Two-Way ANOVA tests

Contents at a glance

Agilent Technologies
genespring_support@agilent.com
Phone (866) 744-7638

I. Definition and Applications

Whereas one-way analysis of variance (ANOVA) tests measure significant effects of one factor only, two-way analysis of variance (ANOVA) tests (also called two-factor analysis of variance) measure the effects of two factors simultaneously. For example, an experiment might be defined by two parameters, such as treatment and time point. One-way ANOVA tests would be able to assess only the treatment effect or the time effect. Two-way ANOVA on the other hand would not only be able to assess both time and treatment in the same test, but also whether there is an interaction between the parameters. A two-way test generates three p-values, one for each parameter independently, and one measuring the interaction between the two parameters.

II. Two-Way ANOVA prerequisites

Before running a two-way ANOVA test, experimental data must meet these prerequisites:

1) The parameters tested need to be part of the experimental design.
2) Parameters need to be defined appropriately in the Experiment Parameter window (under **Experiments** menu). Your experiment should contain at least two parameters, and each set of replicates should share common parameter values (see Fig 1).

Treatment type	Time
	hours
no	yes
N/A	no
Treated	0
Treated	0
Treated	2
Treated	2
Untreated	0
Untreated	0
Untreated	2
Untreated	2

Fig 1: Sample parameter window settings. There are two parameters, treatment and time, defining four groups of replicate samples with common parameter values (treated vs untreated, time 0 or 2 hours).

3) Two-way ANOVA is most powerful when the experiment has the same number of replicates in each group defined by the pair of parameters. This is called a "balanced design" (see Fig. 2). However, two-way tests can also be applied to "proportional design" experiments, where the proportion of samples across each parameter group is retained (see Fig. 3).

	Control	Treated
Time 0	4	4
Time 2	4	4
Time 4	4	4
Time 8	4	4

Fig. 2: Example of a balanced design replication: there are 4 replicate samples for each group.

	Control	Treated
Time 0	3	4
Time 2	6	8
Time 4	3	4
Time 8	9	12

Fig. 3: Example of proportional design replication.

Experiments with mild deviations from a proportional design may still be analyzed, but experiments with a highly disproportional design cannot be analyzed using two-way ANOVA. (GeneSpring will display an error message if attempting to run a Two-Way ANOVA on such a data set). Two-way tests can also be analyzed on data with only one replicate per group or condition. However, the interaction between the factors cannot be tested.

In summary, two-way ANOVA tests are best to use when your experiment was designed to measure two different factors, or when you wish to test two factors at the same time.

III. How to use the Two-Way ANOVA tool

After defining the parameters in the Experiment Parameter window, select “Statistical Analysis (ANOVA)” from the **Tools** menu. Select the “2-Way Tests” tab in the middle of the window (Fig. 4).

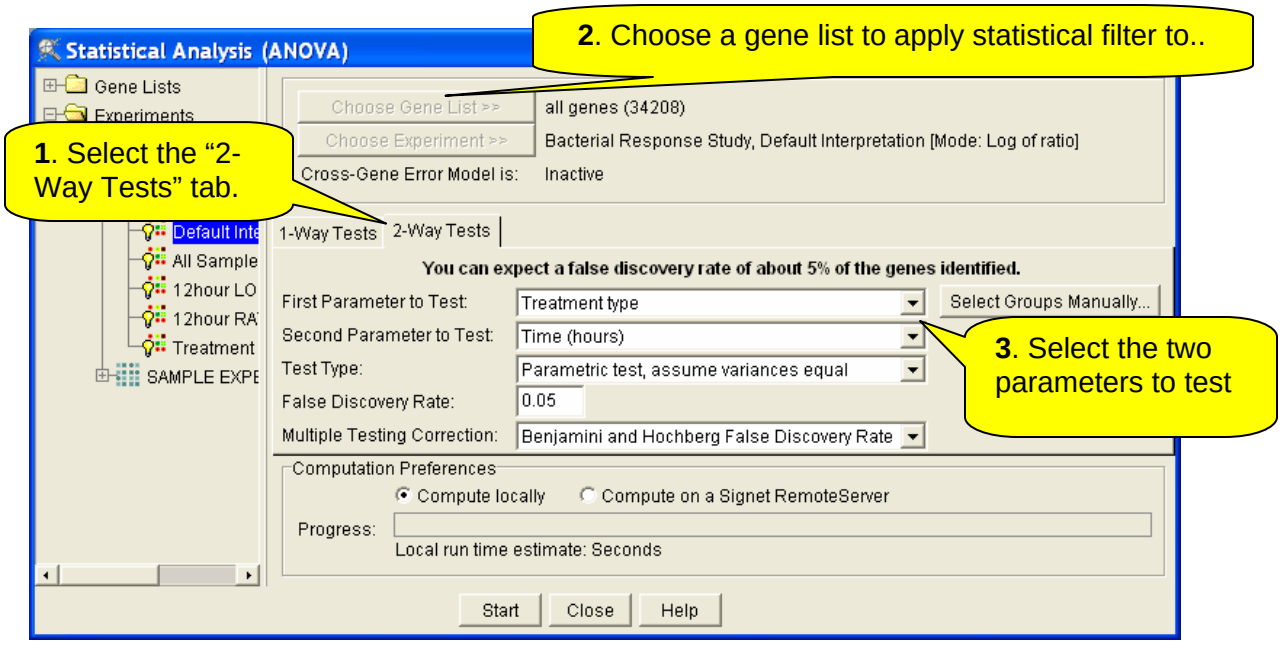

Fig. 4: Statistical Analysis window. Make the choice of one-way or two-way tests (i.e. ANOVA) from the tabs (balloon 1).

Next, set up the analysis as follows:

1. Select the gene list you wish to apply the analysis to from the Gene List folder in the left-hand panel (Navigator), and click “Choose Gene List”.
2. The experiment and associated interpretation selected by default is the one displayed in the main GeneSpring window. If you wish to change the experiment settings, select it from the Experiments folder on the right-hand panel, and click “Choose Experiments”.
3. Select the first parameter to test from the first pull-down menu.
4. Select the second parameter to test from the second pull-down menu.
5. Choose a statistical test from the “Test Type” pull-down menu.

A. Parametric test, assume variances equal.

Parametric tests are the most widely used tests in statistical analysis. The parametric test offered in the two-way ANOVA window assumes normal distribution and equal variances between groups. Even if you are not certain

about both these criteria, the parametric test assuming equal variance is a very robust test, and should be safe to use.

B. Non-parametric test

For those who prefer to use non-parametric tests, the GeneSpring Two-Way ANOVA tool offers an option to use one. Non-parametric tests have the advantage that there is no underlying assumption of normality. However, using that test will only calculate p-values for each parameter effect, but no interaction p-value. If you wish to use Two-Way ANOVA to determine whether there is an interaction between your two parameters, choose the parametric test instead. Another disadvantage of the non-parametric test is that you often need more replicate samples per condition to achieve the same power as a parametric test.

C. Multiple Testing Corrections

A choice of multiple testing corrections is offered to control for occurrence of false positives that arise by virtue of performing many tests. False positive genes are those that are unduly found to be significant when they are not.

The purpose of multiple testing correction is to keep the overall error rate/false positives to less than the user-specified p-value cutoff, even if thousands of genes are being analyzed.

Test Type	**Type of Error control**	**Genes identified by chance after MTC**
Bonferroni	Family-wise error rate	If testing 10 000 genes with p-value cutoff = 0.05, expect **0.05** genes passing filter to show significance by chance
Bonferroni step-down (Holm)		Same as above
Westfall and Young Permutation		Same as above
Benjamini and Hochberg	False Discovery Rate	If testing 10 000 genes with p-value cutoff = 0.05, expect **5%** of genes to show significance by chance.

D. Recommendations

The default correction set for multiple testing is the Benjamini and Hochberg False Discovery Rate. This procedure is the least stringent of all the methods mentioned above, but it provides a good balance between identification of many

of the statistically significant genes and protection against false positives (Type I error).

For more specific information on Multiple Testing Correction, consult the Analysis Guide on the topic.

IV. How to interpret the results?

Filtering using the Two-Way ANOVA tool offers the possibility to save three gene lists:

1. A list of genes found to be significant across the first parameter
2. A list of genes found to be significant across the second parameter
3. A list of genes found to show a significant interaction between the two parameters.

	Gene Name	Treatment type p-value	Time p-value	Interaction p-value
1	AA827680	1.34e-5	0.00692	5.67e-8
2	AA837363	3.04e-6	0.00692	2.6e-6
3	AA810021	7.37e-6	0.00712	2.58e-8
4	AA833800	0.00312	0.00813	5.23e-5
5	AA827203	0.00276	0.0124	0.000168
6	AA831677	4.09e-5	0.0443	0.00013
7	AA290624	0.156	0.047	0.0196
8	AA789125	0.000853	0.047	2.15e-7
9	AA744684	0.00272	0.182	5.08e-6
10	AA810101	0.00585	0.21	8.9e-5
11	AA808080	0.00281	0.383	0.000872
12	T71976	3.39e-6	0.383	5.67e-8
13	AI057022	0.0281	0.383	0.0996

Copy to Clipboard | Save Lists... | Display in Venn Diagram | Cancel | Help

Fig 5: Table of Two-Way ANOVA results. Four columns are displayed: the gene name, the first parameter p-value, second parameter p-value and parameter interaction p-value for each gene found to be significant in at least one column. P-values shown in gray are those that did not pass the specified cut-off.

Interpreting those different results is critical to understanding the results of the Two-Way ANOVA test.

A. Individual factor's p-values

Each parameter is first tested individually, independently from the other parameter. Two-Way ANOVA first performs two equivalent tests of ANOVA. Genes that are found to be significant in at least one test are displayed in the resulting table (Fig 5).

a) Genes that have a p-value less than the specified cutoff in the first column (Treatment in Fig 5 example) are found to be significant among groups defined by the first parameter.

b) Genes that have a p-value less than the specified cutoff in the second column (Time in Fig 5 example) are found to be significantly affected by the second parameter.

These two gene lists can be saved by selecting "Save Lists" in the test results window.

B. Interaction p-values

Genes with a p-value less than the specified cutoff in the third column (Interaction p-value) show a significant interaction between the two parameters tested (in this example, Treatment and Time). A list of genes showing interaction can be saved by selecting "Save Lists" in the test results window.

The relationship of a gene's expression across two parameters can be classified depending on whether it is a function of one of the parameters, both independently or both with interaction effects. The idealized schematics below provide examples of each. See the sets represented as part of Venn Diagram sets (Fig 7) as well.

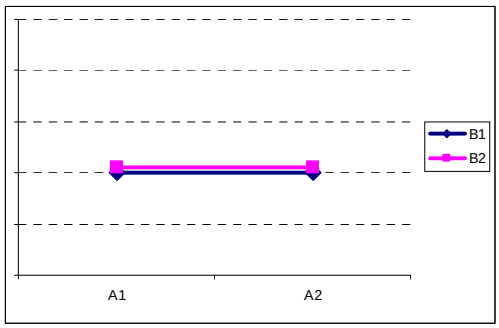

Fig 6a: No effect, either parameter

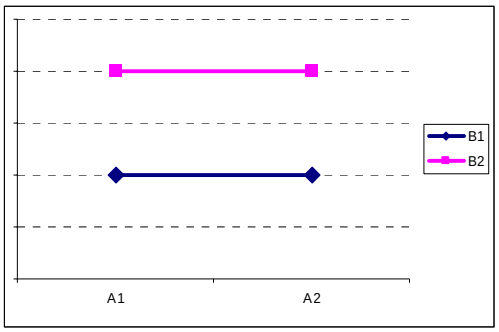

Fig 6b: No parameter A effect, parameter B effect

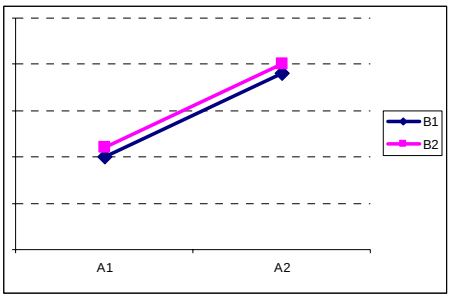

Fig 6c: Parameter A effect, no parameter B effect

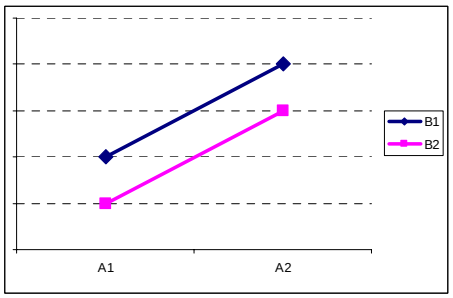

Fig 6d:Parameter A effect, Parameter B effect

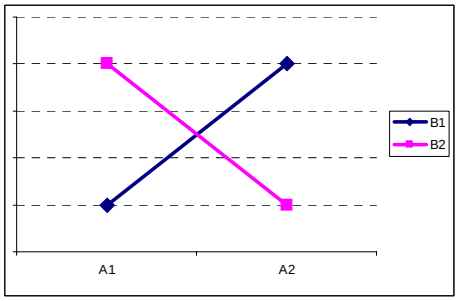

Fig 6e: No main effects, interaction effect

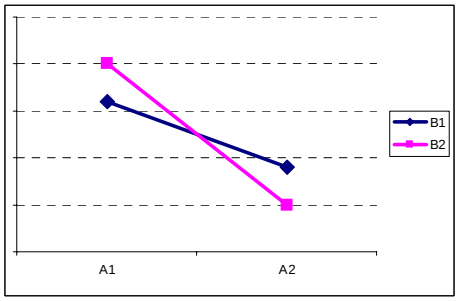

Fig 6f: A effect, no B, slight interaction

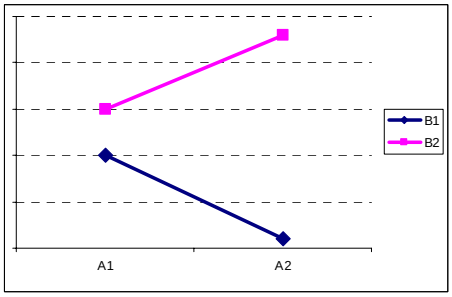

Fig 6g: No A, B effect, large interaction

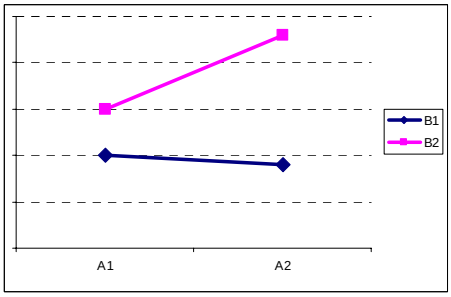

Fig 6h: A effect, B effect, interaction effect

Fig 6: Both parameter Treatment (A) and Time (B) are represented simultaneously in the same scheme. Treatment, A1 and A2, are represented on the X-axis. Time points, B1 and B2 are color-coded, in blue and pink respectively.

In Fig 6a, neither a change in treatment nor a change in time leads to a change in expression.

Fig 6b and 6c represent examples where either time or treatment show an effect in measurement, but not both. Fig 6d represents a case where both treatment and time show an effect. Since both parameters show a similar effect, there is no interaction between the two.

Fig 6g shows an example where time has a large effect on measurement, but not treatment. However, since time seems to have a different effect between the two treatments, we say that there is an interaction between the treatment and time. Note that the more the two profiles deviate from paralleling each other, the greater the interaction effect.

All three gene lists can also be compared using the Venn Diagram, to observe overlap between lists. A Venn Diagram can be displayed by clicking on “Display In Venn Diagram” in the test results window. The Venn Diagram will be displayed in the main GeneSpring window.

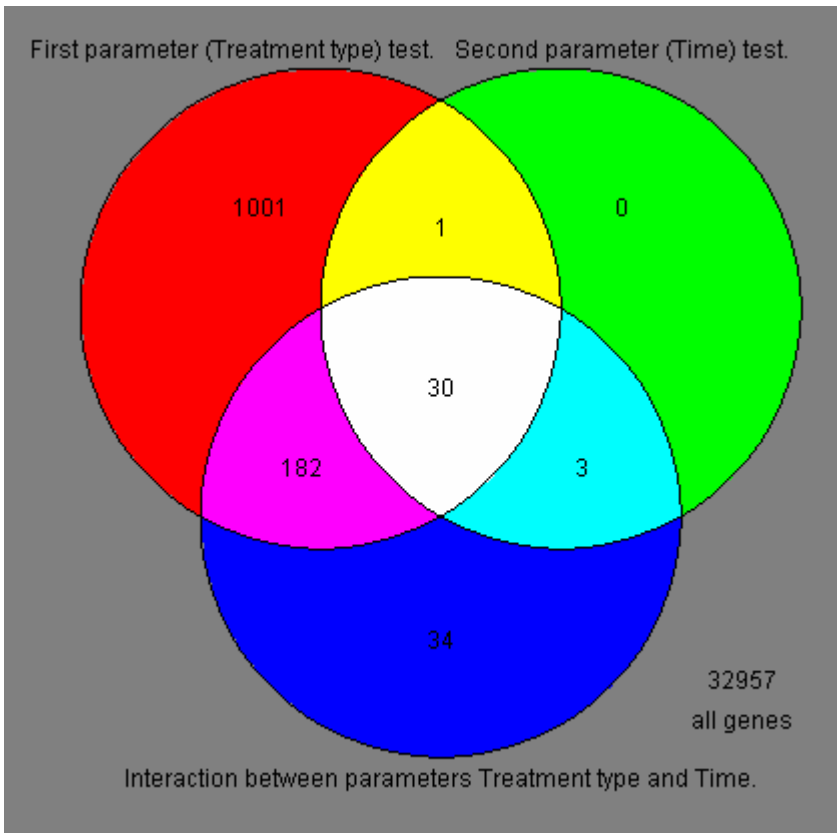

Fig 7: A Venn Diagram view allows you to save gene lists that intersect between different categories, such as genes that are affected by Treatment type and show an interaction between both Treatment and Time.

V. Technical details.

For each gene, 2-way ANOVA results are calculated using a sum of squares decomposition. The total sum of squares for a data set is a measure of the variability among all the data. The idea behind an analysis of variance (ANOVA) is to divide this total variability into variability between groups and variability within groups (also referred to as error variability). If the variability between groups is large compared to the variability within groups, as determined via a statistical test, we conclude that there are significant differences between groups.

In the case of a 2-way ANOVA, the total variability is divided up into four components: variability among the levels of each of the two factors, variability due to interaction of the two factors, and variability within cells (error variability). Three separate statistical tests are performed (based on the F statistic), comparing each of the first three sources of variability (variability due to first factor, variability due to second factor, and variability due to interaction) to the error variability. In each test, the resulting p-value allows us to determine whether that specific effect is significant.

VI. Frequently asked questions

Q. How can genes show no significant p-value for either parameter, but a significant interaction p-value?

A. Sometimes, the effect of individual parameters cannot be measured, even though parameter effect could be observed on their own. Fig 6e illustrates such an example, where the effect of neither drug type or drug concentration can be measured. However, it is clear from the diagram that both drugs have a very different effect at different concentrations. When one is decreasing gene expression measurement at a certain concentration, the other drug increases that same gene expression at the same concentration. Therefore both effects annihilate each other, but an interaction between drug type and concentration is found to be significant.

Q. How do I ensure my data is balanced for 2-way ANOVA?

A. Design the experiment to have a balanced design in the beginning. Ensure that you have equal replicates in each group. See Fig 3 for an example of a design replication experiment. A proportional design is also possible, but not as powerful.

Q. What is the difference between one-way and two-way ANOVA?

A. Two-way ANOVA analyzes the effect on gene expression of two parameters at the same time. To use the two-way ANOVA, you would need to have at least two

genespring_support@agilent.com
Phone (866) 744-7638

parameters to analyze in your data. One-way ANOVA only analyzes one parameter effect on gene expression at a time.

Q. Why aren't post-hoc tests available?

A. Post-hoc tests for 2-way ANOVA are not currently implemented in GeneSpring. Although it is possible to test for differences in the levels of a factor following a 2-way ANOVA, it is not generally advisable if interactions are present. For genes where no interactions are present, the 1-way ANOVA tool can be used to test for differences in a single factor.

VII. Literature

Zar, Jerrold H. (1999) Biostatistical Analysis, 4th ed. Upper Saddle River, NJ. Prentice-Hall.