

Data Set:

Velocity of Rat2 cells (um/hr):

73.44, 96.48, 33.696, 60.84, 48.96, 41.04, 31.86, 44.64, 17.892, 43.92, 31.428, 23.976, 39.6, 27.576, 30.996, 32.472, 34.056, 40.68, 25.128, 28.332, 33.156, 22.356, 32.508, 44.64, 27.504, 31.572, 54, 51.84, 34.848, 37.8 [n=30]

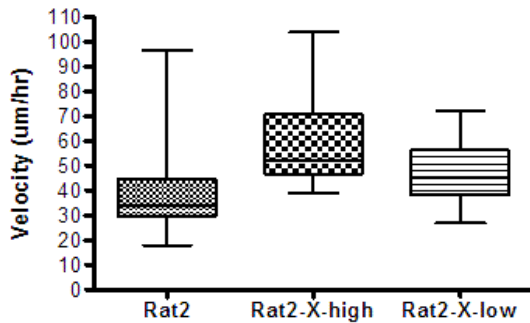
Velocity of Rat2 cells with protein X highly over-expressed (X-high, um/hr):

69.48, 70.56, 46.44, 43.2, 74.52, 52.2, 103.68, 45, 62.64, 38.88, 46.8, 76.68, 49.32, 52.2, 51.84, 70.56, 81.72, 79.56, 59.76, 40.32, 52.2, 50.76, 42.84, 55.08, 50.4, 57.24, 41.04, 90.36 [n=28]

Velocity of Rat2 cells with protein X lowly expressed (X-low, um/hr):

60.48, 37.08, 45.72, 51.48, 41.76, 26.784, 47.88, 59.76, 66.6, 68.4, 72, 32.328, 39.96, 59.04, 66.6, 34.128, 48.24, 32.688, 35.82, 53.28, 44.28, 38.88, 41.76, 44.28, 53.28, 42.48, 43.56, 52.56, 52.56, 35.46 [n=30]

Set alpha=0.05. Use Prism® 4 to do one-way ANOVA and Dunnett’s procedure:



	Rat2	Rat2-X-high	Rat2-X-low
	Y	Y	Y
Number of values	30	28	30
Minimum	17.89	38.88	26.78
25% Percentile	29.66	46.62	37.98
Median	33.88	52.20	45.00
75% Percentile	44.64	70.56	56.16
Maximum	96.48	103.7	72.00
Mean	39.24	59.12	47.64
Std. Deviation	16.07	16.55	11.83
Std. Error	2.933	3.128	2.161
Lower 95% CI	33.24	52.70	43.22
Upper 95% CI	45.24	65.54	52.06

Parameter	Value	Rat2-X-high	Rat2-X-low	Data Set-D
	Y	Y	Y	Y
Table Analyzed				
Data 1				
One-way analysis of variance				
P value	P<0.0001			
P value summary	***			
Are means signif. different? (P < 0.05)	Yes			
Number of groups	3			
F	12.90			
R squared	0.2329			
Bartlett's test for equal variances				
Bartlett's statistic (corrected)	3.583			
P value	0.1667			
P value summary	ns			
Do the variances differ signif. (P < 0.05)	No			
ANOVA Table	SS	df	MS	
Treatment (between columns)	5750	2	2875	
Residual (within columns)	18940	85	222.9	
Total	24690	87		

- a)
Rat2 is a fibroblast cell line suitable for single cell tracking. Our lab's interest is protein X's function in cell migration. Two stable Rat2 cell lines with protein X expressed in different levels were constructed by retrovirus infection and FACS. Basically, we infected Rat2 cells with EGFP-X, and sorted the fluorescent cells into two groups: high green signal and low green signal. Our hypothesis is that over-expression of protein X in Rat2 cells affects cytoskeleton and influences cell migration.
- b)
As in this experiments, all the environment, like cell culture condition, genetic background of cell line (before retrovirus infection), cell-confluence during tracking experiments, parameter-setting of tracking program, *etc.*, is the same. There is only one dependent variable: velocity of migration; one independent variable: levels of extrinsic expressed protein X - none, low or high.
- c)
The required assumptions for one-way ANOVA are:
Independent groups -- The effect of protein X on cell velocity is independent during the experiments.
Continuous interval -- The measurement of velocity is continuous.
Random sampling -- Cells were picked up randomly from a 5-field-movie.
Normal population distribution -- With large n (n=28-30 here), sampling distribution of means will be normal. I also check the values in each group by leaf-and-stem method, which shows that the distribution is normal presumably.
Equal variances of groups -- Based on the calculation within the groups, S.D. of each group are 16.07, 16.55 and 11.83. Bartlett's test gives $p=0.1667$, which means there is no significant difference between groups' variances.
- d)
Null hypothesis: The population means cell migration velocities of Rat2 cells with X-high, X-low and control Rat2 cells are equal.
Alternative hypothesis: There is at least one difference among the population means cell migration velocities of Rat2 cells with X-high, X-low and control.
- e)
I will do Dunnett multiple comparisons following ANOVA. I am interested in whether different expression level of X in cell will have different effect on cell migration. It is a many-to-one procedure, and Dunnett Test is suitable for this.
- f)
Numerator degree of freedom for F is 2; denominator degree of freedom is 85.
- g)
Value of F-critical is 3.104.
- h)

p value of observed result is <0.0001 .

i)

Reject the null hypothesis.

j)

Dunnett's Multiple Comparison Test	Mean Diff.	q	P value	95% CI of diff
Rat2 vs Rat2-X-high	-19.88	5.067	$P < 0.01$	-28.73 to -11.02
Rat2 vs Rat2-X-low	-8.396	2.178	$P > 0.05$	-17.10 to 0.3051

k)

If I only look at means of these three groups, 39.24(control) 47.64(X-low) 59.12(X-high), I will get the conclusion that protein X is an enhancer of cell migration and this effect is dose-dependent.

After ANOVA and Dunnett test, I think there is another possibility: protein X negatively regulates cell migration; there are two pools of protein X, inactive and active forms; lowly expressed extrinsic X stores in the inactive pool and does not affect cell migration; highly expressed X works as a dominant negative by some intra-molecular inhibition, which brings all the active form of both endogenous and extrinsic X to inactive form. One thing I have not mentioned before is that Rat2 cell has endogenous protein X. The expression level of EGFP-X-low is about the same as endogenous X, and the EGFP-X-high is about 3 fold as endogenous. I prefer this interpretation, because it fits with my other biochemistry data.

l)

Data were analyzed first with a one-way ANOVA, conducted with the PC-based program PRISM (GraphPad, San Diego, CA). Alpha was set at 0.05. The results of this analysis indicated that the mean cell migration velocity varied significantly greater among the three groups [$F(2,85)=12.90$, $p<0.0001$]. Specific comparisons among groups were then conducted with Dunnett's test. These results indicated that only highly over-expressed protein X will significantly enhance Rat2 cells migration ($p<0.01$).

m)

An alternative hypothesis is that the highly over-expressed protein X enhances cell migration velocity to two fold, while the lowly expressed protein X in Rat2 does not affect cell migration. I want to detect this difference.

Based on previous data and the hypothesis, I have these means: control=X-low=**39.24**um/hr; X-high=**78.48**um/hr. So the S.D. is **22.66**. **14.93** is the pooled S.D., which will be used as an estimate of the S.D. of the population.

Effect size, $f=22.66/14.93=1.52$.

As the sample size required for one-way ANOVA of $f=0.6$, $N=3$, $\alpha=0.05$, $\beta=0.20$ is 10, I would need less than 10 sample size to detect the difference.