

Obtaining Optimal Results from Oligo GEMarrays Part One: Focus on cRNA Amount and Image Acquisition

The Oligo GEMarray from SuperArray Bioscience is designed for performing gene expression profiling that focuses on specific biological pathways, genes related to a particular disease state, or genes that are otherwise functionally similar. The development of the platform has already optimized many aspects of microarray performance. However as with any biological assay system, the end user or researcher needs to optimize a few experimental parameters in order to extract the maximum amount of useful information from their actual experiment. First, a sufficient amount of labeled target must be used in the hybridization, a factor determined not only by how much is added but also the initial synthesis yield. Some arrays require more labeled target than others. Second, the exposure time must also be varied to generate an image with a low background, high positive call, and wide dynamic range. Finally, choosing the appropriate normalization factor based on the observed results is also very important. All three of these factors are interdependent on one another. This article discusses each of these factors in turn with a particular focus on the amount of labeled cRNA target and on image acquisition.

Anticipated Raw Data:

The initial result (or the real raw data) from an Oligo GEMarray experiment is a grayscale digital image of the array spots. Generation of the image involves either an integrated imager, such as a CCD camera or a laser fluorescent scanner, or a flatbed scanner digitized image of developed X-ray film. The relative level of gene expression is directly related to the spot intensity. Dark spots represent highly expressed genes, light spots represent genes expressed at a lower level, and blank spots indicate a lack of expression or at least an undetectable level of expression. Figure 1 displays a few hypothetical array images containing only four spots. The spots represent either no, low, medium, or high levels of gene expression. The three simulated array images represent a result with too little cRNA or an under-exposure (Figure 1A), an ideal experimental result (Figure 1B), and a result with too much cRNA or an over-exposure (Figure 1C). The goal of optimizing the microarray experimental parameters is to obtain an image as close to the ideal as possible.

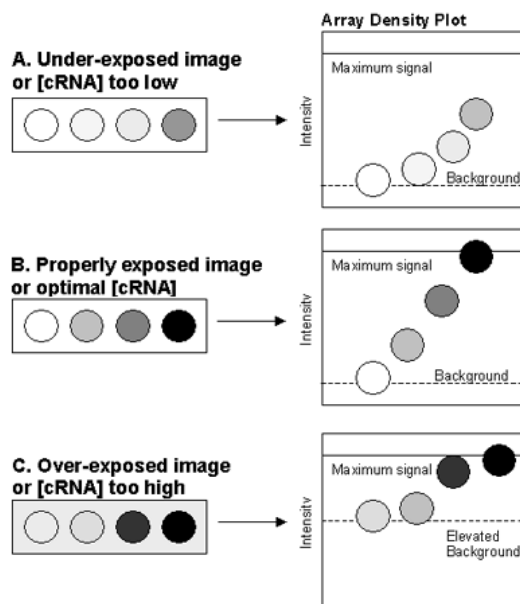


Figure 1: This figure displays three hypothetical microarray images with spots of varying intensities and results using various exposure times or amounts of labeled cRNA target in the hybridization.

Optimizing Labeled cRNA Target Yield:

The first critical parameter for optimal cRNA target yield is the quality of the RNA preparation. Follow all recommendations made by the RNA isolation kit manufacturer very closely. Small errors or mistakes in these protocols drastically affect the quantity and quality of the resulting total RNA. When isolating RNA from transformed cultured cells for use with the Oligo GEArray, we recommend using a spin-column based method, specifically the ArrayGrade™ Total RNA Isolation Kit ([GA-013](#)). For RNA isolation from tissues, we recommend using an extraction-based method first (specifically TRIzol from Invitrogen) followed by a clean up step with the above spin-column based method. Also, perform all RNA quality control checks recommended by the microarray labeling method before proceeding with cRNA target synthesis. For example, determine the A260/A280 and A260/A230 ratios to insure a lack of potential contamination by proteins or salts, respectively. Also, characterize the total RNA preparation by agarose gel electrophoresis to insure its integrity. Ethidium-bromide staining should reveal two sharp 28S and 18S ribosomal RNA (rRNA) bands with roughly a 2:1 intensity ratio. Analysis by the Agilent 2100 BioAnalyzer may also substitute for analysis by agarose gel.

First time users in particular should use as much RNA in the cRNA target synthesis as possible while remaining within the specifications set forth in the microarray user manual. For most applications where ample RNA is available (such as cultured cells), use the maximum amount of RNA recommended for the Oligo GEArray, 3 µg. The lower end of the RNA range is meant for applications where large quantities of RNA are difficult to isolate (such as small samples of difficult animal tissues or patient biopsies), and also require more optimization (to be discussed elsewhere), such as the length of the synthesis reaction (from 4 up to 18 hours). Under the ideal circumstances, a four-hour incubation should be sufficient to obtain adequate yields of cRNA (10 to 30 µg) from larger amounts of RNA (3 µg) starting material. Applications using smaller amounts of input RNA (as little as 500 ng) should extend the reaction up to 18 hours (overnight) for yields of up to 8 µg.

Throughout the labeled cRNA target synthesis procedure, use certified RNase-free reagents and lab ware particularly pipette tips and tubes. Prepare dilutions to determine cRNA yield by A260 in RNase-free TE buffer pH 8.0 and not water. The dilution needs to be well buffered because the UV absorbance of nucleic acid changes dramatically with pH. Also, perform all recommended quality control checks of the labeled cRNA target listed in the use manual before proceeding with microarray hybridization including agarose gel characterization and determination of the labeling efficiency.

Determining How Much To Use cRNA During Hybridization:

Low-Density Tetra-Spotted versus High-Density Single-Spotted Oligo GEArrays

The amount of cRNA or target used in a hybridization experiment is one of the most critical factors determining the overall signal level on array image. The more cRNA used, the more intense the signal on the array image. When an insufficient amount of target is used in the hybridization, the overall signal will be low and false negatives will be more likely (Figure 1A). When too much cRNA is used, it may cause a high background and more signal saturation (Figure 1C). However, an optimized amount of cRNA target can yield an ideal image (Figure 1B).

The Oligo GEArrays are available in different densities and printing layouts. The majority of the Oligo GEArrays contain approximately 100 genes (catalog numbers 001 through 400). In this type of array, each gene is represented by four repeated dots arranged in a unique square shape called a tetra-spot. SuperArray also offers Oligo GEArrays (catalog numbers 400 through 999) that contain up to 440 genes per array. In this type of array, a single spot represents each gene, and the distances between these spots are much smaller than between the tetra-spots. These differences dictate different considerations in the amount of cRNA used for hybridization and subsequent image acquisition. A previous newsletter article ([Volume 1, Issue 5, Article 2](#)) also discusses this concept.

In an experiment to further demonstrate these differences as well as the importance of using the correct amount of cRNA, we have generated, for testing purposes only, a single spot version of the normally tetra-spotted Oligo GEArray Mouse Cell Cycle Microarray ([OMM-020](#)). Increasing amounts of biotin-labeled cRNA target prepared from a universal source of reference RNA was hybridized to either the tetra-spot or the single spot versions of OMM-020. Figure 2 shows the array images acquired using the same exposure time, and Table 1 summarizes the raw data obtained via densitometry of the array images. The two different versions of the same microarray display different responses to the increasing amounts of cRNA.

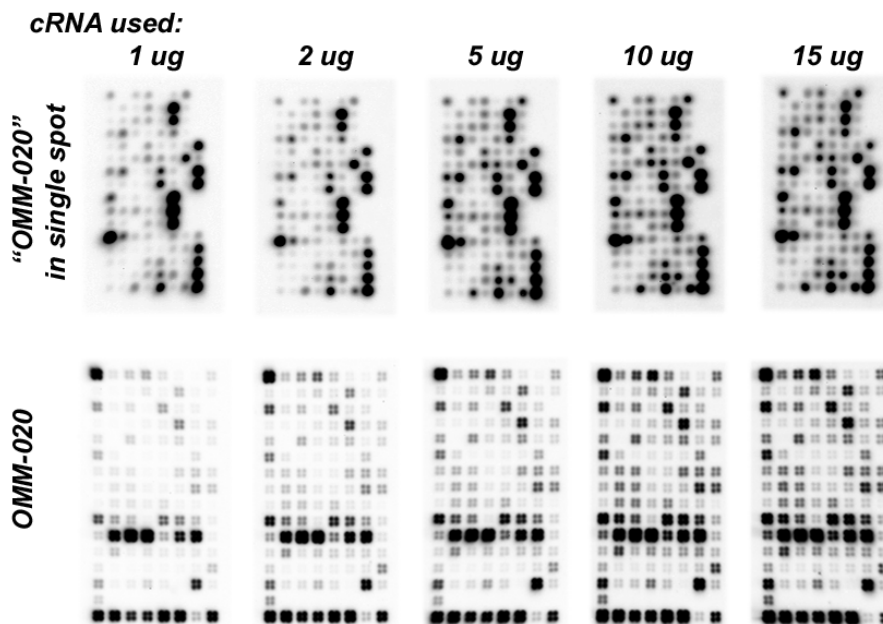


Figure 2: The same microarray spotted in two different formats responds differently to increasing amounts of hybridized cRNA. A single spot version of the normally tetra-spotted Oligo GEArray® Mouse Cell Cycle Microarray ([OMM-020](#)) was generated. Labeled cRNA target was prepared from 3 μg of XpressRef™ Mouse Universal Reference Total RNA ([GA-005](#)) using the TrueLabeling-AMP Kit ([GA-010](#)). Increasing amounts of biotin-labeled cRNA target (from 1 to 15 μg) was hybridized to both versions of the microarray. The image of each microarray is displayed. (The actual size of the single spot version of OMM-020 is about one-quarter that of the standard OMM-020. The former array images were enlarged to match the larger tetra-spotted OMM-020). Table 1 summarizes the densitometry results from these arrays.

Table 1: Summary of Spot Intensity Data from OMM-020 Microarray Images:

	OMM-020 Single Spot					OMM-020 Tetra-Spot				
	1 μg	2 μg	5 μg	10 μg	15 μg	1 μg	2 μg	5 μg	10 μg	15 μg
Total Intensity	1059312	1063576	1598280	1856632	1720853	662522	996107	1030566	1412542	1513423
Maximum Intensity	52464	49057	50671	51228	47991	50749	50970	49403	49707	49645
Minimum Intensity	458	534	930	1392	1272	177	280	411	699	844
Mean Intensity	8276	8309	12487	14505	13444	5176	7782	8051	11035	11824
Median Intensity	3131	3103	6483	7997	8145	1041	2146	2610	4498	6059
Background	547	614	1114	1602	1681	328	395	499	824	1002
Standard Deviation	80	68	260	316	542	52	21	53	74	114
Threshold	787	818	1895	2551	3307	482	458	659	1044	1343
Present Calls	117	119	113	113	109	103	117	118	118	118
Absent Calls	11	9	15	15	19	25	11	10	10	10

Increasing the amount of cRNA in the hybridization solution generally increases the signal intensities on both arrays formats. The background and minimum intensity values increase at roughly the same rate on both arrays with increasing amounts of cRNA hybridized. However, even the microarrays using the lowest amount of cRNA contain saturated spots as evidenced by the observation of roughly the same maximum intensity on all microarrays. To a certain extent therefore, increasing amounts of cRNA raise the background level ever closer to the saturation level and thereby reduce the linear dynamic range of the method, making the results look less like Figure 1B and more like Figure 1C.

Using the same amount of cRNA target, the mean and total intensities on the single-spotted microarray are greater than those of the corresponding tetra-spot arrays. These intensities reach saturated levels when roughly 5 μg of cRNA is used in the hybridization with the single-spotted microarray. However for the tetra-spot array, the mean and total signal intensities continue to increase even when 15 μg of cRNA is used in the hybridization. This result indicates that the single-spotted microarrays more easily saturate with lower levels of cRNA than the tetra-spotted microarrays. These results also show a loss of positive calls at higher levels of cRNA on the single-spotted microarray but an increase to a steady number of positive calls even at the highest level of cRNA on the tetra-spotted microarray. Reciprocal changes occur in the absent calls. An increasing background and an increasing incidence of spots that “bleed” into adjacent areas on the single-spotted microarray most likely causes this phenomenon.

Therefore, to avoid turning ideal Oligo GEArray experiments (Figure 1B) into over-exposed images (Figure 1C), a lower amount of cRNA (2-5 μg) is recommended for the higher-density, single-spotted microarrays (catalog numbers 000-400), and a higher amount of cRNA (5-20 μg) is recommended for the lower-density, tetra-spotted microarrays (catalog numbers 401-999).

Each array image in this experiment has been acquired with the same exposure time. As discussed below, exposure time is another control point where adjustments must be made to optimize the array image and results. For example, an image of the single-spotted microarray hybridized with 15 μg of cRNA re-acquired using a shorter exposure time may not be as over-exposed and yield fewer saturated signals. However, microarray experiments are best compared if the data is acquired under the same conditions, including exposure time. Therefore, a balance is usually struck between the cRNA amount used in hybridization and the exposure time.

Determining The Optimal Exposure Time:

Exposure time to the capture medium, whether X-ray film or a CCD camera, is another critical factor defined by the researcher determining the overall signal level on array image. Longer exposure times generate more intense signals on the array image. When the exposure time is too short, the overall signal will be low and false negatives will be more likely (Figure 1A). When the exposure time is too long, the background will be higher and more signals will reach saturation (Figure 1C). Both situations also narrow the available linear dynamic range of the detection medium and the microarray experiment. However, an optimized exposure time can yield the best image possible (Figure 1B).

In Figure 1B, all positive spots are significantly darker than the blank spot and significantly lighter than the most intense spot yielding an obvious intensity difference between all of these positive spots. The four spots span the complete distance between the background and the maximum signal level that the capture medium permits. Therefore, this array image utilizes the full dynamic range available. The under-exposed image (Figure 1A) does not utilize the upper part of the capture medium’s dynamic range. All of the positive spots are packed in a narrow range close to the background level, and the low expression spot is not significantly darker than the background causing more false negative signals. In the over-exposed image (Figure 1C), all of the spots are much darker. However, the increased background level shrinks the dynamic range between the low and high intensity spots. The high background also masks the genes expressed at a lower level thereby, much like short exposure times, increasing the number of false negatives. Over-

exposure also makes medium expression spots much darker so that they cannot be differentiated from high expression spots.

A good array image should have a minimum background and fully utilize the linear dynamic range offered by the recording medium. A good array image with good intensity distribution can only be obtained by trial and error. A good practice from photography, called bracketing, can be used. Generate a series of exposures with both longer and shorter than the estimated proper exposure time to help guarantee that one of the exposures should be optimal. Therefore when a researcher starts using a microarray for the first time or a new array experimentation condition is used for the first time (e.g. new sample, new labeling or processing method), appropriate testing should be conducted to determine optimal exposure conditions.

The principles discussed here apply to most image acquisition devices, particularly X-ray film or cooled CCD camera imager. Multiple exposure times can be generated in succession without affecting the results. Similar precautions should be observed when setting the laser power or PMT setting of laser fluorescent scanners. Those settings also influence the background and signal saturation of the fluorescent detection method. However, these scans must be obtained correctly the first time due to the photo bleaching caused by the scan itself. It may prove helpful in fluorescent detection to generate a set of replicate test microarray experiments to be used for optimizing detection before real experiments are performed with more precious samples.

Remember that image acquisition and the amount of cRNA used during hybridization work hand-in-hand to define the overall intensity and dynamic range of the array image and results. If the hybridization contains too little labeled cRNA target, it may be difficult to obtain a long enough exposure time for a good array image. Similarly, if the hybridization contains too much labeled cRNA target, it may prove difficult to obtain a short enough exposure time for a good array image.

Normalizing Microarray Results:

Normalizing microarray results is the mathematical operation of dividing background-corrected intensity values by a common factor that has a unique value for each microarray. This factor must have the ability to respond to systematic variation or error in the data from one array to the next. The division corrects the data for that systematic variation or error so that results for each gene can be compared across arrays and therefore across biological experimental conditions.

Microarray researchers use a few different methods for normalization. The most commonly used and recognized method corrects the data based on the relative expression level of an appropriate housekeeping gene. These genes are usually expressed well enough for the microarray to detect and maintain a constant expression despite the occurrence of other biological changes. However, not every housekeeping gene is suitable for every experimental study. Some of these genes may change their expression level during certain biological changes or responses. For example, the expression of ACTB increases throughout development, and GAPD may be a better choice for such studies. Conversely, in certain nutrition models, the level of GAPD changes with nutrient status, and ACTB may be a better choice.

For this very reason, the Oligo GEArrays include several (from 4 to 9) different housekeeping genes so that the researcher may choose the appropriate one for each experiment or study. The chosen housekeeping gene must not change its raw expression level (spot intensity) more than the array-to-array reproducibility of the microarray method (10 percent in the case of the Oligo GEArray). If it changes by more than this amount, its expression may very well be regulated by the experimental conditions under study. Also, the expression level of the chosen gene must not lie too close to the saturation or background level of the microarray. Only if the gene has a moderate level of expression can the microarray reliably detect whether the expression is actually changing or not.

Other factors used for normalizing microarray results, particularly when all of the available housekeeping genes fail to meet both of the above criteria, include the median and interquartile values of the microarray. The median value is the simple statistical parameter of the median of all the intensity values on the microarray. (For example, see Table 1.) This value changes with the systematic variation on the array. Plots of intensity values in rank order for several different arrays results clearly demonstrate this notion. The resulting sigmoid-shaped curve, which looks much like a traditional hybridization "COT" curve, and the median intensity value shifts in concert up or down in this plot from array to array. (Data not shown.)

The interquartile is also a median value, except the calculation only uses the middle 50 percent of all of the intensity values on the array. It ignores one-quarter of the least intense signals and one-quarter of the most intense signals. This method is particularly useful for low-density microarrays such as the Oligo GEArrays. If the microarray results contain a disproportionate number of low intensity (genes expressed at a very low level) or high intensity spots (genes expressed at a very high level), results normalized to a simple median skew the gene expression levels down or up, respectively. The interquartile removes the influence of the more extreme intensity values, uses only the genes theoretically expressed in the linear dynamic range of the microarray experiment, and prevents skewed gene expression levels.

SUMMARY:

The amounts of cRNA used for hybridization and the image acquisition settings are two important experimental parameters that can affect GEArray results. The investigator needs to use enough labeled target to generate good signal intensity. However, more cRNA is not always better if signal saturation and bleeding becomes a problem. The cRNA range used in this example should only be considered as a reference or starting point for your own experiment, since different cRNA preparations may have different qualities. The quantity and quality of the initial total RNA samples also affect the quantity and quality of the labeled cRNA target. The quality of all materials generated for the microarray experiment should always be checked as described. Finally, using the proper factor for normalization also helps optimize the microarray result by correcting for systematic error between microarrays and experimental conditions.