CHAPTER III

WHOLE-ANIMAL MICROARRAY EXPERIMENTS TO IDENTIFY UNC-4 TARGET GENES

Introduction

As described in Chapter II, genetic results have ruled out a role for the UNC- 4 target genes *acr-5* (acetylcholine receptor subunit), *del-1* (DEG/ENaC cation subunit), and *glr-4* (glutamate receptor subunit) in synaptic specificity. The key to uncovering the mechanism UNC-4 employs to insure proper synaptic specificity lies in identifying other downstream genes. Since my experiments have revealed that *unc-4* controls multiple genes (i.e. redundant pathways), these genes are unlikely to be revealed in suppressor screens. Thus, we opted for a microarray-based strategy with which to query the entire genome for all *unc-4* regulated genes.

Two types of DNA microarrays were developed in the 1990s, spotted arrays and synthesized arrays (e.g. Affymetrix, see Chapter IV). Here I describe spotted arrays, which contain nucleic acids (genomic fragments, cDNAs, oligos) which are 'printed' onto a glass slide. mRNA from two experimental conditions (i.e. wildtype vs. mutant) are converted into differentially labeled cDNAs (i.e. green and red) and hybridized to the array. Spots that have similar intensities between the red and green channel are genes that are unchanged between the two conditions. Spots that are differentially affected are brighter in one channel vs. the other. Thus the readout of the experiment provides a global snapshot of gene expression.

In 2001, when I began these experiments, Stuart Kim's lab at Stanford was pioneering the use of printed microarrays in *C. elegans*. His lab generated arrays containing most (~17,500) of the predicted genes in *C. elegans* and published gene-expression profiles of *C. elegans* development and of the germline (Reinke et al. 2000; Jiang et al. 2001). With these efforts proving fruitful, we decided to adopt a similar whole-animal genomics approach to identify UNC-4 regulated transcripts. For these experiments, mRNA was extracted from wildtype (wt), *unc-4 (e120),* and *unc-37 (e262)* animals. UNC-4 target genes were sought by identifying transcripts with elevated expression levels in *unc-4* and *unc-37* mutants vs. wt. After performing the hybridizations, scanning, and data analysis, we identified 10 significantly ($P \leq 0.001$) upregulated genes. This number, however, is less than the expected random frequency (17.5). We therefore concluded that the sensitivity of this whole-animal experimental design was insufficient to reliably detect *unc-4* regulated transcripts.

Materials and Methods

Preparing large populations of synchronized L2 animals for RNA extraction (from Jennifer Ross and David Zarkower)

To generate a sufficient amount of mRNA to use for microarray experiments, I had to grow nematodes in large batches of liquid culture; each microarray hybridization required 10 μg of mRNA. Thus, 20 μg is the minimum amount of mRNA needed from each batch of wildtype animals and 10 μg from the mutant animals. However, more RNA is actually required for other controls, and in case of an error while labeling or

hybridizing. The following protocol generates, on average, 1 ml of packed L2 animals; from those I extracted ~90 μg of mRNA.

Growing bacteria used to feed worms in liquid culture

Nine liters of media (5X stock: 133 g tryptone, 267 g yeast extract, 88 ml 50% glycerol, 200 ml 5X Phosphate buffer [115.5 g KH_2PO_4, 627 g K_2HPO_4, 1L milli-Q H_2O], 1.912L milli-Q H2O) were inoculated with 2L of starter culture (*E. coli* NA22 in 2XYT, grown overnight) and grown in an 11L fermenter at 37ºC for 4-6 hours. The bacteria were pelleted by repeated centrifugation in 500 ml bottles. Each pellet was resuspended in 20 ml M9/Glycerol (~10% glycerol). Individual drops of bacterial resuspension were flash-frozen in liquid nitrogen. Known gram amounts were transferred into 15 or 50 ml conical tubes and stored at -80ºC until needed.

Preparing synchronized populations of L2 larvae for RNA extraction

All spins are done at 5000 rpm for 5 min in Beckman rotor 7.5 at 4ºC. Worms were grown on six-eight 150 mm 8P Plates at 20ºC. When most of the food was consumed, worms were washed off using M9. Worms were added to 1L S-media [5.8g NaCl, 50 ml 1M $KHPO_4$ (136.1g KH_2PO_4, adjusted to pH 6.0 with solid KOH, in IL), 950 ml milli-Q H_2O; autoclave. Then add (per 500ml): 1.5 ml MgSO4, 1.5 ml 1M $CaCl_2$, 5 ml trace metals solution (0.346g $FeSO_4$-$7H_2O$; 0.930g Na_2EDTA; 0.098g $MnCl_2$-$4H_2O$; 0.144g $ZnSO_4$-$7H_2O$; 0.012g $CuSO_4$-$5H_2O$ per 500 ml, autoclaved, wrapped in foil), 5 ml KCitrate pH 6.0 (21.02g citric acid, monohydrate, adjusted to pH 6.0 with solid KOH, in 100 ml), 5 ml antibiotic/antimycotic (Invitrogen 15240-062)]. Add 5-7g bacteria. Shake

at 20ºC and monitor growth, adding bacteria and antibiotic/antimycotic as needed. After 3-5 days most of the animals should be gravid adults. Harvest by pouring culture into two 500 ml glass cylinders and leave at 4ºC for 2 hours (to allow worms to settle to the bottom). Remove most of the liquid with a pipette. To capture worms, remove cotton from the tip of a 25 ml pipette. Place pipette into pipettor upside-down and pipet worm suspension; transfer to two 50 ml conicals. Wash in 2X cold M9; pool worms. Sucrose float (30 ml cold M9, 20 ml 70% sucrose). Harvest floating worms with a pasteur pipet (remove thin tip with diamond-tipped pen). Wash in M9 3X. Perform a hypochlorite extraction of the embryos as follows: Resuspend sucrosed adults in 75 ml bleach solution (15 ml Chlorox, 3.75 ml 10N NaOH, 56.26 ml water). Transfer to 100 ml beaker (with stir bar) and stir fast 5-7 minutes (until solution turns dark yellow and many eggs have been released). Filter through a 53 μm Spectra nylon mesh (Fisher 08-670-201) to remove carcasses and dauer larvae from the embryos. Pour filtrate into four 50 ml conicals tubes and spin for 2.5 min. Pool eggs, wash 3X M9, spin 5 min. Hatch eggs overnight by shaking at 20ºC in S-media with no-food (to ensure egg prep worked, also drop embryos into a 60 mm plate). Harvest L1s by pouring into 500 ml glass cylinder and incubate for 30 min at 4ºC. Remove most of the liquid and transfer L1s (using pipet tip without cotton, as above) into four 50 ml conicals. Spin, then pool L1s and wash 3X M9. Sucrose float and wash as above. Add L1s to 500 ml S-media, then add 5 g bacteria and shake at 20ºC for 3-5 days. When the population contains mostly gravid adults, harvest adults, bleach, allow eggs to hatch, and harvest L1s as above. Pool L1s and resuspend in 10-15 ml M9. Using a transfer pipet drop 500 μl-1 ml per 150 mM 8P plate (depends on density of worms). Grow at 25ºC for 18-20 hours for mid-L2 (as evidenced by post-deirid

division). Harvest L2s by washing off the plates in cold M9; collect in 50 ml tubes, Wash and pool L2s. Sucrose float and wash as above. Remove most of the supernatant and resuspend worms in remaining solution. Transfer worms to microcentrifuge tubes (500 μl aliquots). Wash 1X M9. Remove an aliquot of worms to check on the synchrony of the population, then remove all of the supernatant and flash freeze in liquid nitrogen. Store at -80ºC.

Extracting total RNA from frozen L2 larvae

Estimate size of frozen worm pellet (usually around 500 μl). In a 50 ml conical tube place 1 ml TriZOL reagent for every 100 μl frozen worms. Homogenize worms by grinding with an RNase-free mortar and pestle under liquid nitrogen. Use an RNase-free spatula to transfer worm powder into TriZOL; mix gently to dissolve powder. Incubate for 5 minutes at RT. Aliquot 1 ml TriZOL/worm suspension into microcentrifuge tubes. Add 0.2 ml chloroform/tube. Mix by inversion then incubate at RT for 2-3 minutes. Centrifuge for 4-5 min (no more than 12000 x*g*) at 4ºC. Transfer the clear, upper aqueous layer to fresh tubes. Precipitate with equal volume (~0.5 ml) isopropanol/tube. Mix by inversion, incubate at RT for 10 min. Centrifuge (no more than 12000 x*g*) for 10 min at 4ºC(RNA pellet should be clear and gel-like). Remove supernatant and wash with 1 ml 75% RNase-free EtOH. Mix by vortexing, spin at 7500 x*g* for 5 min at 4ºC. Carefully remove supernatant and air dry pellet (do not overdry!). Resuspend each pellet in 200 μl DEPC-treated water, then pool (1 ml total). Determine concentration using a spectrophotometer. Store at -80ºC.

Isolating mRNA from extracted total RNA

Follow the manufacturer's directions for the Ambion Poly-A Pure kit, except start with 2.0 mg total RNA and resuspend final mRNA in 45-50 μl DEPC-treated water. (Typical yield ranges from 80-100 μg). Run 5 μl on formaldehyde gel to check RNA integrity and mRNA enrichment. Alternatively, use an Agilent Bioanalyzer (Vanderbilt Microarray Shared Resource) to check quality of mRNA. Dilute to 1 μg/μl and store in 10 μg aliquots.

Making labeled cDNA for microarray hybridization

To 10 μg of mRNA, add 1.5 μl anchored dT primer (Operon, 22mer, 5'T_{22} VN 3', 2 μg/μl). Vortex to mix, spin down, and then incubate at 70ºC for 10 min. Quench in ice for 8 min. To RNA/anchored dT reaction add: 9.6 μl Trimix (0.6 μl 25 mM dATP/dCTP/dGTP/10 mM dTTP, 6 μl SuperScript buffer, 3 μl 0.1M DTT), 2.9 μl water, 3 μl Cy*-dUTP (can use either Cy3 or Cy5), 1 μl RNase Inhibitor, 2 μl SuperScript II (Invitrogen). Incubate in 42ºC water bath for 1.5 hours. Add 1.5 μl 1N NaOH to degrade RNA; incubate at 65ºC for exactly 10 min. Add 1.5 μl HCl to neutralize NaOH. Use QiaQuick PCR Purification kit (Qiagen #28106) to purify cDNA. Follow manufacturer's protocol except: mix together two RNA species to be hybridized together on the array (e.g. N2-Cy5 vs. unc4-Cy3) and elute 2X 30 μl. Speed vac eluates to evaporate all liquid (use high heat, usually takes 1.5-3 hours). Resuspend colored pellet (without pipeting up and down) in 2.5 μl Tris pH 7.4, 3 μl 20X SSC, 2 μl yeast tRNA, 0.5 μl 10% SDS. Heat at 100ºC for 2 min. Spin for 10 min at max speed to collect labeled, denatured cDNA.

Hybridizing labeled cDNA to microarrays

Each array consists of a glass slide spotted with *C. elegans* exon-enriched genomic DNA. Each array is labeled with a sticker at one end. The spots begin approximately 2 mm away from this sticker. Place 28.5 μl of probe mix in the center of the array, slightly biased toward the sticker. Angle a 20x44 coverslip and drop on slide (to reduce the chance of bubbles). Position coverslip 2 mm away from sticker. Place array in hybridization chamber (2 black metal pieces held together by screws). Place equal sized drops of 3X SSC on array (on both sides of coverslip). Put top on and screw tightly (to ensure no water leaks in). Put chamber in 65ºC water bath for 20 hours. Following hybridization, wash arrays. First, place array in glass slide tray. Drop tray into glass chamber containing 200 ml 3X SSC, 0.2% SDS. Mix up and down until coverslips fall off, incubating for a total of 1 min. Incubate arrays in 200ml 0.2X SSC for 1 min. Finally, transfer tray to 0.1X SSC for 1 min. Spin slide tray in eppendorf rotor for 2 min @ 600 rpm. Transfer arrays to slide box until ready to scan.

Scanning hybridized arrays

Use GenePix Pro software connected to a scanner from Axon labs. Place slide in scanner, face down with sticker in front. First perform a low resolution scan. Look at PMT histograms to see if one channel has a higher intensity than the other. Rescan after adjusting PMTs so that histograms overlay. Highlight the area containing the spots and perform a high resolution scan. Save as 2 channel TIFF.

Masking scanned array images

Open 2 color TIFF in GenePix Pro. Click on New Block button and overlay on spots. Computer automatically resizes circles to encompass spots, discarding those it deems unusable. Manually change spots to include or exclude spots the computer misread. Save mask as *.gps file.

Normalizing microarray data

In GenePix, open *.gps file. Use analyze function to create a *.gpr file containing raw intensity values. Upload raw values to the Stanford Microarray Database (SMD), which normalizes data (channel 1 vs. channel 2).

Retrieving data from SMD

Using the username and password, log on to the SMD (http://genome-www5.stanford.edu/). To retrieve data, click JOHNWANG under experimenter and *C. elegans* for the next two options. Click 'data retrieval and analysis', select the experiment, and continue to filtering. Select '$\log_2$ of R/G Normalized Ratio (Mean)'. Filter using Flag = 0 (denotes good spots as identified by masking) and Failed = 0 (genes on array whose PCR products were confirmed by gel electrophoresis). Click retrieve data. It will be displayed in the browser; save as a tab-delimited file and open in Excel.

Statistically analyzing normalized data

We performed a Student's t-test in Excel to ask if the null hypothesis (that the $\log_2$ ratio experimental/baseline = 0). To summarize, we first calculated the mean $\log_2$ ratio,

the standard deviation, and the t-statistic (divide mean by standard deviation). Determine n using script (=SUM(ISNUMBER(Cell 1) + ISNUMBER (Cell 2)…). Generate a P-value using the following script: =IF (T-stat>0, TDIST (Tstat, n-1, 2), 1-TDIST (-Tstat, n-1, 2)). Sort via confidence limits using: =IF (P-value $\leq$ 0.0005, 1, 0) and =IF (P-value $\leq$ 0.005, 1, 0).

Results

mRNA isolated from wildtype, *unc-4*, and *unc-37* mutants

mRNA was isolated from three independent batches of synchronized mid-L2 larvae per genotype (N2, *unc-4, unc-37*). We chose this timepoint because UNC-4 function is required between the L2 and L3 stages to maintain synaptic inputs to VA motor neurons (Miller et al. 1992). We therefore reasoned that the target genes involved in synaptic specificity should be expressed during this period in development. RNAs were evaluated on an Agilent bioanalyzer to check quality. All but one (N2 #3) sample had the expected profile: two large rRNA bands and a broad distribution of mRNAs (Figure 3.1). A fourth extraction from wild-type animals was performed to replace N2 #3, which was degraded.

mRNA samples hybridized to *C. elegans* spotted arrays

To identify UNC-4 target genes, I hybridized three N2 samples vs. three *unc-4* or *unc-37* samples (Table 3.1). Most of the array hybridizations worked well and had low background (Figure 3.2A). The *unc-37* hybridizations were repeated due to a high level

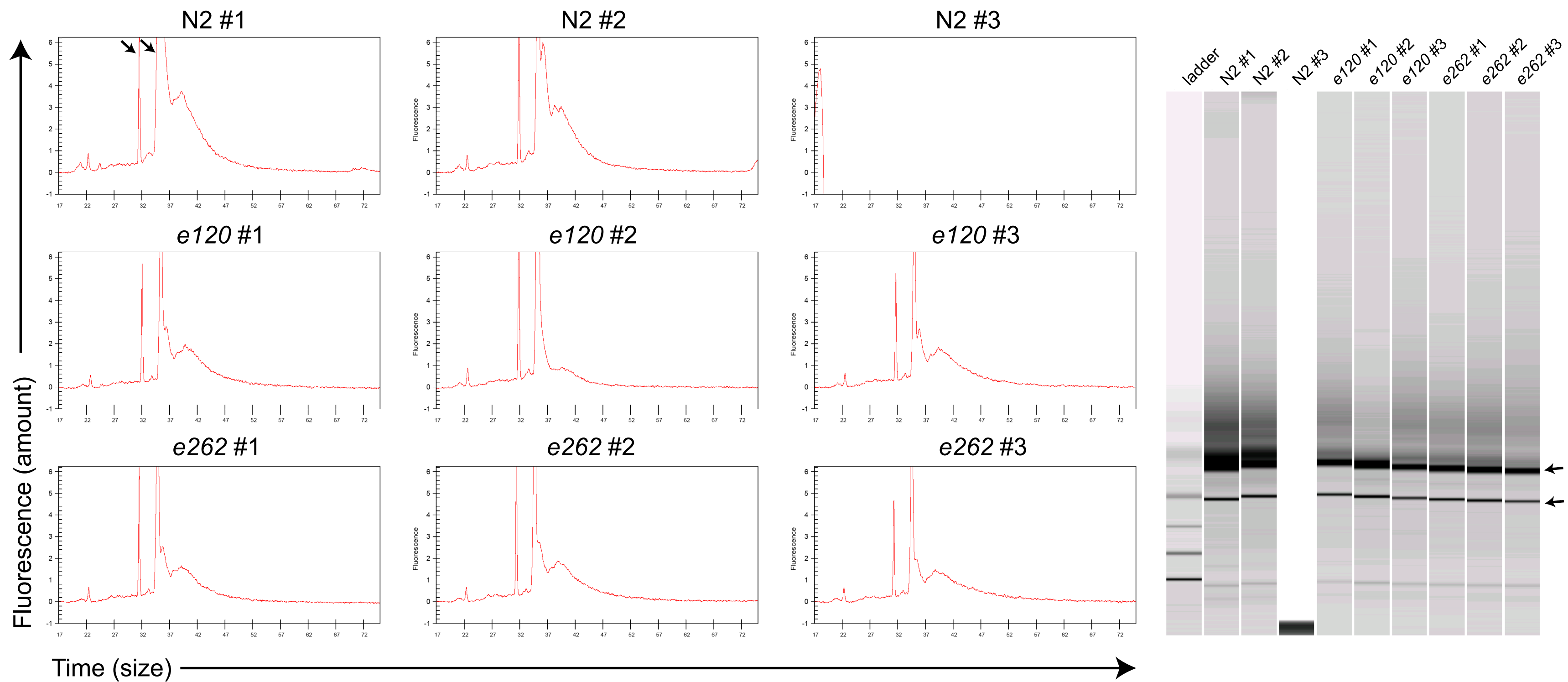

Figure 3.1 Bioanalysis results from mRNA extracted from N2, *unc-4* and *unc-37* staged L2 preparations.
Left. Graphs represent amount (fluorescence) and size (time) of RNA. All samples show a similar pattern, with highly abundant rRNA bands (arrows) and a broad distribution of less intense mRNA bands. Sample N2#3 is degraded.

Right. A representation of the graphs in 'gel' form. In this view it is easy to distinguish the rRNA bands (arrows) and the smear of mRNA.

of background and/or loss of spotted DNA on part of the array (Figure 3.2C).

A close-up of one block of the array shows that most spots are yellow, while others are green or red (Figure 3.2B). The red and green spots are indicative of genes for which transcript levels differ between the two samples hybridized to the array. The scanned arrays were masked and the raw data loaded into the Stanford Microarray Database (SMD) for normalization (see Materials and Methods).

Analysis of microarray data

I downloaded normalized, filtered data from SMD (see Materials and Methods). Josh Stuart, a statistician in Stuart Kim's lab, assisted in the analysis. He tested the arrays for correlations to see if biological replicates were comparable (e.g. *unc-4* vs. N2). A correlation of 1 is a perfect match. Josh found that the data are moderately correlated. However, N2 vs. *unc-4* repeat #3 did not match with the other two *unc-4* repeats; it seemed as if the dye was switched from what was initially annotated. In addition, Josh's analysis suggested that N2 sample #4 appears to be an outlier compared with the other 2 N2 samples. I therefore decided to remove any hybridizations performed using this N2 mRNA (*unc-4* #3, *unc-37* #3).

I averaged the $\log_2$ ratios (mutant intensity/wildtype intensity) between two *unc-4* hybridizations and five *unc-37* hybridizations to identify transcripts which were significantly elevated in both mutant backgrounds. I then used Student's t-test to ask if these mean $\log_2$ ratios were significantly different from 0 (which would happen if the intensities between mutant and wildtype were almost identical). We identified 38 genes that were significant at the 99% limits ($P \leq 0.01$) and 10 at 99.9% ($P \leq 0.001$) (Table 3.2).

Table 3.1. Hybridizations performed on *C. elegans* chip.		
Name	**Cy5 labeled**	**Cy3 labeled**
unc-4 #1	N2 #1	*e120* #3
unc-4 #2	N2 #2	*e120* #2
unc-4 #3	N2 #4	*e120* #1
unc-37 #1	N2 #1	*e262* #1
unc-37 #2	N2 #2	*e262* #2
unc-37 #3	N2 #4	*e262* #3
unc-37 #1b	N2 #1	*e262* #1
unc-37 #2b	N2 #2	*e262* #2
unc-37 #3b	N2 #1	*e262* #3

There were 17,500 genes on the *C. elegans* chip at the time of these experiments. By chance, at the 99.9% limits 17.5 genes will be found significant and 175 at the 99% limits. Since the number of significant genes are fewer than the expected false positives (e.g. 10 vs. 17.5 at $P \leq 0.001$), and the fact that the three known target genes (*acr-5, del-1, glr-4*, see Chapter II) were not included in this group, we concluded that these microarray data were unlikely to reveal authentic *unc-4* target genes.

Discussion

Because forward genetic screens failed to identify downstream UNC-4 targets that could suppress the backward movement defect, we opted to take a genomics approach. The availability of DNA microarrays allowed us to take advantage of the fully sequenced genome. Our experimental design was simple: compare wildtype transcript levels to those from *unc-4* and *unc-37* mutant backgrounds; genes that were significantly upregulated in the mutants vs. wildtype would be considered candidate target genes. However, because *unc-4* is expressed in ~5% of cells at the L2 stage, the sensitivity of the experiment might not be sufficient to detect a 2-fold change in expression levels in such a small number of cells.

As evidenced by the data presented above, the experimental design was indeed insufficiently sensitive. None of the known *unc-4* targets genes (*acr-5, del-1, glr-4*) were upregulated in the mutants compared to wildtype. Furthermore, the number of genes found to be significantly different fell below the threshold calculated from the expected false positive rate, thereby indicating that these genes are likely noise.

However, the trip to Stanford was not a total loss. While there I spoke to Peter

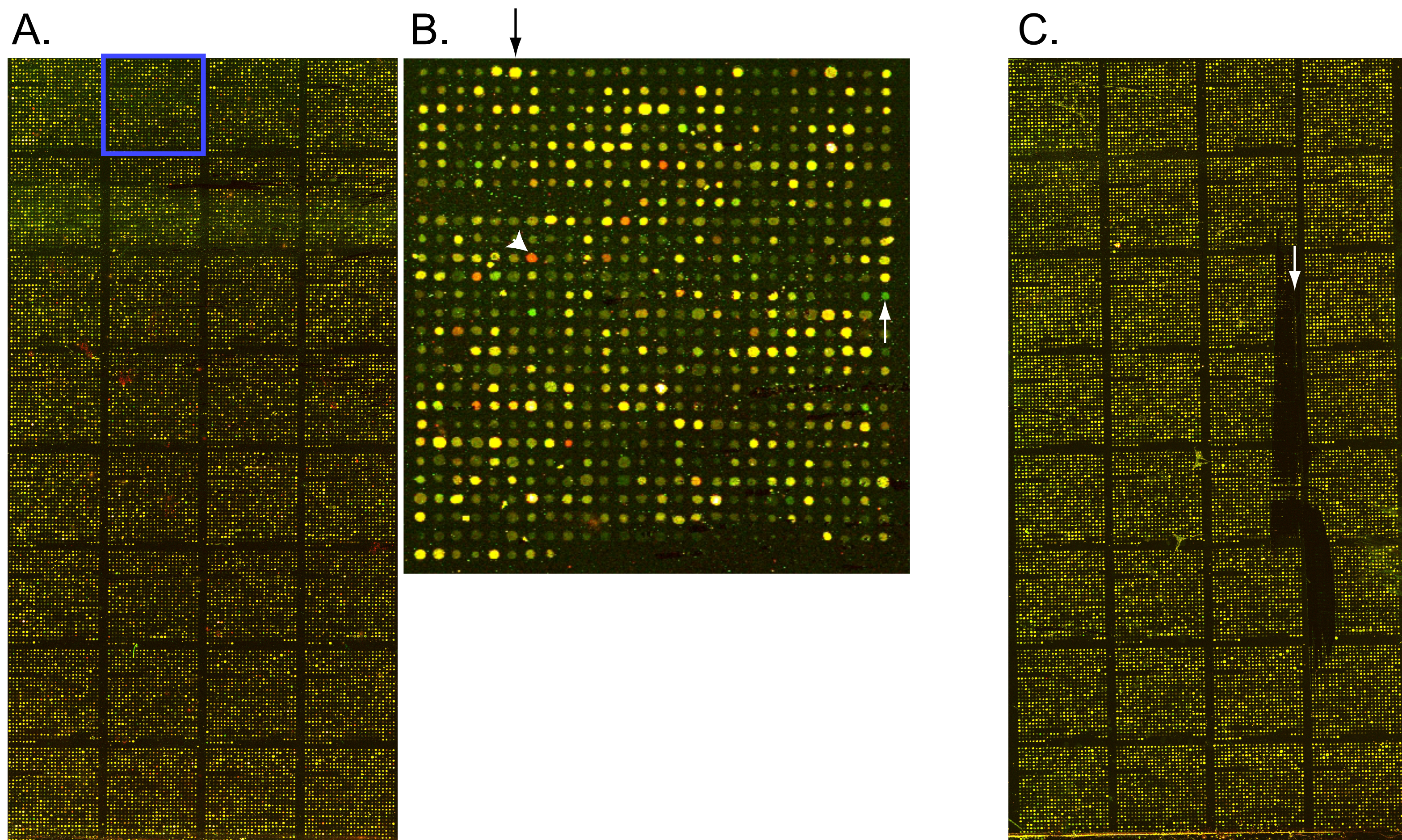

Figure 3.2 Representative images of scanned whole-animal micorarray experiments.
A. Example of a good hybridization, with low background and no smudges or lost spots.
B. Close-up of boxed area in A, showing yellow (black arrow), red (white arrowhead), and green (white arrow) spots.
C. Example of a poor hybridization. A large section of the chip has been scraped away (white arrow).

Table 3.2 List of genes significantly upregulated in *unc-4* and *unc-37* mutants.

Cosmid Name	Function	P-value
R11A5.7	Putative carboxypeptidase	4.2E-5
T07F10.1	Aminopeptidase	4.5E-5
ZK662.2	unknown function	8.1E-5
W03G1.7	Putative acid sphingomyelinase	1.4E-4
C06E7.4	unknown function	1.9E-4
F14H8.1	Oxysterol-binding (OSBP) protein	1.9E-4
B0213.15	P450 heme-thiolate protein	2.5E-4
T06D8.9	unknown function	4.4E-4
F47F6.3	G-protein coupled receptor	8.2E-4
F13B12.4	cysteine synthases	8.8E-4
K01D12.9	unknown function	0.001
F57H12.3	unknown function	0.001
C33E10.10	unknown function	0.001
F14H12.8	unknown function	0.002
C02E7.7	unknown function	0.002
ZC123.3	unknown function	0.002
K01A2.7	Collagen	0.002
R12E2.15	unknown function	0.002
Y56A3A.9	unknown function	0.002
M02G9.1	unknown function	0.002
4R79.3	unknown function	0.003
B0218.8	C-type lectin family	0.003
W04A4.3	unknown function	0.003
K02A4.1	Transaminase protein	0.003
F10D11.6	unknown function	0.003
Y116F11A.	unknown function	0.003
T05H4.7	Chitinase	0.004
K04G11.2	unknown function	0.004
C30H6.5	unknown function	0.004
F13C5.3	unknown function	0.004
T07C4.4	Putative antibacterial agent	0.004
PDB1.1	unknown function	0.004
T07C4.1	Orotidine-5'-phosphate decarboxylase	0.004
F33D11.3	Collagen	0.005
C10F3.3	Adenylyl cyclase	0.005
F45D11.O	unknown function	0.005
F54D5.15	unknown function	0.005
W05F2.3	unknown function	0.005

Roy about a new method he was developing to isolate cell-specific transcripts. He used an epitope tagged version of the poly-A binding protein, PAB-1, to isolate mRNAs from muscle cells (Roy et al. 2002). PAB-1 binds the poly-A tails of most mRNAs, thus capturing the transcriptome of a cell. As described in the next two chapters, I adopted this cell-specific strategy with the goal of comparing transcripts specifically from VA motor neurons in wildtype vs. *unc-4* and *unc-37* mutant backgrounds.