

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
МИКОЛАЇВСЬКИЙ НАЦІОНАЛЬНИЙ АГРАРНИЙ УНІВЕРСИТЕТ**

**Факультет технології виробництва і переробки продукції
тваринництва, стандартизації та біотехнології**

Кафедра біотехнології та біоінженерії

Генетичні ресурси сільськогосподарських тварин

Методичні рекомендації
для практичних робіт
для здобувачів першого (бакалаврського) рівня вищої освіти ОПП
«Біотехнології та біоінженерія» спеціальності 162 «Біотехнології та
біоінженерія» денної форми здобуття вищої освіти



Миколаїв
2024

УДК 636.082:575

Г34

Друкується за рішенням науково-методичної комісії факультету технології виробництва і переробки продукції тваринництва, стандартизації та біотехнології Миколаївського національного аграрного університету від «20» листопада 2024 р., протокол № 3.

Укладачі:

О. С. Крамаренко - канд. с.-г. наук, доцент кафедри біотехнології та біоінженерії, Миколаївський національний аграрний університет.

Рецензенти:

П. А. Ващенко – д-р с.-г. наук, старший науковий співробітник, професор кафедри технології виробництва продукції тваринництва, Подільський державний аграрний університет;

С. С. Крамаренко – д-р біол. наук, професор, професор кафедри біотехнології та біоінженерії, Миколаївський національний аграрний університет.

З М І С Т

| | |
|--|----|
| Вступ | 4 |
| Методика математико-статистичного аналізу імуногенетичних даних для великої рогатої худоби | 7 |
| Розрахунок популяційно-генетичних параметрів для імуногенетичних даних з використанням програми GenAIEx | 12 |
| Етап 1. Розрахунок вибірових частот антигенів та показників генетичної мінливості в популяціях | 17 |
| Етап 2. Оцінка міжгрупової (міжпопуляційної) генетичної диференціації за антигенними факторами | 27 |
| Етап 3. Візуалізація груп у просторі генетичної мінливості з використанням Аналізу Головних Координат (РСоА) | 34 |
| Етап 4. R- та Q-аналіз зчепленого успадкування антигенних факторів | 36 |
| Список використаної та рекомендованої літератури | 49 |

ВСТУП

Дисципліна «Генетичні ресурси сільськогосподарських тварин» розрахована на підготовку бакалавр з біотехнології та біоінженерії із освітньої спеціальності 162 «Біотехнології та біоінженерія» і займає провідне місце в системі навчання. Дисципліна є базовою для вивчення спеціальних курсів.

Дисципліна вивчає походження сільськогосподарських тварин та їх еволюцію, біорізноманіття і генофонд свійських тварин, генетичні основи їх доместикації; процеси породоутворення, класифікації та характеристики порід свійських тварин, їх породне районування, напрямки селекційно-племінної роботи із різними видами та породами сільськогосподарських тварин; структури, характеристики, значення та основні напрямки діяльності племінних підприємств України.

Мета дисципліни – вивчення походження сільськогосподарських тварин, їх еволюції і селекції; процесів породоутворення, напрямків селекційно-племінної роботи; структур, характеристик, значення та основних напрямків діяльності племінних господарств України.

Завдання дисципліни: сформувані у здобувача вищої освіти систему теоретичних та практичних навичок з питань оцінки та використання генетичних ресурсів сільськогосподарських тварин.

Предмет дисципліни: походження сільськогосподарських тварин, їх еволюція і селекція; процеси породоутворення, напрямки селекційно-племінної роботи; структури, характеристики, значення та основні напрямки діяльності племінних господарств України.

Об'єкт дисципліни: генетичні ресурси сільськогосподарських тварин.

Інтегральна компетентність

Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми, що характеризуються комплексністю та невизначеністю умов у

біотехнології та біоінженерії, або у процесі навчання, що передбачає застосування теорій та методів біотехнології та біоінженерії.

Загальні компетентності:

К04. Навички використання інформаційних і комунікаційних технологій;

К07. Прагнення до збереження навколишнього середовища;

Спеціальні (фахові, предметні) компетентності:

К11. Здатність використовувати ґрунтовні знання з хімії та біології в обсязі, необхідному для досягнення інших результатів освітньої програми;

К13. Здатність працювати з біологічними агентами, використовуваними у біотехнологічних процесах (мікроорганізми, гриби, рослини, тварини, віруси, окремі їхні компоненти);

К14. Здатність здійснювати експериментальні дослідження з вдосконалення біологічних агентів, у тому числі викликати зміни у структурі спадкового апарату та функціональній активності біологічних агентів;

К24. Здатність дотримуватися вимог біобезпеки, біозахисту та біоетики.

Додаткові компетентності:

К26. Здатність розробляти та застосовувати на практиці нові біотехнології, що дозволяють підвищити ефективність тваринництва.

Програмні результати навчання:

ПР11. Вміти здійснювати базові генетичні та цитологічні дослідження з вдосконалення і підвищення біосинтетичної здатності біологічних агентів з урахуванням принципів біобезпеки, біозахисту та біоетики (індукований мутагенез з використанням фізичних і хімічних мутагенних факторів, відбір та накопичення ауксотрофних мутантів, перенесення генетичної інформації тощо);

ПР14. Вміти обґрунтувати вибір біологічного агента, складу поживного середовища і способу культивування, необхідних допоміжних робіт та основних стадій технологічного процесу;

Додаткові програмні результати:

ПР25. Вміти розробляти та застосовувати на практиці нові технології, що дозволяють підвищити ефективність тваринництва: техніку трансплантації і мікрomanipуляцій на ембріонах домашніх тварин, отримання кормових засобів (білок, амінокислоти, вітаміни) мікробіологічним синтезом.

МЕТОДИКА МАТЕМАТИКО-СТАТИСТИЧНОГО АНАЛІЗУ ІМУНОГЕНЕТИЧНИХ ДАНИХ ДЛЯ ВЕЛИКОЇ РОГАТОЇ ХУДОБИ

Імуногенетичні дані, отримані після проведення серологічних досліджень для великої рогатої худоби мають деякі особливості, що інколи ускладнюють їх аналіз із популяційно-генетичної точки зору. Останній підхід потребує наступної інформації:

- оцінок частот фенотипів, генотипів та алелей за певними локусами (тобто, для відповідних еритроцитарних антигенів груп крові);
- оцінок рівня генного різноманіття у самому широкому розумінні (наприклад, за М.Неєм);
- визначення ступеня генетичного інбридингу;
- оцінки рівня генетичної диференціації між групою популяцій або інших зоотехнічних одиниць;
- встановлення рівня генетичної подібності між групами популяцій або інших зоотехнічних одиниць (наприклад, за М.Неєм, Л.Животовським і т.і.);
- та інші.

На відміну від результатів аналізу білкового поліморфізму крові чи молока, або молекулярно-генетичних маркерів із кодомінантним типом успадкування (наприклад, мікросателітів, тощо), при проведенні імуногенетичного аналізу великої рогатої худоби дуже важко отримати повну генотипову формулу у вигляді A/B , де A та B – різні алельні форми для певного локусу. Для цього необхідно мати результати не лише

досліджуваної тварини (пробанда), і його мати та батька. Такий повний генетичний аналіз часто не проводиться, тому результатом імуногенетичної експертизи тварини найчастіше є лише її **фенотип**, тобто, формульний запис, який складається із переліку антигенів крові, що були відмічені при постановці серологічного аналізу.

Другою важливою особливістю імуногенетичного аналізу великої рогатої худоби є дуже висока кількість алелей, зареєстрованих для деяких антигенних систем. Наприклад, за 12 генетичними системами антигенів великої рогатої худоби на теперішній час виділено більше 100 антигенів, в яких враховано більше, ніж 500 алелей (більшість яких входять до систем *B* та *C*). Наявність великого алельного різноманіття, можливість прояву у особини декількох алелей за однією системою призводить до ситуації, коли в популяції неспоріднених тварин кожна зареєстрована алель (насамперед, за системою *B*) є унікальною, що унеможлиблює проведення популяційно-генетичного аналізу за тим планом, що було викладено вище.

Тому нами було запропоновано новий, більш адекватний і технічно спрощений, підхід до аналізу імуногенетичної мінливості великої рогатої худоби. Він полягає в наступному.

В цілому нами під час аналізу було виявлено наявність 53 антигени (табл. 1). В такому разі всю імуногенетичну формулу повністю (тобто, у випадку, коли зустрічаються всі 53 антигени) можна записати наступним чином:

$$A_1 A_2 Z' B_2 G_2 G_3 K I_1 I_2 O_1 O_2 P_2 Q T_1 T_2 Y_2 A'_1 \dots U U' H'' U'' Z.$$

Ми пропонуємо перейти від такої загальноприйнятої форми запису імуногенетичного фенотипу будь-якої тварини до розглядання його у вигляді складного бінарного запису, де для кожної алелі, яку було відмічено у групі тварин, що аналізується, необхідно ставити символ “1”, якщо ця алель серологічно фіксується, або символ “0”, якщо аглютинація не

відбулася. В цьому випадку, наприклад, для особини, що має за системою *A* лише одну алель – *A*₂, формула буде мати наступний вигляд: 0 1 0 ... Аналогічно кодується наявність/відсутність кожного із 53 антигенів (алелей), що зареєстровані у кожній аналізованій тварини.

Таблиця 1

Перелік антигенів великої рогатої худоби, виявлених під час дослідження

| Системи | | | | | | | | |
|-----------------------|-------------------------|-----------------------|----------|----------|----------|----------|-----------------------|----------|
| A | B | C | F | J | L | M | SU | Z |
| <i>A</i> ₁ | <i>B</i> ₂ | <i>C</i> ₁ | <i>F</i> | <i>J</i> | <i>L</i> | <i>M</i> | <i>S</i> ₁ | <i>Z</i> |
| <i>A</i> ₂ | <i>G</i> ₂ | <i>C</i> ₂ | <i>V</i> | | | | <i>H</i> ' | |
| <i>Z</i> ' | <i>G</i> ₃ | <i>E</i> | | | | | <i>U</i> | |
| | <i>K</i> | <i>R</i> ₁ | | | | | <i>U</i> ' | |
| | <i>I</i> ₁ | <i>R</i> ₂ | | | | | <i>H</i> '' | |
| | <i>I</i> ₂ | <i>W</i> | | | | | <i>U</i> '' | |
| | <i>O</i> ₁ | <i>X</i> ₁ | | | | | | |
| | <i>O</i> ₂ | <i>X</i> ₂ | | | | | | |
| | <i>P</i> ₂ | <i>C</i> ' | | | | | | |
| | <i>Q</i> | <i>L</i> ' | | | | | | |
| | <i>T</i> ₁ | | | | | | | |
| | <i>T</i> ₂ | | | | | | | |
| | <i>Y</i> ₂ | | | | | | | |
| | <i>A</i> ' ₁ | | | | | | | |
| | <i>A</i> ' ₂ | | | | | | | |
| | <i>B</i> ' | | | | | | | |
| | <i>D</i> ' | | | | | | | |
| | <i>E</i> ' ₂ | | | | | | | |
| | <i>G</i> ' | | | | | | | |
| | <i>I</i> ' | | | | | | | |
| | <i>K</i> ' | | | | | | | |
| | <i>J</i> ' ₂ | | | | | | | |
| | <i>O</i> ' | | | | | | | |
| | <i>P</i> ' | | | | | | | |
| | <i>Q</i> ' | | | | | | | |
| | <i>Y</i> ' | | | | | | | |
| | <i>B</i> '' | | | | | | | |
| | <i>G</i> '' | | | | | | | |

У цьому випадку загальна матриця із вихідними даними при запису їх у табличному редакторі MS Excel буде мати наступний вигляд (рис. 1).

З одного боку така форма запису уніфікує дані, отримані для різних тварин, а з іншого боку – вже дозволяє провести розрахунки популяційно-генетичних параметрів, які були надані вище.

В такому вигляді (набір одиниць та нулів) дана форма запису результатів імуногенетичних досліджень певної тварини може бути визначена як її **гаплотип**. Майже аналогічний підхід до визначення гаплотипного різноманіття використовується і у випадку аналізу структури послідовності ДНК за участю певного набору ендонуклеаз рестрикції. В більш широкому розумінні, *гаплотипом* називають комбінацію алелей тісно зчеплених локусів.

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | AA | AB | AC | AD | AE | AF | AG | | | | |
|----|----------------|-----------|----|----|----|----|----|---|----|----|----|----|----|-----------|----|----|----|-----|-----|----|----|-----|----|----|----|-----|----|----|----|----|-----|-----|----|---|---|---|---|
| 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2 | | Система А | | | | | | | | | | | | Система В | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | | A1 | A2 | Z' | B2 | G2 | G3 | K | I1 | I2 | O1 | O2 | P2 | Q | T1 | T2 | Y2 | A'1 | A'2 | B' | D' | E'2 | G' | I' | K' | J'2 | O' | P' | Q' | Y' | B'' | G'' | C1 | | | | |
| 4 | кличка и номер | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 5 | актриса | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 6 | русалка | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | | |
| 7 | добрая | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | | |
| 8 | роботка | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | | |
| 9 | палатка | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 10 | десятка | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | | |
| 11 | лаванда | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 12 | касятка | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 13 | разлука | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 14 | балерина | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 15 | кукла | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | |
| 16 | веха | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | нимфа | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 18 | рассада | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 19 | красуня | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 20 | рыбка | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | |
| 21 | василек | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| 22 | иволга | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 23 | река | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 24 | арка | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 25 | лирика | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 26 | зухра | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 27 | лиска | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Рис. 1. Зовнішній вигляд електронного аркуша MS Excel із результатами імуногенетичного аналізу дослідної популяції великої рогатої худоби

Така складна бінарна форма запису вихідних даних має ряд переваг.

1. Вона відображує увесь набір антигенів, які були виявлені для певної тварини. Тобто, така форма є комплексною, що забезпечує важливу перевагу при аналізі генетичної структури популяції чи групи популяцій. При цьому враховується незалежне успадкування кожного антигенного фактора, а також незалежне успадкування самих систем антигенів.

2. Вона відмічає наявність того чи іншого антигену, хоча при цьому важко визначити – чи є певна тварина гомозиготною чи гетерозиготною за даним геном (антигеном), оскільки немає інформації про генотип її батьків. Але ми можемо використати навіть таку неповну інформацію, якщо враховуватимемо, що система успадкування кожної алелі є домінантна. В цьому випадку, тварин із відсутніми тими чи іншими антигенами можна розглядати як рецесивні гомозиготи і на підставі їх частоти визначити частоту певної алелі.

3. Вона дозволяє проаналізувати і характер зчеплення груп антигенів, які формують алелі. При цьому для матриці, подібної до тієї, що наведено на рисунку 1, можна проводити два різних типи аналізів. R-аналіз дозволяє проаналізувати розподіл окремих тварин (чи їх груп) у багатовимірному просторі виявлених антигенів (чи їх груп) на підставі матриці їх подібності, розрахованої за допомогою міри Хеммінга (див. нижче). З іншого боку, Q-аналіз дозволяє на підставі матриці подібності між окремими антигенами у різних тварин проаналізувати ступінь зчепленого успадкування між антигенами та встановити, вклад яких антигенів (чи їх груп) більш вагомий при диференціації тварин різного походження.

4. При появі нового антигенного фактора він дописується у кінці ряду антигенів відповідної системи у повній бінарній формулі, а для попередніх тварин та їх груп у цьому стовпчику матриці проставляються нулі.

5. Нарешті, такий підхід дозволяє значно автоматизувати як процес накопичення та збереження імуногенетичної інформації (наприклад, за

допомогою редактора електронних таблиць MS Excel), а також використовувати сучасне програмне забезпечення для проведення популяційно-генетичного аналізу імуногенетичних даних.

РОЗРАХУНОК ПОПУЛЯЦІЙНО-ГЕНЕТИЧНИХ ПАРАМЕТРІВ ДЛЯ ІМУНОГЕНЕТИЧНИХ ДАНИХ З ВИКОРИСТАННЯМ ПРОГРАМИ GenAIEx

Всі розрахунки ми проводили з використанням програми “Генетичний аналіз в Excel’е” (GenAIEx v. 6.0), що була розроблена австралійськими вченими-генетиками Р.Пиколлом та П.Смаусе у 2006 році (Peakall, Smouse, 2006). Ця програма має декілька переваг. По-перше, вона являє собою програму типу Add-In, тобто, вбудовується в табличний редактор MS Excel і використовує дані, що набрані у цьому редакторі. А по-друге, ця програма розповсюджується безкоштовно і її можна вільно скачати в Інтернеті на сайті авторів (www.anu.edu.au/BoZo/GenAIEx). Нарешті, по-третє, вона здатна розрахувати популяційно-генетичні параметри, що нас цікавлять, на підставі бази імуногенетичних даних, яка записана у бінарній формі як, наприклад, наведено на рисунку 1.

Але для проведення необхідних розрахунків по-перше необхідно встановити програму. Для цього необхідно відкрити новий аркуш MS Excel. У головному Меню знайти опцію “**Сервіс**” і у ньому вибрати процедуру “**Надстройки**” (рис. 2).

Далі необхідно клацнути кнопку “**Обзор**”. Знайти папку “GenAIEx_6” , відкрити її, обрати файл **GenAIEx6.xla** та клацнути кнопку **ОК** (рис. 3).

Програма GenAIEx 6 з’явиться у переліку Надбудов на вашому комп’ютері. У віконці напроти цієї програми повинна стояти галочка. Для того, щоб її запустити необхідно натиснути на кнопку **ОК** (рис. 4). Тоді у переліку головного Меню з’явиться нова опція “**GenAIEx**”, а програма

почне працювати, внаслідок чого на фоні аркуша MS Excel з'явиться її сторінка (рис. 5), яка зникне через декілька секунд.

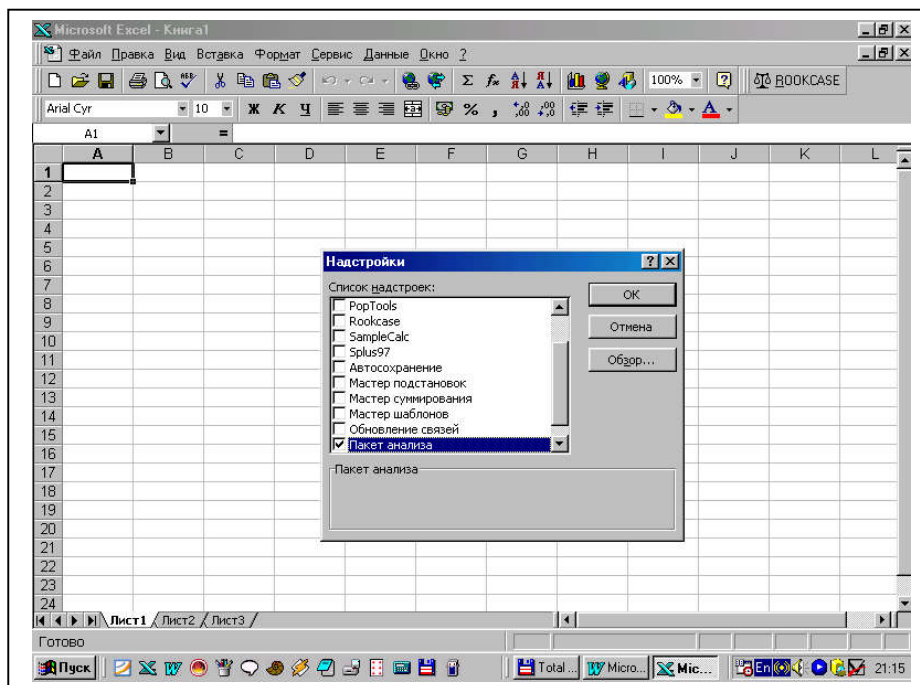


Рис. 2.

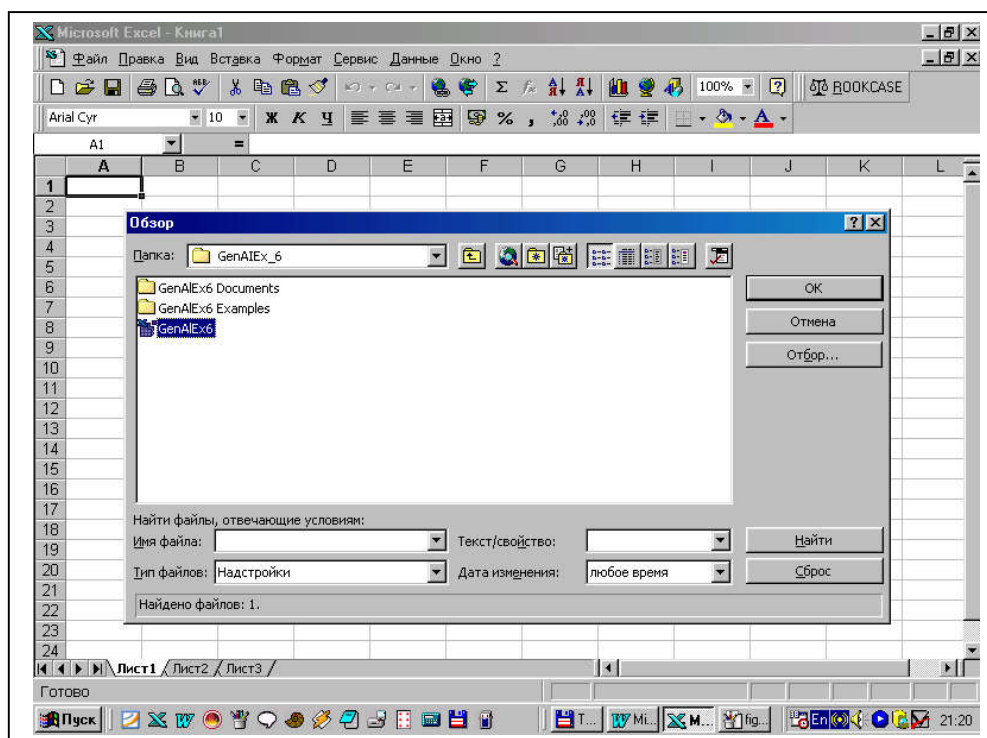


Рис. 3.

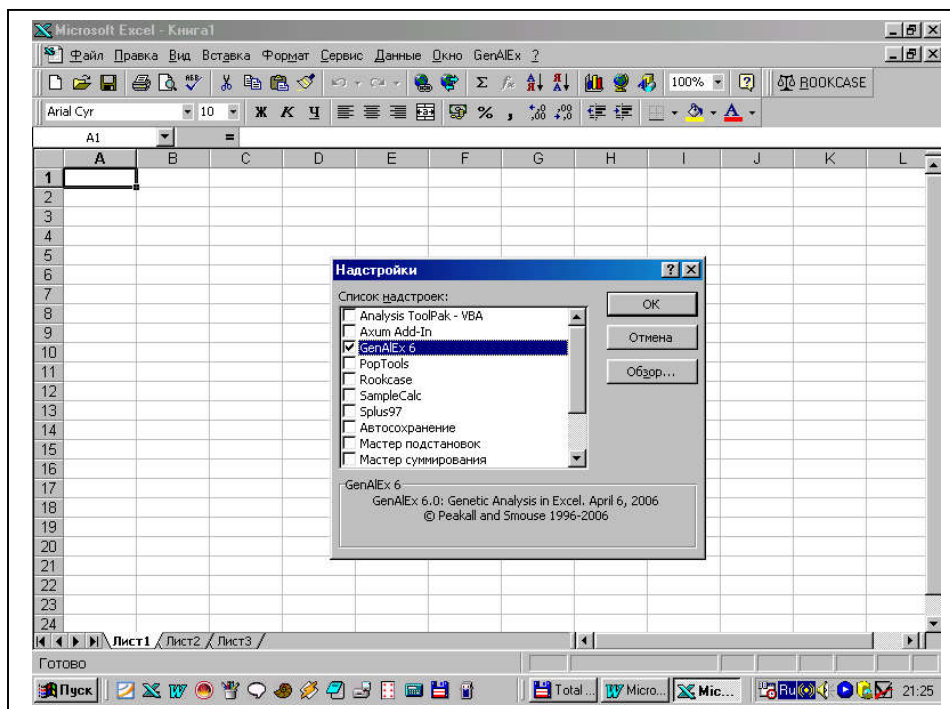


Рис. 4.

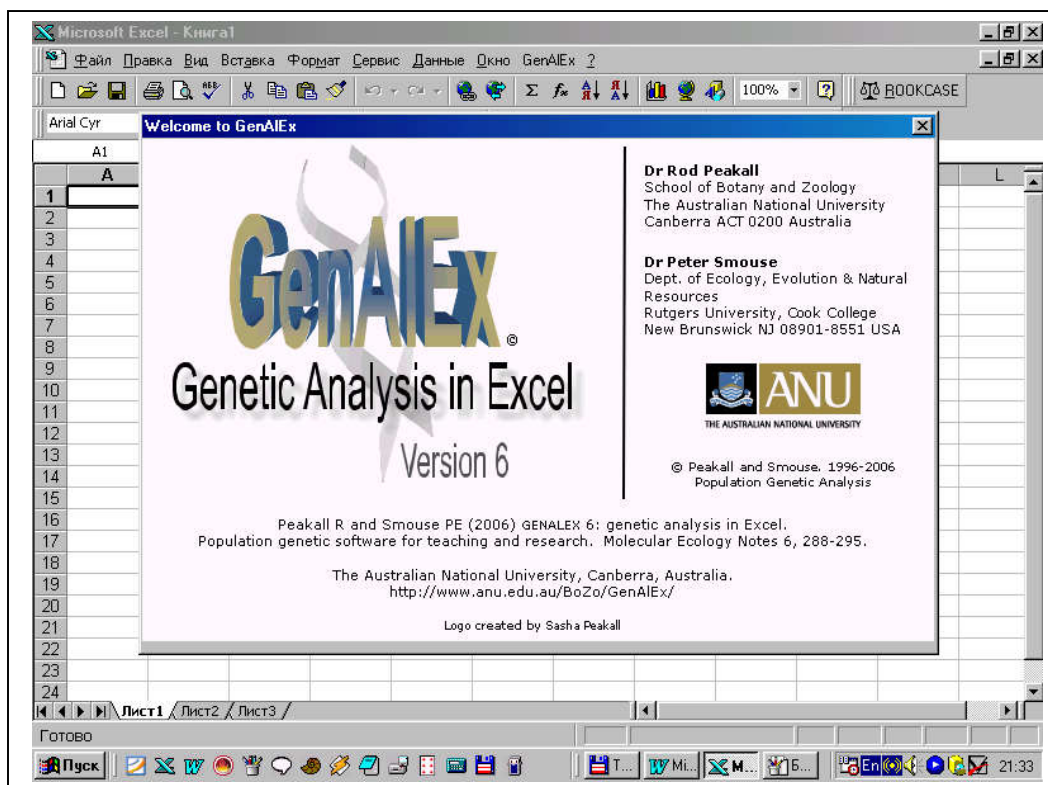


Рис. 5.

Далі необхідно сформувати аркуш з вихідними даними так, щоб їх могла прочитати програма. У загальному випадку на такому аркуші повинні бути наступні елементи (рис. 6):

| Sample | Pop | Locus1 | Locus2 | Locus3 | Locus4 | Locus5 | Locus6 | Locus7 | Locus8 | Locus9 | Locus10 | Locus11 | Locus12 | Locus13 |
|--------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|---------|---------|---------|
| 1 | Pop1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | Pop1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| 3 | Pop1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| 4 | Pop1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| 5 | Pop1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 6 | Pop1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| 7 | Pop1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| 8 | Pop1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 1 |
| 9 | Pop1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 1 |
| 10 | Pop1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 |
| 11 | Pop2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 |
| 12 | Pop2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| 13 | Pop2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 |
| 14 | Pop2 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 15 | Pop2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| 16 | Pop2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 1 |
| 17 | Pop2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 |
| 18 | Pop3 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| 19 | Pop3 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 20 | Pop3 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| 21 | Pop3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 22 | Pop3 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| 23 | Pop3 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 24 | Pop3 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 1 |
| 25 | Pop3 | 2 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 |
| 26 | Pop3 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Рис. 6.

| Клітка аркуша | Інформація |
|---------------|--|
| A1 | - кількість локусів, що включено до аналізу (у нашому випадку – загальна кількість еритроцитарних антигенів: 53), тобто, загальна кількість стовпчиків матриці вихідних даних; |
| B1 | - загальна кількість особин, які використовуються при аналізі, тобто, загальна кількість строчок матриці вихідних даних: 120; |
| C1 | - кількість груп (популяцій), що використовуються в аналізі; |

| | |
|------------|---|
| D1 – H1 | - чисельність кожної групи (в нашому випадку їх п'ять, тому використано п'ять кліток; зрозуміло, якщо груп більше, то використовуються додатково клітки праворуч); |
| I1 | - кількість сукупностей груп (в програмній мові – регіонів); в загальному випадку – це наступна клітка після останньої, що містить об'єм останньої групи; |
| J1 | - загальна кількість особин в кожній із сукупностей груп; в нашому випадку таких сукупностей одна, тому її чисельність дорівнює загальній чисельності досліджених тварин; але, якщо б таких сукупностей було б більше – їх чисельності необхідно було б занести послідовно в клітки праворуч; |
| A2 | - пояснення для набору вихідних даних (в нашому випадку – це дані, що використані в дослідженнях М.І.Гиль (2008), тому ми дали назву – Gill Data); |
| D2 – H2 | - імена для груп тварин (згідно програмної мови їм надаються імена – Pop1, Pop2, Pop3 і т.д.); якщо груп тварин (популяцій) більше, то використовуються наступні клітки праворуч; |
| J2 | - ім'я для сукупності груп (згідно програмної мови для таких сукупностей прийнято ім'я Region); якщо таких сукупностей груп більше, то використовуються наступні клітки праворуч; |
| A3 | - заноситься слово “Sample”, тобто “проба” англійською; |
| A4 – A123 | - послідовно заносяться номери тварин (для нашого випадку від 1 до 120); |
| B3 | - заноситься слово “Pop”, тобто скорочення від “популяція” англійською; |
| B4 – B123 | - послідовно заносяться кодові позначення належності кожної тварини до певної групи (для нашого випадку використовуються позначення Pop1, Pop2, ..., Pop5); таким чином зрозуміло, що тварина із порядковим номером, наприклад, №22 належить до третьої групи (має код Pop3) і т.д. |
| C3 – BC3 | - ім'я для кожного локусу (антигену); згідно програмної мови використовуються позначення Locus1, Locus2, ..., Locus53 (у нашому випадку використовується 53 антигени); |
| C4 – BC123 | - код для результатів серологічної реакції для кожної тварини по кожному антигену: у випадку, коли реакція |

| | |
|--|---|
| | негативна, тобто, антиген відсутній – заносимо цифру “1”, а коли реакція позитивна, тобто, антиген присутній – заносимо цифру “2” (згідно програмної мови цифра “0” означає відсутні або пропущені дані); таку матрицю з даними легко можна отримати із матриці, що наведено на рисунку 1, додавши до кожного значення одиницю. |
|--|---|

Після цього аркуш, що містить готову для аналізу вибірку даних, необхідно перейменувати і надати йому ім'я “D”, від англійського “Data” – дані.

Тепер наші імуногенетичні дані повністю готові до популяційно-генетичного аналізу. Продемонструємо послідовність етапів проведення такого аналізу.

Етап 1. Розрахунок вибірових частот антигенів та показників генетичної мінливості в популяціях.

Для цього необхідно (при відкритому аркуші із даними) знайти у Головному Меню опцію “**GenAIEx**” та клацнути по неї. В меню, що містить перелік всіх основних опцій програми, треба обрати опцію “**Frequency...**” (**Розрахунок частот**) та клацнути по неї (рис. 7).

З'явиться закладка “**Allele Frequency Data Parameters**” (**Параметри частот алелей**). В цій закладці автоматично буде заповнено майже всі клітки, але необхідно самому обрати “**Data Format**” (**Формат даних**), обравши тип “**Haploid**” (**Гаплоїдні**) (рис. 8). Після цього клацнути по кнопці **OK**.

З'явиться нова закладка “**Haploid Frequency Options**” (**Опції для розрахунку гаплоїдних частот**). Вона містить перелік показників, які можна вказати програмі для розрахунку. Насамперед, це:

- **Frequency by Pop** – розрахунок частот для окремих популяцій (із можливістю зобразити вихідні дані у графічній формі);
- **Haploid Diversity by Pop** – генетичне (гаплоїдне) різноманіття для окремих популяцій;

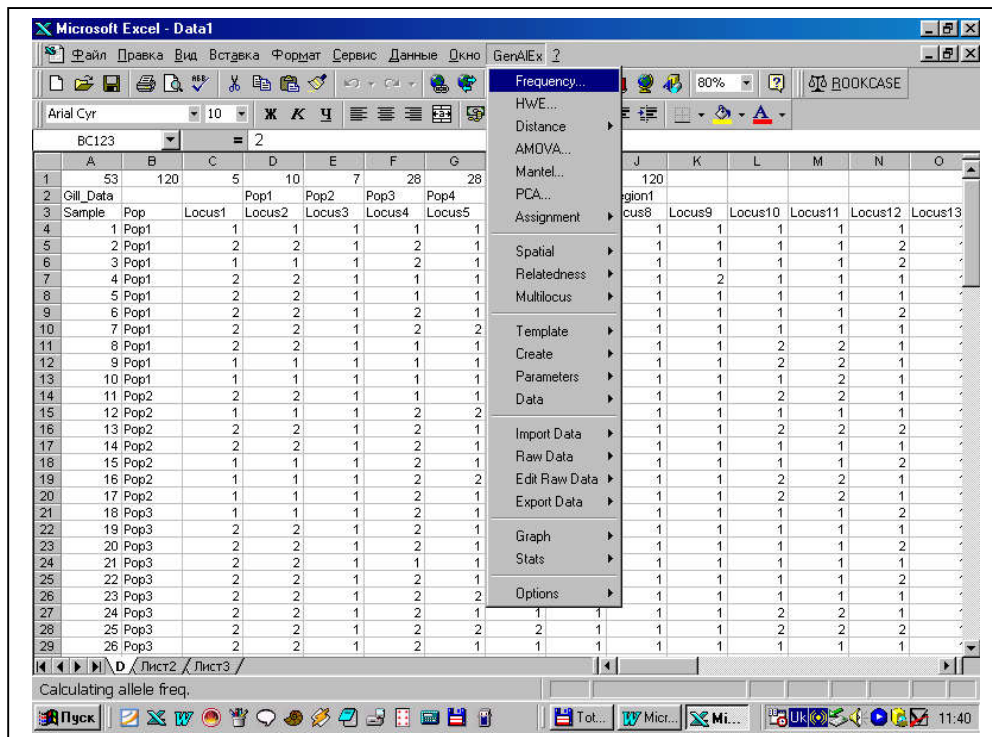


Рис. 7.

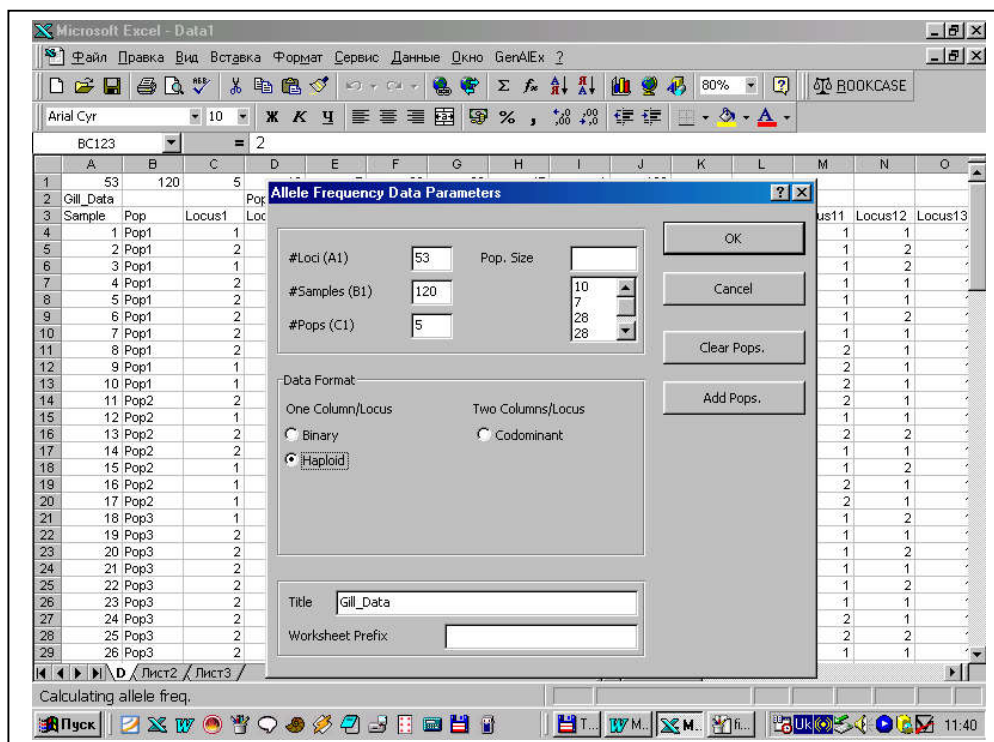


Рис. 8.

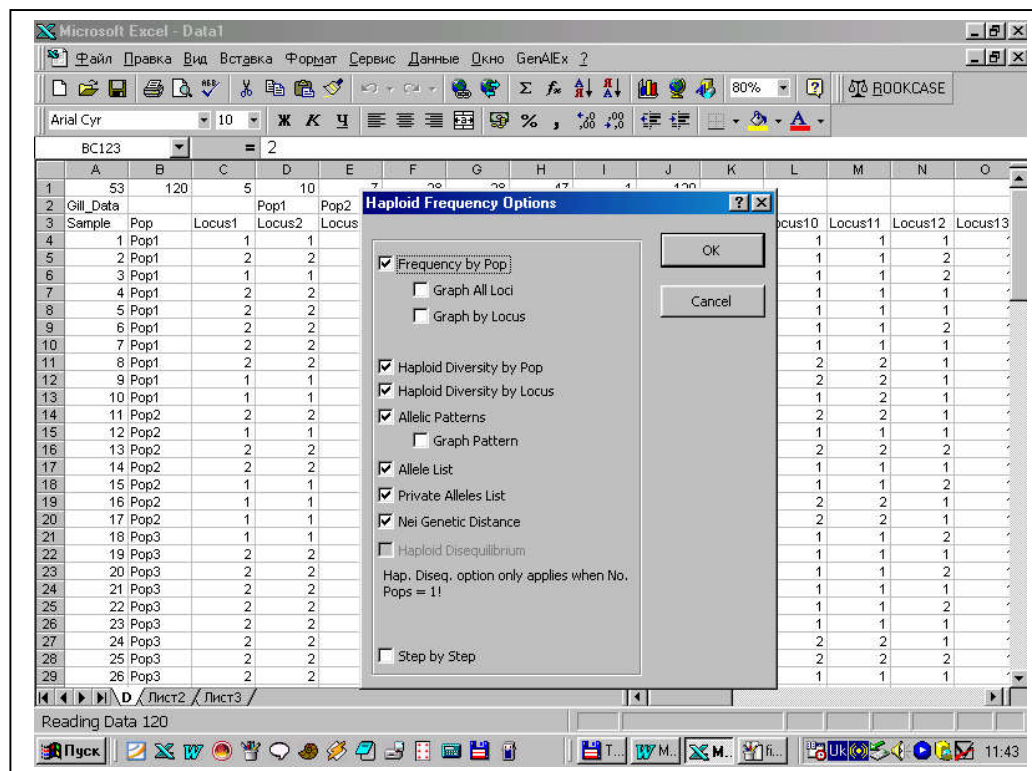


Рис. 9.

- **Haploid Diversity by Locus** – генетичне (гаплоїдне) різноманіття для окремих локусів;
- **Allelic Patterns** – типи алелей;
- **Allele List** – список алелей;
- **Private Alleles List** – список унікальних для кожної популяції (“приватних”) алелей;
- **Nei Genetic Distance** – генетичні відстані за М.Неєм.

Для подальшого аналізу необхідно обрати важливі для нас показники, як показано на рисунку 9, та клацнути по кнопці **OK**.

Програма почне розраховувати (якщо аркуші із вихідними даними мають правильний формат), а всі результати формувати на окремих нових аркушах.

Аркуш “AFP” має назву “Allele Frequencies by Pop for Haploid Data” (Частоти алелей в популяціях для гаплоїдних даних). Він містить відповідні результати та має наступний вигляд (рис. 10).

За допомогою цих результатів можна визначити, що, наприклад, у Популяції №1 (Pop1; в нашому випадку це - червона степова порода) частота антигену A_1 (Locus1) становить 0,600 (для гаплоїдної алелі “2”, тобто, для випадку наявності цього антигену).

The screenshot shows an Excel spreadsheet with the following data:

| Allele Frequencies by Pop for Haploid Data | | | | | | | |
|--|--------|-----------|-------|-------|-------|-------|--|
| Data Sheet | | D | | | | | |
| Data Title | | Gill_Data | | | | | |
| No. Loci | | 53 | | | | | |
| No. Samples | | 120 | | | | | |
| No. Pops. | | 5 | | | | | |
| Haploid Allele Frequencies by Populations | | | | | | | |
| Locus | Allele | Pop1 | Pop2 | Pop3 | Pop4 | Pop5 | |
| Locus1 | 1 | 0,400 | 0,571 | 0,536 | 1,000 | 1,000 | |
| | 2 | 0,600 | 0,429 | 0,464 | 0,000 | 0,000 | |
| Locus2 | 1 | 0,400 | 0,571 | 0,536 | 0,321 | 0,553 | |
| | 2 | 0,600 | 0,429 | 0,464 | 0,679 | 0,447 | |
| Locus3 | 1 | 1,000 | 1,000 | 1,000 | 0,929 | 1,000 | |
| | 2 | 0,000 | 0,000 | 0,000 | 0,071 | 0,000 | |
| Locus4 | 1 | 0,600 | 0,143 | 0,250 | 0,857 | 0,638 | |
| | 2 | 0,400 | 0,857 | 0,750 | 0,143 | 0,362 | |
| Locus5 | 1 | 0,900 | 0,714 | 0,643 | 0,821 | 0,787 | |
| | 2 | 0,100 | 0,286 | 0,357 | 0,179 | 0,213 | |
| Locus6 | 1 | 0,700 | 0,714 | 0,643 | 0,750 | 0,638 | |
| | 2 | 0,300 | 0,286 | 0,357 | 0,250 | 0,362 | |
| Locus7 | 1 | 0,900 | 0,714 | 0,750 | 1,000 | 1,000 | |
| | 2 | 0,100 | 0,286 | 0,250 | 0,000 | 0,000 | |
| Locus8 | 1 | 1,000 | 1,000 | 0,929 | 1,000 | 0,936 | |

Рис. 10.

З іншого боку, частоту власно алелі A_1 можна розрахувати за формулою:

$$pA_1 = 1 - \sqrt{1 - 0,600} = 0,368, \quad (1)$$

враховуючи домінуючий тип успадкування антигенних факторів (див. вище).

Аркуш “HDP” має назву “Haploid Diversity by Population” (Гаплоїдне різноманіття у популяціях) (рис. 11). Він містить наступну інформацію:

- **Pop** – ім’я популяції;
- **Locus** – ім’я локусу;
- **N** – об’єм групи (популяції);
- **Na** – кількість зареєстрованих для кожного локусу в кожній популяції алелей (в нашому випадку цифра “2” в цьому стовпчику означає, що даний антиген був відмічений в даній популяції, а цифра “1” – що такий антиген не був відмічений у жодній тварини даної популяції);
- **h** – показник гаплоїдного генетичного різноманіття.

| Pop | Locus | N | Na | h |
|------|---------|----|----|-------|
| Pop1 | Locus1 | 10 | 2 | 0,480 |
| | Locus2 | 10 | 2 | 0,480 |
| | Locus3 | 10 | 1 | 0,000 |
| | Locus4 | 10 | 2 | 0,480 |
| | Locus5 | 10 | 2 | 0,180 |
| | Locus6 | 10 | 2 | 0,420 |
| | Locus7 | 10 | 2 | 0,180 |
| | Locus8 | 10 | 1 | 0,000 |
| | Locus9 | 10 | 2 | 0,180 |
| | Locus10 | 10 | 2 | 0,320 |
| | Locus11 | 10 | 2 | 0,420 |
| | Locus12 | 10 | 2 | 0,420 |
| | Locus13 | 10 | 1 | 0,000 |
| | Locus14 | 10 | 2 | 0,320 |
| | Locus15 | 10 | 2 | 0,320 |

Рис. 11.

Останній показник розраховується за наступною формулою:

$$h = 1 - \sum_{i=1}^2 P_i^2, \quad (2)$$

де P_i – частка тварин у даній популяції, що має та не має певний антиген, відповідно.

Наприклад, якщо у Популяції №1 (Pop1; в нашому випадку це - червона степова порода) частота наявності антигену A_1 (Locus1) становить 0,600, а частота його відсутності – 0,400, то оцінка гапloidного генетичного різноманіття для неї буде складати:

$$h = 1 - (0,600^2 + 0,400^2) = 1 - 0,520 = 0,480.$$

Наприкінці цього аркушу наведено також оцінки для середнього по популяції гапloidного генетичного різноманіття (H – **Mean Population Diversity**) та очікуваної варіанси (Ve) кількості локусів, за якими дві випадковим чином обрані особини відрізняються.

Аркуш “HDL” має назву “**Haploid Diversity by Locus**” (Гапloidне різноманіття для локусів) (рис. 12).

| | | Locus1 | Locus2 | Locus3 | Locus4 | Locus5 | Locus6 | Locus7 | Locus8 | Locus9 | |
|----|----------------------------|-----------|--------|--------|--------|--------|--------|--------|--------|--------|--|
| 1 | Haploid Diversity by Locus | | | | | | | | | | |
| 2 | | | | | | | | | | | |
| 3 | Data Sheet | D | | | | | | | | | |
| 4 | Data Title | Gill_Data | | | | | | | | | |
| 5 | | | | | | | | | | | |
| 6 | No. Loci | 53 | | | | | | | | | |
| 7 | No. Samples | 120 | | | | | | | | | |
| 8 | No. Pops. | 5 | | | | | | | | | |
| 9 | | | | | | | | | | | |
| 10 | Diversity by Locus | | | | | | | | | | |
| 11 | | | | | | | | | | | |
| 12 | | Locus1 | Locus2 | Locus3 | Locus4 | Locus5 | Locus6 | Locus7 | Locus8 | Locus9 | |
| 13 | Pop1 | N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | |
| 14 | | Na | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | |
| 15 | | h | 0,480 | 0,480 | 0,000 | 0,480 | 0,180 | 0,420 | 0,180 | 0,000 | |
| 16 | Pop2 | N | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | |
| 17 | | Na | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | |
| 18 | | h | 0,490 | 0,490 | 0,000 | 0,245 | 0,408 | 0,408 | 0,408 | 0,000 | |
| 19 | Pop3 | N | 28 | 28 | 28 | 28 | 28 | 28 | 28 | 28 | |
| 20 | | Na | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | |
| 21 | | h | 0,497 | 0,497 | 0,000 | 0,375 | 0,459 | 0,459 | 0,375 | 0,133 | |
| 22 | Pop4 | N | 28 | 28 | 28 | 28 | 28 | 28 | 28 | 28 | |
| 23 | | Na | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | |
| 24 | | h | 0,000 | 0,436 | 0,133 | 0,245 | 0,293 | 0,375 | 0,000 | 0,000 | |
| 25 | Pop5 | N | 47 | 47 | 47 | 47 | 47 | 47 | 47 | 47 | |
| 26 | | Na | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | |
| 27 | | h | 0,000 | 0,494 | 0,000 | 0,462 | 0,335 | 0,462 | 0,000 | 0,120 | |

Рис. 12.

Він містить ту ж саму інформацію, що й попередній аркуш; змінено лише форму подання цієї інформації.

Аркуш “АРТ” має назву “Allelic Patterns for Haploid Data” (Алельне різноманіття для гапліодних даних) (рис. 13). На цьому аркуші наведено деякі показники генетичного різноманіття окремих популяцій, а саме:

- N_a – кількість алелей на локус;
- $N_a \text{ Freq. } > 5\%$ – кількість алелей на локус, частота яких більше 5% (95% межа поліморфізму);
- N_e – ефективна кількість алелей;
- I – інформаційний індекс Шеннона-Уївера;
- *No. Private Alleles* – кількість унікальних (“приватних”) алелей, тобто, алелей, що зустрічаються лише в одній популяції;
- *No. LComm Alleles (<=25%)* – кількість спільних алелей, які зустрічаються в менш, ніж 25% популяцій;

| Allelic Patterns for Haploid Data | | | | | | |
|--|-------|-----------|-------|-------|-------|--|
| Data Sheet | | D | | | | |
| Data Title | | Gill_Data | | | | |
| | | | | | | |
| No. Loci | 53 | | | | | |
| No. Samples | 120 | | | | | |
| No. Pops. | 5 | | | | | |
| | | | | | | |
| Mean Allelic Patterns Across Populations | | | | | | |
| | | | | | | |
| Mean values | | | | | | |
| Population | Pop1 | Pop2 | Pop3 | Pop4 | Pop5 | |
| N_a | 1,811 | 1,623 | 1,849 | 1,679 | 1,774 | |
| $N_a \text{ Freq. } \geq 5\%$ | 1,811 | 1,623 | 1,755 | 1,604 | 1,604 | |
| N_e | 1,475 | 1,430 | 1,519 | 1,326 | 1,381 | |
| I | 0,430 | 0,362 | 0,443 | 0,311 | 0,346 | |
| No. Private Alleles | 0,019 | 0,000 | 0,000 | 0,019 | 0,019 | |
| No. LComm Alleles (<=25%) | 0,000 | 0,000 | 0,000 | 0,000 | 0,000 | |
| No. LComm Alleles (<=50%) | 0,075 | 0,038 | 0,132 | 0,113 | 0,132 | |
| H_e | 0,286 | 0,246 | 0,298 | 0,201 | 0,227 | |
| | | | | | | |
| Standard Error (SE) values | | | | | | |
| Population | Pop1 | Pop2 | Pop3 | Pop4 | Pop5 | |
| N_a | 0,054 | 0,067 | 0,050 | 0,065 | 0,058 | |
| $N_a \text{ Freq. } \geq 5\%$ | 0,054 | 0,067 | 0,060 | 0,068 | 0,068 | |
| N_e | 0,045 | 0,053 | 0,051 | 0,046 | 0,050 | |

Рис. 14.

Аркуш “PAS” має назву “**Summary of Private Alleles by Population**” (Інформація про унікальні алелі у популяціях) (рис. 15). На цьому аркуші наведено інформацію про те, у якій популяції була зафіксована унікальна (“приватна”) алель, яка ця алель (у нашому випадку завжди це буде алель “2”, тобто, наявність антигену) та з якою частотою вона зустрічається.

Аркуш “PAL” має назву “**List of Samples with One or More Private Alleles**” (Перелік проб, що мають одну чи декілька унікальних алелей) (рис. 16). На цьому аркуші наведена наступна інформація:

- **Sample** – номер проби (тобто, тварини) у загальному списку;
- **Pop** – номер групи (популяції), до якої належить ця тварина;
- **Locus1-Locus53** – мультилокусний генотип для кожної з цих тварин;
- **No. Loci with Private Alleles** – кількість локусів із унікальними алелями;
- **Loci with Private Alleles** – ім’я локусу, для якого відмічено наявність унікальної алелі.

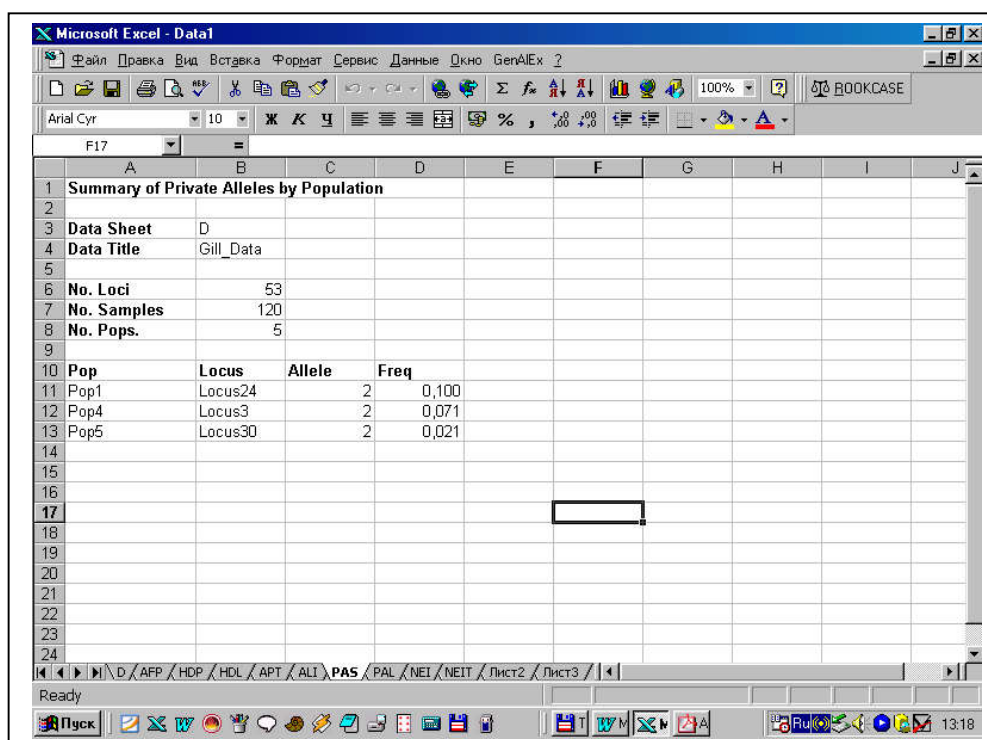


Рис. 15.

The screenshot shows an Excel spreadsheet with the following data:

| Sample | Pop | Locus1 | Locus2 | Locus3 | Locus4 | Locus5 | Locus6 | Locus7 | Locus8 | Locus9 | Locus10 |
|--------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| 10 | Pop1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 55 | Pop4 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 56 | Pop4 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 116 | Pop5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |

Рис. 16.

Аркуш “NEI” містить дві матриці з результатами (рис. 17):

- матриця “**Pairwise Population Matrix of Nei Genetic Distance**” містить для кожної пари груп (популяцій) оцінки генетичної відстані М.Нея;
- матриця “**Pairwise Population Matrix of Nei Genetic Identity**” містить для кожної пари груп (популяцій) оцінки генетичної ідентичності М.Нея.

Аркуш “NEIT” має назву “**Pairwise Population Nei Genetic Distance Values As Table**” (Попарні оцінки генетичної відстані М.Нея у табличній формі) містить оцінки генетичної відстані М.Нея, але представлені у вигляді таблиці (рис. 18). Крім того, наведено також об’єм кожної із груп (популяцій).

The screenshot shows two matrices in an Excel spreadsheet. The first matrix, titled "Pairwise Population Matrix of Nei Genetic Distance", has the following values:

| | Pop1 | Pop2 | Pop3 | Pop4 | Pop5 |
|------|-------|-------|-------|-------|-------|
| Pop1 | 1,000 | | | | |
| Pop2 | 0,063 | 1,000 | | | |
| Pop3 | 0,039 | 0,035 | 1,000 | | |
| Pop4 | 0,062 | 0,142 | 0,078 | 1,000 | |
| Pop5 | 0,045 | 0,108 | 0,058 | 0,019 | 1,000 |

The second matrix, titled "Pairwise Population Matrix of Nei Genetic Identity", has the following values:

| | Pop1 | Pop2 | Pop3 | Pop4 | Pop5 |
|------|-------|-------|-------|-------|-------|
| Pop1 | 1,000 | | | | |
| Pop2 | 0,939 | 1,000 | | | |
| Pop3 | 0,961 | 0,966 | 1,000 | | |
| Pop4 | 0,940 | 0,867 | 0,925 | 1,000 | |
| Pop5 | 0,956 | 0,898 | 0,944 | 0,981 | 1,000 |

Рис. 17.

В кінці відмітимо, що для отримання більшості перелічених вище популяційно-генетичних показників для всіх груп (популяцій) в цілому, необхідно зробити деякі зміни на аркуші із вихідними даними (у першій та другій строчці), як це показано на рисунку 19.

Етап 2. Оцінка міжгрупової (міжпопуляційної) генетичної диференціації за антигенними факторами.

Одним із поширених методів оцінки міжгрупової (міжпопуляційної) диференціації, який придатний для аналізу імуногенетичних даних у вигляді бінарної матриці, є метод аналізу молекулярної мінливості (Analysis of MOlecular VAriation) – AMOVA. Він базується на розрахунку відстаней між окремими особинами у багатовимірному просторі, що формують антигенні фактори, та використанні в наступному алгоритмі дисперсійного аналізу

Р.Фішера, тобто, алгоритму розкладання мінливості на внутрішньо- та міжгрупову компоненти.

| Pop1 | Pop2 | Nei Genetic Distance | #Pop1 | #Pop2 |
|------|------|----------------------|-------|-------|
| Pop1 | Pop2 | 0,063 | 10 | 7 |
| Pop1 | Pop3 | 0,039 | 10 | 28 |
| Pop2 | Pop3 | 0,035 | 7 | 28 |
| Pop1 | Pop4 | 0,062 | 10 | 28 |
| Pop2 | Pop4 | 0,142 | 7 | 28 |
| Pop3 | Pop4 | 0,078 | 28 | 28 |
| Pop1 | Pop5 | 0,045 | 10 | 47 |
| Pop2 | Pop5 | 0,108 | 7 | 47 |
| Pop3 | Pop5 | 0,058 | 28 | 47 |
| Pop4 | Pop5 | 0,019 | 28 | 47 |

Рис. 18.

| | A | B | C | D | E | F | G | H | I |
|----|-----------|------|--------|--------|--------|---------|--------|--------|--------|
| 1 | 53 | 120 | 1 | 120 | 1 | 120 | | | |
| 2 | Gill Data | | | Pop1 | | Region1 | | | |
| 3 | Sample | Pop | Locus1 | Locus2 | Locus3 | Locus4 | Locus5 | Locus6 | Locus7 |
| 4 | 1 | Pop1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 5 | 2 | Pop1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 |
| 6 | 3 | Pop1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 |
| 7 | 4 | Pop1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| 8 | 5 | Pop1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| 9 | 6 | Pop1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 |
| 10 | 7 | Pop1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 |
| 11 | 8 | Pop1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 |
| 12 | 9 | Pop1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 |

Рис. 19.

Мірою генетичної диференціації між популяціями є оцінка показника Φ_{st} , що є аналогом показника F_{st} , запропонованого у 1951 р. С.Райтом (Wright, 1951). При цьому, проводиться розрахунок показника як для всіх

груп в цілому, так й для кожної пари груп окремо. Для визначення рівня значущості отриманої оцінки використовується методика, що має назву “взяття повторних вибірок”, або resampling-процедура.

Провести подібний аналіз для імуногенетичних даних, що ми маємо, можна наступним чином.

Необхідно відкрити аркуш із вихідними даними (див. рис. 6). Далі, в Головному меню обрати надбудову **GenAlEx**, відкрити в ньому меню **AMOVA** (рис. 20) та без будь-яких змін клацнути по кнопці **OK**. В наступному меню (рис. 21) обрати тип вихідних даних “**Haploid**” (**Гаплоїдні**) та клацнути по кнопці **OK**. У меню, що відкриється, зробити наступні зміни (рис. 22) та клацнути по кнопці **OK**. Програмі для проведення розрахунків потрібен деякий час – у випадку великих баз даних це може зайняти навіть годину – тому необхідно залишити комп’ютер у спокою, поки програма не закінчить роботу. Це стане зрозуміло, коли з’являться три нових аркуша у файлі – **AM**, **PW** та **PWT**.

Аркуш “**AM**” має назву “**Results of Analysis of Molecular Variance**” (**Результати аналізу молекулярної мінливості**) містять результати аналізу молекулярної мінливості у вигляді таблиці дисперсійного аналізу (рис. 23). В цій таблиці наведено наступні показники:

- **N0** – середній зважений об’єм груп;
- **SSTOT** – сума квадратів дистанцій між всіма об’єктами всіх груп;
- **Pop** – ім’я кожної групи (популяції);
- **n** – об’єм кожної із груп (популяцій);
- **SSWP** – сума квадратів показників відстаней між об’єктами для кожної групи (популяції);
- **Summary AMOVA Table** – таблиця із результатами аналізу молекулярної мінливості;

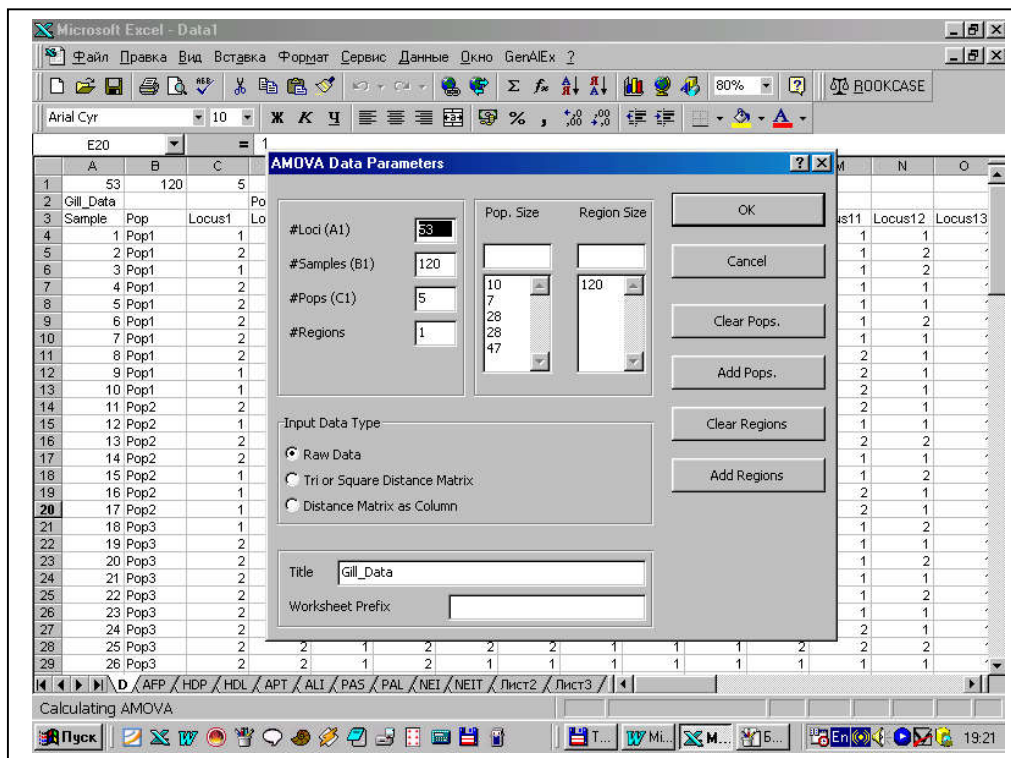


Рис. 20.

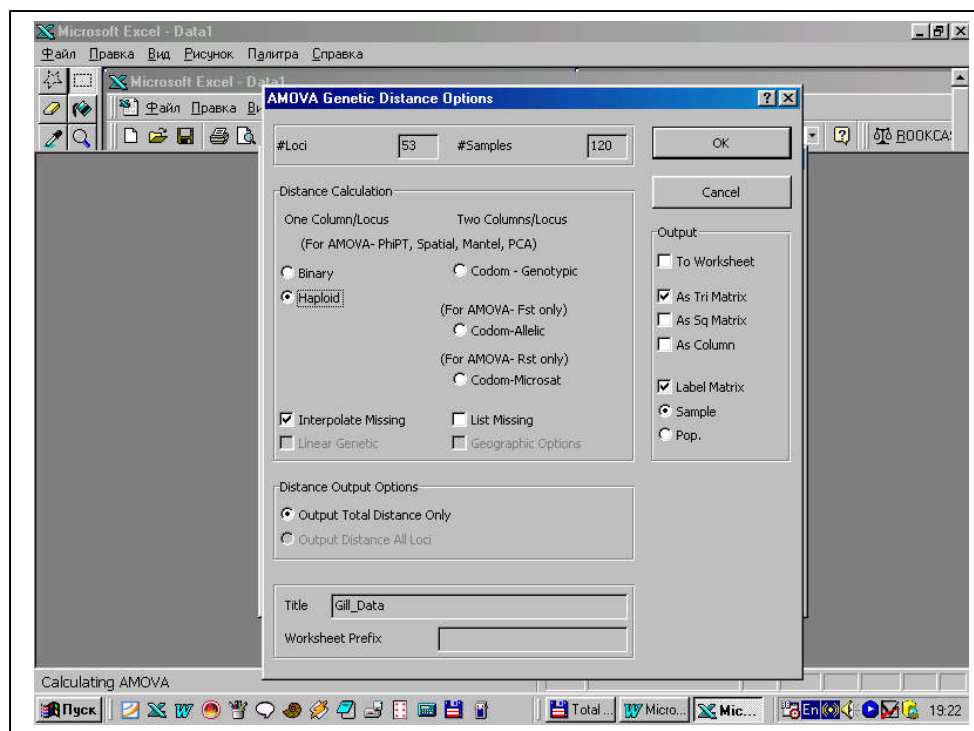


Рис. 21.

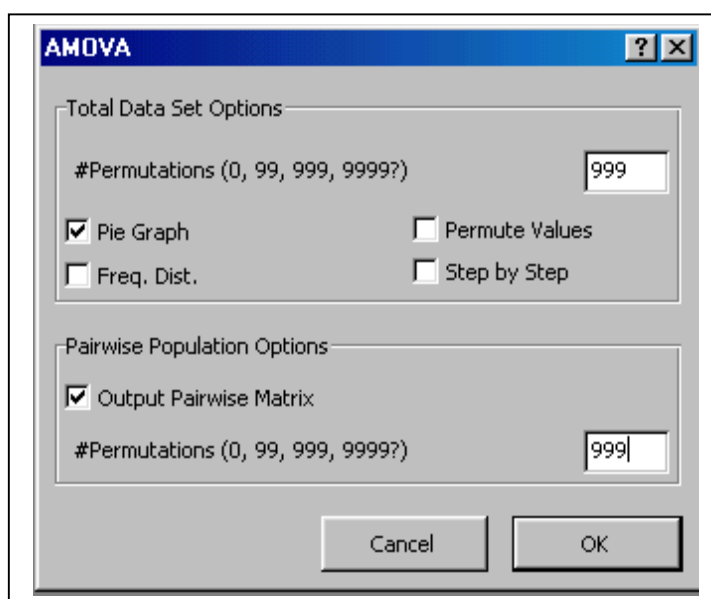


Рис. 22.

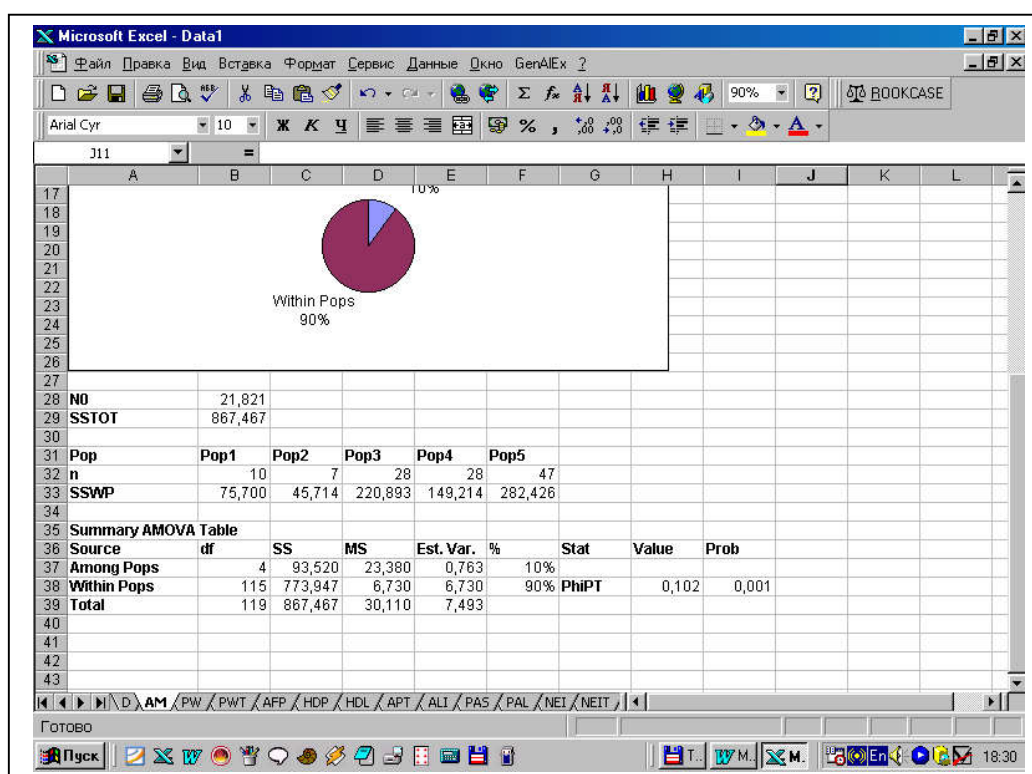


Рис. 23.

- **Among Pops** – показники, які пов’язані із міжгруповою компонентою мінливості;
- **Whitin Pops** – показники, які пов’язані із внутрішньогруповою компонентою мінливості;
- **df** – число ступенів свободи;
- **SS** – сума квадратів дистанцій;
- **MS** – середній квадрат дистанцій;
- **Est.Var.** – очікуваний середній квадрат дистанцій;
- **%** - очікуваний середній квадрат дистанцій у відсотках;
- **Stat** – ім’я показника генетичної диференціації;
- **PhiST** – показник міжгрупової генетичної диференціації (Φ_{st});
- **Value** – оцінка показника Φ_{st} ;
- **Prob** – рівень значущості показника Φ_{st} .

Аркуш “PW” із назвою “**Pairwise Population PhiPT Values**” (Оцінки Φ_{st} між парами популяцій) містить результати аналізу молекулярної мінливості у вигляді парних відмінностей між окремими групами, що включено до аналізу, у матричній формі (рис. 24). При цьому власне оцінки показника генетичної диференціації розташовані у нижній лівій половині матриці, а оцінки рівня їх значущості – у верхній правій половині матриці. (Головна діагональ, природно, містить лише одні нулі.)

Аркуш “PWT” має назву “**Pairwise Population PhiPT Values and Estimates of Nm As Table**” (Оцінки Φ_{st} та руху генів між парами популяцій у табличній формі) і містить результати аналізу молекулярної мінливості у вигляді парних відмінностей між окремими групами, що включено до аналізу, у табличній формі (рис. 25). Крім того, для кожної

пари популяцій наведено оцінку руху генів (Nm), тобто, середню кількість мігрантів, якими обмінюються популяції за одну генерацію.

| Pop1 | Pop2 | Pop3 | Pop4 | Pop5 |
|-------|-------|-------|-------|-------|
| 0,000 | 0,165 | 0,164 | 0,001 | 0,003 |
| 0,024 | 0,000 | 0,406 | 0,001 | 0,001 |
| 0,020 | 0,000 | 0,000 | 0,001 | 0,001 |
| 0,115 | 0,272 | 0,159 | 0,000 | 0,001 |
| 0,069 | 0,198 | 0,119 | 0,038 | 0,000 |

PhiPT Values below diagonal. Probability values based on 999 permutations are shown above diagonal.
Warning! Negative pairwise PhiPT converted to zero.

Рис. 24.

| Pop1 | Pop2 | PhiPT | Nm | #Pop1 | #Pop2 | Prob | No. PW Pm |
|------|------|-------|---------|-------|-------|-------|-----------|
| Pop1 | Pop2 | 0,024 | 9,991 | 10 | 7 | 0,165 | 999 |
| Pop1 | Pop3 | 0,020 | 12,218 | 10 | 28 | 0,164 | 999 |
| Pop2 | Pop3 | 0,000 | #DIV/0! | 7 | 28 | 0,406 | 999 |
| Pop1 | Pop4 | 0,115 | 1,915 | 10 | 28 | 0,001 | 999 |
| Pop2 | Pop4 | 0,272 | 0,669 | 7 | 28 | 0,001 | 999 |
| Pop3 | Pop4 | 0,159 | 1,327 | 28 | 28 | 0,001 | 999 |
| Pop1 | Pop5 | 0,069 | 3,373 | 10 | 47 | 0,003 | 999 |
| Pop2 | Pop5 | 0,198 | 1,012 | 7 | 47 | 0,001 | 999 |
| Pop3 | Pop5 | 0,119 | 1,856 | 28 | 47 | 0,001 | 999 |
| Pop4 | Pop5 | 0,038 | 6,341 | 28 | 47 | 0,001 | 999 |

Рис. 25.

Етап 3. Візуалізація груп у просторі генетичної мінливості з використанням Аналізу Головних Координат (РСоА).

Для більш детального аналізу взаємовідносин між окремими групами, що включено до аналізу, у просторі генетичної мінливості можна також використати Аналіз Головних Координат (РСоА). Для проведення цього аналізу вихідні дані повинні бути представлені у вигляді матриці відстаней. Для цього можна використати матрицю з показниками міжгрупових генетичний відстаней М.Нея (рис.17), або матрицю з показниками міжгрупових оцінок показника генетичної диференціації – Φ_{st} (рис. 24).

Для отримання результатів РСоА необхідно відкрити аркуш із матрицею вихідних даних, знайти у Головному меню надбудову GenAlEx, відкрити її та обрати опцію РСА. У меню, що відкриється (рис. 26), не треба робити ніяких змін, а лише клацнути по кнопці **ОК**. Програма здійснить необхідні розрахунки, результати яких буде надано на новому аркуші – “РСА” – “Principal Coordinates Analysis (PCA)” (Аналіз Головних Координат) (рис. 27а). На цьому аркуші наведена наступна інформація:

- **Percentage of variation explained by the first 3 axes** – відсоток мінливості, що пояснюється першими трьома осями;
- **Axis** – номер осі (від 1 до 3);
- **%** - власне оцінки відсотку мінливості;
- **Cum%** - накопичені оцінки відсотку мінливості.

Далі, на аркуші наведено графік розподілу центроїдів кожної із груп (популяцій), що включено до аналізу, у просторі перших двох осей.

Нижня частина аркуша містить таблицю “Eigen Values by Axis and Sample Eigen Vectors” – “Власні оцінки для кожної осі та власні вектори для груп” (рис. 27b).

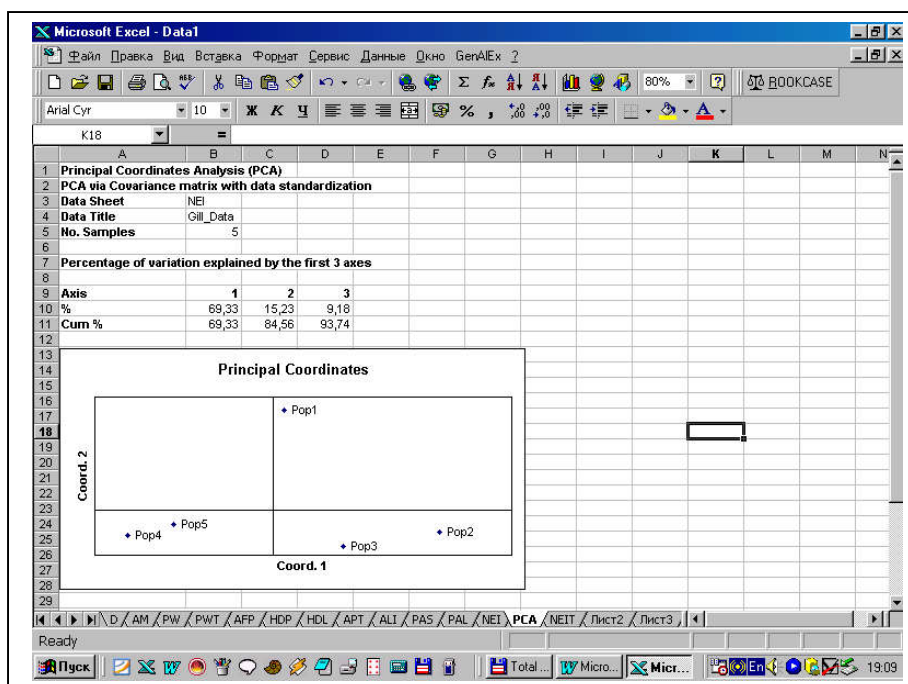


Рис. 27а.

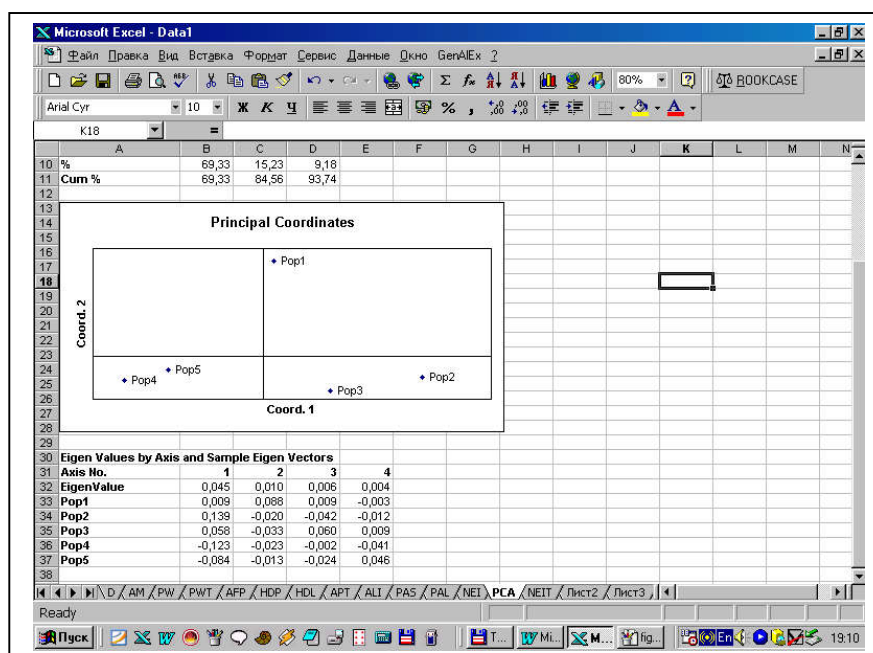


Рис. 27б.

Аналогічний аналіз можна провести для кожної матриці, що містить оцінки відстаней, виміряні будь-яким методом розрахунку генетичної

відмінності, між парами груп (популяцій) за умови, що вихідні дані будуть мати прийнятний для програми GenAIEх формат (див. рис. 17 та рис. 24).

Етап 4. R- та Q-аналіз зчепленого успадкування антигенних факторів.

Як ми вже вказали вище, матриця із мультилокусної формули генотипу будь-якої тварини, що записана у бінарній формі (рис. 1), може бути також використана для аналізу асоціації між окремими антигенними факторами, що може свідчити про їх зчеплене успадкування. При цьому, такий аналіз можна провести у двох варіантах. Аналіз між окремими пробами (тваринами) чи їх групами у просторі антигенних факторів (R-аналіз), або аналіз між окремими антигенними факторами на підставі їх наявності у окремих тварин (Q-аналіз). У першому випадку розраховується матриця відстаней між окремими особинами (чи групами), а у другому – між окремими антигенними факторами.

У будь-якому випадку, обидва типи аналізу можна провести з використанням модулю “**Cluster Analysis**” (**Кластерний аналіз**) програми “STATISTICA”.

Для проведення R-аналізу попередньо необхідно зробити деякі попередні розрахунки, насамперед, розрахувати координати центроїдів кожної вибірки у реальному 53-вимірному просторі антигенних факторів, що використовуються в аналізі. Для цього використаємо опцію “**Breakdown & one-way ANOVA**” (“**Класифікація та однофакторний дисперсійний аналіз**”) модуля “**Basic Statistics and Tables**” (“**Описова статистика та таблиці**”) (рис. 28). В меню, що відкриється необхідно обрати змінні. В якості факторної змінної (**Grouping**) необхідно обрати змінну, що містить інформацію щодо належності кожної тварини до однієї із п’яти груп (популяцій). А в якості залежної змінної (**Dependent**) - обрати 53 стовпчики,

що містять інформацію щодо наявності/відсутності певного антигенного фактора (рис. 29).

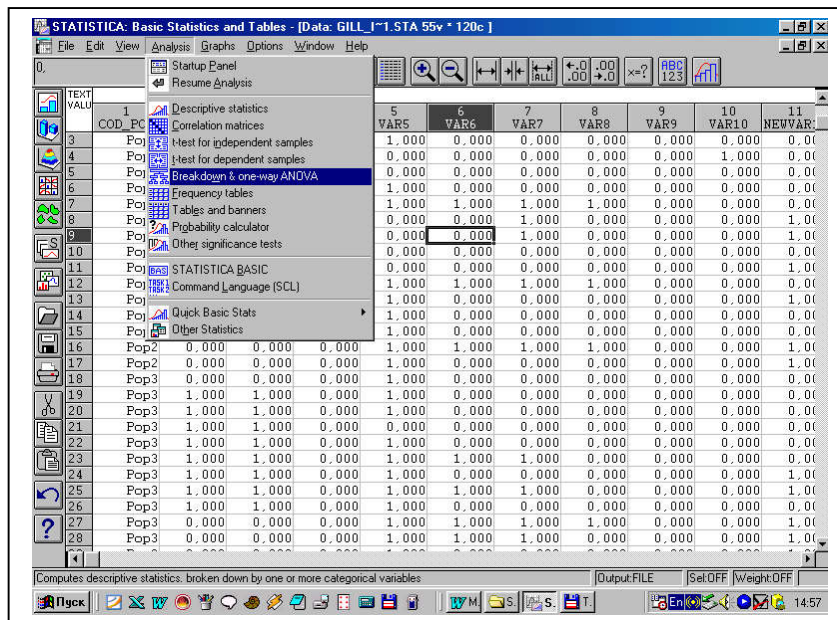


Рис. 28.

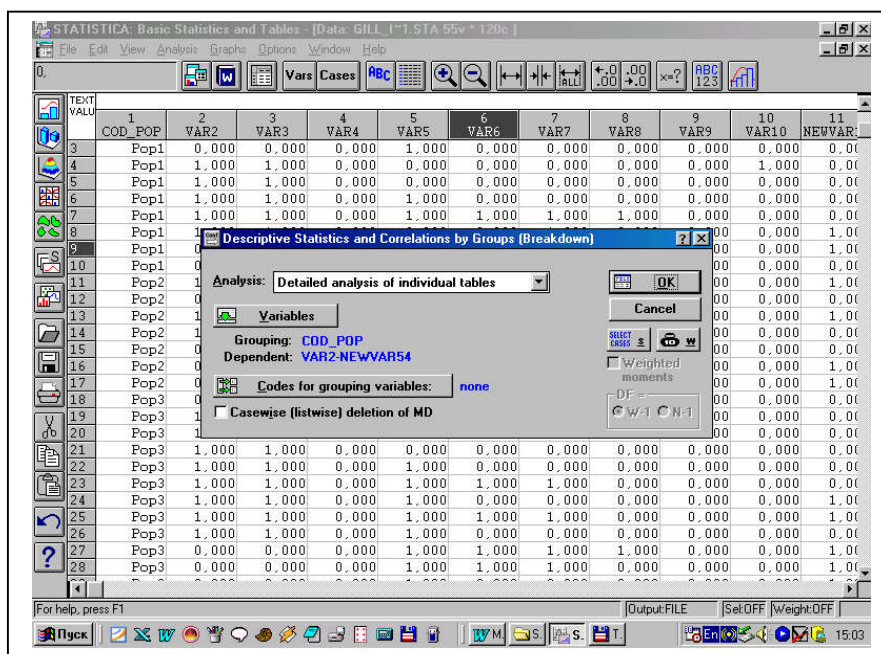


Рис. 29.

Клацнути по кнопці **OK**, внаслідок чого відкриється нове віконце (рис. 30). В ньому необхідно клацнути по кнопці **“Summary table of means”** (**“Таблиця із середніми арифметичними”**). Тоді відкриється таблиця із

середніми значеннями по всіх 53 стовпчиках для кожної із п'яти генотипових груп (рис. 31).

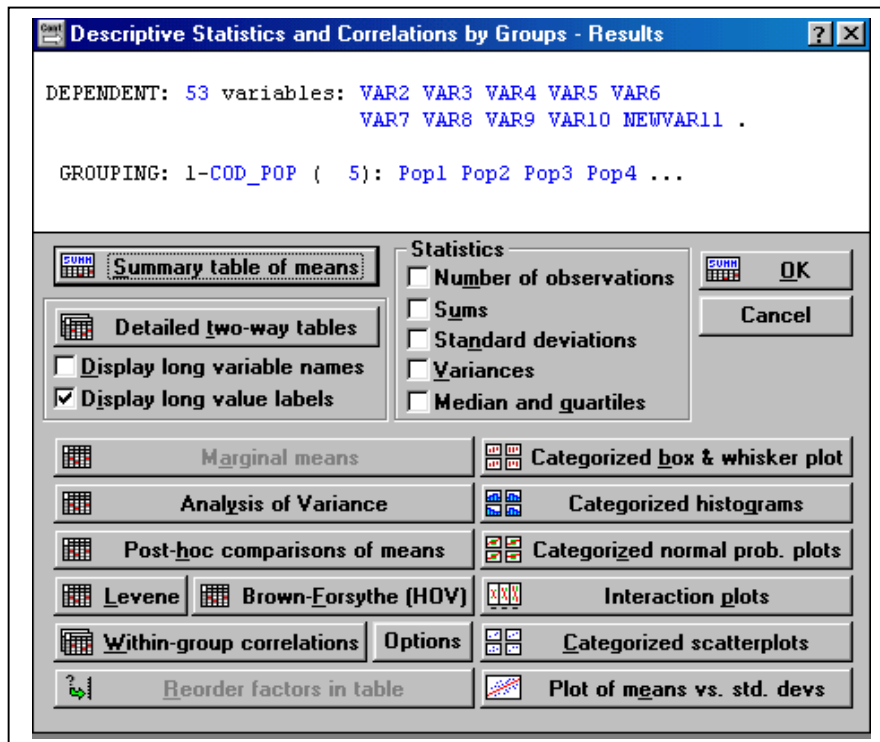


Рис. 30.

| COD_POP | VAR2 | VAR3 | VAR4 | VAR5 | VAR6 | VAR7 |
|----------|----------|---------|----------|---------|---------|---------|
| Pop1 | ,600000 | ,600000 | 0,000000 | ,400000 | ,100000 | ,300000 |
| Pop2 | ,428571 | ,428571 | 0,000000 | ,857143 | ,285714 | ,285714 |
| Pop3 | ,464286 | ,464286 | 0,000000 | ,750000 | ,357143 | ,357143 |
| Pop4 | 0,000000 | ,678571 | ,071429 | ,142857 | ,178571 | ,250000 |
| Pop5 | 0,000000 | ,446809 | 0,000000 | ,361702 | ,212766 | ,361702 |
| All Grps | ,183333 | ,516667 | ,016667 | ,433333 | ,233333 | ,325000 |

Рис. 31.

Для того, щоб ці значення використовувати в подальшому, необхідно їх зберегти як дані. Для цього обрати наступний режим роботи із файлами – “Save as Data” (“Зберегти як дані”) (рис. 32).

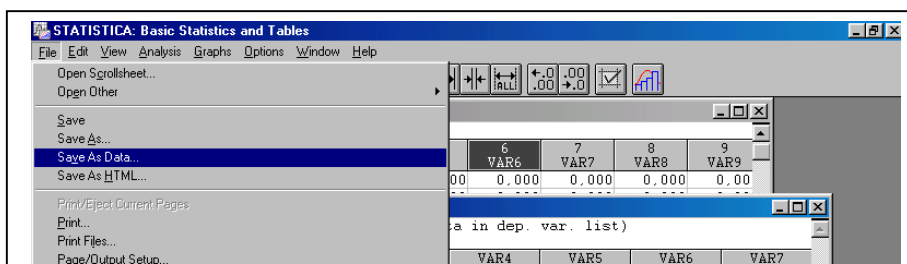


Рис. 32.

На запитання програми щодо ім'я нового файлу із середніми арифметичними ми надамо наступне: **Gill_immuno_means.sta** (рис. 33).

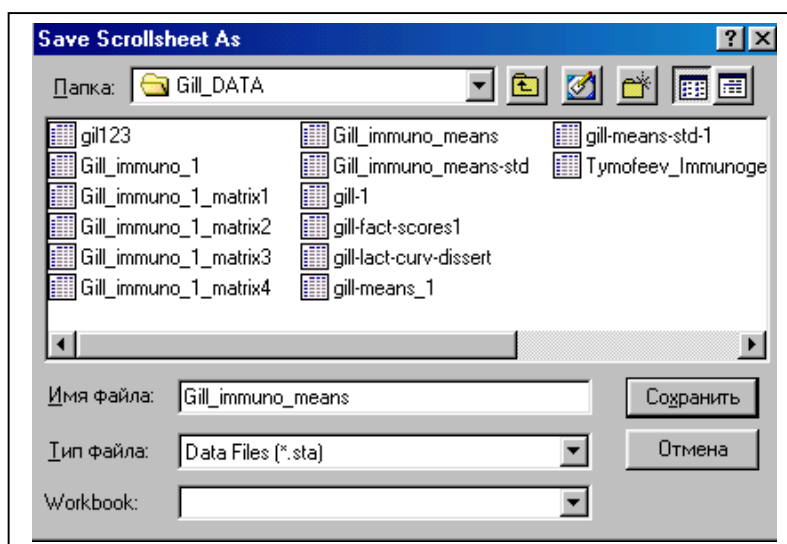


Рис. 33.

Тепер відкриємо той файл із середніми по групах, який ми зберегли. Він потребує ще деяких змін. По-перше, необхідно видалити останню строчку таблиці (**All Grps**), що містить середні арифметичні значення по всім строчкам (тобто, для 120 тварин). Для цього використовуємо опцію

“Строки” (Cases) та обираємо для останньої строчки функцію “Видалити” (Delete) (рис. 34).

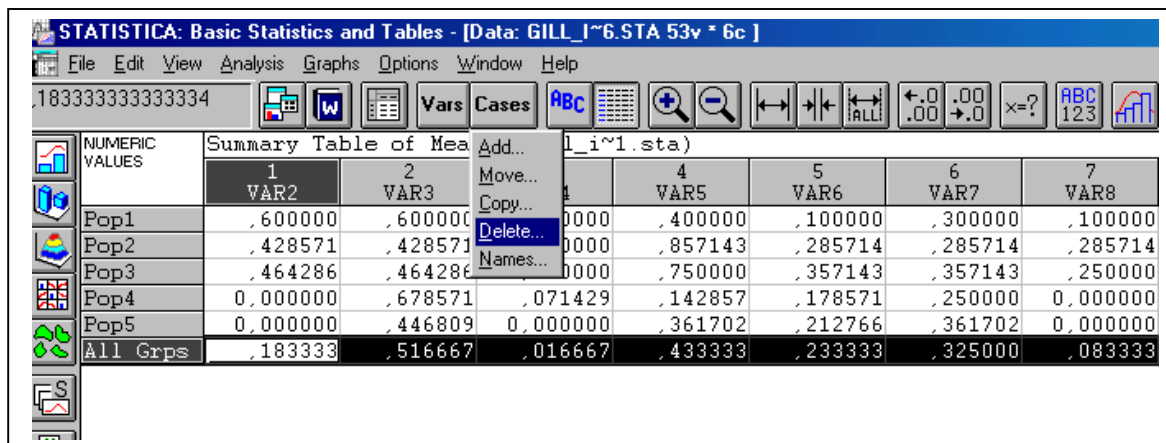


Рис. 34.

По-друге, необхідно стандартизувати дані, замінивши їх нормованими відхиленнями (z-мітками). Це дозволяє використовувати такі стандартизовані дані для процесу кластеризації, оскільки вони у такому випадку мають однаковий розподіл (із середнім арифметичним, що дорівнює нулю, та середнім квадратичним відхилення, що дорівнює одиниці). Для цього, виділимо всі стовпчики таблиці із середніми, клацнемо по ім'ям будь-якого стовпчика правою кнопкою миші (ПКМ). В меню, що відкриється оберемо опцію “**Fill/Standardize Block**” (Стандартизація Блоку даних), й оберемо процедуру “**Standardize Column**” (Стандартизація Стовпчиків) (рис. 35). Цей файл необхідно також зберегти окремо; назвемо його **Gill_immuno_means_std.sta** (рис. 36).

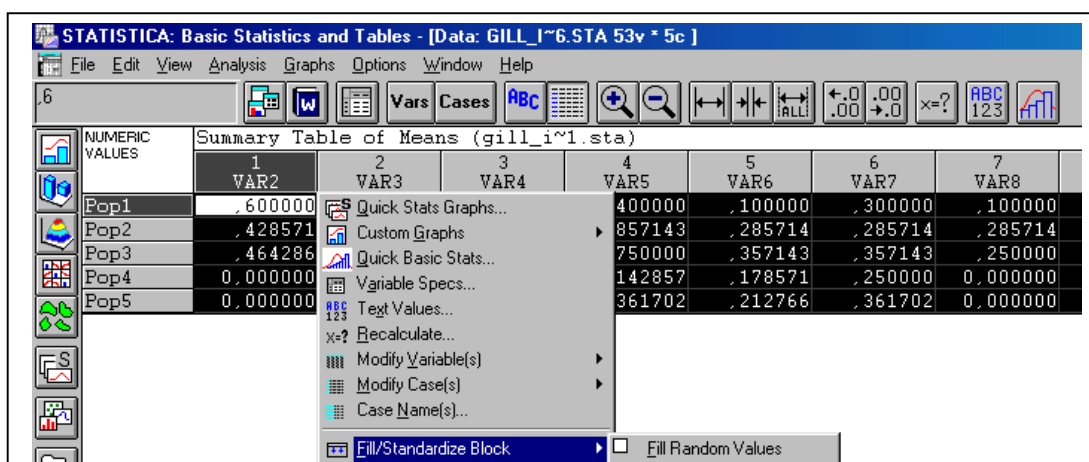


Рис. 35.

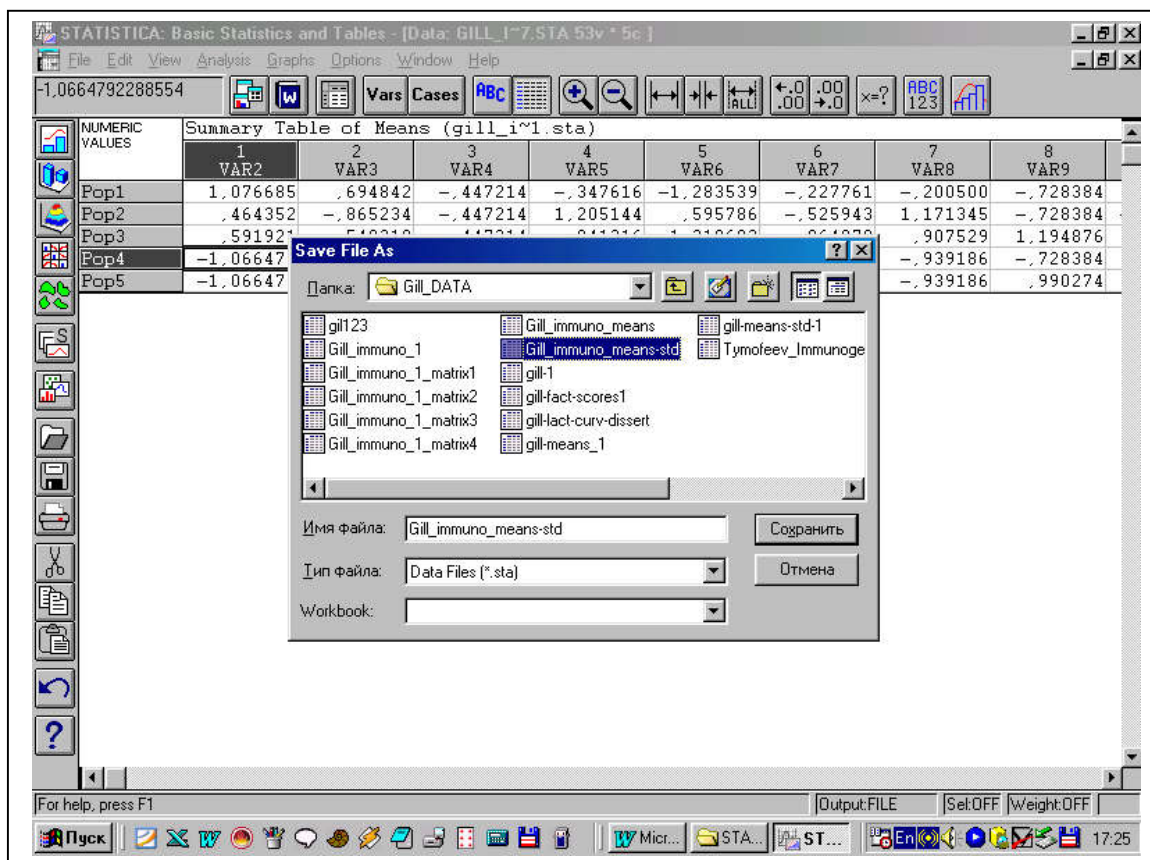


Рис. 36.

Зараз, після проведення необхідних підготовчих дій, файл із даними вже придатний для його кластеризації. Для цього відкриємо модуль “Кластерний аналіз” (Cluster Analysis) та оберемо метод кластеризації “Joining (tree clustering)” (Побудова дерев кластеризації). У вікні, що відкриється зробимо наступні установки (рис. 37) та клацнемо по кнопці **ОК**.

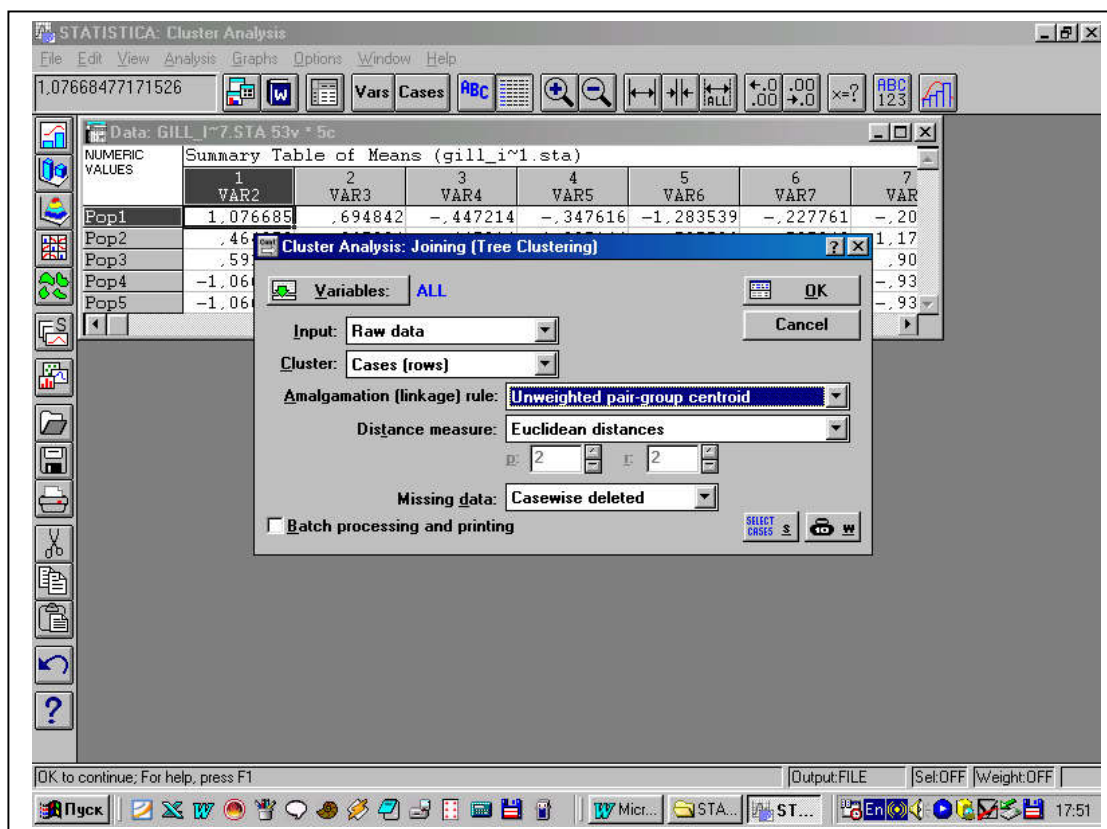


Рис. 37.

Вікно, що відкриється, дає змогу візуалізувати результати кластерного аналізу у вигляді дендрограми (дерева кластеризації) у вертикальному (**Vertical icicle plot**) чи горизонтальному (**Horizontal hierarchical tree plot**) вигляді (рис. 38). Можна просто клацнути по кнопці **OK** та отримати дендрограму (рис. 39). Але нам необхідно не стільки дендрограма, скільки матриця евклідових відстаней між окремими популяціями. Її ми будемо використовувати у процедурі багатовимірного неметричного шкалювання. Для її отримання ми клацнемо по кнопці “**Save distance matrix**” (рис. 38).

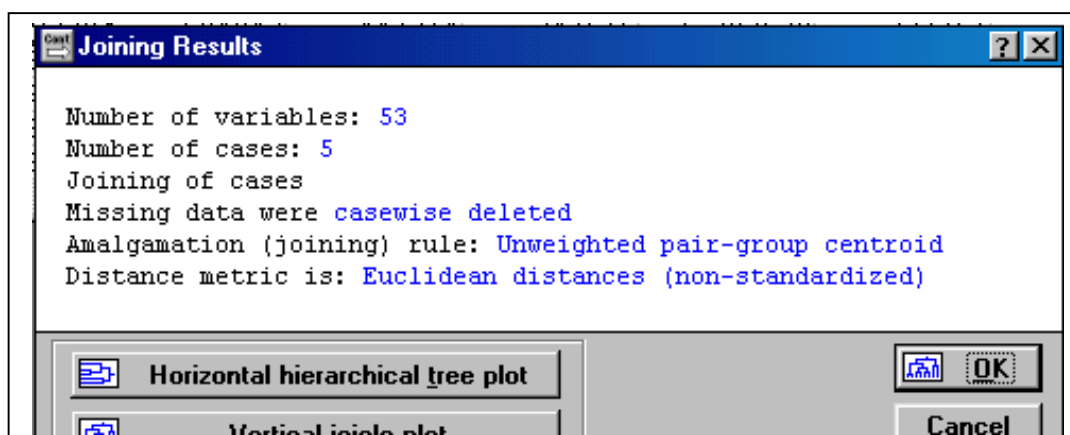


Рис. 38.

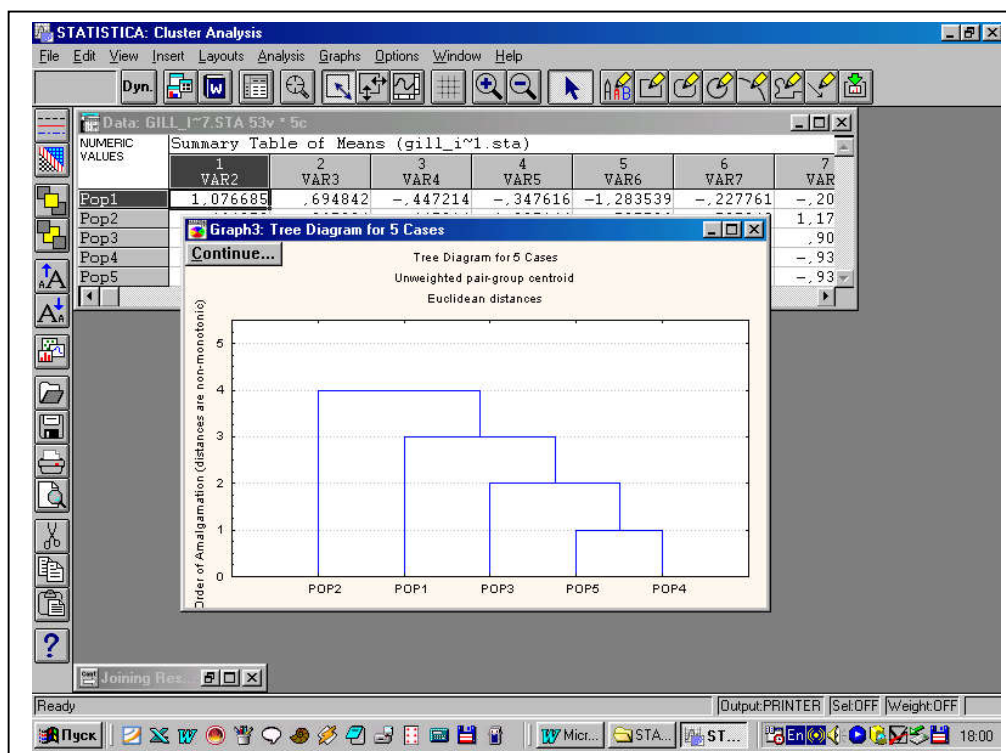


Рис. 39.

Збережемо цю матрицю під іменем **Gill_immuno_1_matrix_pop.sta**. Тепер відкриємо модуль “Багатовимірне неметричне шкалювання” (**Multidimensional Scaling**) (рис. 40). А в цьому модулі відкриємо файл із матрицею.

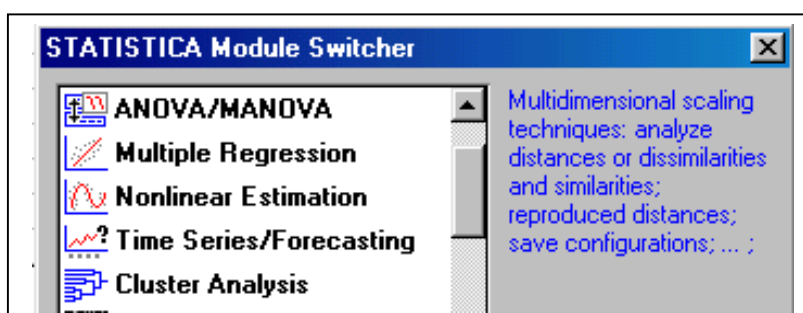


Рис. 40.

Зробимо наступні установки у вікні, що відкриється (рис. 41) та клацнемо по кнопці **OK**.

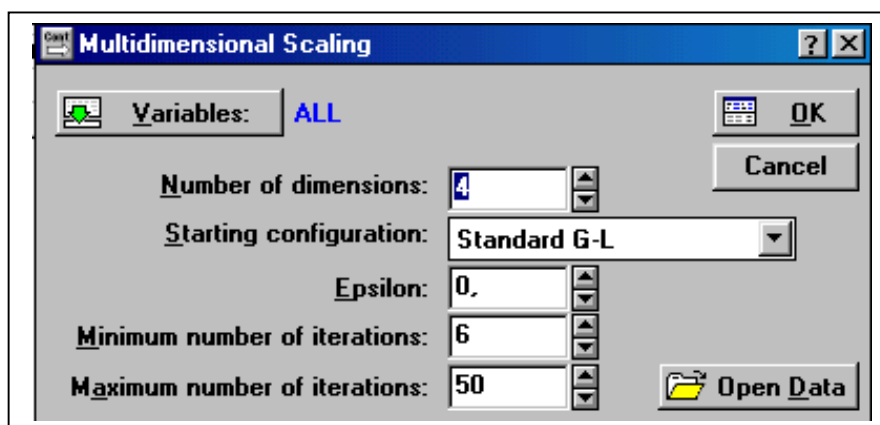


Рис. 41.

Почнеться ітераційна процедура пошуку оптимального розміщення об'єктів у новому просторі (рис. 42). Коли вона закінчиться необхідно клацнути по кнопці **OK** для того, щоб отримати результати (рис. 43).

| Parameter Estimation | | | | | | |
|----------------------|---------|------------|------------|------------|------------|----------|
| iter. | [dim=4] | D-star | D-star | D-hat | d-hat | |
| s: t: | cosin | step | raw stress | alienation | raw stress | stress |
| 0 | 0 | | ,0264024 | ,0324933 | | |
| 1 | 1 | ,200 | ,0118653 | ,0217931 | | |
| 1 | 2 | ,978 ,731 | ,0025585 | ,0101064 | | |
| 1 | 3 | -,480 ,264 | ,0005945 | ,0048783 | | |
| 1 | 4 | ,914 ,633 | ,0001797 | ,0026789 | | |
| 1 | 5 | -,410 ,267 | 0,000000 | 0,000000 | | |
| 0 | * | | 0,000000 | 0,000000 | 0,000000 | 0,000000 |

Рис. 42.

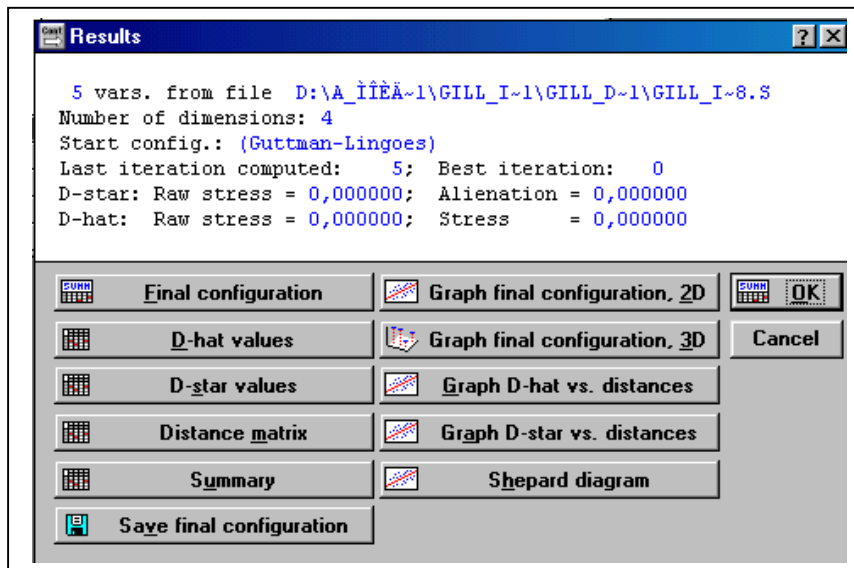


Рис. 43.

Для візуалізації об'єктів (популяцій) у новому багатовимірному просторі необхідно клацнути по кнопці **“Graph final configuration, 2D”** (Двовимірний графік фінальної конфігурації) й ми отримуємо шуканий графік (рис. 44).

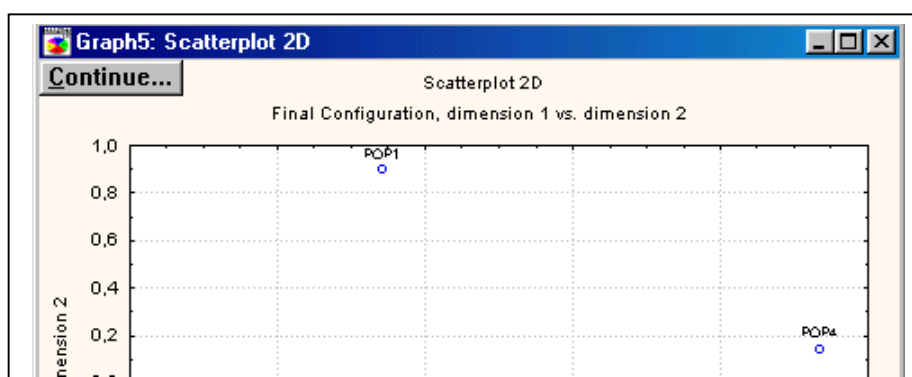


Рис. 44.

Після внесення необхідних змін та формування, цей графік вже набуває закінченого вигляду (рис. 45).

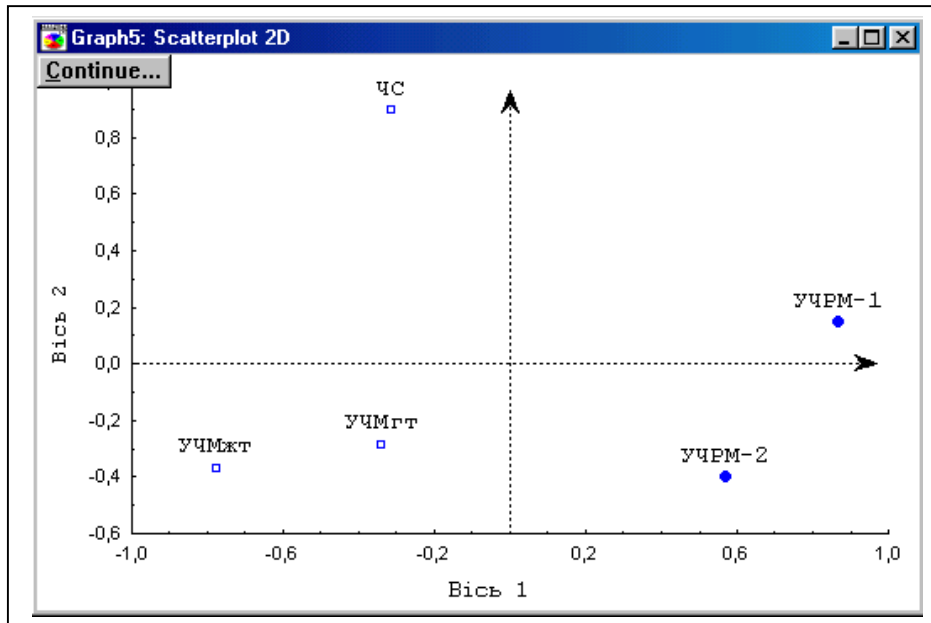


Рис. 45.

Q-аналіз також проводиться на вихідній базі даних (файл **Gill_immuno_1.sta**), але має деякі відмінності. Відкриємо цей файл у модулі “Кластерний аналіз” та зробимо необхідні установки (рис. 46).

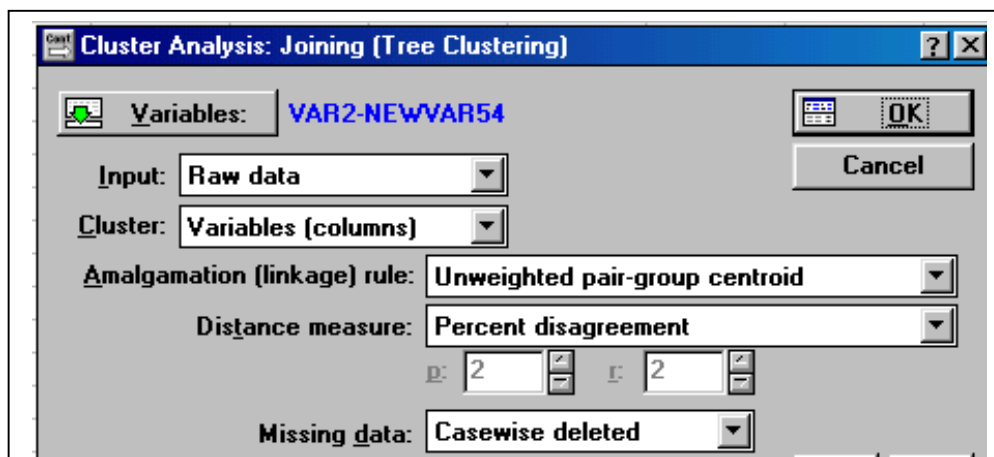


Рис. 46.

Матрицю із відстанями між 53 еритроцитарними антигенами збережемо під іменем **Gill_immuno_1_matrix_antig.sta**. Далі відкриємо цю матрицю в модулі “**Багатовимірне неметричне шкалювання**”, зробимо необхідні установки (рис. 47) та клацнемо кнопку **OK**.

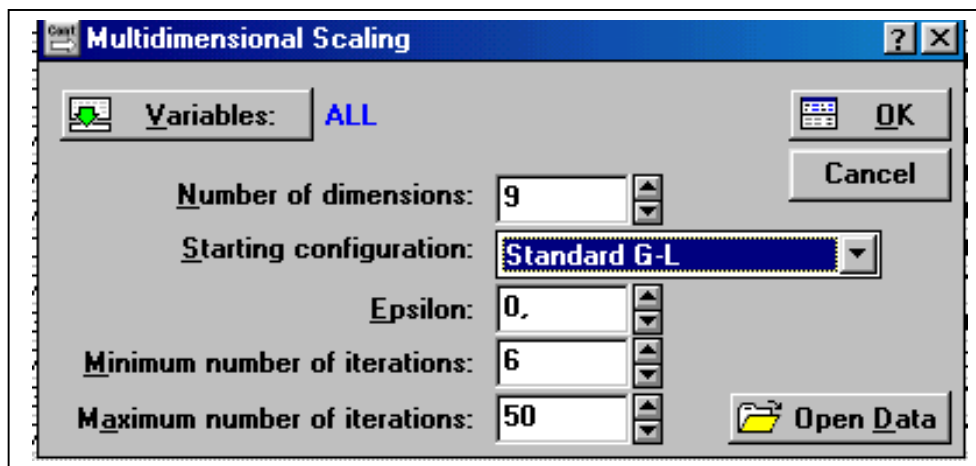


Рис. 47.

Після ітераційної процедури пошуку оптимальної конфігурації ми можемо візуалізувати отриману топологію об'єктів, побудувавши двовимірний графік (рис. 48).

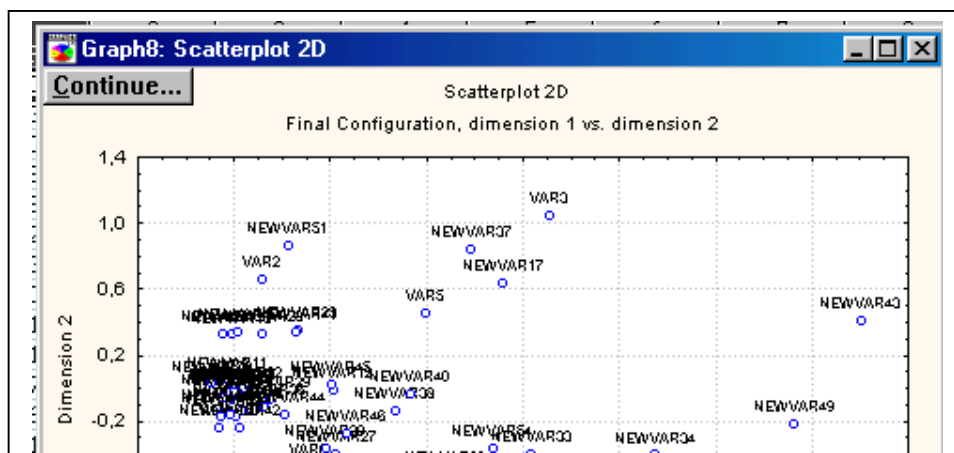


Рис. 48.

Після внесення необхідних змін та формування, цей графік вже набуває закінченого вигляду (рис. 49).

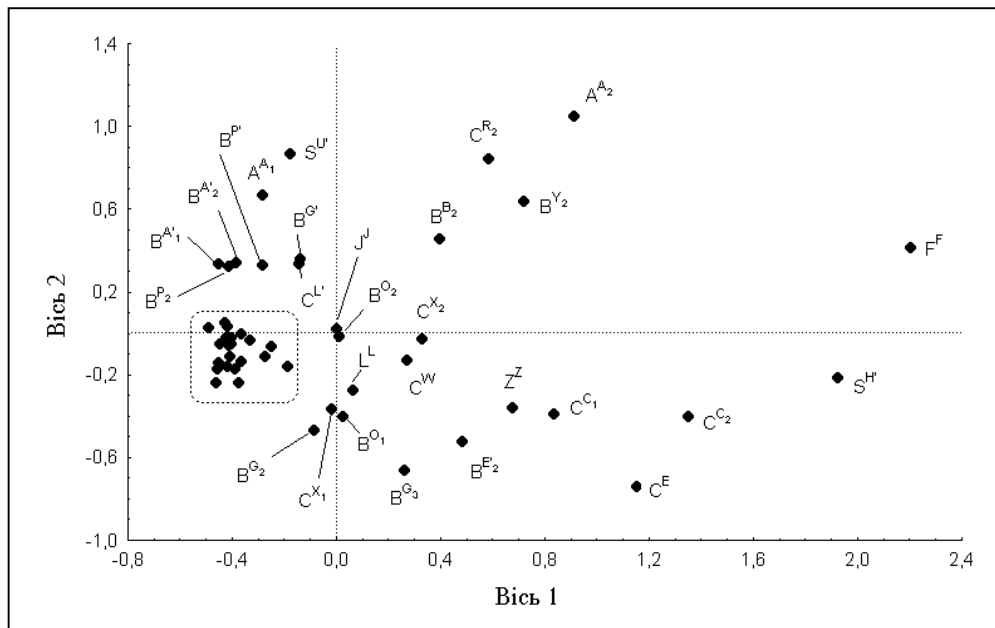


Рис. 49.

СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Вдовиченко Ю., Жарук П. Генетичні ресурси овець в Україні. *Вісник аграрної науки*. 2019. Т. 97, № 5. С. 38-44.
2. Войтенко С., Сидоренко О. Збереження генофонду та підвищення продуктивності худоби білоголової української породи. *Вісник аграрної науки*. 2021. Т. 99, № 2. С. 41–51.
3. Войтенко С. Л., Порхун М. Г., Сидоренко О. В., Ільницька Т. Є. Генетичні ресурси сільськогосподарських тварин України на початку

- третього тисячоліття. *Розведення і генетика тварин*. 2019. Вип. 58. С. 110-119.
4. Гладій М. В., Полупан Ю. П., Ковтун С. І., Кузубний С. В., Вишневський Л. В., Копилов К. В., Щербак О. В. Наукові та організаційні аспекти розведення, генетики, біотехнології та збереження генофонду у тваринництві. *Розведення і генетика тварин*. 2018. Вип. 56. С. 5-14.
 5. Дзіцюк В. В., Типило Х. Т., Гузеватий О. Є. Цитогенетика сільськогосподарських і домашніх тварин : монографія. Київ : Аграрна наука, 2021. 127 с.
 6. Кругляк О. В. Генетичні ресурси молочного скотарства України. *Економіка АПК*. 2018. № 1. С. 33-39.
 7. Методологія оцінки генотипу тварин за молекулярно-генетичними маркерами у тваринництві України : монографія / К. В. Копилов, О. М. Жукорський, К. В. Копилова та ін.; за наук. ред. акад. НААН М. В. Гладія. Київ : Аграрна наука, 2015. 208 с.
 8. Почукалін А. Є., Прийма С. В., Різун В. Забезпеченість генетичними ресурсами скотарства України. *Вісник Сумського національного аграрного університету. Серія «Тваринництво»*. 2022. № 1. С. 59-64.
 9. Селекційно-генетичний моніторинг у конярстві / за ред. І. В. Ткачової. Київ : Аграрна наука, 2018. 204 с.
 10. Сідашова С. О., Ковтун С. І. Генетичні ресурси племінних молочних стад: селекційний потенціал кращих корів та ефективність їх відтворення. *Розведення і генетика тварин*. 2018. Вип. 55. С. 209-219.
 11. Супрун І. Генетичні ресурси рисистого конярства в Україні. *Вісник Сумського національного аграрного університету. Серія «Тваринництво»*. 2020. № 3. С. 67-76.
 12. Хмельничий Л. М., Павленко Ю. М. Генетичні маркери в селекції та збереженні генофонду бурої худоби Сумського регіону. *Вісник Сумського національного аграрного університету. Серія «Тваринництво»*. 2021. № 3. С. 3-6.
 13. *The Genetics of the Pig* / Edited by M. Rothschild, A. Ruvinsky. CABI Publishing, 2011. 520 p.

Навчальне видання

Генетичні ресурси сільськогосподарських тварин

Методичні рекомендації

Укладач: **Крамаренко** Олександр Сергійович

Формат 60 × 84/16. Ум. друк. арк. 2,0.
Тираж 10 прим. Зам. №523.

Надруковано у видавничому відділі
Миколаївського національного аграрного університету
54020, м. Миколаїв, вул. Георгія Гонгадзе, 9

Свідоцтво суб'єкта видавничої справи ДК № 4490 від 20.02.2013 р.