Ideas in Optimization. McGraw-Hill, Advanced Topics in Computer Science (1999)
4. De Giovanni, C., Nicoletti, G., Landuzzi, L., Astolfi, A., Croci, S., Comes, A., Ferrini, S., Meazza, R., Iezzi, M., Di Carlo, E., Musiani, P., Cavallo,F., Nanni, P., Lollini, P.-L.: Immunoprevention of HER-2/neu transgenic mammary carcinoma through an interleukin 12-engineered allogeneic cell vaccine. Cancer Res., **64:11** (2004) 4001-4009.
5. Goldberg, D.E.: A comparative analysis of selection schemes used in genetic algorithms. Foundations of Genetic Algorithms, Morgan Kaufmann, G. Rawlins eds. (1991)
6. Lollini, P.-L., De Giovanni, C., Pannellini, T., Cavallo, F., Forni, G., Nanni, P.: Cancer immunoprevention. Future Oncology, **1:1** (2005) 57-66
7. Mitchell, M.: An Introduction to Genetic Algorithm. The MIT Press (1996)
8. Motta, S., Lollini, P.-L., Castiglione, F., Pappalardo, F.: Modelling Vaccination Schedules for a Cancer Immunoprevention Vaccine. Immunome Research, submitted (2005)
9. Pappalardo, F., Lollini, P.-L., Castiglione, F., Motta, S.: Modelling and simulation of cancer immunoprevention vaccine. Bioinformatics, **21:12**, (2005) 2891-2897