Comment on "Response to PNAS paper by Kerby Shedden and Stephen Cooper" by Raymond Cho and David Lockhart.

By Stephen Cooper

A response has been posted on a web-site responding to the publication of the paper:

> Kerby Shedden and Stephen Cooper, "Analysis of cell-cycle-specific gene expression in human cells as determined by microarrays and double-thymidine block synchronization" ***Proc. Natl. Acad. Sci. USA***, (2002) 99:4379-4384,

which in turn was an analysis of a paper published in Nature Genetics in 2001:

> Cho, et al. Transcriptional Regulation and Function in the Human Cell Cycle" ***Nature Genetics***, 2001 27:48-54.

The response by Cho and Lockhart may be read at:

http://www.salk.edu/docs/labs/chipdata/response.html

but in order to make it easier for the reader to understand the comments presented here (hereafter referred to as "Comments"), the original Cho/Lockhart Response (hereafter referred to as "Response") is printed out in full below. Line numbers have been added to the Response for ease of reference.

Some General Thoughts

Before any detailed comments, a few general points should be noted. Although I am not a professional statistician, I have learned enough about statistics to understand that there may be many ways of looking at the same problem, and different approaches to various issues of statistical analysis. This means that reasonable and alternative approaches may be in opposition and come to different conclusions. It is one of the purposes of the Comments presented here to highlight where these differences arise, and so perhaps guide future studies of the cell cycle and microarray analyses.

One reason for the statistical disagreement is the relative paucity of data that we are dealing with. And by paucity of data I do not want to slight the thousands of genes that were actually analyzed by Cho, et al, and that many different chips have been used to analyze the time-course of gene expression in two independent experiments on human cells. This is an enormous amount of work and by any standards a great deal of data. What I refer to as the paucity of experimental data is related to the fact that we are discussing only two replicate experiments, and within those two experiments only points taken every two hours for 24 hours. If there were twenty experiments, and points taken every 10 minutes, it may be that statistical analysis would not even be necessary to see a particular result as being reproducible and reflecting the actual pattern of gene expression. It might be obvious that with additional experiments, and finer time points, a particular gene could show a reproducible pattern in, say, 19 out of 20 experiments, and no curve fitting or fancy statistical analysis would even be necessary. We could then agree on the results and conclusions. The results could be clear and obvious. The data could speak for itself. But we have to play with the cards dealt, and until there are additional and confirmatory experiments we can only analyze the data at hand.

Fundamental Biological Problems

It is important to note that the points raised in the Response are related entirely to the statistical analysis. Aside from the statistical problem, there is another problem that is, in fact, even more important (and which was not dealt with in detail by Shedden and Cooper). That problem is the biological problem. At the end of this discussion I will emphasize this biological problem, which is the proposal that the cells under discussion were not even synchronized. This point was actually raised, in an oblique manner, in our paper in the section on the "Definition of a Synchronized culture": which reads:

> **Definition of a synchronized culture.** The central experimental problem faced when analyzing data proposing cell-cycle-specific expression using synchronized cultures is the question: "Are the cells actually synchronized?" It is important to have criteria for analyzing cells as a function of cell age during the division cycle. An important criterion—and perhaps the single most important criterion—is that the cells should exhibit synchronized cell divisions. Without synchronized cell divisions it is difficult to know how to evaluate a synchrony experiment. No cell counts or frequency of mitosis data are provided along with the microarray data to allow one to know whether the mammalian cells studied using the microarray technology exhibited synchronized divisions (1). [note: numbers refer to referenced papers]
>
> It has been proposed that no synchronization method where all cells are treated identically can synchronize cells (3-5). This stricture applies to the double-thymidine block of Rao and Johnson (6) that was used to synchronize mammalian cells for cell cycle analysis (1). A truly synchronized culture is one that displays synchronized cell divisions and a synchronized passage through the cell cycle. We are unaware of reproducible evidence that cells subjected to this double-thymidine block are actually synchronized. No evidence related to synchronization is presented in the microarray paper (1). If the cells were not synchronized, then of course the results under examination would not be valid. In addition, we have pointed out that the lack of peak decay (Fig. 4) supports the proposal that the cells are not actually synchronized.
>
> *For the statistical analysis presented here, however, we have given the data the benefit of a doubt, and have assumed that the cells were actually synchronized* [emphasis added]. We have also assumed that the proposal (1) that the interdivision times were 12 h, and that over 24 h of analysis two complete cell cycles were studied. As it turns out, even if the cellular interdivision time were different (e.g., as short as 10 h, or as long as 14 h), our results (data not shown) would not be any different. The proposal that there is periodic cell-cycle-specific variation in the expression of a large number of genes should therefore be viewed with caution.

Reading through these three paragraphs we see the biological problem that must be dealt with.. The key biological problem is whether or not the cells studied were synchronized and are

actually models of the normal, unperturbed cell cycle. But in order to even begin to discuss the statistical analysis of the data we must necessarily, and somewhat reluctantly, give the synchronization procedure the benefit of a doubt. We do this although no data was actually presented by Cho, et al., that the cells are actually synchronized. There are, as will be seen below, both statistical (experimental) and theoretical reasons to suspect that the cells studied by Cho, et al., were not actually synchronized in the first place.

A corollary problem related to the question of whether or not the cells studied were actually synchronized is the problem of the introduction of artifacts, periodicities, and alterations to the normal cycle by the double-thymidine treatment. As will be seen below, the Response argues that when the data is looked at in a certain way, there are reproducible cell-cycle-patterns of gene expression. However, the point made here in this Comment is that even if reproducible patterns are found, there will still remain the question, possibly outside of the domain of statistical analysis, of whether or not the data does reflect the normal, artifact-free, pattern of gene expression during the division cycle.

While we are on the subject of biological problems, it should be noted that an additional problem is the use of "primary" cells. It does not appear as if the original cells analyzed were cloned and grown from a single isolated cell. This means that the cells studied could be derived from many different cell types. This could lead to a large variability in cell-cycle interdivision times. This variability would obviate and make superfluous any measurements of gene expression during the cell cycle.

Contrasting Belief Systems

Perhaps a more subtle but equally important reason for different conclusions (from the same data) are the different belief systems brought to bear on the problem. This is seen in line 19 of the Response (see below), where it is stated "In our published analysis we assumed the existence of cell cycle-regulated genes." This belief system is articulated even more strongly in the introductory part of the original Nature-Genetics paper where it is stated:

> Proper regulation of gene activity during the cell cycle is likely to govern critical aspects of diverse processes……..The pleiotropic effects of [these mutations] suggest that a considerable proportion of all genes experience differential regulation during the cell cycle.

The origin of the Shedden/Cooper PNAS critique stems from (but is not based on) a different assumption, that there are very few proteins whose expression varies during the cell cycle. More strongly, this alternative view (currently, and admittedly, a minority viewpoint) proposes that the control of the passage of a cell through the cell cycle is based on the continuous and phase-independent accumulation of various control molecules. At a more experimental level it is proposed that there are no genes expressed specifically in the G1-phase of the cell cycle. While there may be genes expressed, or proteins present, in a cell-cycle specific manner, these genes are related to the S/G2/M phases of the cell cycle. To put this idea in a more personal way, when I read the original paper with the postulation of numerous G1-phase specific gene expression patterns I realized that I had to test this proposal. That Shedden and I were able to find the problems we did find indicates the strength of the alternative viewpoint.

A complete presentation of these ideas is available from a number of papers that can be read at the web-site

www.umich.edu/~cooper

where clicking on the appropriate papers will enable the reader to understand the arguments and experiments supporting this alternative view of the cell cycle.

I will now comment on various points made in the Response, referring the line numbers in each case.

Statistical Analysis in the Response

The key critique of the PNAS paper is contained in the comment (lines 119-126)

> So how is it that Shedden and Cooper conclude that the observed periodic patterns are the product of random chance? The answer is that they failed to treat our data sets as independent replicates of the same measurement, *instead treating the two experiments as if they were completely unrelated*. In other words, instead of asking the appropriate question, "Is the occurrence of periodic patterns for the same gene in both data sets observed to be beyond that expected by random chance?", they asked "Is the occurrence of a periodic pattern for a gene in *one* data set likely beyond random chance? **Twice!**"

Let me simplify, if I may, this critique in terms of a numerical example. Imagine that one performed one experiment on 6000 genes and measured, over time, gene expression using microarrays. After accumulating the data, the gene expression patterns were tested for goodness of fit to some pattern such as a sine wave. Let us say some threshold was used such as 0.5 (PVE; proportion of variance explained by a sine wave), and with this threshold 1000 genes were found to exhibit cyclical expression. Now a replicate experiment is performed. Again the gene expression patterns are determined and again 1000 genes are above the threshold. Now a comparison of which genes are above the threshold is made. Let us say that 500 of the genes are found to be above the threshold in both experiments.

Using the assumption that there are expected to be numerous genes that are expressed in a cell-cycle specific manner, one can feel good about this result and propose that most or even all of the 500 genes are expressed in a cell-cycle-specific manner. One might suspect a few false positives which could reduce the number, but then there are the false-negatives which did not, by accident repeat, and this may balance out the numbers.

But let us imagine a third experiment being performed. There are two possibilities. A result fitting with the idea of numerous cell-cycle-specific gene expression patterns would be that perhaps 490 of the original 500 repeat. But if the result was that 250 or perhaps only 100 of the original 500 repeated, it can be seen that this result would suggest that it is only because of the large numbers of genes studied does one get a number of cycle-specific gene expression patterns.

This example shows why the limitation of the number of experiments is so problematic. In another paper I have dealt with the problem of how large microarrays may lead to artifacticious results (to appear in Trends in Genetics). This paper can be seen at

http://www-personal.umich.edu/~cooper/ESP.pdf

So the question raised here is whether, among the genes that are proposed to have reproducible patterns of cyclic gene expression, these genes are really expressed in a cell-cycle specific manner (as proposed in the Response). Or, as we have proposed, there is a strong possibility that

statistical variation and experimental noise can give rise to the apparent patterns of cycle-specific expression.

At this time there may be no absolute and final answer. It is possible that one should look at the two experiments together in the manner proposed in the Response. But an alternative view is that one should take each experiment and test that experiment to see whether it was a satisfactory experiment. If the experiment fails the test, combining that experiment with another one that also failed the test does not seem, to me, to be appropriate. The analysis in Figure 1 of the PNAS paper indicates that each experiment by itself has problems. That is why I feel that combining experiments is not a valid approach to understanding gene expression during the cell cycle.

I will now move on to various other points made in the Response.

Some Minor Points

A comment is required about the phrase "necessarily imperfect synchronization." (Line 86). This phrase implies, or has embedded in the language, the idea that the cells are actually synchronized. Separate from the statistical issues raised in the PNAS paper, or in the Response, I take issue with the assumption that the cells are synchronized. It has been proposed that no batch treatment can ever synchronize a cell (see my paper at

http://www-personal.umich.edu/~cooper/cp/index.html

for a discussion of this point), . Just because cells may be arrested using a particular inhibition protocol does not mean that the cells are synchronized. No evidence of synchrony is presented in the Nature-Genetics paper. A complete analysis of the inherent problems with double-thymidine blocks are presented in a paper currently in preparation, and this analysis is presented in an abbreviated form at the end of this discussion. (See Tables 1 and 2 below and text related to these tables.)

Similarly, I take issue with the assumption inherent in the phrase "greater ease of cell synchronization" (line 111) which implies that the yeast cells studied by the Stanford group (in Spellman, et al, 1998) were actually synchronized. As discussed below, and in a paper to appear in Nucleic Acids Research, it is not clear that the yeast cells were models of the cell cycle.

Peak Decay

An important point of the Response relates to the problem of peak decay. First of all, to clarify the original proposal made in the PNAS paper, the decay analysis is unrelated to any "requirement for periodicity." (see lines 162-175). It is not correct to associate the peak decay problem with any analysis of the periodicity of genes; that is, whether genes exhibit periodic expression behavior. The peak decay analysis has only to do with the randomization expectations as a synchronized culture passes through many division cycles. It is generally observed, and theoretically expected, that synchrony will decay with time. One may need only to refer to the second law of thermodynamics to understand this expectation. Regarding the amplitude reduction expected, there may be a valid point made in the Response. It is true that it is not clear how to express the expected decay in numerical terms. But what is expected is that if we had good measurements of all genes, that all genes would show decay. Having a second peak with a higher value can only be due to artifacts in the measurement, statistical error, or as noted in the Response, some suppression of expression in the first cycle. But we find that genes may exhibit an increase in the second peak as likely as a decrease; this is what is unexpected. The numerical expectation is related to the "width" of the differences between peaks, and we make

no mention of this. We only point out that the frequency with which peaks are increased in the second cycle is just as likely as peaks decreasing in the second cycle.

While on this subject of peak decay, a further word must be given to the proposal, made in lines 172-173 that

> … there could very well be some suppression of the peak amplitude of the first cell cycle following arrest.

If this proposal were true, it would surely place the entire analysis of human cells using double-thymidine block synchronization in jeopardy. If this suppression were taking place, it would mean that the cells were not normal and no further analysis should be pursued.

In the original Response a calculation of peak decay figures for 33 genes that Cho and Lockhart propose as giving repeat cyclicity in two experiments is presented. This is Figure C (one must go to the original Response and click on the link to Figure C there in order to see it. It is a two-peaked histogram, with admittedly approximately 2/3 of the genes (66 values from 33 genes in two experiments) showing positive values (i.e., peak decay). The original figure can be seen at…

http://www.salk.edu/docs/labs/chipdata/figC.html

If one looks at their Figure C different people will see different things. It is like the half-glass of wine—to some it is half full, to others, half empty. The fact that among 33 genes proposed by Cho and Lockhart to be a good selection of genes reproducibly expressed during the cell cycle there are 1/3 which have second peaks higher than the first can be taken as evidence weakening the entire group of 33 genes. Thus, one could argue that if 1/3 of the data results do not give the expected pattern, it is possible that 22 of the remaining 44 genes with a peak decay observed could be the result of statistical variation.

Again, as noted above, it is hard to know what to conclude at this time, because more data would certainly help. As I see it, the burden of proof of cell-cycle-specific gene expression is on those who propose it, and what the PNAS paper, and this Commentary are saying is that if one had alternative ideas the original data is certainly problematic enough to warrant skepticism regarding patterns of gene expression during the cell cycle.

Prior Biological Results

One of the supporting assumptions of the Nature-Genetics paper is that others have found cyclic variations in other systems and this supports the proposal made here. For example, in the response it is stated “Pattern searching methods that employ the Fourier transform have been tested previously for finding cell cycle-regulated transcripts in yeast by Pat Brown and his group at Stanford.” (lines 105-107) We are well aware of the yeast data (Spellman, et all, 1998) and have analyzed this data in another paper (see web site for manuscript form of this paper). Simply put, as with the double-thymidine block method used on the human cells, where I assert and will explain below why, the cells are not actually synchronized. Similarly, I do not accept three out of the four methods used by the yeast group to synchronize their cells (two temperature arrests, and alpha-factor inhibition). Besides the fact that any periodicities produced may be artifacts of the synchronization method, there are theoretical reasons to believe that the cells are not synchronized in that they are not representative during outgrowth of successive cells of different ages during the normal cell cycle.

To generalize the point made here, although there are numerous papers proposing cycle-specific gene expression, many of these papers are subject to the same biological critique made of the Nature-Genetics work, that the cells studied were not actually synchronized, and that there are periodicities introduced by the methods.

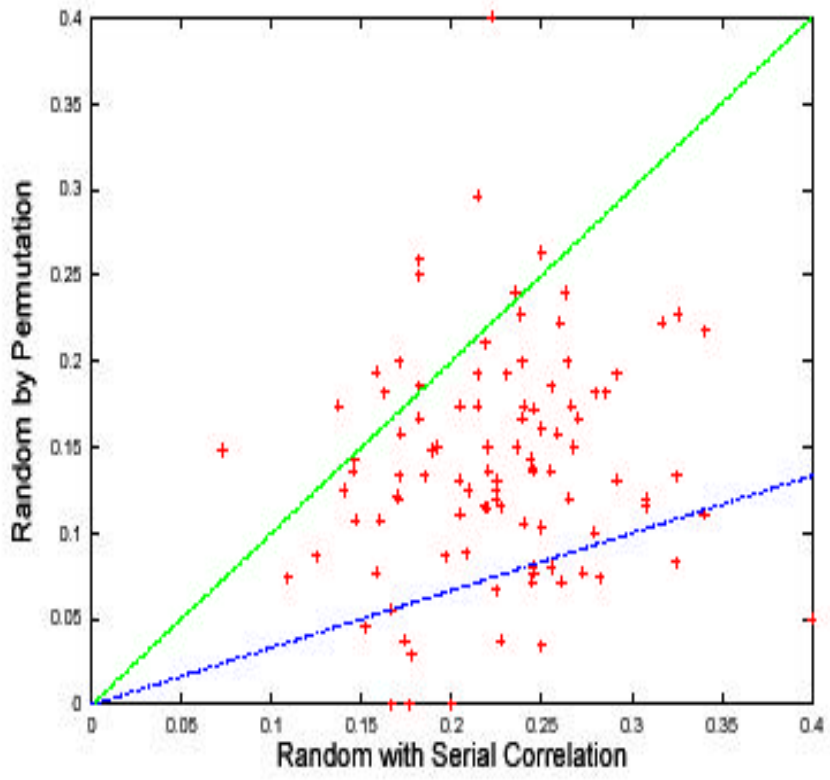
Detailed Statistical Analysis

I now turn to the key statistical arguments made in lines 47-61. The main statistical argument in the Response is that if you soften up the criterion for cyclicity, you get a strong excess of cyclic genes in real data compared to randomized data.

In the Response, the following three-stage analysis was performed. The Authors:

(i) Identified pairs of genes that are highly correlated between N2 and N3.
(ii) Visually assessed cyclicity in these genes.
(iii) Permuted the data and repeat d(i) and (ii).

They found 32/156 (cyclic/reproducible) in the actual data, and 2/27 (cyclic/reproducible) in the permuted data. The important thing is the ratio of these ratios, which is approximately 3:1. The response to this is that there are many different types of random variation. Permutation is a simple way to randomize the data that anybody can understand. But it is not accurate in all respects. We were aware of this from the outset, but the inaccuracies of permutation are conservative from our point of view, and since we wanted to write a paper for a general audience, we chose to ignore them. We believe that the 3:1 excess found in the rebuttal can be explained by the failure of the permuted data to capture the internal dynamics of real microarray data.



The reason that permutation fails to be realistic is that even if there were no cyclicity, there will still be serial correlations between different time points in a single gene. If the cell cultures were sampled very densely (say every 10 minutes instead of every two hours), then these serial correlations would all be positive (if some chance occurrence causes CDK7 to be up at 2:10 then it is likely to still be somewhat high at 2:20). When you sample at 2 hour intervals, however, the situation is more complex. There is ample evidence that the data are serially correlated, but the correlation is generally negative rather than positive. This is something like the "regression to the mean" metaphor.

To generate more realistic random (non-cyclic) data, two artificial data sets were generated from a Gaussian distribution that has the same mean and covariance as experiment N2. Then genes were selected that had correlation greater than .7 between the two artificial experiments.

Among these, we determined the number of genes such that the larger of the two Fourier PVE's is at least .5 (which is a low enough threshold to satisfy pretty much any of the genes that were posted to the web). Next the two Gaussian data sets were permuted, and repeated the correlation/Fourier step. This gives two proportions: the proportion of reproducible genes that are cyclic in Gaussian data, and the proportion of reproducible genes that are cyclic in permuted data. These are to be compared to the 32/156 and 2/27 proportions cited above. The figure above (previous page) shows a scatterplot of the two proportions over 100 replications.

You'll see that 89/100 times, the Gaussian data produces more cyclic genes than the permuted data (points below the green (upper) line), and that 20/100 times the Gaussian data produces at least 3 times more cyclic genes that the permuted data (points below the blue (lower) line).

This seems to suggest that the 3:1 excess found in the rebuttal is explainable by the differences between the two models for randomness. One cannot argue with confidence that the apparent cyclicities are not due to the intrinsic serial dynamics that exist in a culture of cells prepared according to their protocol.

A less important issue is the ratio 156/27, which just addresses the issue of reproducibility. This excess can also be explained by serial correlation.

The Problem of Peak Location

Another important point was made in the PNAS paper was not addressed at all by the Response, In the PNAS paper it was shown that even for those genes with high cyclicity there was an extremely poor, in fact negligible, correlation between the phase location of the peaks in the two experiments. The reproducibility of results in two experiments should cover not only the cyclicity values, but also the phase locations. The fact that the phases did not reproduce is evidence that one should not combine the two independent experiments.

Why Double-Thymidine Block does not Synchronize

I will now return to the analysis alluded to above, and present an explanation why it is not clear that the cells used in the Nature-Genetics analysis were actually synchronized.

The pattern of mass increase, DNA replication, and cell division for an exponentially growing, unperturbed, steady-state culture is presented in Table 1. The purpose of this table is to orient the reader for the analyses presented in Table 2. Each of the initial cells (top row is time zero) should be followed *down* each column in order to follow the history of each initial cell. The top line in Table 1 represents cells at various stages during the division cycle. By following each cell down each column one can see what happens to each individual cell after time zero. When a cell

divides, the mother-cell mass is apportioned equally to the two daughter cells. Cells divide at mass 2.00 but the daughter cells are not illustrated immediately after this division but 0.07 (one-fourteenth) of a cell cycle later. At this time, during steady state growth, the cells are of size 1.05.

During steady state growth (Table 1) across any row, the DNA pattern as well as the cell mass pattern, is invariant. There are always five cells with a G1-phase amount of DNA, four in S phase, four in G2 phase, and one M phase cell. Mass increases within each row from 1.05 to 2.00, although the columnar location of the largest and smallest cells varies during the period of steady-state growth.

Table 1. Steady-state growth

G1	1.05	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00
G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**
G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10
G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16
G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22
S	1.35	S	1.41	S	1.49	S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28
S	1.41	S	1.49	S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35
S	1.49	S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41
S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49
G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56
G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	S	1.64
G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	S	1.64	G2	1.72
G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	S	1.64	G2	1.72	G2	1.81
M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	S	1.64	G2	1.72	G2	1.81	G2	1.90
G1	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28	S	1.35	S	1.41	S	1.49	S	1.56	S	1.64	G2	1.72	G2	1.81	G2	1.90	G2	2.00
L																										H	

Table 1. Steady State Growth. Each vertical column illustrates the DNA condition and cell mass for cells growing exponentially without inhibition. The cell at the upper left just started G1 phase a short time earlier. It moves through G1-, S-, and G2-phases (moving down the column) to divide at mitosis (M-phase) to produce a cell with a G1-phase amount of DNA (lower left). The first row across the top of the table shows the DNA contents and cell masses for cells throughout the cell cycle at the start of the analysis. These cells have the same DNA contents and cell masses presented horizontally as in the first column presented vertically. The cells across first row each represent a cell during one-fourteenth of the cell cycle. At the start of the analysis (first row across top of table) the first five cells are in G1 phase, the next four cells are in S phase, four cells follow in G2 phase, and one cell represents the final cell at division, in M phase. The numerical values accompanying the DNA descriptions indicate the cytoplasmic mass of cells with the mass increasing exponentially during the cell cycle. A newborn cell has a mass of 1.0, a dividing cell has a mass of 2.00, and a cell of age 0.07 (one-fourteenth of the cell cycle time) has a mass of 1.05. Each of these initial cells (top row) can now be followed down each column during a period of time equivalent to one interdivision time. Each of the cells increases in mass, the cells pass through the DNA phases of the cell cycle, and at M phase the cells divide. The first cell (upper left) goes through G1 phase, enters S phase at size 1.35, enters G2 phase at size 1.64, and divides at size 2.00 to produce a cell with a G1 phase amount of DNA and size 1.05 after passing through a short period (0.07) of the next cell cycle. The bold-face, underlined values indicate that a newborn cell has been produced. The mass in these cells is halved as two daughter cells are produced at mitosis. The cell masses are given on a “per cell” basis. Thus, in the first column, the cell in M phase divides at size 2.00 producing 2 cells of size 1.00; after 0.07 of a cell division cycle the two cells each have a size of 1.05. Below the first column is an “L” and below the last column is an “H”. These letters indicate in this and subsequent tables, the highest and lowest masses present in the cells after a period equivalent to one full interdivision time. At the end of one interdivision time, and during the entire period of growth, the cell masses spread over a factor of 2.00, just as in the original culture. The smallest cell is always 1.05 (note, this cell is age 0.07) and the largest (age 1.00) is mass 2.00.

Table 2 illustrates inhibition of DNA replication as a synchronizing method. During inhibition of DNA synthesis mass increases allowing cells to pass through the G1-, G2-, and M-phases of the cell cycle to produce cells that are arrested either at the start of S phase (italicized S in Table 2), or in mid-S-phase for those cells caught in S phase at the moment of inhibition of DNA replication.

By following each of the initial cells (top row) down each column one can see that after a period of inhibition a collection of cells is produced with varied sizes. The spread in cell sizes is no narrower than that in the original culture (from 1.64 to 3.12, compared to 1.05 to 2.00). Thus, as with the analysis in Table 2, no synchronization has taken place.

Table 2. Inhibition of DNA synthesis

G1	1.05	G1	1.10	G1	1.16	G1	1.22	G1	1.28	*S*	*1.35*	S	1.41	S	1.49	S	1.56	G2	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00
G1	1.10	G1	1.16	G1	1.22	G1	1.28	*S*	*1.35*	*S*	*1.41*	S	1.49	S	1.56	S	1.64	G2	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**
G1	1.16	G1	1.22	G1	1.28	*S*	*1.35*	*S*	*1.41*	*S*	*1.49*	S	1.56	S	1.64	S	1.72	G2	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10
G1	1.22	G1	1.28	*S*	*1.35*	*S*	*1.41*	*S*	*1.49*	*S*	*1.56*	S	1.64	S	1.72	S	1.81	G2	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16
G1	1.28	*S*	*1.35*	*S*	*1.41*	*S*	*1.49*	*S*	*1.56*	*S*	*1.64*	S	1.72	S	1.81	S	1.90	M	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22
S	*1.35*	*S*	*1.41*	*S*	*1.49*	*S*	*1.56*	*S*	*1.64*	*S*	*1.72*	S	1.81	S	1.90	S	2.00	**G1**	**1.05**	G1	1.10	G1	1.16	G1	1.22	G1	1.28
S	*1.41*	*S*	*1.49*	*S*	*1.56*	*S*	*1.64*	*S*	*1.72*	*S*	*1.81*	S	1.90	S	2.00	S	2.10	G1	1.10	G1	1.16	G1	1.22	G1	1.28	*S*	1.35
S	*1.49*	*S*	*1.56*	*S*	*1.64*	*S*	*1.72*	*S*	*1.81*	*S*	*1.90*	S	2.00	S	2.10	S	2.21	G1	1.16	G1	1.22	G1	1.28	*S*	1.35	*S*	1.41
S	*1.56*	*S*	*1.64*	*S*	*1.72*	*S*	*1.81*	*S*	*1.90*	*S*	*2.00*	S	2.10	S	2.21	S	2.32	G1	1.22	G1	1.28	*S*	1.35	*S*	*1.41*	*S*	1.49
S	*1.64*	*S*	*1.72*	*S*	*1.81*	*S*	*1.90*	*S*	*2.00*	*S*	*2.10*	S	2.21	S	2.32	S	2.44	G1	1.28	*S*	1.35	*S*	*1.41*	*S*	*1.49*	*S*	1.56
S	*1.72*	*S*	*1.81*	*S*	*1.90*	*S*	*2.00*	*S*	*2.10*	*S*	*2.21*	S	2.32	S	2.44	S	2.56	***S***	1.35	*S*	*1.41*	*S*	*1.49*	*S*	*1.56*	*S*	1.64
S	*1.81*	*S*	*1.90*	*S*	*2.00*	*S*	*2.10*	*S*	*2.21*	*S*	*2.32*	S	2.44	S	2.56	S	2.69	*S*	*1.41*	*S*	*1.49*	*S*	*1.56*	*S*	*1.64*	*S*	1.72
S	*1.90*	*S*	*2.00*	*S*	*2.10*	*S*	*2.21*	*S*	*2.32*	*S*	*2.44*	S	2.56	S	2.69	S	2.83	*S*	*1.49*	*S*	*1.56*	*S*	*1.64*	*S*	*1.72*	*S*	1.81
S	*2.00*	*S*	*2.10*	*S*	*2.21*	*S*	*2.32*	*S*	*2.44*	*S*	*2.56*	S	2.69	S	2.83	S	2.97	*S*	*1.56*	*S*	*1.64*	*S*	*1.72*	*S*	*1.81*	*S*	1.90
S	*2.10*	*S*	*2.21*	*S*	*2.32*	*S*	*2.44*	*S*	*2.56*	*S*	*2.69*	S	2.83	S	2.97	S	3.12	*S*	*1.64*	*S*	*1.72*	*S*	*1.81*	*S*	*1.90*	*S*	2.00
																H		*L*									

Table 2. Consequences of inhibition of DNA synthesis. The cells in a growing culture (top row) are subjected to inhibition of DNA replication for a period equivalent to one interdivision time. In this case we assume some condition is imposed that specifically inhibits DNA synthesis from occurring. As in Tables 1 and 2, the top row shows the starting cells with a given pattern of DNA contents and cell mass increasing exponentially during the division cycle. Cells in S phase remain in S phase although mass increases. These cells do not divide, as they never complete S phase. Cells in G1 phase increase mass and then presumably "initiate" S phase although no DNA replication occurs as DNA synthesis, by definition of the experiment, is inhibited. In this case the "initiation" of S phase without subsequent DNA synthesis means that the cells have achieved the ability to initiate DNA replication. Cells at the start of S phase (in this nascent synthetic state) are indicated by italics. Cells in the G2 phase complete the G2 phase, perform mitosis, divide, pass through G1 phase, and then arrive at the start of S phase (again indicated by italics). At the end of the period of DNA inhibition all cells are shown "in S phase" or at the "start of S phase", but these cells all have different cell masses, ranging from 1.64 to 3.12. (The "H" and "L" indicate the smallest and largest cells in the culture.) Upon release from inhibition of DNA replication, there may be a rapid synthesis of DNA, but subsequent initiations of DNA replication would be spread out in time (asynchronous) because of the variation in cell masses at the time that DNA replication permitted to occur.

Following this Commentary is the response of Cho and Lockhart. There may be other points we have not touched on, but one final point should be made. Merely showing graphs that have this or that pattern that looks cyclic is not enough to convince a reader holding a different set of assumptions about the control of the cell cycle that there are cell-cycle specific patterns of gene expression, particularly those in the G1-phase. Also, merely showing that some particular pattern has a second peak lower than the first does not eliminate the basic statistical critique made in the PNAS paper. This type of evidence is anecdotal, in that the other side of the argument could print graphs showing the opposite. As noted at the outset, more data is needed, but also there must be an approach to the data that is not based on an assumption that basically assumes what is to be proven. Assuming at the outset that there are numerous genes expressed in a cell-cycle manner certainly is not the way to objectively approach the question. I hope that further work on this subject will be based on full consideration of both approaches to the problem of gene expression during the cell cycle.

Response of Lockhart and Cho

This addendum has been created in response to an article recently published in PNAS by Kerby Shedden and Stephen Cooper, "Analysis of cell-cycle-specific gene expression in human cells as determined by microarrays and double-thymidine block synchronization" ***Proc. Natl. Acad. Sci. USA***, Vol. 99, Issue 7, 4379-4384, April 2, 2002, regarding our earlier article in Nature Genetics, "Transcriptional Regulation and Function in the Human Cell Cycle" ***Nature Genetics***, 2001 Jan; 27(1): 48-54.

We welcome scientific criticism and careful reanalysis of our work, and agree that different interpretations of data are possible. However, we disagree with many of the methods and conclusions presented in the paper by Shedden and Cooper, and we consider it appropriate and helpful to provide a detailed response in the interests of clarity and scientific progress.

Our response is divided into two sections. The first section describes our own additional analyses of the original data that clearly demonstrate the existence of periodic, cell cycle-regulated genes. The second section explains why we reached conclusions in our original manuscript that are so fundamentally different than those reached by Shedden and Cooper based on their reanalysis of our published data.

Raymond J. Cho and David J. Lockhart

Section I: Statistical Verification of Cell Cycle-Regulated Genes

In our published analysis[1], we assumed the existence of cell cycle-regulated genes. We accepted as "periodic" genes with observed fluctuations in mRNA levels that, when replicate data were considered, had a clear pattern of cell cycle stage dependent regulation. Shedden and Cooper assert that using only purely objective statistical analysis, our data fail to establish the existence of periodic, cell cycle-regulated genes.

We would like to stress that the use of expected patterns as guides to find other, similarly behaving genes is accepted practice in current expression analysis[2-5]. However, we also demonstrate here that a statistical analysis of our data even WITHOUT such assumptions supports the existence of many genes with mRNA levels that fluctuate in a cell cycle-dependent manner. As even further proof, we also followed up on several of the genes identified from the array-based measurements by performing quantitative RT-PCR measurements on a third set of totally independent samples (see Chip vs. RT-PCR Data)

1. Is there clear evidence of non-random gene activity in our data?

The simplest way to ask this question is, *what is the likelihood that the observed degree of reproducibility between replicate data was the product of chance?* To answer this question, we first randomized our two data sets in the manner described by Shedden and Cooper. Next, we calculated correlation coefficients between the replicate data sets for

every gene in the real and randomized data sets. We then asked, for a given correlation coefficient threshold, how many genes in the real and randomized data sets display at least that threshold level of consistency. For example, there are 41 genes that yield a correlation coefficient of 0.80 or higher in the real data, compared to just 4 in the randomized data. Finally, we calculated the probability (*P*-value), using the binomial distribution function, that the greater number of consistent genes in the real data is statistically significant. The probability that the difference between real and randomized data sets is the product of random chance is extremely low (Table A). We also show in Table A that this difference holds true over a range of correlation coefficient thresholds.

2. Is there non-random, periodic gene activity in our expession data?

The simplest way to ask this question is: *Are consistent, periodic patterns observed in the two replicate data sets with a frequency above that expected by random chance?* To answer this question, we counted the genes, in the original data and in a pair of randomized data sets, that (1) displayed an arbitrary, high correlation coefficient between replicate data sets (greater than 0.70), AND (2) displayed, by eye, a clearly periodic pattern. The results of this analysis are shown in Figures A and B of this response. There were 156 genes with a correlation coefficient higher than 0.70 in the real, unrandomized data, compared to only 27 in the randomized data. Of these, 32 showed a clearly periodic pattern in the real data compared to 2 in the randomized data. The probability that this difference in the number of selected genes has occurred by chance is very small, estimated by the binomial distribution to be 2 x 10-27. We also examined the data for many genes with correlation coefficients below the 0.70 threshold in our real data sets and found a similar rate of occurrence of periodically oscillating genes. We find that with a correlation coefficient threshold of 0.40, we can conservatively identify hundreds of genes with clearly periodic, cell cycle-phase-dependent mRNA levels.

3. Are the patterns we observe convincingly periodic?

The simplest way to address this question is to examine the actual data (Figure 3 of our original paper and Figure B of this letter) as obvious proof of cell cycle periodicity (see also the RT-PCR results for selected genes). This also helps to shed some light on Shedden and Cooper's methods and their requirements for assessing periodicity: Their methods do NOT identify these obvious periodic patterns as periodic (see Figure C of this letter and corresponding discussion in Section II), indicating, perhaps, that there is something wrong with their underlying analysis or that the requirements for periodicity are unreasonably stringent.

Section II: A Detailed Analysis of the Manuscript by Shedden and Cooper

Our original analysis combined with the reanalysis presented here invites the question of how it is that Shedden and Cooper reached such fundamentally different conclusions.

A reading of Shedden and Cooper's manuscript reveals three pervasive flaws. First, the authors employed arbitrarily defined tests for cell cycle-dependent periodicity: peak

decay and conformance to a sine wave. **These requirements are either patently unreasonable or have been previously invalidated**.

Second, despite the impression given in the text, Shedden and Cooper do not actually use these tests to assess the periodicity of specific genes. Instead, they use these tests to ask the rather different, and less relevant question of whether periodicity **in one of our replicates is predictive of periodicity in the other**. In other words, they use these tests to evaluate perfect reproducibility rather than as a reasonable test of cell cycle periodicity. As we detail below, it is not reasonable to expect that all periodic behavior be perfectly consistent between the two independent data sets for all genes. These are technically demanding and difficult experiments, and noise is expected (especially due to variations in cell handling and necessarily imperfect synchronization). This is exactly why we performed replicates: to determine which observations ARE consistent. If the data were perfect, and there was no noise or other sources of variation, replicates would not be as important. But since there is random noise and other experimental sources of variation, we performed independent replicates and only interpret as "real" those patterns that are, in fact, consistently periodic. The conclusions in our paper require only that: 1) certain patterns are consistent across the two complete cycles and between the two data sets, and 2) that genes display periodic behavior that is unlikely to have occurred by chance.

Finally, when asking whether the patterns found in our real data can also be found in randomized data, Shedden and Cooper **treat our replicate data sets as if they are completely unrelated to one another**. This approach unreasonably ignores the precise reason we performed replicate experiments – to ensure that it is possible to discern consistent, oscillatory patterns above what might be expected by chance.

Specific Assertions in the Shedden and Cooper Paper

(1) Assertion: (Shedden and Cooper, Figure 1) Compared to a randomized data set, there is a negligible excess of periodically behaving genes in our data set.

(1) Comments:

1. *Strict, perfect sine wave periodicity is an unreasonable description of the behavior of cell cycle-regulated genes.* Pattern searching methods that employ the Fourier transform have been tested previously for finding cell cycle-regulated transcripts in yeast by Pat Brown and his group at Stanford. They found that this "pure" analytical approach was very poor at finding oscillatory patterns in real experimental data, which led them to supplement this approach with numerous additional analysis methods[4]. And this was for cell cycle data obtained with yeast, which tend to be much less noisy due to the greater simplicity of the genome and the greater ease of cell synchonization and handling for yeast cells compared to human cells. Furthermore, there is no reason to expect that cell cycle dependent mRNA levels would be perfectly sinusoidal.

2. *Our replicate data sets were unreasonably analyzed as unrelated experiments*. In our analysis above, we clearly demonstrate that relative to two randomized data sets, we find

a large excess of consistent, periodic patterns in our two experiments. Furthermore, we demonstrate that consistent, periodic patterns occur more than 10-fold more frequently in real vs. randomized data (see Figures A and B of this response and our calculations above). So how is it that Shedden and Cooper conclude that the observed periodic patterns are the product of random chance? The answer is that they failed to treat our data sets as independent replicates of the same measurement, *instead treating the two experiments as if they were completely unrelated*. In other words, instead of asking the appropriate question, "Is the occurrence of periodic patterns for the same gene in both data sets observed to be beyond that expected by random chance?", they asked "Is the occurrence of a periodic pattern for a gene in *one* data set likely beyond random chance? **Twice!**" Because of the complexity of the experiments involved, we agree that using only one data set would make it very difficult to reliably establish cell cycle regulation of multiple genes, which is why we performed the experiments in duplicate (triplicate counting the RT-PCR follow-up measurements) across nearly two complete cell cycles each time.

(2) Assertion: (Shedden and Cooper, Figure 2) A Fourier transform fit in one experiment is not predictive of a Fourier transform fit in the other experiment.

(2) Comments:

1. *There is no basis for expecting all genes that appear to cycle in one experiment to cycle in the other*. Real data are noisy because variation can be introduced during cell handling and RNA preparation, and that is exactly why we did independent replicate experiments: to reduce the incidence of false positives. **The appropriate requirement is not that ALL periodic behavior is consistent between the two data sets. Rather, we sought to identify periodic patterns that are reproducible and clearly not the product of random chance.** to illustrate this point, a simple example is helpful. Because of the experimental complexity and the biological complexity of the system, variations due to factors such as cell passage, small differences in cell density, exact environmental conditions and sample handling are likely to affect the expression levels of some genes. As such, there are two ways one could present the results of a study of this type. One might claim that there was perfectly reproducible evidence for every gene that displayed differential expression in any experiment. Alternatively, and more realistically, one would identify those genes that showed reproducible changes in expression in all replicates, and any differences observed in one measurement that were not consistent would be dismissed in the absence of other confirmatory evidence. This latter scenario is by far the more common because experiments are rarely perfect, and measuring consistency across independent replicates is an excellent (and standard) way to reduce false positives that may arise by chance. Shedden and Cooper seem to expect that independent replicates should be identical in all respects, and if they are not, then nothing can be considered real. **They have set up a straw-man argument to disprove a point that was never asserted in our paper, and more importantly, that is absolutely *not* required to conclude that we have identified cell cycle-regulated genes.** Once again, the requirement is not that ALL periodic behavior seen for any gene be consistent between the two data sets, but that it is possible to identify numerous periodic patterns that are clearly not the product of chance.

(3) Assertion: (Fig. 3) Periodically expressed genes do not demonstrate "peak decay".

(3) Comments:

1. *"Peak decay", as defined by Shedden and Cooper, is arbitrary and an unreasonable requirement for periodicity.* Like the "superimposability" test described above, "peak decay" – the difference in peak amplitudes between the first and second cell cycles - is an arbitrary measure of periodicity. First, as confirmed by the cell cycle experts mentioned above, "peak decay" in this particular manner has never been tested, much less validated, on any transcriptional data, conventional or genome-wide. And even if some degree of amplitude reduction might be expected, the EXTENT to which this should be seen is completely unknown (i.e., should there be a 5% reduction, a 10% reduction, a 50% reduction?), making any *a priori* requirement meaningless. Second, there are many known biological reasons why such an idealized definition of peak decay would not fit real data. For example, there could very well be some suppression of the peak amplitude of the first cell cycle following arrest. Third, many of our curves obviously *do* display a decrease in peak intensity between the first and second cell cycles (see Fig. B of this letter).

2. *The genes chosen to analyze "peak decay" were selected poorly.* An interesting question is why the obvious peak decay for some genes, as seen in Figure B of this response, was not identified in Shedden and Cooper's analysis. The answer is that the analysis of Shedden and Cooper is misleading. The genes Shedden and Cooper chose to analyze were not reproducibly oscillating genes between the data sets, and are not ones that we identified as cell-cycle regulated, so they are not expected to display any peak decay. Instead, they chose only genes that displayed periodicity in only one data set (Shedden and Cooper, Figure 3). In other words, they have chosen to seek peak decay for genes that are not expected to display this effect, and that were not identified originally as cel-cycle regulated. Given their choices, any observation of global peak decay would have been very surprising. To prove this point, we repeated the "peak decay" test on our list of 33 genes from Figures A and B (shown in Figure C). We observe two major groups of genes, the larger of which clearly displays peak decay. The genes in the smaller group that show a slightly higher peak in the second cycle may be simply genes whose expression is suppressed following cell cycle arrest.

References

1. Cho, R.J., et al., Transcriptional regulation and function during the human cell cycle. Nat Genet, 2001. **27**(1): p. 48-54.

2. Iyer, V.R., et al., The transcriptional program in the response of human fibroblasts to serum. Science, 1999. **283**(5398): p. 83-7.

3. White, K.P., et al., Microarray analysis of Drosophila development during metamorphosis. Science, 1999. **286**(5447): p. 2179-84.

4. Spellman, P.T., et al., Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell, 1998. **9**(12): p. 3273-97.

5. Cho, R.J., et al., A genome-wide transcriptional analysis of the mitotic cell cycle. Mol Cell, 1998. **2**(1): p. 65-73.