## NUMERICAL SIMULATIONS OF SELECTION USING A SPREADSHEET

• Can use spreadsheet programs (like Excel) to execute repetitive calculations, like the recursion equations we have seen in this course.

– In other words, spreadsheets can be used as a simple programming language.

• Purpose of this demonstration: show how to do this, using one-locus selection as a working example, with one of the more widely used spreadsheet programs: Microsoft Excel.

– Note: many of the commands and procedures in this demonstration are very useful for other purposes, like analyzing data, keeping track of grades, alphabetizing, budgeting, etc.

• **Some Useful Terms and Commands**

**cell location**: The location of a cell is given by a number (row) and letter (column) combination). For example, cell B9 is in the 9th row of the 2nd column (column B).

**function**: Fills cell with the result of a calculation. Functions always start with "=". For example, to compute (1+2+3)*4, type "=(1+2+3)*4" in a cell and press return (enter).

• Excel has a wide variety of built-in functions that are useful for data analysis.

**fixed and absolute cell references**: References to specified cell locations.

• *Relative* references are used for a *variables*. A relative reference refers to a cell that is a fixed *relative* location from the current active cell. When a cell reference is entered by selecting a cell, Excel automatically assumes you want a relative reference.

• *Absolute* references are used for fixed *parameters*. An absolute reference refers to a fixed cell location in the spreadsheet. To create an absolute reference, either place the cursor to the right of the cell location and type "command-t" or type "$" in front of the row and column names. For example, $B$9 is a absolute reference to the 9th cell of column B.

**fill down**: Use this to "iterate" your computations.

• Select the current cell and those below that you want to fill. Type "command-d" or select "fill down" from the "edit" menu. The cells below will be filled with the contents of the topmost cell (including adjusted relative cell references). **Fill right/fill up/fill left** are used similarly.

• NOTE: **copy** and **paste** also adjust cell references and can be used for the same purpose as fill.

**<u>freeze panes</u>**: Use to “freeze” column and/or row headings

- To freeze column headings only: select cell just below the first column heading and then choose “freeze panes” from the “window” menu
- To freeze row headings only: select cell just to the right of the first row heading.
- To freeze row and column headings: select cell in the first row and column.

**• Tutorial: Selection at a Haploid Diallelic Locus**

– Goal: compute frequency of *A* for 50 generations given a selection coefficient *s* and initial value of *p* using the equation: $\Delta p = pq\frac{s}{1 - sq}$

– Steps:

(1) Enter parameter value for selection coefficient, *s* (e.g., set $s = 0.1$).

(2) Set up column for the generation numbers.

- Enter “0” for generation 0.
- Enter the formula “= [*click cell above*] + 1” to generate next generation.
  - – Important: use relative reference since “generation” is a variable.
  - – Simply click cell to enter its location (instead of typing it into the formula).
  - – Press return (enter).
- Select the previous cell and the next 49 cells below; choose “fill down” from the edit menu.
  - – Note: fill down enters the correct relative reference (= 1 cell above) for each cell.

(3) In column B, enter the initial value of *p*. (e.g., set $p = 0.2$).

(4) In column C, enter the <u>formula</u> for *q*: Type “=1 – [*click cell to the left*]” .

- Hint: Type in formula, then replace letters with their respective cell locations.

(5) In column D, enter the formula for $\Delta p$.

• Important: use an absolute reference for *s*.

(6) In the next row, column B, enter the formula for $p' = p + \Delta p$:

• Type "= [*click cell above*]+[*click cell for* $\Delta p$ *in previous row*]" .

(7) Use "fill down" to fill in the remaining row entries.

• Note: Fill Down gives the correct relative cell locations and keeps the absolute cell locations fixed.

(8) Starting with the second row, use "fill down" to fill in the rows for the remaining generations.

• Useful to freeze row headings when examining results.

– With this program in place, it is easy to explore at the effects different selection coefficients, different initial values of *p*, etc.

• Change a parameter or initial value, and Excel updates the rest.

– Can even graph the results if desired.

• Select cells containing information to be graphed, then use "Insert Chart" and the "Chart Wizard".

**• Over- and Underdominance in Fitness at a Diallelic Diploid Locus**

– In this example, use Excel to iterate the equation $\Delta p = pq\dfrac{qt - ps}{1 - p^2 s - q^2 t}$ for various values of the selection coefficients *t* and *s*, and initial frequencies of *A* .

• parameters: $s, t$

• variables: $p, q, \Delta p$

– Compare results of iteration with predicted equilibrium: $\hat{p} = t/(s + t)$

**• Exploring the Adaptive Landscape and Fisher's Fundamental Theorem**

– Use Excel to follow changes in *p* and $\overline{w}$ .

– Utilize the general formula for selection at a diallelic diploid locus: $\Delta p = pq\frac{\overline{w}_A - \overline{w}_a}{\overline{w}}$

where $\overline{w}_A = pw_{AA} + qw_{Aa}$, $\overline{w}_a = pw_{Aa} + qw_{aa}$, and

$$\overline{w} = p^2 w_{AA} + 2pqw_{Aa} + q^2 w_{aa} = p\overline{w}_A + q\overline{w}_a.$$

- parameters: $w_{AA}$, $w_{Aa}$, $w_{aa}$
- variables: $p$, $q$, $\overline{w}_A$, $\overline{w}_a$, $\overline{w}$

– Compare actual value of $\Delta\overline{w} = \overline{w}' - \overline{w}$ with the predicted value based on Fisher's Fundamental Theorem of Natural Selection: $\Delta\overline{w} \approx \mathrm{var}(\alpha)/\overline{w} = 2\left(p\alpha_A^2 + q\alpha_a^2\right)/\overline{w}$, where $\alpha_A = \overline{w}_A - \overline{w}$ and $\alpha_a = \overline{w}_a - \overline{w}$.