Technical Note

UHN Microarray Centre

# Microarray Experiments in a High-Throughput Manner

- Most aspects of the standard microarray protocol can be modified to allow for high-throughput gene expression profiling.
- Such modifications can increase the overall capacity of a core facility or high-volume laboratory and the technical and experimental variability can be reduced.
- The major disadvantage to complete automation of the protocol is the high cost associated with specialised robotics, however for laboratories that already posess such equipment it is possible to put them to use in the microarray workflow.

**Introduction**

New technologies for genomics are helping to speed the drug discovery process. Automation is important in allowing researchers to meet the high-throughput demands of today's research environment. For a core facility, the ability to perform most, if not all, aspects of the experiment in a high-throughput manner would be beneficial. Automation allows for more samples to be processed simultaneously, while reducing the likelihood of simple errors and the technical variability involved with multi-step experiments.

The UHNMAC has evaluated many products and instruments that would allow for high-throughput processing at various steps of the microarray process, including array production, RNA isolation, sample labelling, purification of labelled-cDNA, hybridisation, washing, scanning, and spot quantification.

Many aspects of array production have already been automated. cDNA array production involves PCR amplification of clone libraries, purification of the PCR-amplified products and transferring appropriate volumes to printing plates, and spotting the nucleic acids onto the arrays. Similarly oligonucleotide array production requires the high-throughput generation of thousands of oligonucleotide probes, resuspension in appropriate spotting buffers, normalisation of concentration and then spotting of the products onto the arrays. The UHNMAC has automated this entire process either in-house or through the use of commercial vendors (in the case of oligo production). Once arrays are printed, depending on the substrates used, several post-printing processing steps may be required such as chemical blocking of reactive sites on the slide. Here too automation can play a roll, and not only increases throughput and reproducibility but also reduces exposure of the technician to potentially hazardous chemicals. The UHNMAC Bioengineering group has developed a system that has automated the post-printing array processing allowing for the processing of up to 476 arrays in the time it would take a technician to process one quarter of that amount.

While the production process has been automated in a high-throughput and robust manner, the downstream usage of arrays has largely proceeded in a manual or at best semi-automated manner. This report outlines the work that has been done to perform several aspects of microarray experiments in a high-throughput manner and, more importantly,

integrated manner. In particular RNA isolation, cDNA labelling, and purification have been addressed. Integration with automated hybridisation systems is possible as are several unique solutions. Post hybridsation, washing, and both scanning and spot quantitation can also be automated to further increase throughput.

**Options for Automating Microarray Experiments**

One of the keys to successful full automation of the labelling protocol is to operate in 96-well format as much as possible. This ensures proper translation from one step to another, allows for barcoding and tracking of each step and maximises throughtput.

RNA isolation: RNA isolation is usually carried out using reagents such as Trizol (Invitrogen) or spin-columns (RNeasy, Qiagen; Absolutely RNA, Stratagene). Neither of these methods is ideally suitable to automation and as such we looked at alternatives. Both the RNeasy (Qiagen) and Absolutely RNA™(Stratagene) spin columns are also available in 96-well plate formats that allow for more easy integration into an automated system. Both kits facilitate efficient, high-throughput RNA sample preparation. These kits combine the selective binding properties of a silica gel and fiber matrix-based membrane, respectively, with the speed of vacuum based and/or centrifugation driven filtration. The protocols used are in essense the same as for the spin-column variants. Cells are first lysed under highly denaturing conditions with guanidine isothiocyanate followed by the addition of ethanol. This procedure can be carried out in 6-, 24- or 96-well plates using a robotic liquid handler (in our case a Perkin Elmer Multiprobe). In the many cases the cells can be grown directly in the bottom of the wells of these plates, and the media can be aspirated, and then the lysis buffer added. Ethanol is applied to the lysate using the robotic system and the resultant mixture is applied to the 96-well purification plate. Total RNA binds to the column while contaminants are efficiently washed away. In order to draw solutions through the filter plate, either a vacuum manifold can be used (as is found on many high-throughput automation systems) or a robotically accessible centrifuge (such as that from Velocity 11) can be used. We have utilised both methodologies with success. In order to elute samples, a clean collection plate is placed under the filter plate and high quality RNA is then eluted in a small volume of water for the RNeasy kit or in a small volume of elution buffer for Absolutely RNA™ 96 microprep kit. Both kits are ideal for simultaneous isolation of 96 samples from up to $5x10^5$ cells. We found that both kits provided high quality RNA with sufficient quantities. Comparison of RNA isolated via these methods with our standard spin-column based extractions showed little difference.

Labelling reaction: It is clear that setting up the cDNA labelling protocol using either of the direct or indirect labelling methods involves various pipetting steps. Such steps are easily programmed into any liquid handler that has sufficient pipetting accuracy (1-2 µl volumes). We have automated both of these protocols with relative ease, however there are several approaches that can taken. The biggest issue facing automation is the fact that many of the reagents are used (and supplied) in small, automation-unfriendly quantities. An automation system such as the Multiprobe offers a few solutions to this problem. One way of dealing with the problem is the use of the individually addressable pipetting tips. One can instruct the robot to use only a single tip at a time to dip into a tube of enzyme for example. Further to this is the ability to use the robotic liquid handler as a “repeater pipette”. As such, the robot takes up enough solution from the tube to accommodate all the wells of the plate that it may pipette into (for example if using an eight channel robot – each tip would draw 12 µl of solution if 1 µl was to be added to each well – plus some amount as a “buffer” which can be added back to the tube if required). The robot

then is instructed to pipette an appropriate amount of the reagent into each well – without having to return ot the original source. Another strategy which we found to be highly effective was to create either master source plates or master mixes (or a combination of both). With such a strategy a 96 well plate is used to aliquot out the necessary reagents in larger volumes using each column of the plate for a separate reagent. Alternatively the reagents are all mixed in the appropriate ratio in a single tube/ well or column of wells. This also improves pipetting accuracy which is highly desirable.

The total RNA that is extracted from the previous step is already in 96 well plates and as such, aliquots can be taken and placed directly into another 96-well plate for labelling which helps integrate these two parts of the protocol. In order to allow incubation of the samples, we have found that either a robotically accessible hotplate or thermal-cycler can be used. We found the MJ Dyads work very well for this part of the protocol.

Comparison of our automated labelling methodology with a traditional tube-based method using a waterbath showed little difference and indicates that automated labelling is at least as good as our standard manual methodologies.

Purification of labelled-cDNA: Several of the steps in the labelling protocol require purification of either an intermediate or the final labelled product. Again, it is highly desirable to perform this purification using a 96-well format and to this end, five methods of purification were evaluated: RNeasy 96 kit (Qiagen), Microcon®-96 Retentate Assembly Kit (Millipore), AcroPrep™ (Pall), Montage $PCR_{u96}$ (Millipore) and Montage $PCR_{96}$ (Millipore). All of the kits allowed for vacuum or centrifuge-based filtration. Each kit was evaluated following the manufacturer's protocol. These high-throughput purification methods were compared to the standard method using CyScribe™ GFX™ columns (Amersham). It is important to note that none of the plate based methods have been recommended by the manufacturer for this particular application and as such we did not expect to have success with many, if any, of the kits.

Although the RNeasy 96 kit (Qiagen) is primarily designed for the isolation of total RNA, the manufacturer suggests that it can also be used to purify RNA following labelling reactions. In our experience cDNA purified using this kit produced very weak signal intensities following hybridisation to the array possibly due to low recovery of the cDNA. We suspect that the filter became clogged as only 8 µL of the 30 µL of elution buffer was recovered despite extended centrifugation.

The Microcon®-96 Retentate Assembly Kit is a size-exclusion based filtration kit with a molecular weight cut off of 30kDa. The sample is loaded into Microcon filter units and filtered into one of the U-bottom plates using a centrifuge. The retentate recovery process involves placing a 96-well plate over the Microcon filter, inverting the entire assembly, and re-centrifuging. The retentate then collects in the second plate. As there is no reliable dead stop, variation in the elution volume was observed from column to column. For each of the three hybridisations, the results were different and ranged from average signal intensity to weak signal and high background.

The AcroPrep™ (Pall) filtration plate is another size-exclusion based filtration kit with a molecular weight cut off of 100kDa. Purification of labelled-cDNA with the AcroPrep™ (Pall) resulted in hybridisation images with weak signal intensities and high background on all replicates.

The Montage $PCR_{u96}$ (Millipore) plate enables purification of reaction volumes of 20-100 µL. This vacuum based, size-exclusion separation method effectively removes contaminating

salts, unincorporated dNTPs and primers from labelling reactions in less than 15 minutes. Sample contamination was an issue with these plates as the membrane at the bottom of the plate is not adequately divided. Despite sample cross-over, good signal intensities were seen on all hybridisations.

The Montage $PCR_{96}$ (Millipore) was found to be the most automation-friendly method for the purification of labelled-cDNA. There is no centrifugation or lengthy binding and elution steps required and the 15-minute protocol yields consistently good results. Excellent images (good signal intensities with minimal background) were obtained when using the Montage $PCR_{96}$ (Millipore) kit for labelled-cDNA purification in the Direct and Indirect Labelling protocols. We were very pleased with this result as it effectively allows us to automate all the steps up to the hybridisation of the arrays.

Hybridisation & washing: At this point and time, hybridisation is one area of the protocol that continues to be semi-automated at best. While several excellent solutions exist for automated hybridisation (for example the Advalytix SlideBooster), these stations require manual setup and as such it is not possible to go straight from the 96-well source plate post-labelling onto the hybridisation station. As such this aspect of the protocol was not studied extensively in this study. In preliminary evaluations, we have found that the standard plastic slide box is sufficient for overnight hybridizations of cDNA microarrays. There is really no limit to how many you can set up at once, although 6 hybridisation boxes (2 arrays per hybridisation box; total of 12 arrays) is usually the most one technician handles per day due to the limitaitions of the upstream parts of the protocol. We have also found that our washing protocol, using plastic wash boxes and racks that hold 25 slides, are ideal (and inexpensive) for washing 12 slides (placed in every other slot of the 25-slide rack). As our eventual goal is to automate as much of the process as possible we have looked at ways to improve the way that we perform hybridisation. We are working on several novel approaches that we believe will eventually allow for a fully integrated automation solution.

**Automation, Throughput and Sample Tracking**

Automation provides the potential for higher throughput, as well as greater reproducibility and fewer errors. We have taken the approach of adapting as much of the protocol as possible to a 96-well format as this is currently the most amenable to automation. It is important to note that often automation does not truly increase throughput. In many cases we have seen that a good technician can easily outpace a robotic system. Where the benefit really comes into play is that while the technician can often work faster, the robot is not prone to distraction or fatigue.

Adaptation of the microarray protocol to a 96-well format can in fact increase the throughput of manual methods that use more traditional means. While many technicians feel comfortable processing a dozen samples at a time, a lot of manual pipetting steps mean that scaling up much beyond this level is difficult. Using 96-well plates along with multichannel pipettors can have a tremendous impact and can increase throughput by as much as 8-fold if done properly.

Another advantage of automation is the increased ability for sample tracking. Most liquid handling systems today are capable of taking advantage of barcoding to allow for automated logging of what samples were processed, what sample plates or reagents were used et cetera. As throughput increases, such process control become paramount.

**Automation Post-Hybridisation**

While automation of the microarray protocol has many clear benefits, the utility of implementing such a solution is only beneficial if it does not create a bottle neck further downstream. All

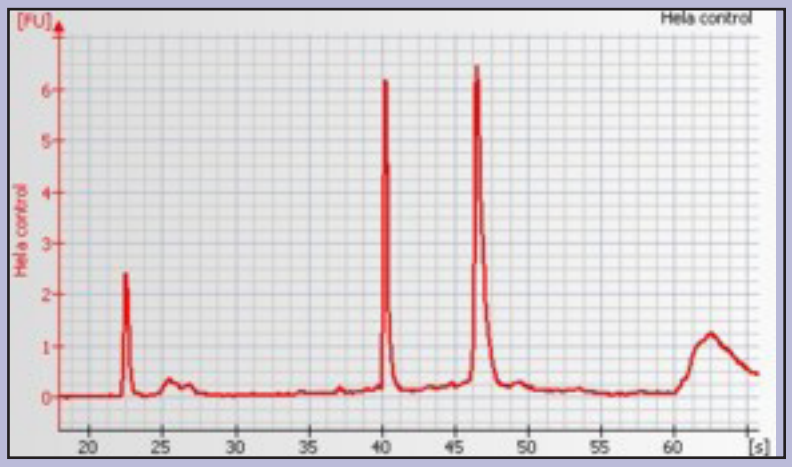
aspects of the process from production to analysis must be balanced for any benefit to be realised. Fortunately many of the downstream aspects of the microarray experiment work-flow have been automated over the past few years. Many scanners currently on the market (Agilent, Molecular Devices/ Axon, PerkinElmer) have autoloader systems that allow for "load and go" scanning. The issue however is that most such systems handle a maximum of 48 slides at a time (and more often only 12). Critical to this automation is the ability of the user to "trust" the scanner to choose settings that are appropriate for image acquisition. Each of the scanner manufacturers employ different methodologies to select the optimal scanning settings and it is up to the user to determine which if any are reasonable.

How then, does one balance the workflow? Automation of microarray processing can easily produce 384 labelled samples in a day (even more if overnight runs are used). One would then be required to set up this many hybridisations (using as many as 48 automated hybridisation stations) and to then use 8 scanners with 48 slide carousels in order to process all of the slides. Clearly this is not a practical solution, however it does not mean that automation is a pointless exercise. It is rare that a laboratory would have to process 384 standard microarrays in one day (outside of a commercial setting). However, drug discovery, diagnostic and validation applications could require throughput in this range. What separates many of these applications from discovery-based microarray experiments is that frequently fewer genes need to be profiled in parallel. As such, it is often possible to multiplex assays to allow processing of 12, 16 or even 24 samples on a single slide. In such situations it becomes immediately apparent how one might balance upstream and downstream automation of the microrarray procedure. With a "12-up" design for an array only 32 individual slides would need to be processed at a time, making it quite feasible to perform an experiment with 384 samples in a day. With improving technologies companies such as Agilent and Illumnia now produce arrays with 4 to 8 full copies of the human genome represented. Such formats would easily allow for 96 samples to be processed in a day again helping to balance the entire workflow.

Once images are acquired several "quantitation" software packages allow for automated processing of the images in order to output intensity values and various quality metrics. This can often be done in a hands-off, walk-away manner that allows for overnight processing – allowing for optimisation of computer resources. This final step in the processing of the microarrays allow for front to back automation of the experiment, increasing throughput and reducing variability.

Despite all of this, it is important to remember that acquisition of data (from RNA extraction to image quantification) is only the beginning – the data analysis and mining that occurs during the bioinformatics stage of the experiment can take a great deal of time. Automation does allow the researcher to get the full dataset into to the hands of the bioinformatician earlier – and to decrease variance, thus increasing statistical power which is beneficial to all involved.

A




B

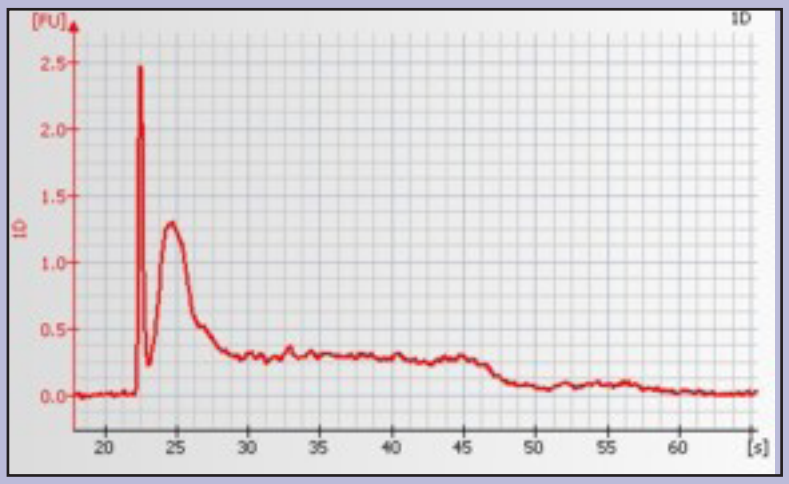



Figure 1. As determined by the Agilent 2100 Bioanalyzer, high quality total RNA was obtained from Qiagen's RNeasy 96 kit (image A) and Stratagene's Absolutely RNA™ 96 microprep kit (image B). The electropherograms clearly show distinct 18S and 28S rRNA peaks with no signs of RNA degradation or genomic DNA contamination.