Patterns of Cluster Formation and Evolutionary Activity in Evolving L-Systems

Jan T. Kim and Kurt Stüber

Max-Planck-Institut für Züchtungsforschung
Carl-von-Linné-Weg 10
5000 Köln 30
Germany

Email: kim@vax.mpiz-koeln.mpg.dbp.de
stueber@vax.mpiz-koeln.mpg.dbp.de

Abstract.

We introduce the simulator LindEvol, in which L-systems evolve under the regime of a genetic algorithm. The L-systems control the growth of individual plants that inhabit a world they share. The fitness of an individual depends largely on interaction and competition with others, hence intrinsic adaptation [Packard, 1989] takes place. Many emergent phenomena can be observed in LindEvol simulations. In this paper we focus on the evolution of structures in the population, which is investigated by distance distribution visualization.

Complex wave patterns are found to emerge in distance distributions during LindEvol runs. Control experiments show that the formation of such patterns is not special to the LindEvol fitness function. The interplay between mutation and selection in a genetic algorithm can give rise to complex structures in distance distributions under the regime of various fitness functions.

We also measured the evolutionary activity in LindEvol runs, as described in [Bedau and Packard, 1992]. Evolutionary steps, which can readily be observed in the fitness curves, are found to be accompanied in many cases by the formation of new waves of evolutionary activity, or by changes in the pattern of the waves.

1. The LindEvol model.

1.1 SIMULATION OF PLANT GROWTH. Plant growth takes place in a two-dimensional world that is organized as a lattice with the topology of a cylinder surface (i. e. there is a "bottom" and a "ceiling" but no left nor right border).

A plant consists of an ensemble of contiguous cells. A cell occupies exactly one square, a square can hold only one cell. A cell can have two states, it is either energy rich or energyless.

The simulation of a generation (a "year") consists of a couple of "days". Each day starts with the simulation of light by "photons". One photon travels down each vertical row of lattice sites in each day. If it encounters a cell, it is absorbed with a probability of 50%, causing the cell to assume the energy rich state. Energy rich cells can absorb photons too, in this case the photon disappears without any effect. If the photon is not absorbed it moves to the square below and the state of the cell remains unchanged. Photons also disappear if they fall below the bottom of the world.

After the simulation of light the genome-controlled plant growth takes place. Each energy rich cell may divide into two energyless cells. The genome consists of a set of rules that determine whether division takes place depending on the local structure of the plant in the nine cell neighborhood of the energy-rich cell (see fig. 1). The sites in the nine cell neighborhood of a cell C are assigned numbers i according to the following scheme:

5 6 7
3 C 4
0 1 2

The local structure is mapped to an eight-bit-number by the function

$$S(C) := \sum_{i=0}^{7} q(i) * 2^i$$

where q(i) := 1 if a cell is located on square i and q(i) := 0 otherwise. S(C) is then searched in the even-numbered positions of the genome, beginning on the left (character positions start with 0). If it is found, the cell C divides and the last three bits in the following character give the number of the square where the daughter cell will be located. If that square is already occupied, no daughter cell is produced and the energy is wasted. If S(C) is not found in an even-numbered position, no division occurs.

A character in an even-numbered position and the subsequent character form a "gene", The first character acts as the regulatory part, and the second character as the structural part of the gene. If there is more than one gene referring to the same input state in a genome, the leftmost one is used. The others are "inactive genes".

The genetic system used in LindEvol was inspired by Conway's "game of life" [Gardner, 1970]. It constitutes a context-sensitive, non-deterministic

L-system [Rozenberg and Salomaa, 1986]. Each gene is a rule, the regulatory part being the left side and the structural part being an implicit form of the right side. As a difference from traditional L-systems, rules in LindEvol work on two-dimensional cell patterns instead of one-dimensional strings.

Furthermore, the genetic encoding of plant growth patterns was inspired by biomorpha [Dawkins, 1986]. LindEvol is an attempt to extend this idea by constructing an artificial environment in which "biomorph" artificial plants evolve under selective influences that are defined by interactions among each other rather than by user-defined fitness criteria.

The simulation of plant growth is started with "plants" consisting of one energyless "germ cell". These are equidistantly placed on the bottom line of grid sites. After 30 days have been simulated, the energy-rich cells in each plant are counted. This number is the fitness value of the genome that controlled the growth of the plant.

The plants are processed in a random sequence that is changed every day, in order to avoid positional effects. The cells of a plant are processed in the order in which they were generated.

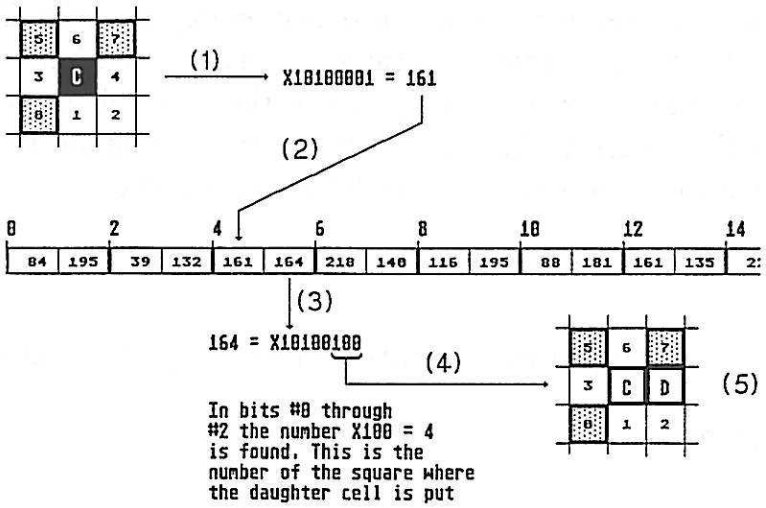

Fig. 1: Control of plant growth by the genome. The activity of the energy-rich cell C is to be determined. Neighboring cells of the same plant are symbolized by shaded squares, open squares indicate free positions. The local structure of the plant around an active cell C is mapped to an 8-bit-number, in which each cell of the local structure corresponds to one bit (1). This number is now searched from left to right in the even-numbered positions of the genome (2). It is found in position #4. The subsequent position #5 contains the code for the action of cell C, bits 0 through 2 of this character give the number of the position where the daughter cell is placed (4). The energy-less daughter cell is placed in the specified position, the active cell is also energy-less after the division (5).

1.2. THE GENETIC ALGORITHM. The genetic algorithm [Goldberg, 1989] used in this model operates with strings of 8-bit characters as genomes. The population size, the selection rate and the mutation rates are control parameters that are constant during a simulation and have to be specified for each individual run. The selection rate is the fraction of the population that

does not survive the selection step. There are three mutation rates, namely the replacement rate, the insertion rate and the deletion rate (see below). The population size has been set to 50 in all runs. The steps of the genetic algorithm are implemented as follows:

1. A start population of random genomes is created.
2. The fitness of each genome is determined using the model of plant growth described above.
3. Selection: A ranking list of the genomes is created according to the fitness values determined in step 2. A constant number of surviving genomes is drawn from the top of the list – there is no randomness in the selection step. Each genome that does not survive is replaced by a copy of a randomly selected surviving genome. Because genomes are only being replaced, the population size is constant. In each replacement process, all surviving genomes are selected with the same probability.
4. Mutation: The simulation of the mutation of a genome consists of three steps. In the first step each character of the genome may be replaced by a randomly selected one. The probability for a replacement to occur per character and generation is the replacement rate mentioned above. In the next step, insertions take place. Insertions can occur between contiguous characters in the genome or at the ends of the genome. Insertions are modeled by inserting two randomly selected characters. The probability per site and per generation for an insertion to occur is the insertion rate. In the last step, deletions take place. They are modeled by removing two adjacent characters from the genome. All pairs of characters in the genome are sites where deletions can occur. The deletion rate is the probability per site and generation for a deletion to occur.
5. Repeat steps 2, 3 and 4 until some condition for finishing the run is fulfilled.

We chose to insert / delete two characters per event because we expected this to lead to a higher speed of evolution of high fitness values. Insertions or deletions of one character would cause frameshifts that affect the whole genome downstream of the site of the event, whereas the insertion or deletion of two characters would have only local effects.

The conception of the LindEvol model matches the outlines given in [Wilson, 1989] in many respects. Because space for growth and light intensity are modelled by the co-evolving plants themselves, adaptation is largely intrinsic [Packard, 1989] in LindEvol.

2. Analysis tools.

In order to get further insight into the processes in LindEvol at the level of population structures, we designed a distance distribution analysis system. A distance distribution is the distribution of values in the distance matrix of a population. Distance distributions are plotted for edit distances and for phylogenetic distances. The phylogenetic distance between two individuals is the number of generations that have passed since their last common ancestor existed. The phylogenetic distance between two individuals that originated from different ancestors in the initial population is undefined.

To calculate phylogenetic distance distributions, the phylogenetic history of LindEvol is recorded. This tree is structured into clusters. By a cluster of depth n we denote a set of individuals that are descendants from a common ancestor n generations ago. The size of a cluster is the number of individuals in that cluster.

Time series of distance distributions form a surface over the time by distance plane. This landscape is visualized by plotting points with values above a given threshold in black and all others in white on a time by distance diagram. The threshold value is 16 for the edit distance distributions and 64 for the phylogenetic distance distributions.

Gene usage is monitored in the LindEvol runs as defined in [Bedau and Packard, 1992]. Because LindEvol genomes are lookup tables like the genomes of the "strategic bugs" developed by Bedau and Packard, the evolutionary activity analysis can be applied to LindEvol. Minor extensions have been added to deal with insertion and deletion events: At insertions, a new usage counter is created for the newly inserted gene, and initially set to zero. Upon deletions, the usage counter of the deleted gene is removed. The usage distribution is sampled every generation and plotted as described for distance distributions with a threshold value of 1.

From the gene usage distributions we calculated the evolutionary activity at $u_0 = 100$. Not all waves in the usage distributions show up as peaks in the evolutionary activity curve because in LindEvol genes can be used more than once in a generation. Usage peaks caused by such genes can grow larger than the u_0 value without exactly assuming u_0. Then, the corresponding wave crosses u_0 without leaving a peak in the evolutionary activity curve.

All three distribution series exhibit marked waves. By a wave we mean a peak that is present in a succession of distributions. Typically the peak's location moves a bit from one distribution to the next one. The rate at which the peak moves is called the slope of the wave. The length of a succession in which a peak is present is called the length of the wave. By the height of a wave we mean the height of the peaks that form it, and likewise, the sharpness of a wave is the sharpness of the peaks that form that wave.

3. Materials and methods.

All programming work was done on an Atari ST computer in GFA basic (available from Richter, Gevelsberg, FRG). A few procedures were re-programmed in 68000 assembly language. Simulations were run on an Atari TT. Diagrams were printed on a Kyocera F-1200 laser printer, using the built-in Prescribe printer language.

4. Results and discussion.

4.1. ORGANIZATION OF THE SIMULATION RUNS. Runs were performed with low mutation rates (replacement rate = 0.003, insertion rate and deletion rate = 0.001) and with high mutation rates (replacement rate = 0.03, insertion rate and deletion rate = 0.01). With each set of mutation rates, one run was done with moderate selection (0.5) and one with strong selection (0.8). Also, a control with no selection was done at each set of mutation rates. All runs were done with populations of 50 individuals, populations were initialized with randomly assembled genomes of 20 genes length. The runs are given the labels:

selection rate	low mutation rates	high mutation rates
0.0	run 1	run 4
0.5	run 2	run 5
0.8	run 3	run 6

As an additional control, runs were done with randomized fitness values. In these, plant growth is simulated, but the fitness values that are assigned to the genomes are not calculated based on the simulation. Instead, random fitness values between 0 and 999 are used. In these control runs, selection takes place in the sense that in each generation a fraction of the population is discarded and some surviving genomes are replicated, but the fitness function upon which this selection is based is the built–in random function of the GFA basic language rather than the amount of energy absorbed by the plant. Random selection controls have been run at low mutation rates with moderate selection (run 7) and at high mutation rates with strong selection (run 8).

4.2. EVOLUTION OF FITNESS VALUES. In the control experiments without selection (runs 1 and 4), no significant evolution towards higher average fitness values is observed. In run 4, elevated maximal fitness values become more frequent in later generations. This is caused by the growth of average genome length (see section 4.3), within a certain range of lengths, longer random genomes have higher fitness values on the average.

A significant evolution towards higher fitness values is observed in all runs with selection. In the initial phase, the maximal fitness typically has a sharp ceiling that often displays several steps upward. In this phase, plants

exhibit finite growth, i.e. they produce a finite maximal number of cells. At the end of a year, all cells of such a plant can be saturated with energy, giving rise to a fitness of the ceiling value. The genetic system of LindEvol also allows infinite growth, in which plants continue to produce cells as they collect energy infinitely. When infinite growth is evolved, the maximal fitness shows no sharp ceiling value anymore, it starts oscillating around some mean value that is in turn subject to further evolution.

Fig. 2 gives an impression of the growth shapes that evolve during LindEvol runs. A more detailed discussion of LindEvol phenotypes will be published elsewhere.

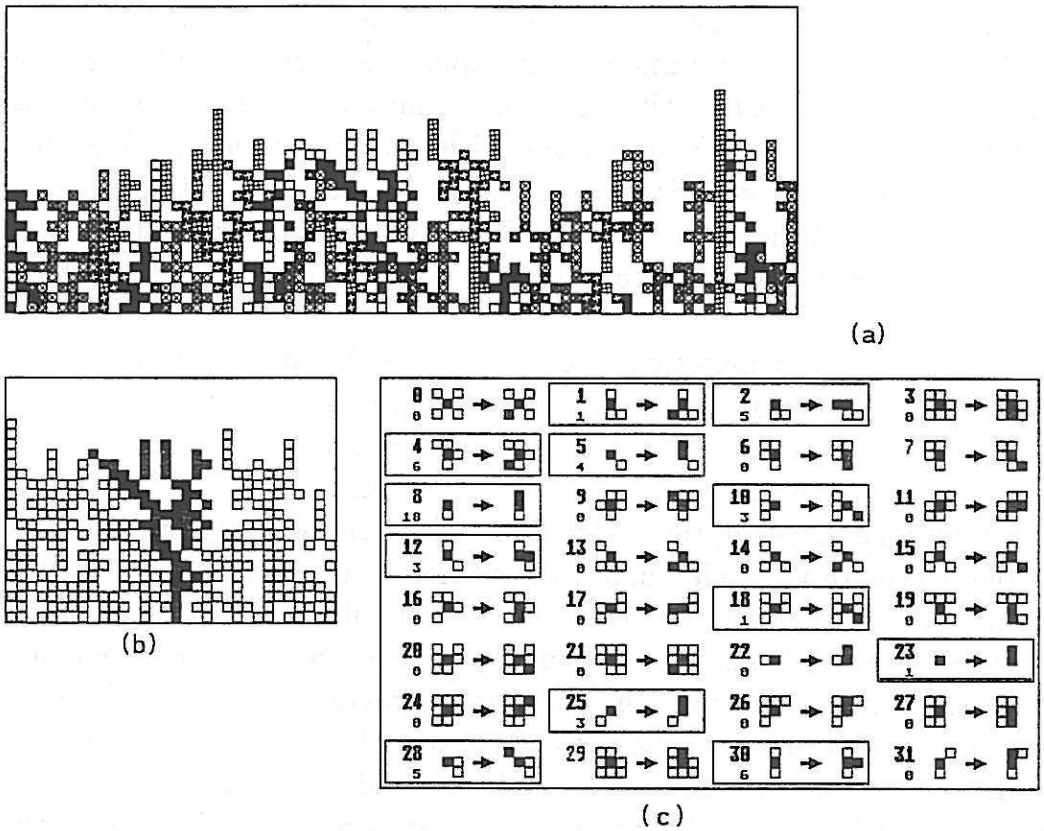

Figure 2: Sample plants and genome. (a) shows a part of a LindEvol world at the end of the 30001th year. Cells are shown as squares. Cells from different plants have different symbols. Closed boxes with white symbols are energy rich cells, open boxes with black symbols are energyless cells.
In (b), a plant also seen in (a) is highlighted. (c) shows the genome of the plant highlighted in (b). Closed boxes are energy rich cells of the highlighted plant, gray boxes are energyless cells of the highlighted plant, and open boxes are cells from other plants. In (c), the rule table is graphically displayed. At the left side of the individual rules, the black center box is the active cell that divides if its local structure matches the pattern of the surrounding open boxes. At the right side of a rule, the position of the daughter cell produced by the division is shown. A darker daughter cell indicates that its position is already occupied in the left side, and that therefore expression of that gene inevitably wastes energy.

Evolutionary steps and gradual changes can both be observed in the fitness values, e.g. in run 2, a very marked step occurred just before generation

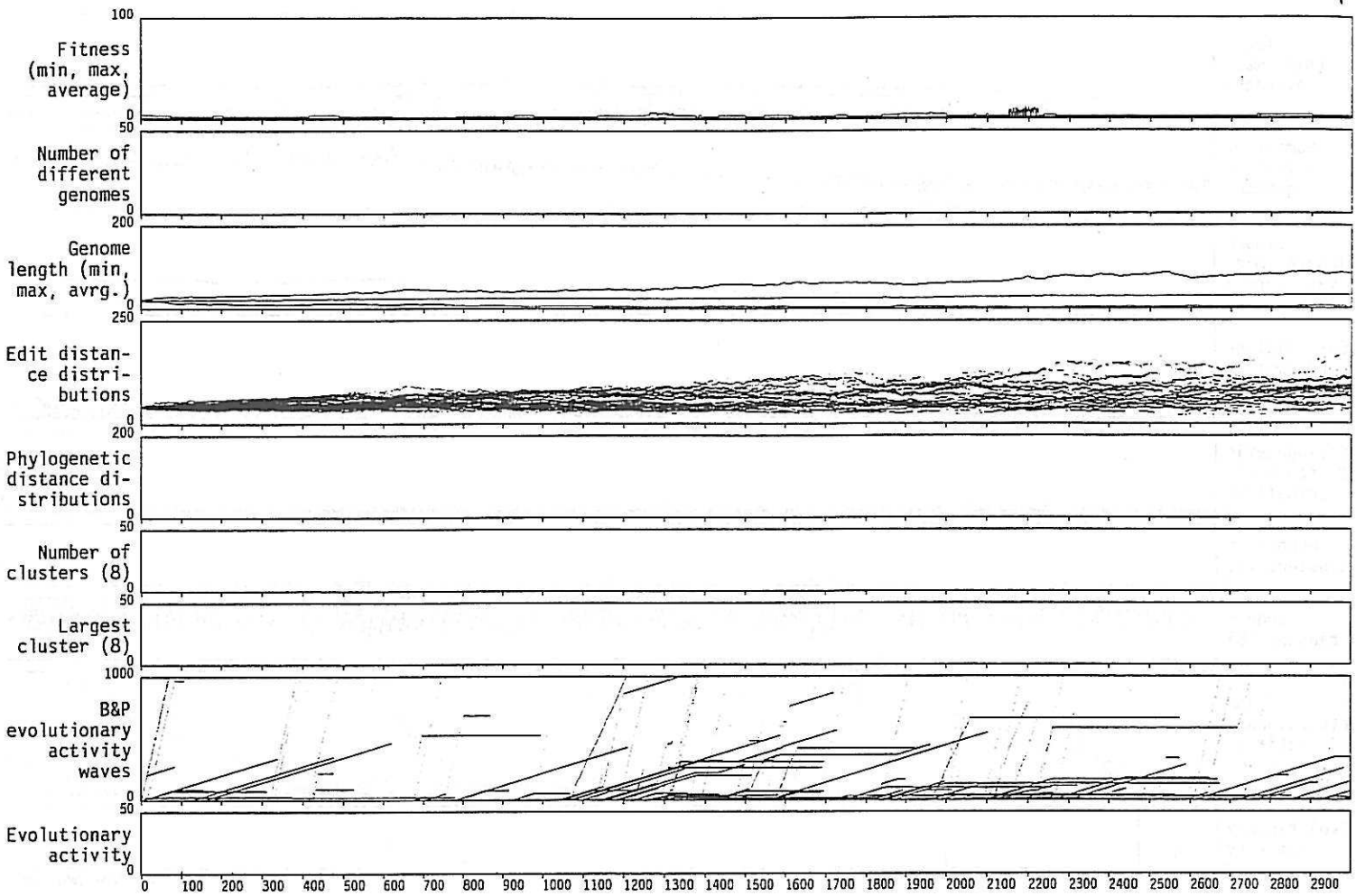

Figure 3: run 1, low mutation rates (0.003, 0.001, 0.001), no selection

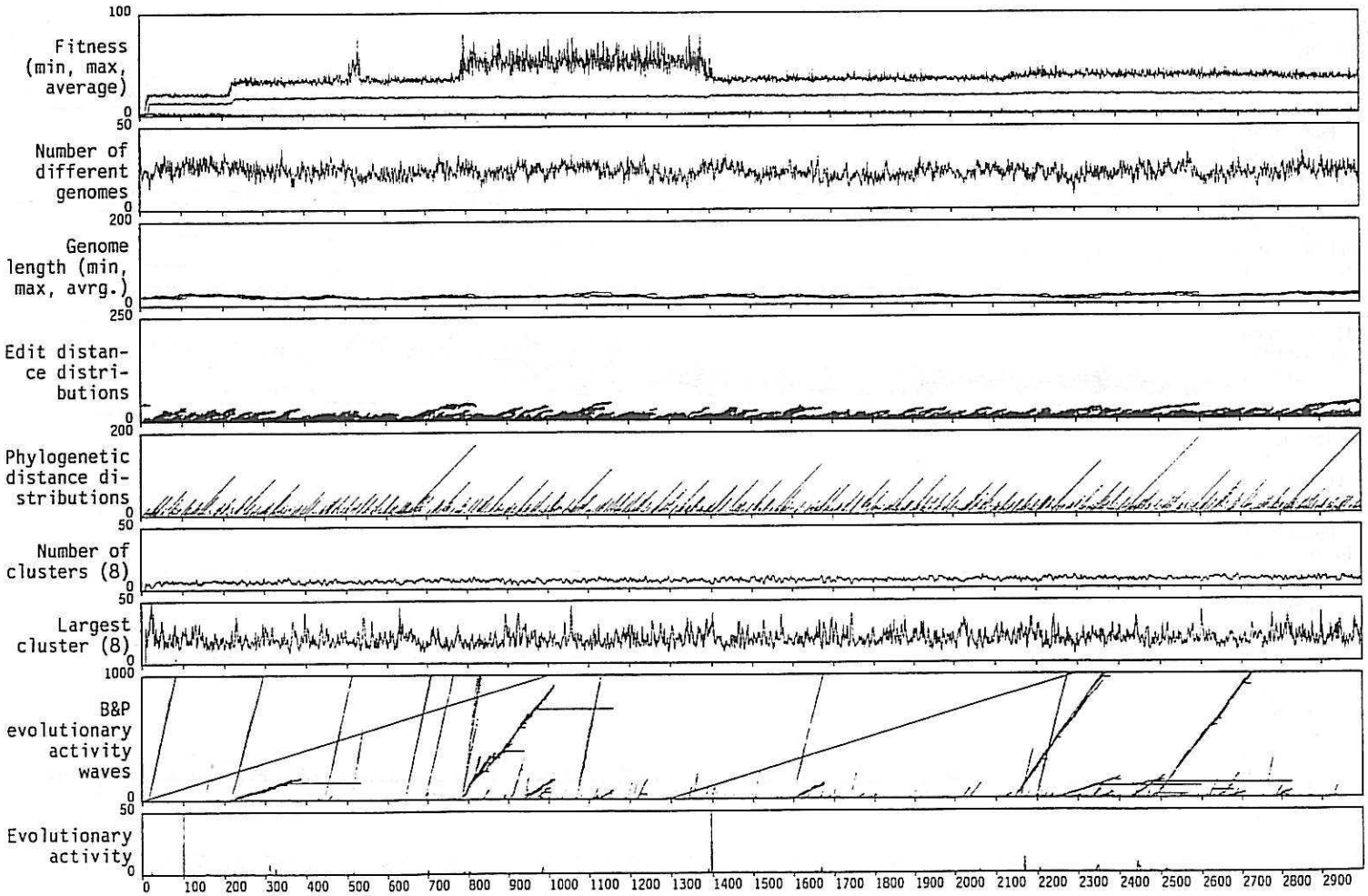

Figure 4: run 2, low mutation rates (0.003, 0.001, 0.001), selection rate = 0.5

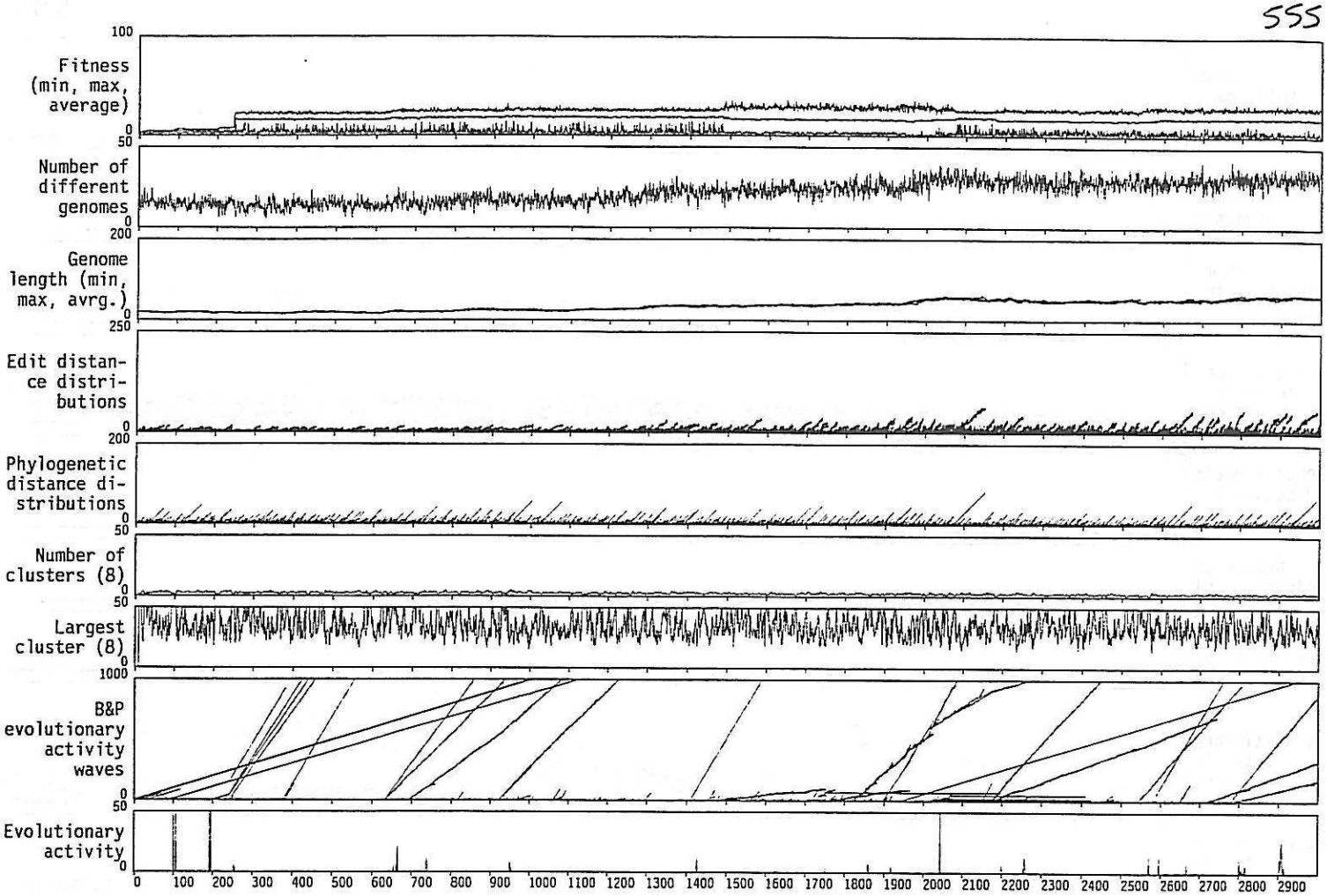

Figure 5: run 3, low mutation rates (0.003, 0.001, 0.001), selection rate = 0.8

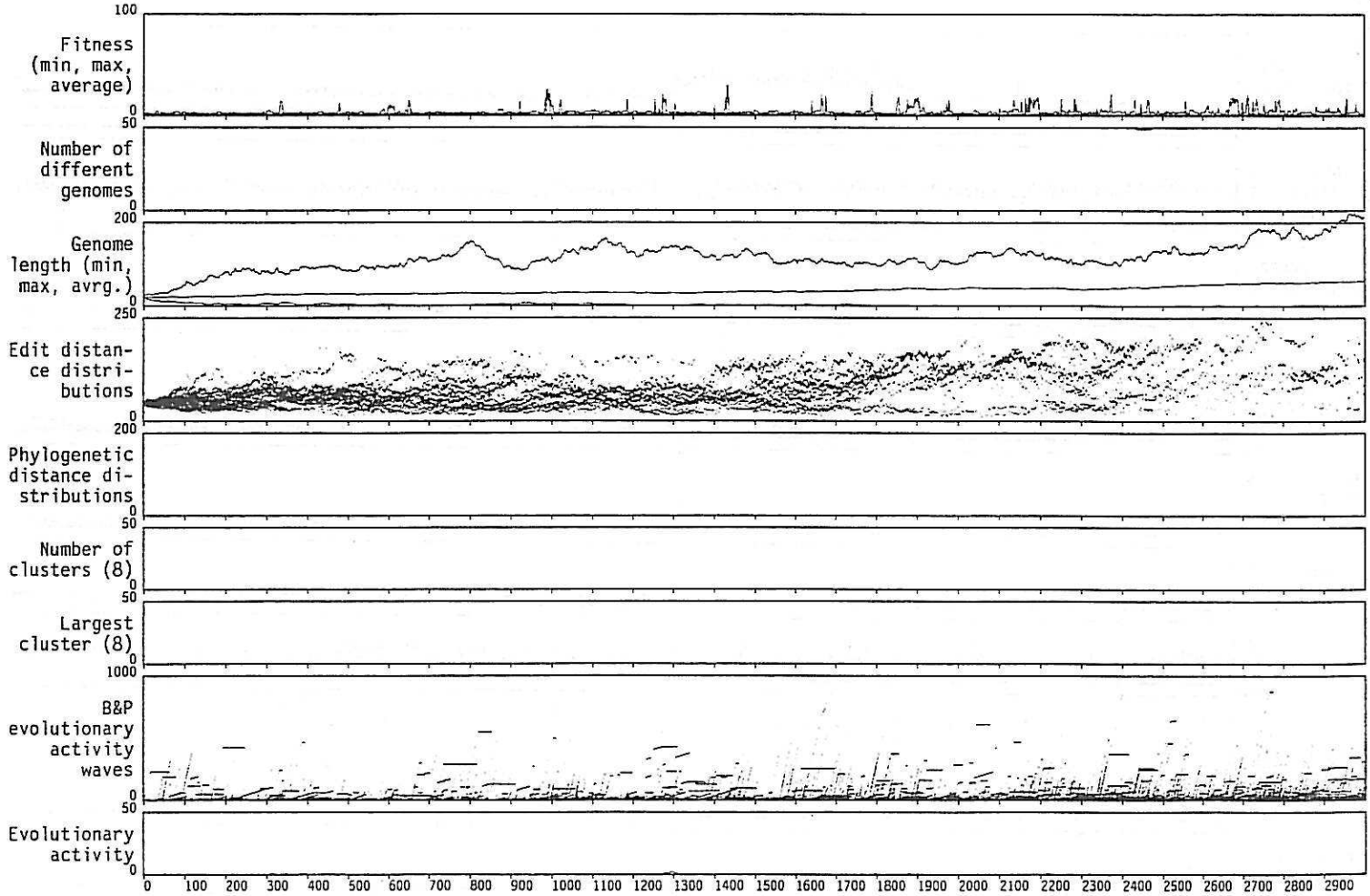

Figure 6: run 4, high mutation rates (0.03, 0.01, 0.01), no selection

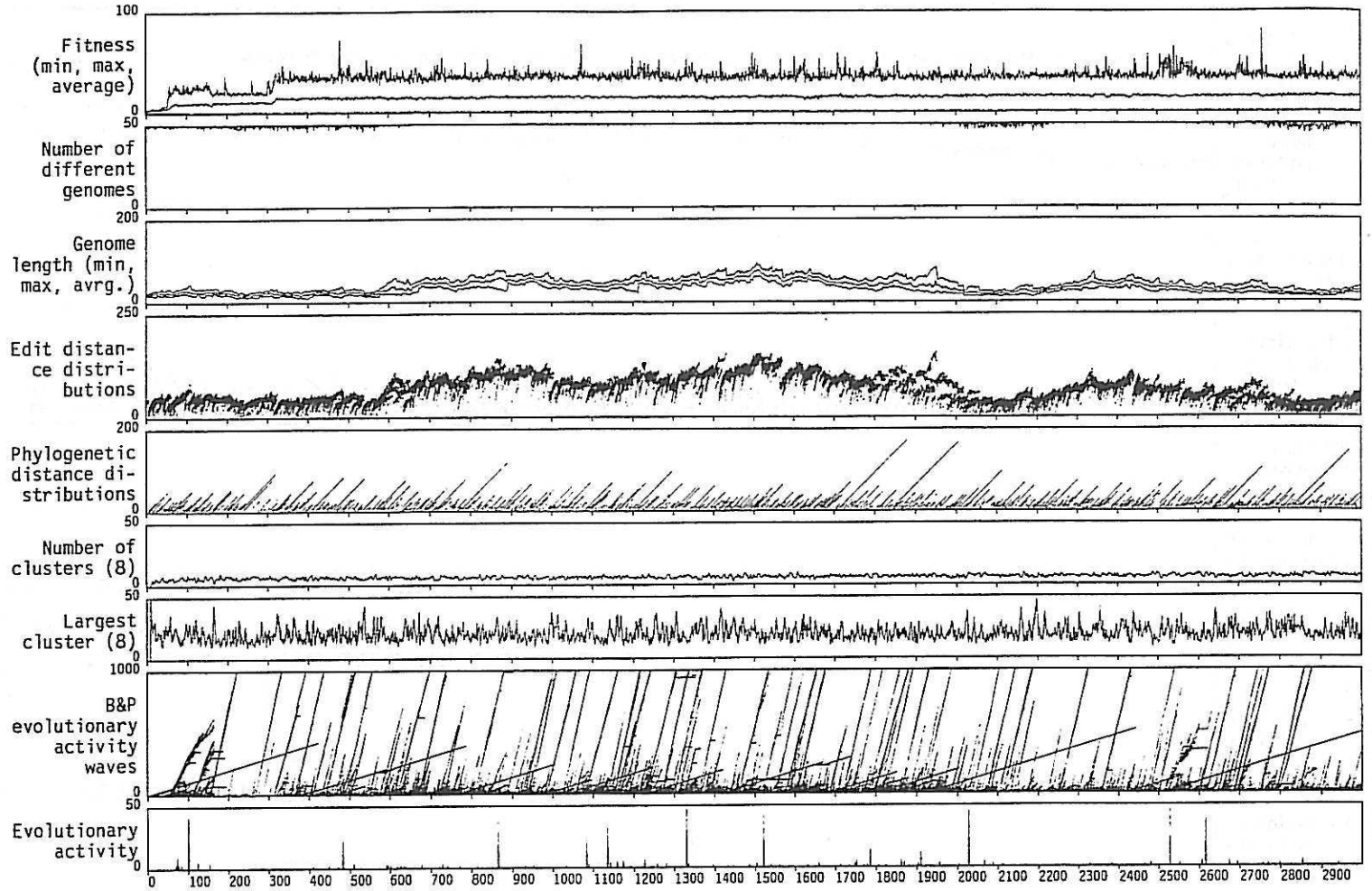

Figure 7: run 5, high mutation rates (0.03, 0.01, 0.01), selection rate = 0.5

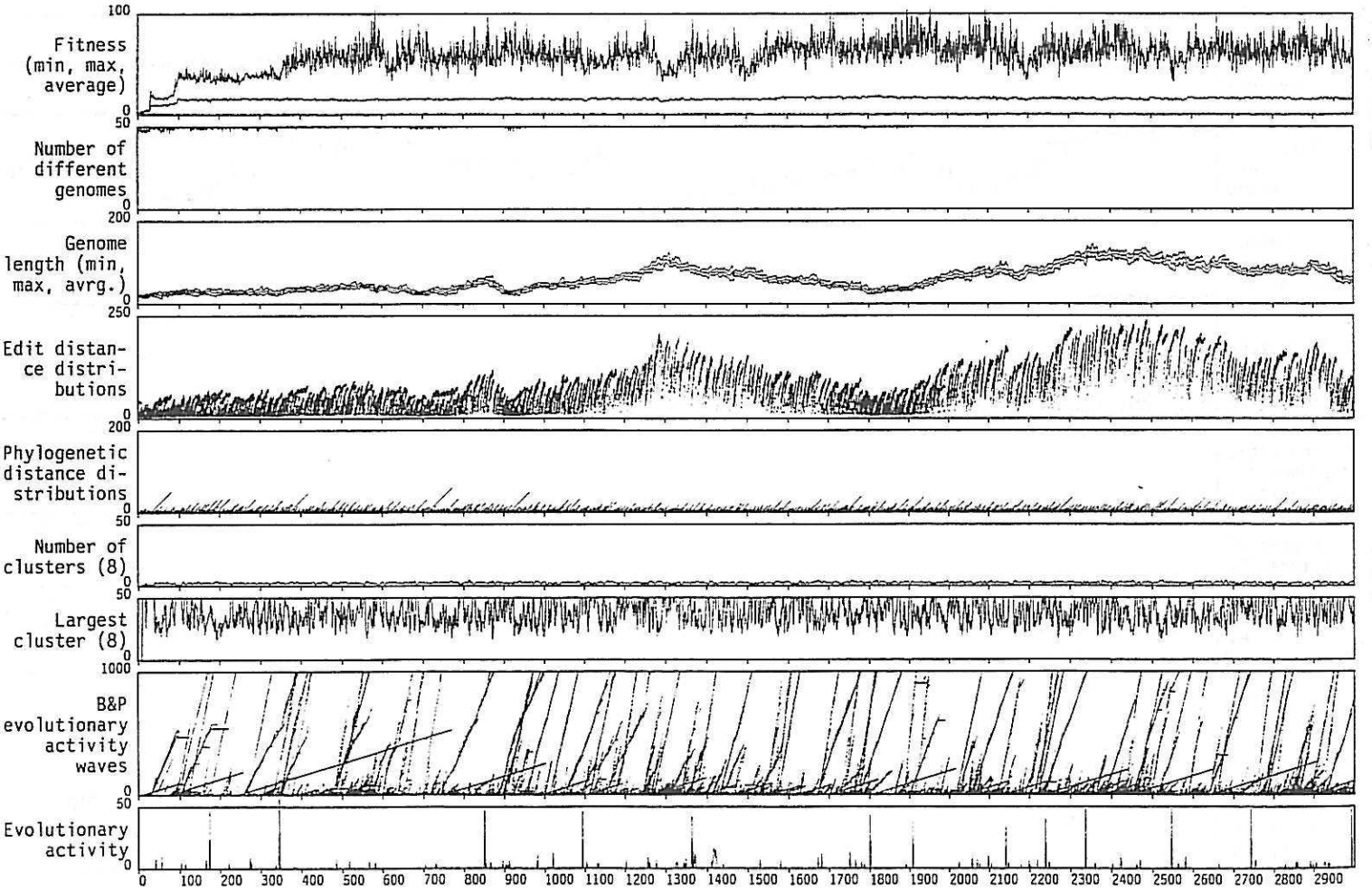

Figure 8: run 6, high mutation rates (0.03, 0.01, 0.01), selection rate = 0.8

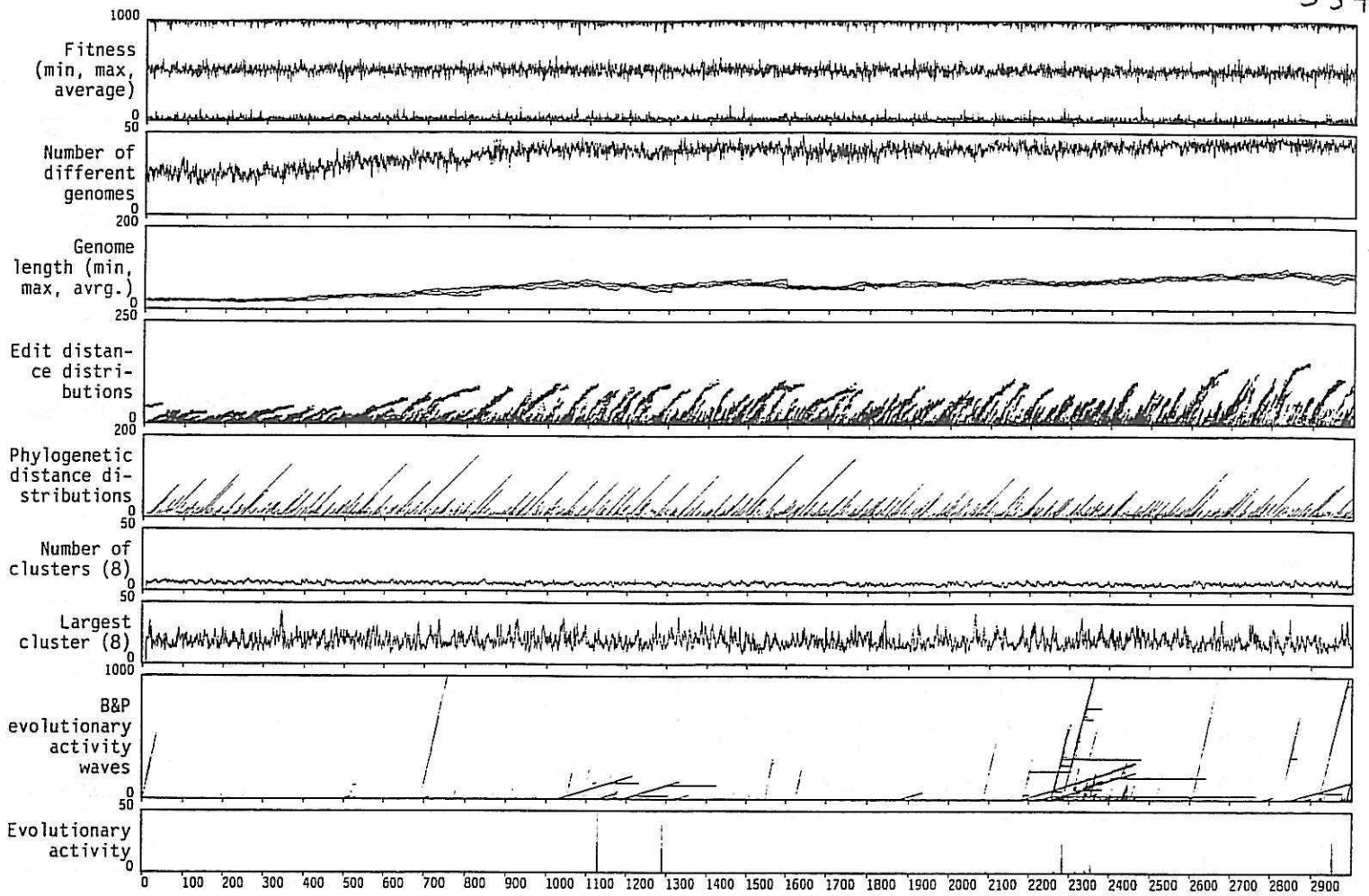

Figure 9: run 7, random fitness control, low mutation rates (0.003, 0.001, 0.001), no selection

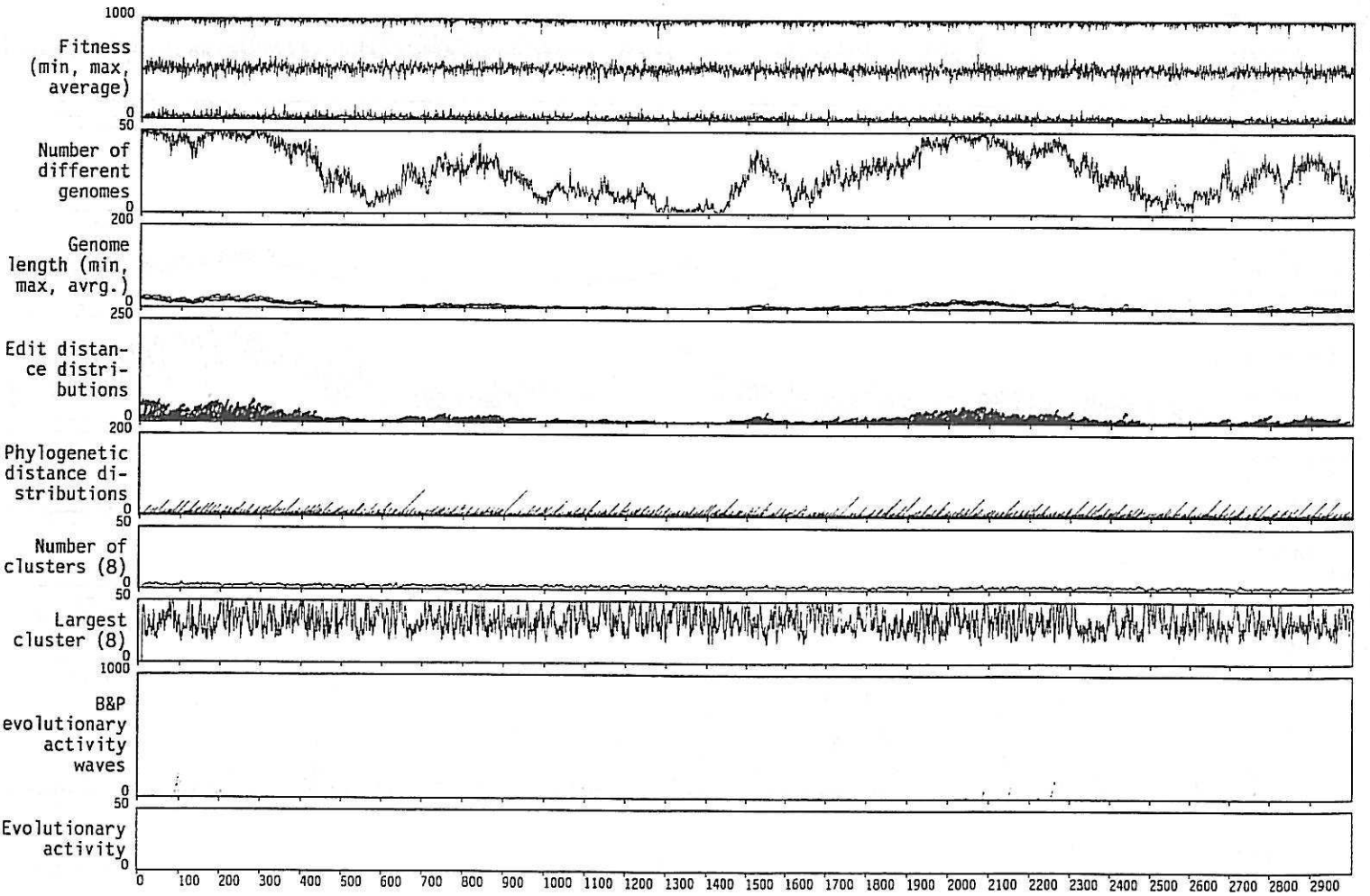

Figure 10: run 8, random fitness control, high mutation rates (0.03, 0.01, 0.01), selection rate = 0.8

800, and a slight, gradual increase of maximal fitness values between generations 2000 and 2100.

In run 3, instances of evolution towards and away from cooperativity can be seen. Near generation 1500, the maximal fitness values increase slightly. This is accompanied by a decrease of average fitness values. Here, evolution has favored genomes that can produce high fitness plants at the cost of abandoning a strategy of avoiding colliding growth in order to avoid the waste of energy associated with collisions. Interestingly, the strategy that evolved near generation 1500 is not permanently favored over more cooperative strategies with lower maximal fitness values. Cooperativity returns around generation 2100, causing the patterns in the fitness curves to become similar to what they were like before generation 1500.

4.3. EVOLUTION OF GENOME LENGTHS. In the controls with no selection, a random diversification of genome lengths and an increase of average genome length is observed. The latter is due to the fact that there are two more potential insertion sites than there are deletion sites in a genome.

Random diversification is not observed when selection is active, regardless of the fitness function involved. Genome lengths increase during most runs, but this is not always the case, as seen in run 2. Increasing genome length may also result from random drift, as shown by the results of runs 7 and 8, so one should be careful when drawing conclusions from increasing genome length.

4.4. CLUSTER NUMBERS AND SIZES. The curves displaying the number of clusters and the size of the largest cluster were expected to indicate evolutionary steps. Once a significantly improved mutant appeared, it should spread through the whole population rather quickly. The cluster originating from that mutant should become exceptionally large, and the number of clusters should decrease during such a takeover. However, no evolutionary step has been found to be accompanied by such a process. The number of clusters and the size of the largest cluster seem to oscillate randomly, unaffected by evolutionary steps or other events. The characteristics curves of the number of clusters and the size of the largest cluster in run 2 and run 6 show no significant difference from those in the control runs with random fitness values performed with the same set of control parameters, runs 7 and 8, respectively.

4.5. DISTANCE DISTRIBUTIONS. In the control runs with no selection, the edit distance distributions have only one very high peak around 40. This peak broadens in the course of the simulations. This reflects the diversification of genome lengths in the population, which is due to insertions and deletions.

In all runs with selection, the wave pattern in the phylogenetic distance distribution exhibits a close relation to the pattern in the edit distance distribution. However, differences are seen in the characteristics of the waves.

Firstly, the waves in the edit distance distributions are fuzzy, whereas the waves in the distance distributions are sharp. Secondly, the waves in the edit distance distributions are bent, their slope typically decreases as they rise upwards, whereas the waves in the phylogenetic distance distributions all travel upwards at the same, constant speed.

Waves in both distance distributions are caused by diverging clusters in the population. Because the number of generations between the current generation and the generation in which members of the diverging clusters departed from a common ancestor increases by one each generation, the corresponding wave in the phylogenetic distance distributions travels upwards at this rate constantly until one of the clusters disappears.

As the clusters diverge, the edit distances increase too, but not in a precise, deterministic way. The edit distances between individuals of the two diverging clusters are distributed around an expectation value, which leads to the fuzzy appearance of the waves. This expectation value increases monotonically each generation. It increases more slowly if the mean distance between two clusters already is already larger because the probability that a character which two strings have in common is affected by a mutation is greater the more common characters the strings share. Thus, the rate at which the mean distance between two clusters increases slows down as the clusters diverge, asymptotically approaching a limit given by the genome lengths.

However, both the increase rate and the limit can be influenced by selectional pressures. Therefore, the waves in the edit distance distributions do not grow with a more or less smoothly decreasing rate, but sometimes increases of the wave slope are observed, such as in run 2 near generation 340. In the control runs with randomized fitness values, such events do not occur because there are no directed selective pressures.

A genome can take over the whole population in 6 generations when the selection rate is 0.5, and in as few as 3 generations when the selection rate is 0.8. Such an event would show up as a distance distribution in which the highest peak is at 6 or 3 respectively. However, such distributions are never found in any of the runs 2, 3, 5 and 6 except at the start of the runs where the maximal phylogenetic distance is limited by the number of generations that have passed. Evolution in LindEvol is not a succession of simple takeovers of mutants that are superior to their ancestors. The dynamics of evolutionary changes are more complex in LindEvol, as we have already mentioned in the discussion of the cluster number and size curves.

One reason why genomes do not spread at the maximal rates mentioned above is the non–deterministic nature of the fitness function. However, this alone cannot account for the lengths of the waves seen in the phylogenetic distance distributions. In run 2, the longest waves in the phylogenetic distance distributions are even longer than those in the control, run 7, which has been performed with the same set of control parameters. This indicates that the LindEvol fitness function does not drive the system towards uniformity any

more strongly than the random fitness function, which cannot be expected to select for uniformity at all. The LindEvol fitness function actually stabilizes diversity in those phases where the long waves in the phylogenetic distance distributions are observed.

4.6. GENE USAGE AND EVOLUTIONARY ACTIVITY. Looking at the usage distribution diagrams, four types of waves can be distinguished:

1. Waves with slope 1. In run 2, such a wave originates near generation 0, another one near generation 1300. These are caused by the germ cell genes and other genes which are used exactly once per generation. If the germ cell gene is not used, this is lethal to the genome. Therefore, the germ cell gene waves are high and sharp.

2. Steep waves, i.e. waves with a slope much greater than one. These waves arise if a gene is used several times in a generation. Typically, such multiple usage of genes occurs in infinite growth. Because the number of times such a gene is used depends on photon absorption, which is non–deterministic, and it is not invariably lethal if such a gene is not used the maximal number of times, these waves have a more fuzzy shape. When an evolutionary step to infinite growth occurs, the number of times a gene is used may switch from one to many times, and the corresponding wave switches from type 1 to type 2. A pronounced instance of such an event can be seen in run 3 near generation 240.

3. Horizontal waves. These are caused by once active genes that became inactive, but were not lost from the genomes. In most cases, they finally are lost after some time, but in some cases they are activated again, causing the wave to start rising upwards again. Such an instance can bee seen in run 2 near generation 2430.

4. Complex waves. These have a varying slope and a branched structure. In run 2, such waves originate near generations 780 and 2130.

In the controls with no selection, surprisingly many evolutionary activity waves are seen. However, none of them is a complex wave. The evolutionary activity curve shows that none of the waves are significantly high. The formation of high waves is only possible if genomes, and hence usage counters, are replicated, which is not the case in runs with no selection.

The control runs were expected to reveal what amount of usage accumulates without any directed selection, as suggested by Mark Bedau [Bedau, 1992]. However, it was found that usage levels do not remain below some rather low threshold in any of the controls, but instead, distinct waves that travel upward are observed in all control experiments.

As cells are produced during plant growth, new local cell structures arise in the plant, which in turn activate new genes. This way, the pattern of gene usage is controlled by the genome itself to a large extent. Therefore, evolutionary activity waves are observed even in the control runs. If a genome is amplified, whether by random drift or by directed selection, its set of usage

counters is amplified in this process, and the corresponding waves become higher as a result. The waves end when the genes that give rise to them disappear from the population, which can be due to competition, due to random drift or due to mutations. The latter two factors are much stronger in run 8 than they are in run 7, therefore only a few, short waves are seen in the usage distribution diagram of run 8.

In runs 2, 3, 5 and 6, the evolutionary activity waves are much more frequent than in the controls with random fitness values, and much higher than in the controls with no selection. Complex waves are only seen in these runs. They are most prominent in the runs with low mutation rates.

In run 3, a complex wave originates near generation 1820. This complex wave is especially interesting because its branching near generation 2020 appears to be correlated to the long wave that is seen in both distance distributions. It is assumed that the branch of the wave was caused by different per generation usage frequencies in individuals of the diverging clusters. However, other branchings of complex waves do not appear to be closely linked to waves in the distance distributions. More research will be necessary to improve our understanding of complex waves.

In runs 2 and 3, the activity waves caused by germ cell genes are typically very long. This indicates that under a regime of low mutation rates, a germ cell gene becomes strongly stabilized once it appears, and successful plant growth patterns evolve based on this germ cell gene. But the spreading of a new germ cell gene is not completely impossible with low mutation rates. In both runs, corresponding activity waves originate in a later phase. However, it cannot be excluded that these are caused by conservative mutations in the right side of the germ cell gene, i. e. mutations that leave the least three bits of the right side, which are the only relevant ones, unchanged. Because a conservative mutation is completely neutral, it is possible that such a mutant takes over the whole population by random drift.

In the runs with high mutation rates, germ cell genes are changed much more frequently. Here, competition between plants with different germ cells can take place. Especially in run 5, it is regularly seen that horizontal waves start rising up again, indicating that a gene temporarily became inactive and was re–activated again. This strongly suggests that genomes evolve to encode alternative growth patterns based on alternative germ cell genes or other key genes, and that evolution sometimes shifts between these alternative patterns.

5. Conclusions and outlook.

In LindEvol, the fitness landscape is shaped by complex interactions among the co-evolving agents. These interactions give rise to a unique, unpredictable history of evolution in each run. We have never observed a LindEvol remaining converged in an optimum of the fitness function. It seems

that the effects which evolution towards an optimum has on the intrinsically modelled fitness landscape always create new optima and open up new pathways for evolution. Hence LindEvol evolution can be considered to be open-ended. Evolutionary steps and gradual phyletic changes can be observed. Especially with moderate selection, cooperation between plants can emerge. All these things are commonly considered to be landmarks of biological evolution.

Evolutionary steps are remarkably complex processes in LindEvol. They are not caused by superior mutants that simply outcompete and replace the previous community, as shown by the finding that evolutionary steps are not accompanied by peaks in the largest cluster curve and by minima in the number of clusters curve. Further investigations on this issue may be interesting, as they may deepen our understanding of the way genetic algorithms work.

Breakdowns of waves in distance distributions do accompany evolutionary steps, but these breakdowns are not noticeably different from others that are not accompanied by any perceivable evolutionary step. Evolution in LindEvol, does not proceed via simple replacements of individuals and species by improved mutants.

It is one of the most fascinating properties of biological evolution that it gives rise to a vast number of species, not only to just one species. This means that evolution does not proceed by replacing populations and species by improved mutants, they coexist and co-evolve during long periods. The distance distributions clearly show that clusters coexist for extended periods, indicating that co-evolution of several species is successfully modeled by LindEvol.

The phylogenetic distance and the edit distance are strongly correlated in LindEvol due to the way mutation is implemented. However, this correlation is disturbed in some cases by selective influences. It is hoped that further investigations of such processes will lead to new insights about the mechanisms underlying the "molecular clock".

6. Acknowledgement.

Thanks are due to Luzie Ursula Wingen for inspiring discussions and for proofreading this paper.

Literature.

Bedau, M. A. and Packard, N. H. 1992. "Measurement of Evolutionary Activity, Teleology, and Life." In *Artificial Life II*, edited by C. G. Langton, 431 – 462. Santa Fe Institute Studies In The Sciences Of Complexity, Proceedings Volume X. Addison-Wesley, Redwood City, CA.

Bedau, M.A. 1992. Personal communication.

Dawkins, R. 1970. *The Blind Watchmaker*. Longman House, Harlow, Essex.

Gardner, M. 1970. "Mathematical Games: The Fantastic Combinations of John Conway's New Solitaire Game 'Life'." *Scientific American* 223(4): 120 – 124.

Goldberg, D.E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, New York.

Kauffman, S. A. and Johnson, S. 1992. "Co-Evolution To The Edge Of Chaos: Coupled Fitness Landscapes, Poised States And Co-Evolutionary Avalanches." In *Artificial Life II*, edited by C. G. Langton, 325 – 370. Santa Fe Institute Studies In The Sciences Of Complexity, Proceedings Volume X. Addison-Wesley, Redwood City, CA.

Packard, N. H. 1989. "Intrinsic Adaptation in a Simple Model for Evolution." In *Artificial Life*, edited by C. G. Langton, 141 – 155. Santa Fe Institute Studies In The Sciences Of Complexity, Proceedings Volume VI. Addison-Wesley, Redwood City, CA.

Rozenberg, G. and Salomaa, A. 1986. *The Book of L.* Springer Verlag, Heidelberg.

Wilson, S.W. 1989. "The Genetic Algorithm and Simulated Evolution." In *Artificial Life*, edited by C. G. Langton, 156 – 166. Santa Fe Institute Studies In The Sciences Of Complexity, Proceedings Volume VI. Addison-Wesley, Redwood City, CA.