available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Statistical analysis of data from retroviral clonal experiments in the developing retina

Stan Pounds[a,*], *Michael A. Dyer*[b,c,*]

[a]*Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN 38105, USA*
[b]*Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA*
[c]*Department of Ophthalmology, University of Tennessee Health Sciences Center, Memphis, TN 38105, USA*

ARTICLE INFO

Article history:
Accepted 30 August 2007
Available online 11 September 2007

Keywords:
Retrovirus
Retinal development
Lineage analysis
Clonal analysis

ABSTRACT

Retroviral lineage studies have been widely used over the past decade to study retinal development in vivo and in explant culture [Donovan S.L., Dyer, M.A., 2006. Preparation and Square Wave Electroporation of Retinal Explant Cultures, Nature Protocols 1, 2710–2718; Donovan, S.L., Schweers, B., Martins, R., Johnson D., Dyer, M.A., 2001. Compensation by tumor suppressor genes during retinal development in mice and humans, BMC Biol 4 , 14; Dyer M.A., Cepko, C.L., 2001. p27Kip1 and p57Kip2 regulate proliferation in distinct retinal progenitor cell populations, J. of Neurosci 21, 4259–4271; Dyer M.A., Cepko, C.L., 2000. p57 (Kip2) regulates progenitor cell proliferation and amacrine interneuron development in the mouse retina, Development 127, 3593–3605; Dyer, M.A., Livesey, F.J., Cepko C.L., Oliver, G., 2003. Prox1 function controls progenitor cell proliferation and horizontal cell genesis in the mammalian retina, Nat Genet 34, 53–58]. These approaches can provide important data on the proliferation, cell fate specification, differentiation and survival of individual neurons and glia derived from single infected retinal progenitor cells. In some experiments, these parameters are compared in retinae from animals with different targeted deletions or transgenes. Alternatively, the effect of ectopic expression of virally encoded transgenes may be studied at the level of individual retinal progenitor cells in vivo and in explant culture. One of the challenges with interpreting retroviral lineage studies is determining the statistical significance of differences in the proliferation, cell fate specification, differentiation of survival of retinal progenitor cells between experimental and control samples. In this study, we provide a clear step-by-step guide to the application of statistical methods to retroviral lineage analyses actual data sets. We anticipate that this will serve as a guide for future statistical analyses of retroviral lineage studies and will help to provide a uniform standard in the field.

1. Introduction

In vivo retroviral lineage studies in the developing retina were first described by Turner and Cepko (1987). They developed a replication-incompetent retroviral shuttle vector that could be used to produce retroviral stocks of sufficient titer for in vivo retinal infection using newborn rat pups (Turner and Cepko, 1987). In the initial study, a β-galactosidase reporter gene was used and subsequent versions of these viral vectors incorporated human placental alkaline phosphatase and nuclear

* *Corresponding authors.*
E-mail addresses: Stanley.pounds@stjude.org (S. Pounds), Micheal.Dyer@stjude.org (M.A. Dyer).

0006-8993/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2007.08.074

β-galactosidase reporter genes (Dyer and Cepko, 2000; Dyer and Cepko, 2001b; Fields-Berry et al., 1992) (Figs. 1A, B). Alkaline phosphatase is ideally suited for visualizing the morphology of infected retinal neurons and glia because it labels their membrane (Figs. 1C, D) and nuclear β-galactosidase is ideally suited for studies focused on retinal progenitor cell proliferation because the number of nuclei in clones of cells derived from infected retinal progenitor cells can be readily identified (Figs. 1E, F). These and similar techniques have been used over the past decade to study retinal development in vivo and in explant culture (Donovan and Dyer, 2006; Donovan et al., 2006).

The advantage of performing lineage studies in the developing retina is that the daughter cells from individual infected retinal progenitor cells do not migrate laterally so clonal boundaries can be unambiguously identified (Fields-Berry et al., 1992). More recent versions of these replication-incompetent retroviral vectors have incorporated an IRES-reporter gene configuration which allows investigators to ectopically express different genes along with the reporter gene (Fig. 1C) (Cepko et al., 1998; Furukawa et al., 1997). In this way, researchers can study the effect of ectopic gene expression on retinal progenitor cell proliferation, cell fate specification, differentiation or survival in vivo using clonal analysis. Versions of these viral vectors that express Cre recombinase have also been used to conditionally inactivate floxed genes in single infected retinal progenitor cells and similarly analyze the effects on proliferation and development of the daughter cells (Zhang et al., 2004). Some investigators have used retroviruses to study the development of the retina in mice carrying targeted deletions of different genes (Dyer and Cepko, 2001a; Dyer et al., 2003). By combining retroviral-mediated ectopic expression with mouse strains carrying floxed or knockout alleles of genes believed to be important for retinal development, this system provides a great deal of flexibility in studying the genetic basis of retinal development.

In this article, we will discuss the statistical consideration for experiments that fall into two different classes. The first type of experiment utilizes different retroviral vectors in the left and right eyes of mice with the same genetic background. For example, one eye of a mouse carrying a floxed allele of a

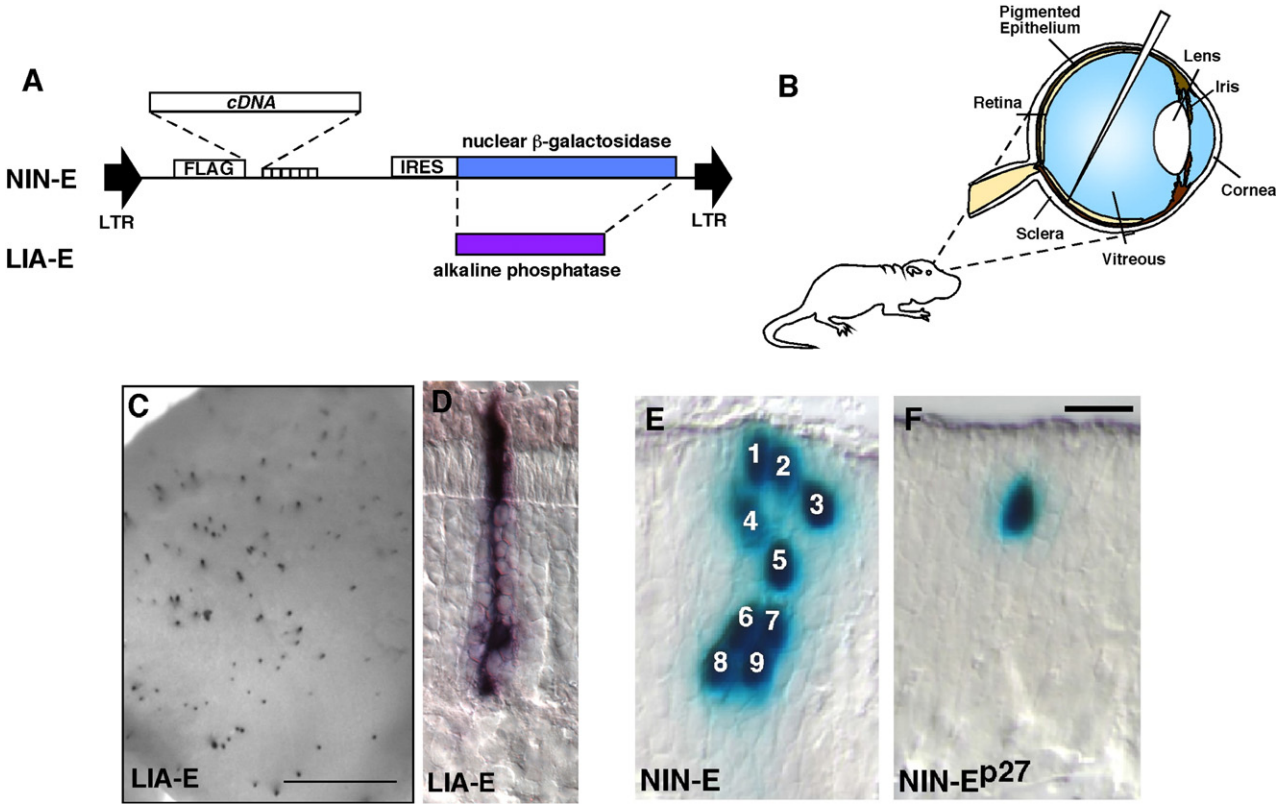

Fig. 1 – Retroviral lineage analysis in vivo and in explant culture. (A) A series of replication incompetent murine retroviral vectors were developed that are suited for expressing a reporter gene such as alkaline phosphatase or nuclear β-galactosidase along with a cDNA of interest. These vectors also encode a FLAG epitope tag and a 6×His tag for protein analysis. (B) For in vivo retroviral infection of the developing retina, P0 mouse pups receive a subretinal injection of 0.5 μl of viral stock with a titer of approximately $0.5–1.0 \times 10^7$ infectious particles per ml. 3 weeks later after retinal development is complete, the retinae are isolated and stained for alkaline phosphatase expression as whole tissue samples (C). Cryosectioning of these stained retinae reveals the neuronal morphology of individually infected retinal cells that were derived from single infected retinal progenitor cells. (D) A rod photoreceptor labeled with alkaline phosphatase is shown. (E, F) For analyzing retinal progenitor cell proliferation in retinal explant cultures, we prefer the NIN-E retrovirus that encodes nuclear β-galactosidase. Even large clones can be readily scored for the size of the clones. (F) An example is shown of a virus that ectopically expresses p27 which induces premature cell cycle exit and as a result smaller clones. Abbreviations: IRES, internal ribosome entry site. Scale bars: 1 mm in panel C and 10 μm in panel F.

gene may receive a Cre-expressing retrovirus in the left eye and a control virus lacking Cre recombinase in the right eye (Fig. 1D). The other type of experiment applies the same virus to both eyes in each mouse but the mice belong to different groups with different genetic backgrounds such as +/+, +/– and –/– for a particular targeted gene (Fig. 1E). These two types of experiments present a number of important challenges for the data analysis. For each type of experiment, these questions include how to summarize data and how to perform statistical comparisons of the number of clones and the number of cells per clone (i.e. clone size). We provide some statistical guidance on these critical questions.

Here, we analyze one data set (Table 1) in which the two eyes from each mouse received different retroviral vectors and a second data set (Table 2) in which the retinae from littermates with different genotypes received intraocular injections of the same retroviral vector. The data used in this paper are actual retroviral clonal data generated for an ongoing study on the role of cell cycle regulators in retinal development. These data sets were selected because they illustrate the importance of using the proper statistical methods for retroviral clonal analysis. We analyze both data sets with various statistical strategies to illustrate good and poor approaches.

1.1. *Statistical models for inference*

Statistical inference procedures are based on statistical models. In biology, model organisms are often used to help understand human disease. For instance, various mouse models of human cancers have been developed. Similarly, statistical models are used in data analysis to help determine whether observed trends are not attributable to chance alone. The applicability of conclusions drawn from preclinical models to human disease depends upon how well the biological model represents the human disease. Likewise, the validity of an inference suggested by a specific data set depends on the appropriateness of the statistical model and the analytic methods that are employed. Therefore, the result of data analysis (such as a *p*-value) should be interpreted in light of whether the underlying model used in the analysis is appropriate for the specific application at hand.

1.2. *Assumptions used in classical statistical methods*

One common element of the models underlying the classical methods (*t*-test, analysis of variance, rank-sum test, Kruskal–Wallis test) used to compare the means or medians of one variable across two or more groups is an assumption that all observations are statistically independent. Biologists commonly use the term "independent" to mean "non-deterministic." For example, a biologist may state that height and sex are independent variables in humans. Generally, men tend to be taller than women but some women are taller than some men. Hence, sex does not determine height and a biologist may use the term "independent" to indicate that the two variables are not deterministically related. However the statistical definition of

Table 1 – Data collected by infecting each eye with a different virus

Mouse ID	1		2		3		4		5		6		Total	
Treatment	Cont	Exp	Cont	Exp	Cont	Exp	Cont	Exp	Cont	Exp	Cont	Exp	Cont	Exp
No. cells														
1	9	9	11	7	6	6	6	1	5	1	6	8	43	32
2	3	5	7	2	1	3	5	1	4	7	3	4	23	22
3	8	5	4	4	3	4		1	1		4	3	20	17
4		2	2	5	1	7	5	3	4	2	3	1	15	20
5	1	2	8	1	2	1	7		1	2		1	19	7
6	1	5	2	4	3	2	1	1	1	1	2	1	10	14
7	1	1	4	2	1	1	4	2	3	2	1	1	14	9
8	2	1	3	2	2	2	3		2	1	2	2	14	8
9	1		2						4	1			8	1
10	1	1	1				2		1	1	1	1	6	3
11			1	1			1				1		3	1
12							1	1	1		1		3	1
13							3		2				5	0
14				1									0	1
15	1		1						1				3	0
16													0	0
17	1	1											1	1
18			1										1	0
19					1								1	0
20				1									0	1
Summaries														
No. clones	29	32	47	30	20	26	38	10	30	18	25	22	189	138
Mean size	4.28	3.88	4.70	4.90	4.55	3.54	5.55	5.00	5.97	4.44	4.52	3.32	4.97	4.10
Differences														
No. clones	3		−17		6		−28		−12		−3		−51	
Mean size	−0.40		0.20		−1.01		−0.55		−1.52		−1.20		−0.87	

Table 2 – Data collected in comparison of genotypic groups

Group	Wild-type			Knock-out				Total	
Mouse no.	1	2	3	1	2	3	4	Wild-type	Knock-out
No. cells									
1	8	14	34	77	54	17	54	56	202
2	6	15	26	54	37	18	30	47	139
3	8	8	13	36	23	9	23	29	91
4	6	2	21	30	17	7	12	29	66
5	6	8	17	34	23	10	26	31	93
6	2	6	10	22	18	2	9	18	51
7	3	5	11	19	19	5	8	19	51
8	3	7	8	15	19	7	11	18	52
9	1	2	10	12	8	4	6	13	30
10	1	2	10	14	3	6	5	13	28
11	4	1	3	2	2	2	0	8	6
12	2	4	5	6	1	3	2	11	12
13	1	0	3	3	2	2	1	4	8
14	0	0	5		3			5	3
15	2	3	4	4		2	6	9	12
16	1	0	1	2		5		2	7
17	0	0	1					1	0
18	0	1	0					1	0
19		1	1					2	0
20			1	1	1	4	1	1	7
21			1	1	1	4	1	1	7
22				1	1	4		0	6
23	1		1	1		4		2	5
24			1	1		4		1	5
25						4		0	4
26						1		0	1
27		1	1					2	0
28	1							1	0
29								0	0
30								0	0
41	1	1						2	0
51	1							1	0
70	1							1	0
Summaries									
No. clones	59	81	188	335	232	124	195	328	886
Ave. size	8.71	6.05	5.93	4.63	4.48	9.01	4.26	46.46	5.12

"independent" has a very precise probabilistic meaning and is not synonymous with non-deterministic. Two variables are statistically independent if knowledge of the value of one variable provides no knowledge regarding the distribution of the other variable in the population. Thus, height and sex in humans are not statistically independent variables because knowledge of sex gives some information regarding the distribution of height in the population. It is important to recognize that a statistical method assuming independence may yield misleading results if applied to a set of observations that do not meet the statistical definition of independence.

The classical methods to compare means or medians of one variable are different in other assumptions of their underlying models. The *t*-test and analysis of variance assume that data values are normally distributed within each group. Methods that assume a specific probability model for the distribution of data (such as the normal distribution) are called parametric methods. If the parametric model accurately represents the actual distribution of the data files, then the corresponding parametric method is usually the method of choice.

However, in many applications, the data are not accurately represented by the assumed distribution of parametric methods. In these cases, non-parametric methods such as the signed-rank, rank-sum and Kruskal–Wallis tests that operate on ranks of data values instead of the actual data values are usually preferred. These methods do not assume that the data values follow a specific parametric model. Instead, non-parametric methods assume that the distribution of data values within each group has the same basic shape. Thus, the non-parametric procedures are applicable in a wider variety of settings than parametric methods, but may have less statistical power when the parametric models are reasonable. We have not found a parametric model that fits the retroviral clonal data very well (Fig. 2). Therefore, we recommend the use of non-parametric methods in the analysis of retroviral lineage experiments.

1.3. *The importance of variability*

These experiments generate data at a number of levels. For example, one can consider the clone size, the number and size

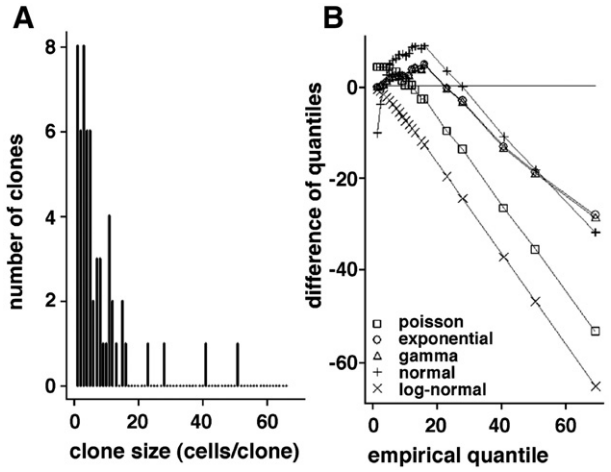

Fig. 2 – Poor fit of several commonly used statistical models. (A) Histogram shows the number of clones of each size for wild-type mouse number 1 in Table 2. (B) For several statistical models, the difference between the fitted model quantile and the empirical quantile of clone size is plotted against the empirical quantile. The lack-of-fit is shown because the differences do not fluctuate around y=0. The models were fit using the maximum likelihood method (Casella and Berger, 1990).

of clones per eye, the number and size of clones per mouse, and the number and size of clones per group. Data can vary considerably at each level. However, it is most import to consider the variability in measurements from mouse to mouse. Otherwise, the results from one unusual mouse may drive the results of the entire analysis. For example, one mouse may be especially prone to infections or receive a slightly greater multiplicity of infection (MOI) due to differences in viral titer or injection efficiency. An analysis that ignores such mouse-to-mouse variation may have its results determined entirely by such an unusual mouse. Thus, it is critical for statistical analysis of clonal data to include the appropriate number of animals to reduce the contribution of individual mice that may be outliers due to these types of variables. Additionally, it is important to include the appropriate number of animals in each group to determine whether any differences between groups could be attributed to chance given the observed degree of variability from mouse to mouse within the groups.

1.4. *Implications for the analysis of clonal data*

Clonal data produce data at a number of levels including the number of clones per eye and the clone sizes. There are many ways to summarize these data to make comparisons between the groups. The primary question for analysis is how to summarize and compare these data in such a way that the implicit underlying statistical models are reasonable for the specific experiment. Therefore, we now describe and comment on appropriate and inappropriate analysis approaches for each type of experiment.

For experiments that test the same vector in animals of different genotypes, one approach would be to average the clone sizes across all eyes belonging to one group and average the clone sizes across all eyes for the other group and simply note which group has a greater average. This analysis can be viewed as descriptive but not inferential. The analysis does not account for variation of the mice or have a clearly defined underlying model of chance to determine statistical significance. From a practical perspective, by ignoring the variation of mice, the result could be driven by one mouse with unusually large or small clones. Another approach would be to apply a *t*-test to the individual clone size observations. This approach is flawed because the underlying statistical model is based on the unrealistic assumption that all clone sizes are independent. The assumption is unrealistic because clones within the same animal are not statistically independent. For example, knowledge of the size of one clone in a mouse will likely provide some information about the distribution of the sizes of the other clones within the same mouse or eye. Observations from the same animal are statistically dependent because the environment of the clones with the same eye or animal is influenced by the same immune system, light exposure, vasculature, diet, etc... Additionally, subtle variations in the preparation and administration of the injection could possibly contribute to eye-to-eye variation in characteristics of clone size distributions. For example, if the reactive gliosis induced by the puncture in the retina to perform the subretinal injection is more extensive in one eye as compared to another eye this could influence the clonal data in a dependent manner for that eye. Also, the actual MOI delivered from the injection could vary from animal to animal. Thus, it is questionable to assume that observations from the same animal are statistically independent.

For experiments that inject two different vectors into the eyes of a set of animals with the same genotype, one could also envision averaging clone size across all eyes receiving one vector and comparing it to the average of clone size across all eyes receiving the other vector. This analysis cannot be considered inferential for reasons similar to those mentioned above (e.g., no underlying model and ignores mouse-to-mouse variation). Similarly, applying a *t*-test to the individual clone sizes is also invalid. Another approach would be to compute eye-level summaries such as the average size and then compare them with the two-sample *t*-test or rank-sum test. These tests assume that all observations are independent and thus the analysis would ignore the dependency of eye-level summaries computed from the same mouse.

The assumption of independence implicit in classical statistical methods is usually reasonable when each data value represents a distinct animal. Thus, analyses using classical methods should be applied to data sets with each animal represented by one data value. This can be accomplished by computing a meaningful summary data value for each animal. We describe a simple and useful way to perform such an analysis for each type of experiment below.

In experiments that use the same virus in both eyes of the animal, one could compute the average clone size for each animal and then use the rank-sum test to compare these values across genotypes. The lack-of-fit of the normal model to the

Table 3 – Minimum possible *p*-value for non-parametric methods

Apply one treatment to each eye in each mouse		Compare two genotypic groups of mice	
N	Minimum *p*-value	Per group N[a]	Minimum *p*-value
2	0.5000	2	0.3333
3	0.2500	3	0.1333
4	0.1250	4	0.0714
5	0.0625	5	0.0444
6	0.0313	6	0.0303
7	0.0156	7	0.0220
8	0.0078	8	0.0167
9	0.0039	9	0.0131
10	0.0020	10	0.0105

[a] Calculations assume that each group has the same sample size.

individual clone sizes (Fig. 2) makes us concerned that the normality assumption of the *t*-test and analysis of variance (ANOVA) is unreasonable. (Theoretically, the central limit theorem implies that the average clone size should be normally distributed if a very large number of clones are observed. However, in practice, it may be unclear how many clones must be observed for the central limit theorem to justify use of the *t*-test or ANOVA in the analysis.) Therefore, we recommend that the rank-based methods be used to compare the animal-specific average clone sizes across genotypes.

In experiments that inject a different virus into each eye, one can compute the average clone size for each eye and then use the between-eye difference as a summary measure for each animal. Then, one can use the signed-rank test to determine whether the mean or median of these between-eye differences is significantly different from zero. We recommend the use of the signed-rank test over the paired *t*-test due to previously mentioned concerns regarding the accuracy of the normal distribution as a model for the data.

2. Results

To illustrate the principles described above, we have collected two data sets that represent the two types of experiments shown in Fig. 1. First we consider various approaches to perform data analysis for infection of left and right eyes with a different retroviral vector (Table 1). The bottom of the table shows the total for the number of clones per eye for the control retrovirus and the experimental retrovirus (e.g. Cre-expressing virus). Examining those totals without considering mouse-level summaries might lead one to infer that the experimental retrovirus is associated with a reduction in the number of clones per eye and a reduction in the number of cells per clone. The total number of clones across eyes with the control virus is 37% greater than that among eyes with the experimental virus. When ignoring mouse-to-mouse variability, the average clone size differs by almost one cell between experimental and control virus. When applied to individual clone sizes, the *t*-test indicates that the average clone size differs significantly between the two viruses ($t=2.13$ with 319 degrees of freedom, $p=0.03$).

However, when carefully considering these variables within each individual mouse, this inference is not so readily apparent. In 2 of the 6 mice (animal numbers 1 and 3), the eye receiving the experimental retrovirus had a larger number of clones than the eye receiving the control retrovirus. In one mouse (animal number 2), the clones in the experimental virus infected eye had a larger number of cells per clone on average than did the clones in the contralateral eye with the control retrovirus. We computed the within-mouse difference between the control and experimental eyes in each mouse and applied the signed-rank test to those differences. Using this approach, we found did not find statistically significant evidence that the experimental retrovirus had any effect on the total number of clones per eye ($p=0.25$). Also, we found that the difference in average clone size between the two vectors was not as significant as previously indicated ($p=0.063$).

Next, we considered data from the other type of experiment in which mice with different genotypes received the same retrovirus in each eye (Table 2). The summary data indicate that there were 886 clones in the knockout mice and only 328 in wild-type mice. In our experience, most experiments that utilize littermates with different genotypes and retroviral injection show this type of asymmetric data distribution. It may reflect the Mendelian ratio of different genotypes or some other secondary consequence of differences in viability of one genotype over another. The average clone size was 5.12 for the knockout mice and 6.46 for the wild-type mice. One might be tempted to conclude that these differences are statistically significant. After all, 886 is almost 3 times greater than 328 and 6.46 is more than 1 cell larger than 5.12. When applied to the sizes of the individual clones, the *t*-test finds a very significant difference between the average clone size of knockout mice and that of wild-type mice ($t=3.7$ with 443 degrees of freedom, $p=0.002$).

However, examining the mouse-specific totals and averages again suggests that these results may not be significant. In particular, a knockout mouse (knockout number 3) had the largest average clone size (9.01 cells per clone) across the entire experiment, thus raising doubts about whether the wild-type animals really tend to have a larger clone size. Additionally, one of the wild-type animals had more clones

Table 4 – Sample size required to achieve 80% statistical power

Type of experiment	Comparison of two vectors	Comparison of two genotypes
Ratio of difference to standard deviation	Total no. of subjects[a]	No. subjects per genotype[a]
0.5	61	69
1.0	17	20
1.5	9	11
2.0	6	8
2.5	6	6
3.0	6	6

[a]Calculations are based on using the sign test for vector comparisons, the rank-sum test for genotype comparisons, and assume that mouse-level data are normally distributed.

than one of the knockout animals (wild-type 3 compared to knockout 3) making it less clear whether the knockout mice tend to have a larger number of clones per eye. To compare the clone number and clone size of the two groups, we applied the rank-sum test to the mouse-specific average clone sizes and number of clones. We do not find significant evidence of a difference between wild-type and knockout mice in terms of the median number of clones per eye ($p=0.12$) or the median clone size ($p=0.40$).

3. Discussion

Retroviral clonal experiments in the developing retina produces detailed data at a number of levels. It is imperative to consider whether the underlying assumptions of the methods used in a statistical analysis are reasonable for this application. In particular, it is essential to account for mouse-to-mouse variation and to not treat observations from the same mouse as statistically independent observations. We note that applying classical non-parametric procedures to mouse-level summary values satisfy both of these requirements. With this approach, it is the number of mice instead of the number of clones that determines whether the experiment has adequate statistical power to identify meaningful effects. We recommend the Rice Virtual Lab in Statistics (http://onlinestatbook.com/rvls.html) as an additional resource for further statistical education.

Knowledge of how the statistical analysis is performed can assist in designing experiments to more effectively use laboratory resources and minimize the number of animals required for each experiment. In particular, it is important to note that the sample size for the statistical analysis is the number of mice, not the number of eyes or the number of clones. Thus, it is important to include enough mice in the experiment so that it is possible to have a statistically significant result. As mentioned above, we recommend non-parametric procedures be used in the analysis of these experiments. Therefore, it is important to include enough animals in the experiment to ensure that the non-parametric procedure used will have adequate statistical power to detect the biologically relevant effects.

Non-parametric procedures have a mathematical lower boundary for the p-value that depends on the sample size (Table 3). Thus the experiment should include at least as many animals as required to make it mathematically possible to achieve the desired level of statistical significance. However, it is important to realize that having a sample size adequate to make it possible to achieve statistical significance does not necessarily mean that it is probable that statistical significance will be achieved. Experiments should be planned to have adequate power to detect practically meaningful effect sizes.

The smallest sample size that gives adequate statistical power depends on the ratio of the difference to be detected to the standard deviation of the mouse-level summary data in the population. In Table 4, we use an approximation based on the normal distribution to estimate the sample size required to have 80% probability to yield a p-value less than 0.05. It is clear that small sample sizes can detect only the largest effect sizes. For example, in Table 1, the standard deviation of the intra-mouse differences of average clone size is 0.6. Thus, to have an 80% probability to detect differences in average clone size of 0.6 cells, 1.2 cells, 1.8 cells, and 2.4 cells as statistically significant at the $p=0.05$ level requires roughly 61, 17, 9, and 6 mice respectively. In Table 2, the standard deviation of average clone size among wild-type mice is 2.3. Thus, to have an 80% probability to detect differences in average clone size of 2.3, 4.6, 6.9, and 9.2 cells as statistically significant at $p=0.05$ requires roughly 69, 20, 11, and 8 mice of each genotype (wild-type and knockout), respectively. Certainly, improvements in experimental techniques that reduce the standard deviation can potentially result in dramatic savings in terms of the number of mice that are required. The accuracy of these sample size calculations depends on how well the normal distribution represents the actual distribution of the mouse-level summary data. Nevertheless, we believe these estimates may be of practical utility until a data set with a very large number of mice is available and can be used to compute more accurate sample size estimates without assuming that the data are normally distributed.

Acknowledgments

This work was supported by grants (to M.A.D.) from the National Institutes of Health, Cancer Center Support from the National Cancer Institute, the American Cancer Society, Research to Prevent Blindness, the Pearle Vision Foundation, the International Retinal Research Foundation, and the American Lebanese Syrian Associated Charities (ALSAC). M. A.D. is a Pew Scholar.

REFERENCES

Casella, G., Berger, R., 1990. Statistical Inference. Brooks, Cole Publishing Company, Belmont.

Cepko, C.L., Fields-Berry, S., Ryder, E., Austin, C., Golden, J., 1998. Lineage analysis using retroviral vectors. Curr. Top. Dev. Biol. 36, 51–74.

Donovan, S.L., Dyer, M.A., 2006. Preparation and square wave electroporation of retinal explant cultures. Nature Protocols 1, 2710–2718.

Donovan, S.L., Schweers, B., Martins, R., Johnson, D., Dyer, M.A., 2006. Compensation by tumor suppressor genes during retinal development in mice and humans. BMC Biol. 4, 14.

Dyer, M.A., Cepko, C.L., 2000. p57(Kip2) regulates progenitor cell proliferation and amacrine interneuron development in the mouse retina. Development 127, 3593–3605.

Dyer, M.A., Cepko, C.L., 2001a. p27Kip1 and p57Kip2 regulate proliferation in distinct retinal progenitor cell populations. J. Neurosci. 21, 4259–4271.

Dyer, M.A., Cepko, C.L., 2001b. Regulating proliferation during retinal development. Nat. Rev., Neurosci. 2, 333–342.

Dyer, M.A., Livesey, F.J., Cepko, C.L., Oliver, G., 2003. Prox1 function controls progenitor cell proliferation and horizontal cell genesis in the mammalian retina. Nat. Genet. 34, 53–58.

Fields-Berry, S.C., Halliday, A.L., Cepko, C.L., 1992. A recombinant

retrovirus encoding alkaline phosphatase confirms clonal boundary assignment in lineage analysis of murine retina. Proc. Natl. Acad. Sci. U. S. A. 89, 693–697.

Furukawa, T., Morrow, E.M., Cepko, C.L., 1997. *Crx*, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. Cell 91, 531–541.

Turner, D.L., Cepko, C.L., 1987. A common progenitor for neurons and glia persists in rat retina late in development. Nature 328, 131–136.

Zhang, J., Gray, J., Wu, L., Leone, G., Rowan, S., Cepko, C.L., Zhu, X., Craft, C.M., Dyer, M.A., 2004. Rb regulates proliferation and rod photoreceptor development in the mouse retina. Nat. Genet. 36, 351–360.