THE ACADEMY OF SCIENCES OF RUSSIA
PUSHCHINO RESEARCH CENTRE
RESEARCH COMPUTING CENTRE

N. L. LUNINA

THE PACKAGE OF EDUCATIONAL PROGRAMS «BIOLOGY»

PC SOFTWARE

PUSHCHINO · 1992

THE ACADEMY OF SCIENCES OF RUSSIA
PUSHCHINO RESEARCH CENTRE
RESEARCH COMPUTING CENTRE

N. L. LUNINA

THE PACKAGE OF EDUCATIONAL PROGRAMS «BIOLOGY»

PC SOFTWARE

PUSHCHINO · 1992

SUMMARY

This issue contains a description of a number of educational programs designed for students studying biology. The program package was primary created for secondary school students, but it can be used in colleges too. The programs can be run at IBM PC compatible and YAMAHA compatible computers.

Introduction

The package of educational programs "Biology" contains 10 programs, 5 of them refer to the field of "Genetics", 3 - to the field of "Biochemistry" and 2 - to the field of "Biosynthesis of protein". This classification is rather conventional, because the programs touch upon the problems of neighboring fields quite often. There are 2 variants of the programs differing in dialog language. The names of the programs with dialog in English are:

BIOANCRO - "Analyzing crossing";
BIOPLINE - "Getting of pure lines";
BIOINGEN - "Interaction of genes";
BIOGNMAP - "Genetic map constructing";
BIODECOD - "Decoding of genetic code";
BIOLIPID - "Lipids";
BIOSTFOR - "Structure formulas of aminoacids";
BIOTRNA - "Secondary structure of a transfer RNA";
BIOSYNTD- "Biosynthesis of protein. Demonstrational program"
BIOSYNTS - "Biosynthesis of protein. Simulational program".

The majority of the programs is created along a similar scheme. Each program starts with outlining the problem to study, the short instruction for work with the program is also given. The learner is required to perform certain operations to check his knowledge. The list of possible operations is usually displayed. It is also pointed out how to quit the program or ask for help (F1). When asking for help, the learner is first supplied with leading considerations. If asked repeatedly, the correct answer is given. But the learner has to input this correct answer himself to continue the program. In the course of working with the programs the count is performed, which considers correct operations, errors, asking for help. At the end of work

a result (a score got related to that maximal possible) together with the learner's name, date, program's name is stored in the file BIO.RES.

1. Genetics

1.1. The outlines

The learner is supposed to be familiar with the outlines of Mendel's genetics:

- there exist alternative pairs of signs;
- the signs may be inherited;
- hereditary signs are transmitted in a discrete way;
- each organism contains 2 factors, responsible for a sign, one of them is expressed outwardly;
- either of the factors may be inherited.

The learner is also supposed to know definitions:

- alternative pairs of signs are called alleles;
- hereditary factors are called genes;
- signs, expressed outwardly are called dominant, those, existing in the organism in a suppressed state - regressive;
- sexual cells, containing a single hereditary factor are called gametes;
- in the course of crossing, a male and a female gametes flow together, zygote is formed;
- zygote, containing the same allele genes, is called a homozygote, different - a heterozygote;
- organism, developing from a homozygote is called homozygous, from a heterozygote - heterozygous;
- gene composition of an organism is called a genotype, its outward realization is called a phenotype;
- crossing of organisms with a single pair of alternative signs is called monohybridous, with two pairs -dihybridous.

To denote units of heredity letter symbolism is used. Hereditary factors, or genes, giving rise to dominant signs are associated with capital letters (A,B,C) and those, for recessive signs - with small letters (a,b,c).

If one cross two organisms, one of which contains only dominant genes (AA) of a given sign, and the other - only recessive ones (aa), then their first-generation progeny will be heterozygous (aA). If the hybrids obtained are crossed to each other, the second-generation progeny will exhibit 4 equiprobable

genotypes: AA, Aa, aA, aa. Organisms of first three types (AA, Aa, aA) will be phenotypically indistinguishable, which give decay in the appearance 3:1.

If two parental organisms with genotypes AABB and aabb (in regard to a certain sign) are taken, their first-generation progeny will be heterozygous with the genotype AaBb. This progeny in turn will give 16 equiprobable genotype combinations in the second generation. Only four of them will be phenotypically distinguishable. Here decay in appearance will be close to a theoretical value 9:3:3:1.

1.2. Analyzing crossing

The program allows the learner to determine the genotype of a plant with known phenotype by means of analyzing crossing. There is an initial sample with unknown genotype. Besides that, there are some individuals to cross to. It is required to perform crossing so (that is to choice such an individual) that it would be possible to determine the genotype of an initial sample from a progeny phenotype.

The program allows the learner to simulate experiments with plants differing in 2 signs - coloration and shape of seeds. It is known, that yellow coloration (A) and round shape (B) dominate; green coloration (a) and wrinkled shape (b) recess.

To become quite clear in the task let us consider an example of analyzing monohybridous crossing. If we have a green-seed pea, a single genotype aa may be supposed. But if a yellow-seed pea is taken, it can has either AA genotype (case 1), or Aa genotype (case 2). If we cross it to a yellow-seed plant, which is homozygous one, the progeny will be completely yellow both in cases 1 and 2. So, this choice of the partner is not effective. The green-seed plant is required. Its genotype is aa, and progeny will be completely yellow in case 1 (only zygotes Aa are formed) and different-seed in case 2 (zygotes Aa and aa are formed).

So, for analyzing crossing an individual with recessive signs should be chosen. This is also valid for dihybridous crossing.

The learner must determine the genotype of a given sample. For this mean he may cross the studied plant to one of four possible plants with different phenotypes and observe progeny F1. If the choice was failure, he may do some more crossings. The progeny's look is displayed each time after the crossing is performed (Pic.1). Having determined the genotype, the learner

should input it in the form "AaBB". If the answer is correct, the Pennet's greed for current crossing is displayed. Otherwise, the message "You are wrong" appears.

The learner may ask for help. Here the following messages may appear:

- to make any crossing move the pointer to any item and press ENTER;
- make another crossing;
- color of progeny F1 is only yellow; shape is only smooth;
  only green; only wrinkled;
  different; different;
- genotype of the studied plant is AaBB.

The first message appears if no one crossing has been done yet. The second one - if not a sample with recessive signs is chosen. The third message indicates that the crossing is correct, but the learner has difficulties in making a conclusion. If in this situation he repeats asking for help, he is given a correct answer.

After the learner executed 8 such tasks, he is informed about a percent of correct answers. The results together with the learner's name are stored in the file "BIO.RES".

1.3. Getting of pure lines

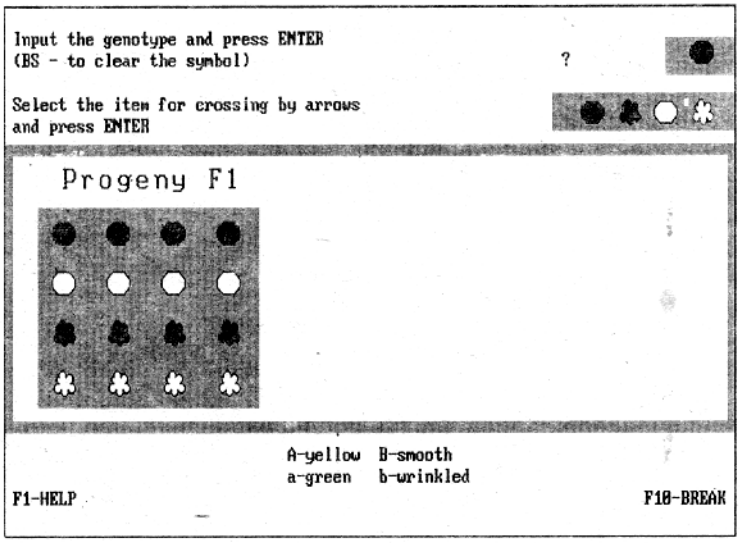
When selecting plants or animals the problem arises to pick up organisms "pure" by a certain sign, that is those with progeny, not decaying by this sign. As a rule, for this mean a progeny F2 (second generation) should be observed. But if studied plants permit self-pollination, the progeny F1 is enough. If progeny F1 contains individuals with genotypes different from those of their parents, than:

first - a parent was heterozygous by this sign,

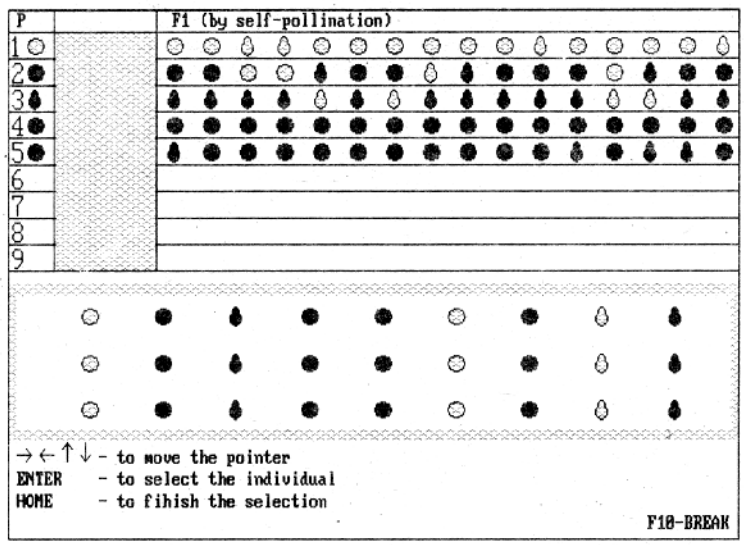
secondly - the expressed sign is recessive.

The program simulates a work with tomatoes, differing in two signs (red and yellow; round and pear-shaped) and permitting self-pollination.

The first part of the program is a search of plants. There is a field with tomatoes of all possible genotypes. The number of possible phenotypes is 4. The learner must choice for experiments a few plants (no more than 9). On measure of choosing he is being informed about F1 progeny's phenotype of each chosen plant (Pic.2). The learner has to:



Pic1. Analyzing crossing



Pic.2. Getting of pure lines.

- find dominant signs,
- decide which plants may be used to get pure lines.

That is why he proceeds his search till he find out 4 plants not giving decay. If he finds these plants right away, he has nevertheless to look for heterozygous plants to make a conclusion about dominance. After that he breaks the regime of searching. The case is possible, when all the 9 plants are looked over, but "pure" plants are not fond. Then searching is ceased automatically.

After quitting the regime of searching the learner must point out the dominant signs. For this mean he should introduce a notation, where capital letters are associated with dominant signs.

The next step is to determine the genotype of the chosen plants and describe it in the introduced notation.

And finally, the learner must point out progeny of which individual may be used to get each of 4 possible pure lines (red round, red pear-shaped, yellow round, yellow pear-shaped). If it appears so, that there are no such individuals among chosen plants, the learner will be returned to the field to continue the choice of plants for getting pure lines.

When determining the genotype of the chosen plants the learner may ask for help. First, he is supplied with leading considerations. If asked repeatedly, the correct answer is given.

He can also ask for help when determining a candidate for pure lines. Here the suitable individual is pointed out.

After the program is finished the learner is informed about a percent of correct answers. The result together with the leaner's name is stored in the file "BIO.RES".

1.4. Interaction of genes

One sign is influenced by two genes. For example, a coloration of a rabbit depends on the gene of coloration C/c (colored/non-colored) and the gene of the pigment distribution along a hair length A/a (uniformly/non-uniformly). A and C dominate. A rabbit without coloration looks white, with uniform coloration looks black and with not uniform coloration - grey. Here some interesting tasks arise.

Task 1. To determine the genotype of a rabbit using known genotype. The number of possible genotypes is 9, possible phenotypes is 3. The learner is given a genotype in the form

Ccaa. He must understand, that the rabbit is colored (C is present), pigment is allocated along the hair uniformly (A is absent), then the rabbit is black. The learner must input "b" ("black") and the black rabbit will appear in the display. Otherwise the message "You are wrong" appears. The learner may ask for help. First, he is provided with leading considerations, if the inquiry repeats, the correct answer is given. In our case, first the message "Coloration (Cc) is uniform (aa)" will be printed, secondly - "black".

Task 2. To determine the genotype using known phenotype. For example, a black rabbit must have a coloration gene C and can not have a non-uniformity gene A, then its genotype may be either CCaa, or Ccaa.

The learner inputs a set of letters. Each letter may be inputted from either latin or russian registers. Letters other than "c" and "a" are not interpreted. Erroneously inputted letter may be erased with a key "BS". If the number of genotypes is more than 1, they may be inputted in arbitrary order. If the genotype is inputted erroneously, all the 4 letters are eliminated. If the genotype is correct, but there exist some more possible variants, a comma appears at the end of notation and the learner must continue inputting. When all possible variants are up, a rabbit with current phenotype is displayed. It may be white, grey or black.

When asking for help, first leading considerations are given. For example, "There is a coloration (must be C). It is not uniform (must be A)". When asked repeatedly, a list of possible genotypes for one and the same phenotype is printed. But the answer, nevertheless, must be inputted by the learner himself. The variants may be ordered arbitrarily.

Task 3. A buck with known genotype is crossed to 4 grey does with various genotypes. The progeny of all 4 crossings is observed. The question is:

- will there be a white baby among progeny?
- will there be a black baby among progeny?

In the case of the positive answer it is required to point out what genotype must have a grey doe for it.

For example, a doe with the genotype CCAA can have neither black nor white babies, since he transmits the gene C to all of them (they will be colored) and the gene A (they will be colored non-uniformly). That is the progeny of this rabbit will be grey.

A buck with the genotype CcAA may have white babies if a doe

will also contain the gene c. But the doe can't have two such genes (otherwise she would be white, but from the task's condition all the does are grey). So, the doe must be heterozygous by the gene of coloration - she has the set Cc. As for the uniformity of coloration, it is not essential for white babies, and the doe may have an arbitrary set of genes A, admissible for a grey individual - Aa or AA. Thus, a buck CcAA will have white babies when crossed to grey does CcAa or CcAA.

The learner may ask for help. As everywhere, first leading considerations are given, then the correct answer is reported. For example, when the learner asks help during answer the question "Can the doe with genotype CCAa have white babies?" first it will be printed "The buck is homozygous by C - all the babies are colored", and secondly - "Not".

Task 4. There are bucks with known phenotype. It is required to determine their genotype. For this mean the learner may cross them to various does, which are displayed. The phenotype of each doe is visible. Besides that it is known, that does of all kinds genotypes are presented. The result of crossing is depicted by a litter of 15 babies, the number of babies of each phenotype is proportional to the theoretically possible one (Pic.3).

For example, if a buck is grey and all his babies are grey too, the only possible genotype he can have is CCAA. Otherwise, being crossed to heterozygous does he would have babies of other coloration.

If a grey buck has white babies, this indicates his heterozygousness by the gene of coloration C - he has a set Cc. To determine his homo- or heterozygousness by A, one need to cross him to black does. A heterozygous buck (Aa) when crossed to black doe (aa) will have black babies, a homozygous one (AA) will not have black babies.

The learner may ask for help. If the job is not finished, the message :"Not all crossings were made" appears. Then the leading considerations are given. For example, "The buck is colored non-uniformly. All the babies are colored. There are non-uniformly colored babies". When the help is asked for repeatedly, the buck's phenotype is reported. In our case it will be CCAa.

1.5. Genetic map constructing

The program is meant for active experimental work on the

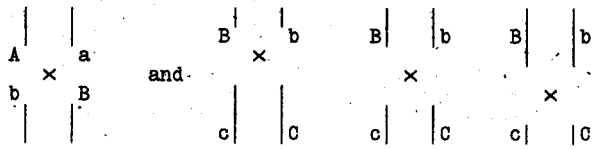
part of a learner that will help him to grasp a difficult notion of "linkage groups" and come to the idea of genes' linear location on the chromosome.

The idea first appeared when analyzing a hereditary transmission of some signs. Let us take a pure line of black (aa) Drosophila females with long wings (BB) and brown males with rudimentary wings (bb). The F1 progeny will be completely brown with long wings (AaBb). Let us cross the F1 females to pure recessive males (aabb). The F2 progeny will exhibit 4 genotypes: AaBa, aaBb. Aabb. aabb. They will look as brown with long wings, black with long wings, brown with rudimentary wings, black with rudimentary wings. It would be natural to expect, that as a result of independent hereditary transmission each group will be equally presented in F2 and will contain 1/4 part of all the progeny's amount. But the experiment reveals a rather different picture:

- brown with rudimentary wings - 41,5%
- black with long wings - 41,5%
- brown with long wings - 8,5%
- black with rudimentary wings - 8,5%

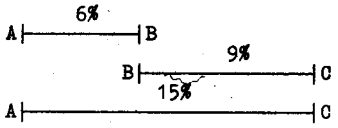
That is both signs of one parent are transmitted in general. The signs which are quite often transmitted in the same combination as in the parental organism are called linked signs. The linked signs form the linkage groups. Above we've considered the linkage group of two signs (body coloration and wings lengths). Usually, linkage groups of a greater number of signs occur. But there remains a small number of descendants with a new combination of signs. They form the so-called recombinant type.

The fact, that some signs are transmitted in the linked form has the following grounds. Genes, corresponding to linked signs, are located on the same chromosome. In the majority of cases a descendant gets one of the parental chromosomes, therefore he inherits either both signs from the mother or both signs from the father. A small number of recombinant types is conditioned by a crossing-over at meiosis. The farther the genes are located on the chromosome - the higher is the probability of crossing-over, and consequently, the greater is the percent of recombinant individuals among the progeny. Let us compare the pairs of signs A and B, located close to each other and B and C, located far from each other.

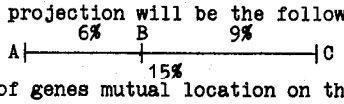


The distance between B and C is approximately 3 times as large, than that between A and C. So, crossing-over there may take place approximately 3 times as frequent. That is why, upon the frequency of recombinant individuals, we may judge of the relative distance between corresponding genes. If the recombinant type BC appears approximately 3 times as frequent than AB, we may expect, that the genes B and C are located on the chromosome approximately 3 times as far from each other, than the genes A and B.

If a linkage group consists of several genes, than, knowing the frequency of recombinant individuals we may form opinion about the mutual location of genes on the chromosome. Thus, if for the genes A and B the recombinant frequency is 6%, for genes B and C - 9% and for genes A and C - 15%, we may conclude, that the gene B is located between the genes A and C, nearer to A. The idea becomes especially obvious if we replace the numbers by the segments with lengths proportional to these numbers. Then the data may be presented as:



The resulted projection will be the following:

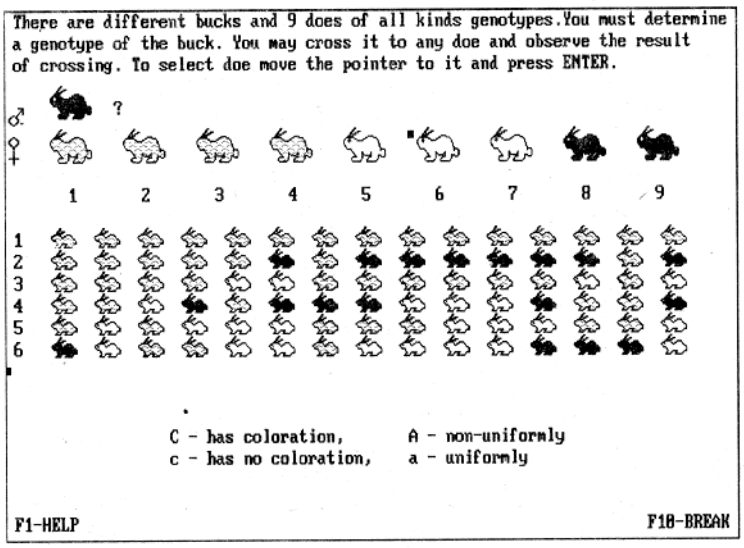


This scheme of genes mutual location on the chromosome is called a genetic map.

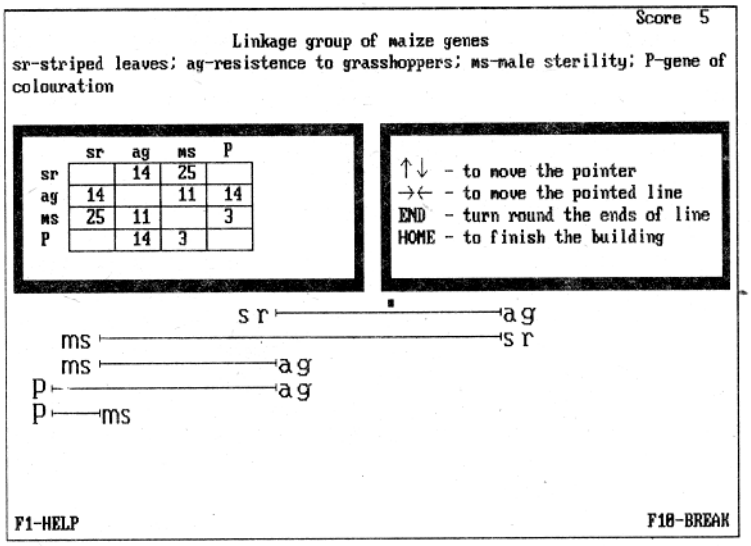
The program offers the learner to analyze several linkage groups. The genetic composition and the matrix of frequencies of recombinations are given for each group. Together with the matrix the segments with lengths proportional to corresponding probabilities are presented. The segments ends are marked with names of corresponding genes.

There is a pointer which moves in the directions up - down and left - right. Besides that, the segment may be turned round.

The learner's task is to dispose the segments so, that the

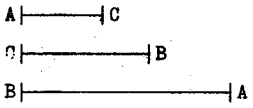


Pic.3. Interaction of genes.



Pic.4. Genetic map constructing.

similar ends were placed under each other, and, consequently, were projected to the same point. For example, for the segments:



the following sequence of operations is prescribed:

1. To move the pointer to the second segment.
2. To move the segment to the right so that the end C were under the end C of the first segment
3. To move the pointer to the third segment.
4. To turn round the segment.
5. To report, that the job is finished.

After that the program checks up on the work done. If the similar ends are really placed under each other, the common projection of the segments is displayed. This projection is just a map of a given linkage group. The learner may dispose the segments differently, in the order B-C-A. If in this case the similar ends are placed under each other, nevertheless, the disposition is recognized as correct. Otherwise the message like "Error for the gene A" appears, and the learner continues working with the same linkage group. After achieving the correct answer, he is offered the next linkage group (Pic.4).

When asking for help, the correct genetic map is displayed.

1.6.Decoding of genetic code

As is known, each protein, synthesized in an organism is coded by a certain gene, which is a site of a ribonucleic acid (RNA). A RNA sequence is composed of 4 kinds of nucleotides. Proteins consist of 20 kinds of aminoacids. Each aminoacid is coded by a sequence of 3 nucleotides (triplet). There exist 64 various triplets. One triplet codes a single aminoacid, some triplets code no aminoacids. The majority of aminoacids are coded by several triplets.

The program simulates the following experiment: there is a site of RNA and a site of a protein (a sequence of aminoacids), coded by it. The reading frame is unknown, that is it is not known whether the reading proceeds from the first, second or the third nucleotide. The learner may change the RNA. The program will change the protein right away. The learner's task is to

decode the table of genetic code, that is to decide which triplets code which aminoacids.

To solve the problem the learner should place the reading frame correctly. The consideration, that one triplet codes a single aminoacid may help here.. If after placing the frame the picture appears such that one triplet codes different aminoacids-

```
AGC UGA CGG UGA AAG
Ala Asp Gly Glu
```

then the frame is wrong. Let us shift it:

```
A GCU GAC GGU GAA AG
  Ala Asp Gly Glu
```

For this choice of the reading frame no conflict situations arise. This means (but doesn't guarantee) that the frame is placed correctly. The program presupposes at least one conflict situation for each wrong frame.

The program has 2 regimes of work: "theorist" and "practician". The theorist may replace any nucleotide of RNA by a needed one. The practician may influence the RNA with one of 3 mutagenous factors: chemical, radiation and ultra-violet which gives rise to random mutations (with different frequency). In both cases a new nucleotide will be read from a modified RNA.

Then the learner must place the data obtained in the table. The place of aminoacid in the table is determined in the following way: the first letter of the coding triplet stands on the left, the second letter - up, and the third letter - on the right. Therefore the data : GCU-Ala, GAC-Asp, GGU-Gly, GAA-Glu must be placed in the table such as:

	U	C	A	G	
G		Ala	Asp Glu	Gly	U C A G

At a round time the learner is supposed to fill in a table's section, containing 16 positions. The protein's site contains 12 aminoacids. The learner will have to make several mutations in RNA to fill in this section. There is a pointer which moves either along the table or, when switched, along the RNA in the "theorist" regime and along the picture with mutagenous factors in the "practician" regime (Pic.5).

When working with the program the learner may ask for help. Here the following messages are possible:

1. AGU-Asp and GGU-Asp.
   Change the frame (INS).
2. Switch the pointer (HOME).
3. Fill in Met on the place of AUG.
4. The mutations are needed.
   Switch the pointer (HOME).
5. Move the pointer to any factor
   and press ENTER.
6. Obtain the triplet AAC in the mRNA.

The program treats the errors. Here the following messages are possible:

1. This place in the table is filled. To clear the place - BS.2.
   The amino acid isn't obtained in the experiment.
3. This triplet isn't obtained in the experiment.
4. The data are not from experiment.

The program proceeds by counting correct operations, errors, asking for help. The results are stored in the file BIO.RES.

2. Biochemistry

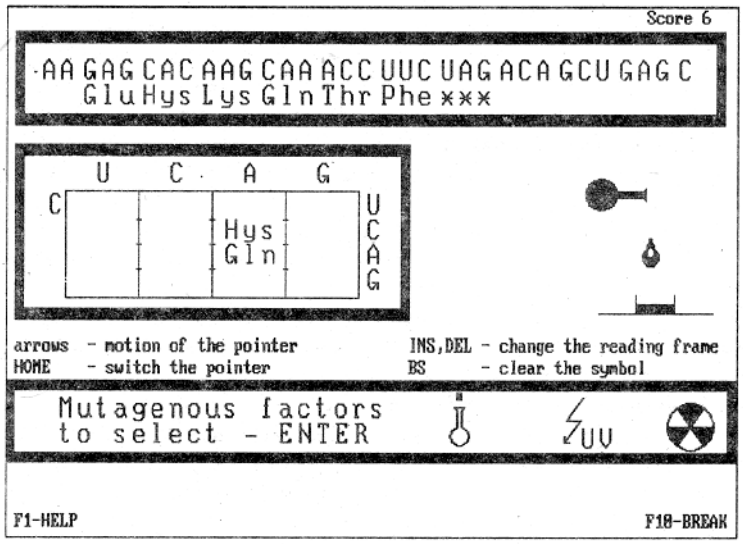
2.1. Lipids

Lipids are important organic substances of a living organism. Their specific structure makes them structural elements of membranes and cell organelles.

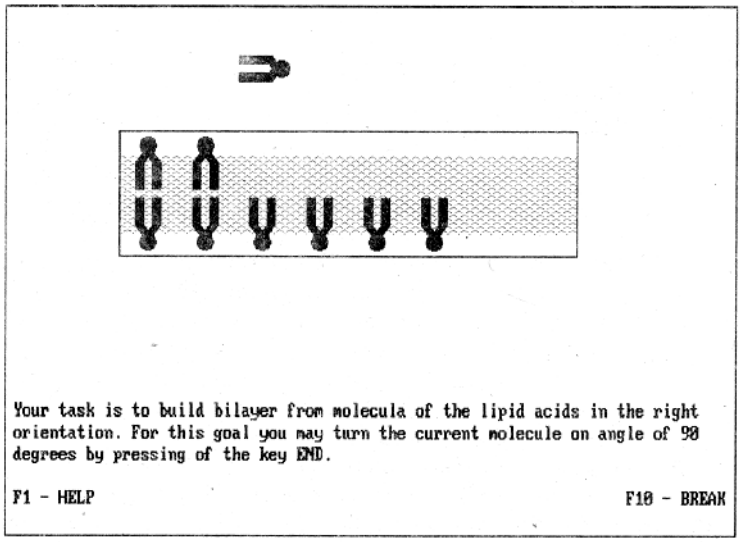
The simplest lipids are fatty acids. A molecular of fatty acid has the polar ends. The hydrocarbonic end is hydrophobous and is always oriented from water; the ionic end dissolves in water. Due to this special feature the molecules of fatty acids are capable to form a monomolecular layer disposed on the separative surface between water and non-polar organic substances, oil, for example.

The other important specimens of lipids are phospholipides. They also have the ionic hydrofillous head and hydrophobous hydrocarbonic ends.

The learner is offered 2 tasks on constructing a monolayer from molecules of fatty acids and 1 task on constructing a bilayer from molecules of phospholipides. At the beginning of work the display shows a separative surface between water and oil, bombarded by arbitrary oriented molecules of a fatty acid. The learner's task is to attach proper orientation to these molecules (ionic end should be oriented to water) at a fall time. In the second task water and oil change their places. In the



Pic. 5. Decoding of genetic code.



Pic.6. Lipids.

third task the molecules of phospholipides move from above and from below and the learner must orient them properly to construct a bimolecular layer (ionic heads should be oriented outside) (Pic.6).

The program calculates a percent of properly oriented molecules in the constructed object.

When asked for help, the program reminds that ionic heads should be oriented to water.

2.2. Structure formulas of aminoacids

Aminoacids are important organic substances in an organism, components of proteins. Their structure has as general as specific features. To master these features the learner should make experiments in the sphere of aminoacid constructing.

For this mean the program gives the learner a constructor with details presented by components of the following structure formulas:

- carbon;
- hydrogen;
- aminogroup;
- carboxyl group;
- hydroxyl group (ion OH^-);
- sulphur;
- groups CH_2 and CH_3.

The constructor contains 24 details. The free links of the details as well as their orientation are pointed out. For example, there are 4 details, consisting of one hydrogen atom with a single free link. This link is oriented up, to the right, down, to the left in various details.

The learner must construct 12 formulas of aminoacids. The work with a current aminoacid starts with informing the learner about the number of each type atoms in the formula. The field for constructing and the constructor are displayed. Each of these objects has a pointer. One of two pointers is active (it is crimson). The learner may:

- switch the active pointer (the key HOME);
- move the active pointer along the object (arrows);
- select a marked detail in the constructor and place it in the marked section of the field (the key ENTER);
- erase a detail from the marked section (the key BS).

In the course of constructing the following parameters are

constantly controlled:

- does the number of each type atoms exceed;
- do the links between parts agree.

When the formula is assembled, that is an object containing the needed number of each type atoms is constructed, the check-up is performed. It is controlled whether the obtained object is an amino acid, in other words does it contain amino-group, carboxyl group and hydroxyl group (Pic.7).

If the learner performed constructing without breaking the rules, but got an object different from really-occurred (isomer), he is informed about a natural layout of the object. The constructing is appreciated as correct.

The correct layout of the formula is also given, when asked for help.

In the course of constructing a count is carried out. For each correctly constructed formula the learner get a score proportional to a number of components in it. The errors (exceeding number of atoms, disagreement of links between components) are penalized. Getting of an object which is not an amino acid is penalized most strongly.

1.3. Secondary structure of a transfer RNA

Ribonucleic acids are important components of all living organisms, all living cells. Nucleic acids participate in protein formation. RNA within but a single cell differ in dimension, composition, functions and location.

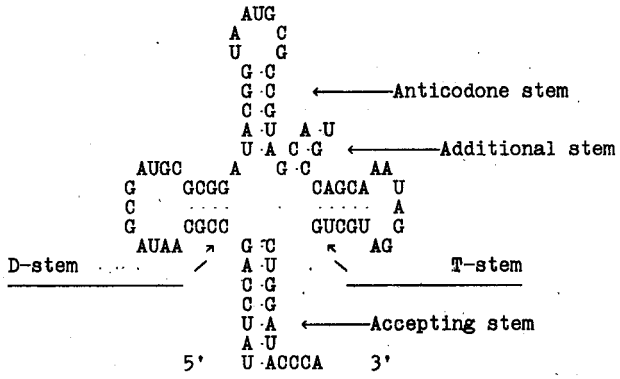
Transfer RNA are the most tiny molecules among Ribonucleic acids. They contain from 75 to 90 nucleic units. The tRNA function is to carry amino acids to ribosomes and place them in certain sites of a polypeptide chain at its biosynthesis. In that way tRNA participates in a translating process, playing the role of a somewhat translator: it translates a sequence of nucleotides to a sequence of aminoacid residua of a protein molecule.

Since the program is intended for children, the model is rather simplified. Thus it is supposed, that the tRNA contains only four types nucleotides: A, U, G, C.

The tRNA molecule presents by itself a single nucleotide chain, twisted to itself. It forms a complex space structure. All the tRNA are constructed along a similar scheme and can be described by the "clover leaf" model, which is based on the principle, that the number of hydrocarbonic links between nitrous

bases tends to be as much as possible.

The "clover leaf" contains five spiral stems, four of which end with loops of unpaired nucleotides. The center is occupied with non-spiral region.



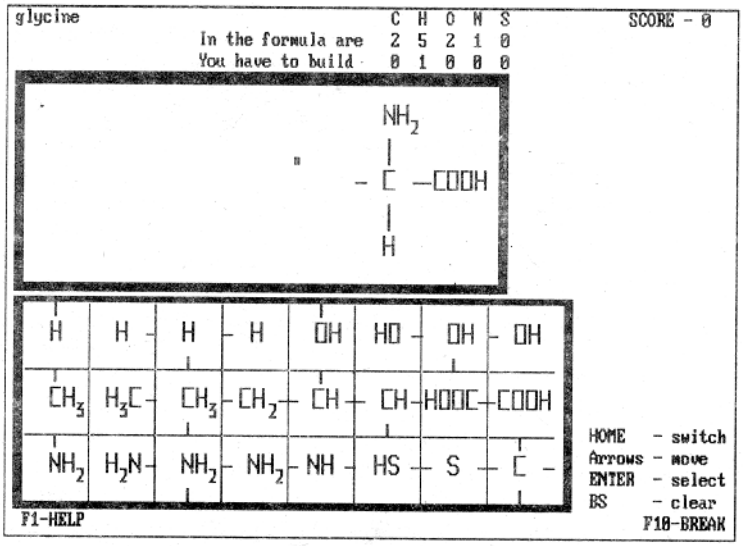
3' and 5'-ends of a nucleotide chain are paired and form the accepting stem. This is the longest spiral site (it is 7 pairs long). It is completed at the 3'-end in the majority of cases with the unpaired sequence CCA. It attaches a corresponding amino acid through its COOH-group. The aminoacid is combined with the corresponding tRNA due to a special enzyme.

An opposition to the accepting enzyme is made by the anticodone stem. It numerates 5 pairs of nucleotides longwise and carries an anticodone loop, consisting of 7 nucleotides. The anticodone loop contains in its middle part an anticodone, consisting of 3 nucleotides, complementary to the codon of given aminoacid in the matrix RNA. For example, the codon 5'-GCC-3' in mRNA is in correspondence with the anticodon 3'-CGG-5'. The latter attaches specific character to the interaction between tRNA and the matrix RNA. Each tRNA has its specific anticodon.

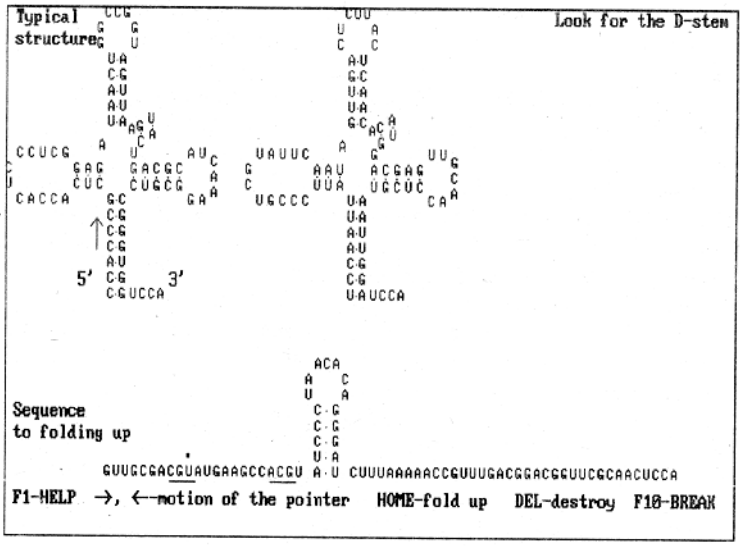
The T-stem carries a loop of 7 nucleotide residua, among which there is a special site, called a pseudouridine loop. Probably it is just this loop which interacts with a ribosome.

The D-stem carries a loop of 8-12 nucleotides. This is a dihidrouridine loop. It is believed, that it participates in the interaction with a specific activating enzyme. All the tRNA contain the 5-th additional stem of varied length. Most likely it participates in equalizing the lengths of various tRNA molecules.

The learner's task is to get a secondary structure of tRNA.



Pic.7. Structure formulas of aminoacids.



Pic.8. Secondary structure of a transfer RNA.

For this mean he may look for the spiral sites in the following way. At the beginning of work all the nucleotides are unpaired. There is a pointer which may be moved along the sequence. If a current nucleotide is the begin of stem (may be linked) then the left part and right part of this stem are underlined, for example:

CCAAUGCAAAUAGCAUUCA

He may links the stem. Then he will see:

```
  AAU
 A   A
  C G
  G C
  U A
  A U
  A U
CC  CA
```

If he doesn't link the stem he continues the search and may find some other stems. Among some sites obtained in such a way the learner must choice one. If later he will find that the choice was ineffective, he may "unlink" this site.

The search is realized by stages. To facilitate the choice the folded RNA is depicted in the screen. Those its site, which is looked for at the moment is being marked by color each time.

Since the last stage of the work is a search of the accepting stem (7 pairs of bases), this site can't be fond if some previous sites are chosen erroneously (Pic.8).

The learner is offered 2 sequences to fold up. At the end of the work he will see 3 folded sequences: the initial sample and two ones, obtained by himself. Then he is asked some questions concerning the general properties of the tRNA secondary structures. To answer these questions he should analyze the screen contents.

The final part of the program concerns the tRNA functions. The screen retains the anticodone loops and the stems of the folded tRNA and demonstrates a table of genetic code. The learner must determine which codons correspond to given anticodons and with which aminoacids will interact these tRNA.

When asking for help at various stages of the work, the following information is given:

- which particular stem is looked for and what is its place on the sequence;
- is it necessary to fold up the stem or the search must be continued;
- does the recently folded stem must be unlinked.

The answers are valued. At the end of work the percent of correct answers is given.

3. Biosynthesis of protein

3.1. Demonstrational program

The program presents the following scheme of biosynthesis:

1. Building of the messenger RNA.
2. Initiation:
   - assembling of a ribosome from two subunits;
   - entrance of mRNA into the ribosome.
3. Elongation:
   - carrying of amino acids into the ribosome by transfer RNA;
   - moving the ribosome along the mRNA.
4. Termination.

The program gives the learner the following controlling opportunities (Pic.9):

Pause - to retain the program;
Enter - to continue the program;
F10 - to break the program.

3.2. Simulational program

The program offers the learner just familiar with the scheme of biosynthesis due to the demonstrating program or some other sources to build up a protein on the basis of randomly generated DNA sequence. For this mean he should:

- point out the stages of biosynthesis in the correct order;
- input the parameters, necessary for each stage: nucleotides, triplets, amino acids, etc.

As the first stage the learner if offered to build up the mRNA. The DNA in the form of the duplicated strand is shown in the screen. The strand is generated randomly, the only condition is that it would start with an indicator and end with one of possible terminators. The strands come apart and the learner is offered to build on the RNA to the lower strand. When building on he should follow the principle of complementarity: A-U, T-A, G-C, C-G. Nucleotide, pointer out correctly, is connected to the RNA. The RNA, built up by learner, goes away from the nucleus and the strands of the DNA come together. Then the list of the biosynthesis stages is displayed.

Next, the assembling of the ribosome is fulfilled. The

ribosome contains two subunits 30S and 50S which are kept apart. One of them is being approached by aminoacids. When it meets a metionin among the aminoacids (the anticodon of its transport RNA is UAC), the subunits close up and the ribosome starts functioning as a single whole. At this stage the learner should activate of metionin by pointing out the anticodon of its tRNA (stage 2).

The next stage is connecting of the mRNA and the ribosome. In the program it looks as entrance of the mRNA into the ribosome. The learner must only point out this stage in-time, after the assembly of ribosome (stage 3).

Then the activation of needed amino acids is realized, that is connecting of amino acids with the tRNA, as a result of which the tRNA-aminoacid complex is formed. This complex is carried to the ribosome where it participates in the growth of the synthesized protein's polypeptide chain. The learner may either activate several amino acids for one making of this stage, or only one and then complete fulfillment of the stage. Amino acids, needed for building a protein are those, with anticodons of transfer RNA complementary to the triplet of mRNA. The learner's task is to name one or several anticodons. The program allows only one activation of the given type aminoacid in the same time. That is why if two molecules of the same aminoacids are needed to build up a protein, they may be obtained only for two making of the stage. The learner must activate the amino acid, use it and activated again (stage 4).

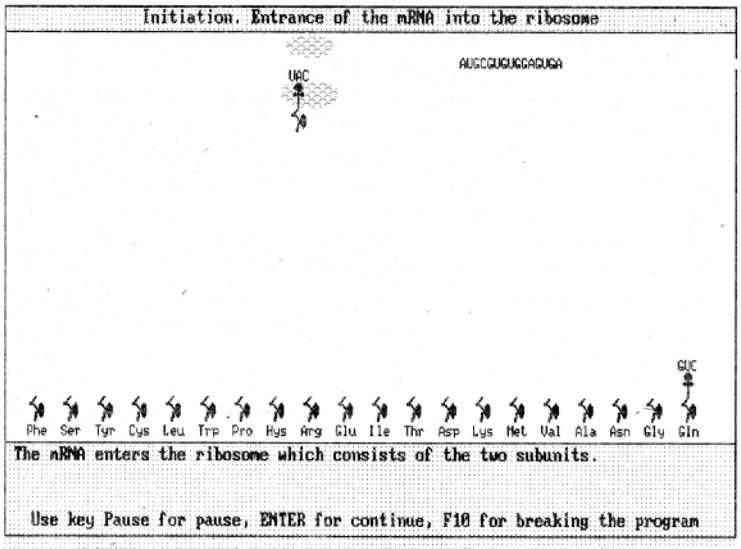
Next, carrying of the activated amino acid to the ribosome is fulfilled. Here the learner must point out the amino acid's code (stage 5).

The next stage is a growth of the polypeptide chain. When the ribosome contains two tRNA, one of which is connected with the growing polypeptide chain, and the other - with just entered amino acid, the polypeptide chain looses connection with the "old" tRNA and connects the "new" amino acid. The "old" tRNA go away from the ribosome after that (stage 6).

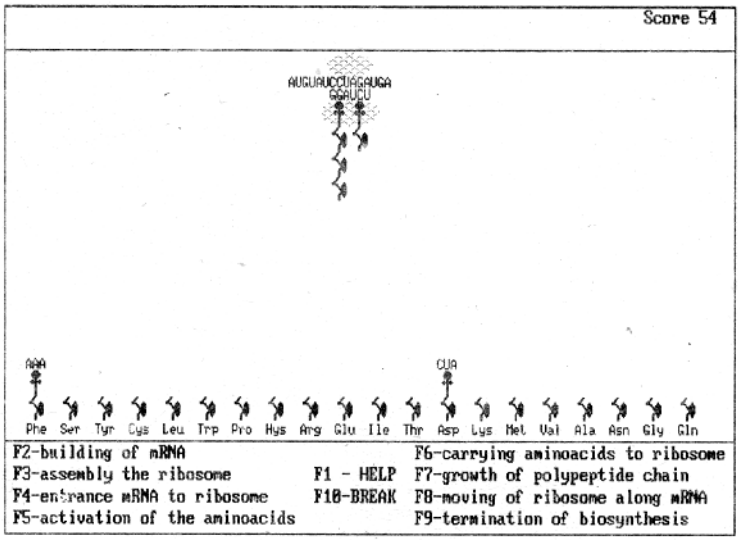
Then, movement of the ribosome along the mRNA by three nucleotides is fulfilled (3 executions of the stage 7).

After that the stages 5-7 should be fulfilled two times more each. If the learner when fulfilling the stage 4 has managed to activate all needed aminoacids, there is no need to repeat this stage. Otherwise stage 4 must be fulfilled once more too(Pic.10).

After the stage 7 is fulfilled 9 times (three series by three



Pic.9. Biosynthesis of protein. Demonstrational program.



Pic.10. Biosynthesis of protein. Simulational program.

times), the stage "termination of the biosynthesis" follows. The screen demonstrates the dispersion of the participating objects (two subunits of the ribosome, mRNA, tRNA).

When working with the program the learner may ask for help. If the inquiry happened between stages, the learner is informed which stage should be the next. If the stage is fulfilled, the learner is reported depending on situation, which nucleotide, anticodon, code of aminoacid should be input, or which key should be used to replace the necessary objects along the screen.

In the course of working with the program the count is performed, which considers correct operations, errors, asking for help. At the end of the program the results are stored in the file BIO.RES.

4. The additional information

The diskette with the package contains 10 programs in the form of executed files (with expansion EXE) and the file with documentation BIODOC.TXT. In the course of work with the programs the file BIO.RES is formed, where the results are stored. The total volume of the programs is 962K; the documentation takes 41K.

CONTENTS

27.II.91 г. Зак.3754Р Тир.215 экз. Изд.№ 268. Уч.-изд.л. 1,6
Цена 80 к. Отпечатано на ротапринте в ОНТИ ПНЦ РАН