Verge Genomics
Statistician Skills Test

1. Instructions to Job Candidate

Dear Job Candidate,

Welcome to the Verge Genomics skills test for statisticians. This test is an opportunity for us to learn more about your technical abilities and communication style. Please read the experimental design in section 2, and answer the questions listed in section 3.

This test has an attached CSV file named "ValidationStudy88_PlateID246.csv"
Please ensure you have this CSV before proceeding with the test.

This task will involve processing the CSV file in an analysis environment of choice, such as R, Python, etc. We anticipate this test can be accomplished in approximately two hours or less. Feel free to copy-paste the questions into a new document and reformat your responses into whatever style you feel best communicates your thought process.

When you are finished, please email your report to shannon@vergegenomics.com
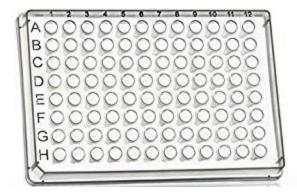
Best Regards,

Victor Hanson-Smith, Ph.D.
Head of Computational Biology
Verge Genomics
South San Francisco, CA

2. Experimental Design

In this experiment drugs are screened for their ability to extend cell survival. For this, we transfer cells into the wells of a 96 well plate and add different compounds to these wells We then track the survival of the cells by taking daily microscopic images of each well for 10 days. The compounds are either drugs suspected to extend survival or a control compound (DMSO). Each microscopy image is computationally processed to identify individual living cells within each well, and to track each cell's survival over time. For each cell, we record the last observed time point (day) at which the cell was observed to be alive. This survival information was compiled into a CSV file (attached to this test), in which each row corresponds to one cell; columns express information about that cell, including in which well it was, with which drug the well was treated with, and the observed survival of the cell.

A 96 Well Plate

A well plate is a flat plate that looks like a tray with multiple wells that are used as small test tubes. The most commonly used well plate has 96 wells, with rows A through H and columns 1 through 12. In our assay, a predefined number of cells is transferred into each of these wells at the beginning of the experiment. A microscopic image is taken of each of these wells daily, for 10 days.

Drugs & Concentrations

The following drugs (VRG drugs, and DMSO as control) were tested:
- VRG0023 3.5 uM
- VRG0037 4.0 uM
- VRG0045 2.0 uM
- VRG0106 2 uM
- VRG0024 0.012 uM
- DMSO as control

96 Well Plate Design

The following table is a map of a stylized 96 well plate, indicating which drugs and concentrations were placed within each well.

VRG0023, 3.5	VRG0023, 3.5	VRG0037, 4	VRG0037, 4	VRG0106, 2	DMSO	DMSO	VRG0106, 2	VRG0045, 2	VRG0045, 2	VRG0024, 0.012	VRG0024, 0.012
VRG0023, 3.5	VRG0023, 3.5	VRG0037, 4	VRG0037, 4	VRG0106, 2	DMSO	DMSO	VRG0106, 2	VRG0045, 2	VRG0045, 2	VRG0024, 0.012	VRG0024, 0.012
VRG0023, 3.5	VRG0023, 3.5	VRG0037, 4	VRG0037, 4	VRG0106, 2	DMSO	DMSO	VRG0106, 2	VRG0045, 2	VRG0045, 2	VRG0024, 0.012	VRG0024, 0.012
VRG0023, 3.5	VRG0023, 3.5	VRG0037, 4	VRG0037, 4	VRG0106, 2	DMSO	DMSO	VRG0106, 2	VRG0045, 2	VRG0045, 2	VRG0024, 0.012	VRG0024, 0.012
VRG0024, 0.012	VRG0024, 0.012	VRG0045, 2	VRG0045, 2	VRG0106, 2	DMSO	DMSO	VRG0106, 2	VRG0037, 4	VRG0037, 4	VRG0023, 3.5	VRG0023, 3.5
VRG0024, 0.012	VRG0024, 0.012	VRG0045, 2	VRG0045, 2	VRG0106, 2	DMSO	DMSO	VRG0106, 2	VRG0037, 4	VRG0037, 4	VRG0023, 3.5	VRG0023, 3.5
VRG0024, 0.012	VRG0024, 0.012	VRG0045, 2	VRG0045, 2	VRG0106, 2	DMSO	DMSO	VRG0106, 2	VRG0037, 4	VRG0037, 4	VRG0023, 3.5	VRG0023, 3.5
VRG0024, 0.012	VRG0024, 0.012	VRG0045, 2	VRG0045, 2	VRG0106, 2	DMSO	DMSO	VRG0106, 2	VRG0037, 4	VRG0037, 4	VRG0023, 3.5	VRG0023, 3.5

Important Columns in the CSV

The attached CSV file, named "ValidationStudy88_PlateID246.csv", includes the following important columns:

- **cell_id**: a unique integer ID for each cell, unique across the entire experiment
- **well_id** and **well_name**: the unique ID and name of the well. In this experiment, there are 96 wells total.
- **well_column** and **well_row**: the column and row index for this well on the plate. The 96 well plate is arranged as a grid of 8-by-12 wells; see the Plate Design above.
- **well_drugname**: the name of the drug applied to the well.
- **well_drug_concentration** and **well_drug_concetration_units**: the concentration and units (in micromolar) applied to the well.
- **cell_last_tp**: The last time point at which the cell was identified alive through analysis of the microscopy images. There are 10 total timepoints (days). Cells that survived through time point 10 are considered alive at the end of the experiment; cells that survived only to timepoint 9 or lower are considered to have died.
- **cell_event**: TRUE/FALSE: did the cell die before timepoint 10?

3. Questions for the Candidate

1. Which drugs, if any, significantly improved cell survival relative to DMSO control?
   <u>Notes:</u>
   - Cell survival is recorded in the CSV in the column named *cell_last_tp*.
   - Please describe how you assess "significance," and explain any assumptions you made about the data.
   - Please include any code and/or figures you used to process the data and arrive at your answer.

2. What are your thoughts on the plate design and our statistical power in this study?
   <u>Notes:</u>
   - Plate location is recorded in the CSV in the columns named *well_column* an *well_row*.
   - Please include any code and/or figures you used to process the data and arrive at your response.

3. Could you briefly summarize your results from questions #1 and #2 for a non-statistics audience? Please mention any follow-up experiments or studies that might be suggested by the results. Please write your summary as if it was a paper abstract, broadly understandable by scientists. Please keep your response to less than 200 words.

When you are finished, email your report to shannon@vergegenomics.com