MQP-BIO-DSA-1045

Importance of *Caenorhabditis elegans* First Introns in Gene Regulation

A Major Qualifying Project Report

Submitted to the Faculty of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

in

Biology and Biotechnology

by

Amanda Kent

April 30, 2009

APPROVED:

Marian Walhout, Ph.D.
Program in Gene Function and Expression
UMass Medical School
Major Advisor

David Adams, Ph.D.
Biology and Biotechnology
WPI Project Advisor

ABSTRACT

The purpose of this project was to determine whether first introns in *C. elegans* genes have an important role in gene expression. Gene expression is regulated by transcription factors (TFs) that bind to DNA sites within each gene. The first introns of 14 genes were assayed by a yeast one-hybrid (Y1H) system to see if any TFs bind. The promoters of most of the 14 genes were also tested for TF interactions. Eight of the introns had TF interactors, suggesting that introns are places where TFs bind and are involved in gene regulation. Interestingly, two introns shared TFs with their own promoters. To validate the identified interactions, more tests will need to be performed *in vivo.* Since only a few first introns were studied, more will be studied further using the Y1H to make a general conclusion about their role in gene expression.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would first like to thank Marian Walhout for allowing me to work on my project in her lab at the University of Massachusetts Medical School, in the Program of Gene Function and Expression, and for providing an excellent learning experience. Thanks, to John Reece-Hoyes, for teaching me protocols, for being there every step of the way, for his valuable project discussions and input, and for helping me find the confidence that I needed. Thanks to Inma Barrasa for helping me with all the bioinformatics part of the project. I would also like to thank Efsun Arda, Natalia Martinez, Chris Grove, Ashley Carraher, and Katie Brown for their support in lab and willingness to answer questions and give a helping hand. Thanks to Susie for all the plates she had to pour. And finally I would like to thank Dave Adams for giving me support, helping with the report writing process, and advising my project.

BACKGROUND

REGULATION OF DIFFERENTIAL GENE EXPRESSION

The genome contains many protein-coding genes and RNA non-coding genes which are responsible for gene regulation. The expression of a particular set of genes in different cells at different times is called differential gene expression. The best studied mechanism in controlling differential gene expression is regulation at the transcriptional level (Lee and Young 2000), where transcription factor proteins (TFs) can directly (protein-DNA) or indirectly (protein-protein-DNA) bind to *cis*-regulatory DNA elements of target genes, resulting in activation or repression of target expression. These *cis*-regulatory elements are commonly found in promoters; the regions immediately upstream from the initiation start site, however they also occur in enhancers that may be located far up- or downstream from the target coding region, and sometimes within introns (Hartwell et al. 2004).

Most gene regulation occurs at the initiation of transcription, and when TFs bind, transcription rates can be altered either in a positive or negative manner (Hartwell et al. 2004). For example, TFs can recruit RNA polymerase to the promoter allowing basal levels of transcription, or different TFs can bind to different enhancer elements associated with different genes and interact with the basal factors at the promoter to cause an increase in transcription initiation. Transcription activity can also be repressed by TFs competing and binding to enhancers which can block activator proteins. Identification of interactions between proteins and DNA can contribute to a better understanding of gene expression.

INTRONS AND GENE REGULATION

The human genome's length is 2.9 Gb containing ~27,000 protein-encoding genes (International Human Genome Sequencing Consortium, 2004). The *C. elegans* genome is smaller in length, being 97 Mb with ~19,000 protein encoding genes (The *C. elegans* Sequencing Consortium, 1998). Sequencing many species genomes has revealed that increasing organism complexity is accompanied by a larger non-coding part of the genome. Non-coding DNA is space between genes as well as introns. Introns are removed by splicing to produce mRNA, and were once thought to be irrelevant in the activity of gene expression. The cell puts a tremendous amount of wasted energy into the synthesis of introns into the primary transcript, only to be removed later by splicing. Introns are known to protect the genetic makeup of an organism from genetic damages by outside influences (ex. Radiation) (King, 2009). Introns are important for alternative splicing as they have binding sites for the splicing protein. Alternative splicing can increase the genetic diversity of the genome without increasing the overall number of genes.

With the support of modern research, scientists have discovered many more possible functions of introns, including gene regulation and structural roles. A common characteristic in eukaryotic structure is a longer first intron compared to non-first introns. Figure 1 shows the results of an analysis conducted with the data from GenBank. There is a clear trend of increased length in the first intron for a variety of species (Bradnam and Korf, 2008). One possible reason for this trend is that the first intron could have binding sites for TFs that control expression of the gene. Shown in the Figure below as a

yellow dot, the first intron is not as relatively large for *C. elegans*. In fact, introns in *C. elegans* are often found to be fewer and smaller than many other organisms.

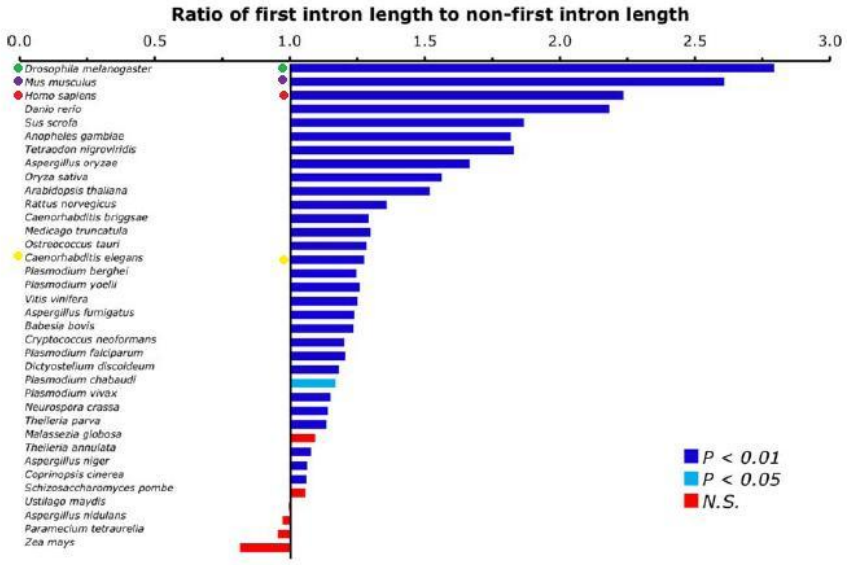
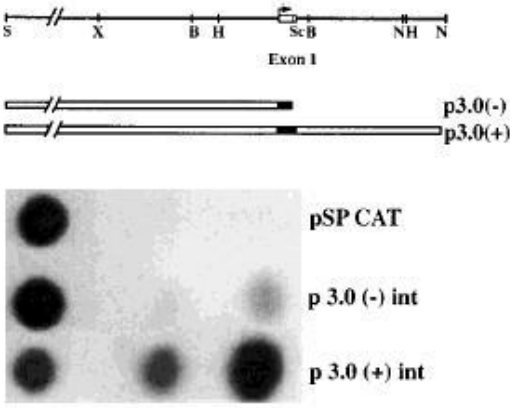


Figure 1. First Introns are the Longest Introns in Most Species. *C. elegans* (yellow dot), human (red dot), *Drosophila* (green dot) and *Mus musculus* (purple dot) are highlighted (Bradnam, and Korf, 2008).

As stated earlier, differential gene expression can be regulated at the transcription level, and since introns encompass a huge part of the genome and are known for having a longer first intron, they could possibly contribute to expression. A specific example of this phenomenon is the discovery that the large first intron in the human *TGM1* gene upregulates transcriptional activity (Polaowska et al. 1999). To determine mechanisms controlling transcriptional activity for *TGM1* during keratinocyte differentiation, the 5' end region was mapped by a CAT reporter assay (Figure 2). The assay was conducted in human keratinocytes, and the first intron increased expression of the CAT reporter concluding that an important cell-type specific transcription regulator lies within the intron.



(Polaowska et al. 1998).

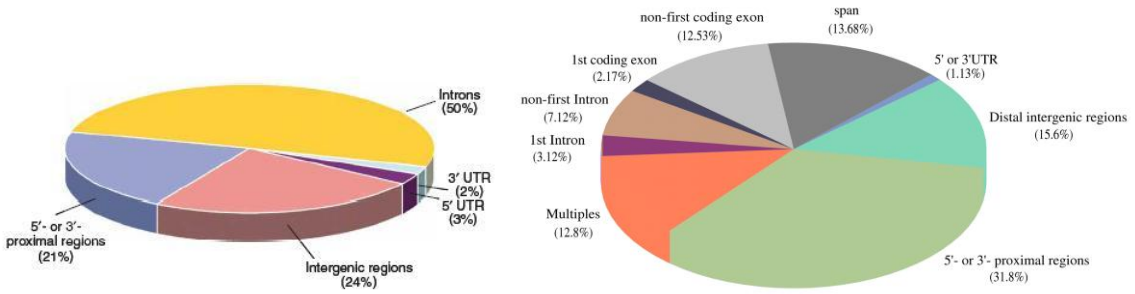
Figure 2. **CAT Assay Driven by Intron-1 of the Human TGM1 Gene.** Cat activity driven by intron-plus and intronless TGM1 in undifferentiated human keratinocytes.

Some more examples of how a gene's intron-1 can regulate expression identify binding sites in the first intron of the *FGFR4* gene using a DNase I hypersensitivity assay (Shah et al. 2002). The TFs Sp1 and HNF1 bind to the *FGFR4* enhancer in the first intron, and the HNF1 protein was confirmed to have a direct regulatory role in *FGFR4* gene activation. Another member of the fibroblast growth receptor family, *FGFR3*, has also shown the importance of the first intron; Sp1 and Sp3 proteins have been identified within the first intron to enhance transcription (McEwen et al 1998), just as the first intron in *FGFR4*. Some more interesting results have shown that the removal of the first intron of *FGFR4* leads to a greater than 85% decrease in transcriptional activity, emphasizing the idea that introns do play a role in gene regulation.

Some introns have also been found to repress transcriptional activity. Transcriptional regulation is essential for controlling *EGFR* overexpression in breast cancer cells and the first intron has been identified to be a negative regulatory element in this activity (Wilson et al. 2002).

To address TF binding in introns generally, genome-wide analyses have determined that in humans, DNase I hypersensitive sites are enriched in introns (Sabo, et al 2006), but in *C. elegans* hypersensitive sites are depleted in introns (Shi et al. 2009). DNase I is an enzyme that cuts naked DNA. When proteins bind to DNA, they protect the DNA from enzymatic cleavage. Genome-wide DNase I assays reveal where nucleosomes are bound to DNA. The DNase I cuts in nucleosome-free regions that are therefore available for TFs to bind. So DNase I hypersensitive sites indicate where TFs are able to bind (Brenowitz M, et al. 2001). Figure 3 shows that fifty percent of the genomic distribution of DNase I hypersensitive sites in humans is introns (Fig. 3a, yellow), and in *C. elegans* about one fourth of DNase I hypersensitive sites were found in introns (Fig. 3b, first intron + non-first intron + span [of exon and intron boundaries]).

a. **b.**



(Sabo, PJ. et al. 2006) (Shi et al. 2009)

Figure 3. Genomic Distribution of DNase I Hypersensitive Sites. Above is a comparison of the percentage of DNase I hypersensitive sites that can potentially bind transcription factors in (**a.**) humans and (**b.**) *C. elegans*.

Despite this genome-wide trend, evidence of TFs binding in the first intron have also been found in *C. elegans.* Chromatin immunoprecipitation (ChIP) experiments in *C. elegans* embryos identified an enhancer located in first intron of the *hlh-1* gene (Figure 4). Experiments found that direct binding of the TF PAL-1 to this enhancer allows for

activation of the body wall muscle regulator (Lei and Liu, 2009). An enhancer element has also been found to contain regulatory sequences necessary for *lim-7* gonadal sheath expression in the first intron (Voutev and Keating, 2008). And in the *unc-54* gene the first intron augments expression levels in body wall muscles (Okkema, et al 1993).

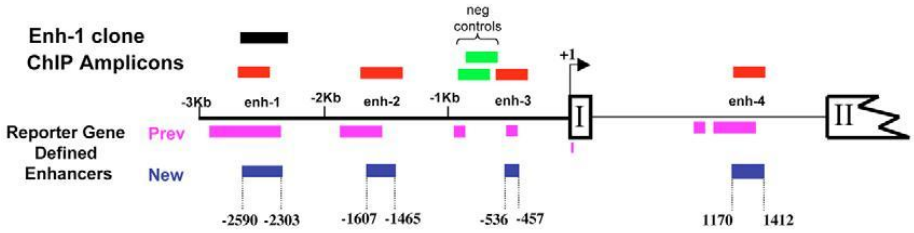


Figure 4. Diagram of the *C. elegans hlh-1* Enhancer Region in Intron-1 Showing Transcription Factor Binding Sites. The regulatory region of *hlh-1* is shown with bodywall muscle enhancer regions (blue boxes) and newly defined muscle enhancers determined by ChIP (red boxes). (Lei and Liu, 2009)

In some instances, introns have been found to have no role in gene expression. In *C. elegans* the first intron of *dat-1* and *cat-2* have been shown to play no role in expression (Flames and Hobert, 2009). But based on the previous studies discussed earlier showing that some TFs do bind to the first intron, a pilot study will be performed in this project on a small set of *C. elegans* first introns to identify protein-DNA interactions to see if any *C. elegans* first introns contribute to gene regulation.

DETECTING DNA-PROTEIN INTERACTIONS

Many methods have been previously used to detect protein-DNA interactions. Some can detect interactions *in vitro*, such as the electophoretic modify shift assay (EMSA). In this assay, DNA and proteins are mixed together and run on an agarose gel. The DNA bound by protein is heavier and migrates slower on a gel than free DNA (Taylor et al. 1994). *In vivo* methods can provide more valid approaches to detect interactions since

the organism itself is tested. Chromatin immunoprecipitation (ChIP) with microarrays is a common approach used to identify protein-DNA interactions *in vivo* (Durand-Dubief and Ekwall, 2009). ChIP can characterize the interactions of individual TFs with one or a couple of their target genes by using specific antibodies that recognize a specific TF. The assay is initiated by formaldehyde-mediated cross-linking of DNA and proteins *in vivo*, followed by cell lysis, immunopurification of specific TF-DNA complexes, and then PCR amplification of the DNA fragments bound by the TF. The amplicons are labeled and hybridized to a genome tiling microarray to reveal the fragment sequences. The major problems with this ChIP microarray assay are that individual antibodies need to be generated for each TF, and it cannot distinguish between direct protein-DNA interactions and indirect protein-protein-DNA interactions, it can only detect strong and ubiquitous interactions.

The yeast one hybrid (Y1H) system used in this project, is another tool used to help understand gene expression by identifying protein-DNA interactions *in vivo*. This assay is gene-centered, meaning it identifies proteins that bind to DNA elements of interest, as opposed to protein-centered networks that identify DNA elements binding to proteins of interest. In this system (Figure 5), a single hybrid protein is generated with a TF (prey protein, purple in Figure 5) fused to a transcription activation domain (AD, red in Figure 5). When the prey interacts with the DNA bait (for example an intron-1 sequence being tested) (blue in Figure 5), the AD activates the reporter gene expression (pink in Figure). Since several TFs contain an activation domain while others do not, to identify a variety of DNA-binding proteins, a strong heterologous AD is added to the

prey hybrid protein so interactions can be detected when the bait and preys are mated (Fields and Song, 1989).

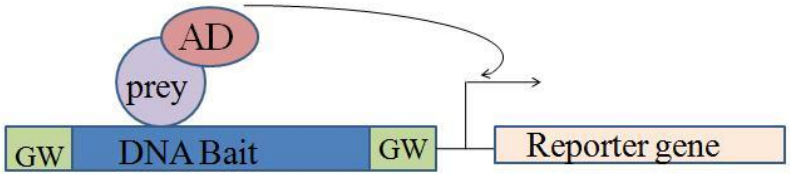


Figure 5. Y1H System with Gateway Recombination Sites (GW). In this assay, a known transcription factor prey (purple) fused to a transcription activation domain (AD) interacts with a binding site in the tested DNA sequence (intron-1 for example) (blue) to transactivate the reporter gene (diagram right).

The two reporter genes *HIS3* and *LacZ* are used in the Y1H system (not shown in Figure 5), and DNA baits are fused to these two reporter genes to form the constructs DNA bait::*HIS3* and DNA bait::*LacZ* that are integrated into the yeast genome. Having two reporters that both must be activated to demonstrate an interaction improves confidence that the interaction occurs, because the TF binds within the DNA bait. Integrating the reporter constructs into the yeast genome ensures that every yeast in the array has the same number of copies of each reporter.

The advantages of the yeast one-hybrid system (Y1H) over ChIP are that it does not require antibodies, and interactions by TFs expressed at low levels or only in a few cells can be identified. The Y1H system unfortunately can only find interactions of one protein binding at a time, so heterodimers are not detected (Field and Song, 1989). Many of the Y1H interactions identified can subsequently be validated when tested *in vivo;* common *in vivo* methods include ChIP and GFP fusion. For example, protein-DNA interactions have been mapped (using Y1H) in a network found to be enriched in TFs expressed in the digestive tract (Deplancke et al. 2006). The interaction found between TF DAF-3 and *Pmdl-1*suggested that *Pmdl-1* was a direct DAF-3 target, and a ChIP

assay with an anti-DAF-3 antibody was performed to confirm this interaction. GFP expression showed that DAF-3 (which promotes *C. elegans* dauer formation) is localized to the *mdl-1* promoter, and represses *mdl-1* gene expression in the pharynx when worms enter the dauer stage.

In another instance, a network of miRNA ↔ TF feedback loops (TFs that bind a miRNA promoter can be regulated by the same miRNA) named "high flux capacity" can identify the high information flow for the TF target and miRNA that are regulated (Martinez et al., *Genes & Development* 2008). One network suggested that *mir-43* and LIN-26 could be co-expressed in the same tissue. Therefore transgenic animals were created for *Pmir-42-44*::GFP, and it was found that *Pmir-42-44* drives GFP expression in embryos and throughout development. Because expression was detected in hypodermal seam cells of larval stages, it was confirmed that LIN-26 and *mir*-43 are coexpressed.

GATEWAY-COMPATIBLE YEAST ONE-HYBRID SYSTEM

The traditional Y1H system is usually used with multiple copies of short DNA fragments, but not with longer DNA fragments because the system is based on restriction-enzyme-based cloning methods, and longer fragment are more likely to have cutting sites that interfere with cloning (Walhout 2006). The Gateway recombination system (Figure 6) uses enzymes that transfer DNA fragments by recognizing 30 bp flanking sequences called "tails" that are virtually never found with DNA fragments to be cloned. The Gateway cloning system allows for rapid cloning of multiple DNA fragments of any size in parallel. Therefore, a gateway compatible Y1H system allows

rapid large scale identification of protein-DNA interactions using both small (DNA elements of interest) and large (promoters) DNA fragments (Deplancke et al. 2004).

Figure 6. The Gateway Cloning Destination Constructs. LR cloning; DNA bait entry clones and Y1H reporter clones are fused to form DNA bait:: *HIS3* and DNA bait::*lacZ* Destination clones flanked with gateway ends (B) (Deplancke et al. 2004).

The *HIS3* Destination construct includes the DNA bait, ampicilin (amp) resistance, origin of replication (ORI), and *HIS3* ORF. The *LacZ* Destination construct contains DNA bait, amp resistance, ORI, a wild type *URA3* ORF, and *Lac Z* ORF (Figure 6). Activational domain (AD)/ TF fusion constructs (including amp resistance and *TRYP1* ORF) are transformed into a different yeast mating strain than the DNA baits. The integrated baits and AD yeast clones are mated, and interactions are observed if the AD (containing TFs) binds to the DNA baits to activate the reporter; blue colonies appear on the Xgal plate (*LacZ* encodes β-galactosidase) and larger colonies appear on the 3 – Aminotriazole (3AT) plate. 3AT is a competitive inhibitor of the HIS3 enzyme; therefore more *HIS3* needs to be expressed for the yeast to grow (Deplancke et al. 2004). The *HIS3, URA3 and TRYP1* ORFs evoke selection of yeast that contain these plasmids.

MARIAN WALHOUT LAB

For a better understanding of differential gene expression at the genome-wide level, transcription regulatory networks need to be mapped and investigated. The overall goal of Marian Walhout's lab in the Program of Gene Function and Expression at UMass Medical School (Worcester, MA) is to map and characterize gene regulatory networks in *C. elegans*, and to understand how these regulatory circuits control animal development, function, and homeostasis. The Walhout lab has used Y1H assays to generate transcription regulatory networks of genes expressed in the digestive tract (Deplancke et al. 2006) and neurons (Vermeirssen et al. 2007) and also of miRNA (Martinez et al. 2008). To do this they have cloned 800 of 940 predicted *C. elegans* TFs (Reece-Hoyes et al. 2005; Vermeirssen et al. 2007) that can be tested individually in parallel whether they bind a DNA bait. They have recently purchased a Singer RoToR HAD robot that is able to generate plates with 1536 features which will increase the throughput of the Y1H screens.

PROJECT PURPOSE

The first intron of a small set of *C. elegans* genes will be tested by the Y1H system to see if the introns interact with known transcription factors. Where possible, the gene promoters, located immediately upstream from the initiation start site, will also be assayed by Y1H. The first introns of 48 genes were chosen from a *C. elegans* database based on size, number of TFs confirmed to bind to the gene promoter, and whether ESTs were present to confirm the intron. TFs are commonly found in promoters so it will be interesting to see 1) if any TFs bind the introns, and 2) if any TFs are shared by the promoter and first intron.

METHODS

GENE SELECTION

The online database WormBase release WS180 (www.wormbase.org) was used to help determine genes and intron sequences for subsequent analysis. First intron sequences containing clearly demarked 5' and 3' splice sites were chosen for approximately 48 genes as primary candidates. Bait strains containing the promoter were already made in the Walhout lab.

PRIMER DESIGN

Forward and reverse PCR primers were designed for the first intron of the 48 selected genes. All primers were designed to have gateway ends (attB4 tails for forward primers, and attB1R for reverse primers compatible with pDONR P4-P1R vector) and their sequences are shown in the Appendix.

VECTOR GENERATION

GENOMIC DNA PCR

First intron sequences were amplified for the 48 chosen genes using polymerase chain reaction (PCR) with *C. elegans g*enomic DNA (isolated in lab) as template. The PCR mix contained 39.8 µL of autoclaved (AC) water, 5.0 µL of 10X $MgCl_2$ Buffer (Invitrogen), 1.5 µL of 50 mM $MgCl_2$ (Invitrogen), 0.5 µL 10 mM dNTPs (Invitrogen), 2.0 µL of 200 µM primer, 0.2 µL of Platinum Taq (Invitrogen), and 1 µL (containing 15 ng/ul) of DNA. The thermocycler protocol was as follows: 94°C for 3 minutes initial

denaturation, followed by 34 cycles of 94°C for 30 seconds for denaturation, 56°C for 30 seconds annealing, followed by 2-3 minutes, depending on the size of the insert at 72 °C for elongation. The cycling steps were followed by a 72°C final elongation step for 5 minutes, and then stored at 4°C.

BP GATEWAY CLONING

For each intron a BP gateway recombination reaction was performed between the PCR amplicons and a pDONR P4-P1R vector (see Appendix) to generate an "Entry vector". 2 µL of the PCR product (approx. 150 ng), 2 µL of pDONR P4-P1R vector (50 ng/µL), and 1 µL of BP Clonase II enzyme (Invitrogen) were mixed and microcentrifuged. The reactions were then incubated at 25°C for at least 12 hours. Next, each BP reaction was transformed into 50 µL of *E.coli* DH5α competent cells (transformation efficiency of >1.0 x 10^8 tranformants/µg), and plated onto LB + 50 ng/µL of kanamycin (Kan) plates. The plates were incubated overnight at 37°C. Each recombination reaction typically produced 100 colonies, of which 4 colonies were selected to be screened for positive clones by PCR.

The colony PCR was a 50 µL reaction containing 44 µl AC water, 5 µL RITA buffer (100 mM TrisHCL, 500 mM KCL and 22.5 mM $MgCl_2$ with pH 8.0), 0.5 µL 10 mM dNTP mix, 0.1 µL M13F and M13R primer (see Appendix), and 0.4 µL of Platinum Taq (Invitrogen). Each colony was picked with a disposable tip, some transferred to a new Kan plate which was incubated at 37°C for 24 hours, and the remainder put in the PCR reaction. The PCR was performed as described above, with an annealing temperature at 56°C for 30 seconds. If the amplicon band size was correct, the PCR

product was sent for sequencing verification. For each intron two putative collect colonies were put into 3 mL of SOC medium supplemented with 3 µL of 50 mg per ml of Kan, and incubated at 37° overnight to be purified by minipreps (QIAGEN) the next day. Glycerol stocks (final 20% v/v) for each colony were made and stored at -80°C. If the sequencing verified correct insert, the miniprep DNA was used as the Entry vector for the LR reaction.

LR CLONING OF DNA BAITS INTO Y1H DESTINATION VECTORS

LR Gateway recombination reactions transferred each intron from the Entry vector to the two Destination vectors; pDEST-*HIS3* and pDEST-*LacZ*, which places each intron upstream of the HIS3 and *Lac Z* reporter genes respectively. 2 µL of the Entry vector (approx. 150 ng/µL), 2µL of either Destination vectors (50 ng), and 1 µL of LR Clonase II enzyme (Invitrogen) were mixed and microcentrifuged. The reactions were then incubated at 25°C for 5+ hours. Next, each LR reaction was transformed into 50 µL of DH5 α competent *E.coli* cells. 2 mL of SOC + 3 µL 50 mg per ml of ampicillin were added and incubated at 37°C overnight. The DNA was purified by using a miniprep protocol (Qiagen) and analyzed on a gel and its DNA concentration determined by spectrophotometry. Only Destination clones with a concentration of at least 200 ng/µL were used for subsequent integration.

YEAST BAIT GENERATION

LINEARIZATION

Once the introns were cloned into Destination vectors, they were ready to be integrated into the genome of the YM4271 strain of yeast. First the Destination clones were linearized by endonuclease digestion. Restriction enzymes (New England BioLabs) AflII or Xhol were used for His clones, and Ncol or Apal were used for Lac clones. To perform the digest, 18 µL of the clone DNA, 2.5 µL of 10X NEB Buffer 4 (New England BioLabs), 2.5 µL of 10 mg/ml BSA protein (New England BioLabs), and 2.0 µL of the appropriate restriction enzyme were microcentrifuged. The digest was incubated at 37°C for 3 hours, and then to make sure it was cut to completion, analyzed on an agarose gel against the uncut DNA.

PREPARATION OF THE YEAST SUSPENSION

A fresh patch of yeast strain YM4271 (MATa, ura3-52, his3-200, ade2-101, ade5, lys2-801, leu2-3, 112, trp1-901, tyr1-501, gal4D, gal8D, ade5::hisG)(Clontech #K1603-1) was prepared by streaking a YEPD plate and incubating at 30°C overnight. Yeast suspensions were prepared in YEPD liquid with an OD_{600} nm of 0.15-0.20 in glass flasks. The flasks were incubated at 30°C with shaking at 200 rpm for about 4-5 hours, until the OD_{600} was between 0.40-0.60. The yeast suspension was transferred to multiple 250 mL tubes and centrifuged at 2000 rpm for 3 minutes, the supernatant was discarded, and the cell pellet was resuspended with 25 mL of AC water per initial 250 mL tube. The suspension was centrifuged for 3 minutes again, the supernatant was discarded again, and the pellet was resuspended in 25 mL of TE/Lithium Acetate (LiAC) (0.1 M Tris-HCL

[pH 8.0], 10 mM EDTA, 1.0 M LiAC, fill up with sterilized water) per initial 250 mL tube. All cell suspensions were combined into one 250 mL tube and centrifuged for another 3 minutes and the supernatant discarded. The final yeast pellet was resuspended with a volume of TE/LiAC equal to 1/200th of the initial YEPD suspension and a volume of denatured salmon sperm DNA (10 mg/mL Invitrogen) equal to 1/200th of the initial YEPD suspension.

INTEGRATION

For each integration reaction, 200 µL of yeast (TE/LiAC, salmon sperm DNA) suspension, 20 µL of each Destination clone, and 1200 µL of TE/LiAC/PEG buffer (0.1 M Tris-HCL [pH 8.0], 10 mM EDTA, 1.0 M LiAC, fill up with 50% [w/v] PEG) was mixed in an eppendorf tube, incubated at 30°C for 30 minutes, then incubated at 42°C for 20 minutes. Following incubation, the mix was microcentrifuged for 3 minutes at 2000 rpm to pellet the cells. The supernatant was siphoned to waste, and 1 mL of AC water was used to resuspend the pellet. The suspension was then microcentrifuged for another 3 minutes, supernatant discarded, and 1300 µL of AC water was used to resuspend the pellet. The resuspension was then spread on a Sc-U-H plate (for double integrants) using glass beads. The integration plates were incubated at 30°C for 5 days.

SELF-ACTIVATION

Self activity is phenomenon where reporter genes are activated in absence of the TF. Independent integrants often show different levels of self activity. The integrant used in Y1H assays need to have lowest possible self activity so that the reporter activity

driven by the TF can be observed above this background. Up to 12 double integrants were transferred to a fresh Sc-U-H plate to be tested for self-activation. The self activity of the *HIS3* reporter was tested by replica plating to Sc-U-H + 40 mM, 20 mM and 10 mM 3AT plates, and growth observed during incubation at 30 °C for five days. The self activity of the *LacZ* reporter was tested by a filter array. The colonies were replica plated to a YEPD plate with a nitro cellulose (NC) membrane overlay, and incubated at 30°C overnight. Two whatman paper filters were placed in an empty 15 cm petri dish and a mix of 6 mL Z-buffer (60 mM Na_2HPO_4 [anhyd.], 60 mM NaH_2PO_4, 10 mM KCL, 1 mM $MgSO_4$. pH adjusted to 7.0 with 10 M NaOH), 11 µL β-mercaptoethanol (SIGMA) and 100 µL of 4% X-gal (w/v dimethylformamide, SIGMA) was added. The NC filter with the yeast colonies was frozen in liquid nitrogen for 10 seconds then placed, yeast side up onto the whatman filters to thaw. The lysed yeast were incubated at 37°C over night so that any *LacZ* protein expressed would generate a blue compound from the clear yeast.

BAIT CONFIRMATION

Before using the bait yeast in the Y1H assays the integrated *LacZ* and *HIS3* clones were checked by yeast PCR to confirm that the introns were cloned upstream of both reporter was the same. 15 µL of Zymolyase solution (2 mg/ml of Zymolyase in 0.1 M sodium phosphate buffer, pH 7.5) was added to an eppendorf tube, and a tiny amount of yeast (about the tip of a pipette tip) was added. The <ZYM> program (37°C for 30 minutes, 95°C for 10 minutes) was chosen on the thermocycler. After the thermocycler program ended, the lysate was diluted by adding 85 µL AC water. 5 µL of this dilution was used in the PCR reaction along with 39 µl AC water, 5 µL RITA buffer, 0.5 µL

10mM dNTP mix, 0.1 µL forward and reverse primer, 0.4 µL of Platinum Taq (Invitrogen). The primers for amplifying from the reporter clones were M13F and His232R for His and P1H1F and LacZ595R for Lac (see Appendix). If the bands were the correct size then the PCR product was sent for sequencing. If sequencing was confirmed, one yeast integrant for each intron was used for the YIH experiment.

THE TF ARRAY

The *C. elegans* genome has about 940 predicited TFs (Reece-Hoyes et al. 2005). For 800 of these TFs, open reading frames (ORFs) have been cloned into pDEST-AD vector tht express the worm TF fused to the activation domain (AD) of the yeast TF GAL4. All the AD-TF cloned have been transformed into the yeast strain Yα1867 (ura3-52, ade2-101, trp1-901, his3-Δ200) that will mate with YM4271. Glycerol stocks of all the Yα1867 transformants are stored in nine 96-well plates at -80°C with and empty AD vector transformed to Yα1867 in the H12 position. Using the Singer RoToR HAD robot, the Yα1867 transformants from four 96-well plates are plated in "quads" (four colonies that from a square) on a 1536-feature Sc-T plate. In this way, three array plates of TF yeast are used in the Y1H assay, the first with the TFs from the original 96-well plates 1 through 4, the second with TFs for plates 5 through 8, and the third with TFs from plate 9.

Y1H EXPERIMENT

On day one, each bait was streaked out to cover one quarter of a 15 cm petri dish onto Sc-U-H plate and grown at 30°C overnight. On day two, all the bait yeast was scraped from the Sc-U-H plate using a toothpick and resuspended in 300 µL of AC in an

eppendorf tube, then the suspension was spread onto a Sc-U-H plate using glass beads to make a lawn, and the plates were incubated at 30°C. Also on day two the Singer RoToR HDA robot was used to generate copies of each of the three transcription factor array plates onto Sc-T plates and incubated at 30°C . On day five, the robot was used to transfer both the bait yeast and TF prey yeast to a YEPD plate so that each "spot" of TF yeast was placed on top of a "spot" of prey yeast and the plated incubated at 30°C for 3 days. Diploids were selected on day 1 by transferring yeast from the mating plate to a Sc-U-H-T plate and incubated at 30°C. On day nine, the diploid yeast were transferred to the array plates: Sc-U-H-T + 3 mM, 5 mM, and 10 mM 3 AT plates, and one Sc-U-H-T + 5 mM 3AT + 160mg/mL Xgal plate and incubated at 30°C. Pictures were then taken of each assay on days 3, 5, and 7.

If a quad was "positive" (ie. blue on the Xgal plate, growth on 3AT plates above background) the identity of the interacting TF was confirmed with AD and Term primers that amplify from the pDEST-AD vector and subsequent sequencing of the PCR product. The identity of the bait was also confirmed by sequencing the PCR product generated from the same yeast lysate using the M13F and His232R primers.

RESULTS

The overall goal of this project was to determine whether the first intron of randomly chosen *C. elegans* genes influences gene expression. The Y1H system is a tool to find any DNA-protein interactions. As discussed in methods, 48 first introns were chosen from the *C. elegans* genome to see whether DNA-protein interactions occur. Due to time constrains, bait strains for only 14 of the 48 introns were generated and analyzed.

GENE SELECTION

Before beginning bench experiments, bioinformatic research was done. The program WormBase release WS180 (www.wormbase.org) assisted in finding an individual gene locus, and identifying its first intron. Figure 7 shows an example of this analysis for gene *lin-32*. The elegans differential gene expression database (EdgeDB) was also used; data obtained by the Walhout laboratory and by other laboratories were previously collected and made available on this site. 48 genes were chosen with the most TFs binding to the gene promoter from a list of promoters assayed by Y1H in the Walhout lab. Data was also collected regarding intron number, sizes, etc. (see Appendix), influencing how genes were chosen.

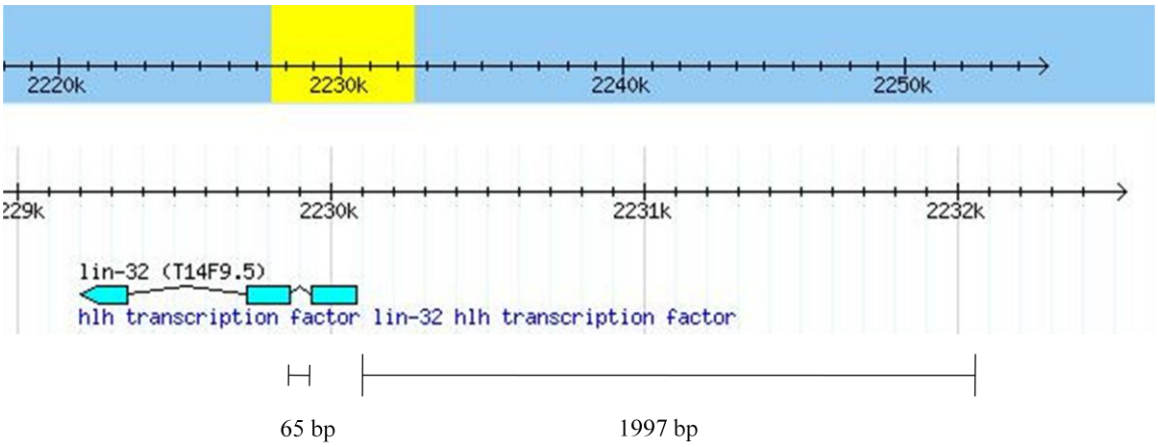


Figure 7. Example of *C. elegans* Gene Annotation for *lin-32*. Annotation taken form Wormbase (release 180). The first intron is 65 bp and the promoter is 1997 bp. The upper map denotes the entire locus. The lin-32 gene is shown in yellow. The lower map shows gene details. Exons are shown in turquoise boxes.

EXPERIMENTAL PIPELINE

Figure 8 shows a pipeline for the number of successful baits through each step of the process. Due to time constraints, only 14 of the 48 identified introns were integrated into bait strains for analysis. The remainder of the introns were transformed into the Destination vectors, and are stored as glycerol stocks at -80°C for future analysis.

PIPELINE

48 Genes and their first introns identified
↓
46 introns successfully amplified
↓
39 introns successfully BP cloned
↓
39 introns successfully LR cloned
↓
14 introns successfully integrated
↓
14 introns for Y1H assay
↓
8 Introns have Protein-DNA interactions

Figure 8: Experimental Pipeline. A flow chart of the number of successful baits through each step of the testing process.

CLONING AND RESULTS FOR *LIN-32*

Results for the first intron of *lin-32* (*FIlin-32*) at each step of the process will be shown and explained. The remainder of gel pictures for the baits can be found in the Appendix. A BP gateway recombination reaction was performed between the first intron PCR amplicons and a pDONR P4-P1R vector. The products were transformed into DH5α and plated on kanamycin plates (pDONR P4-P1R vector has kanamycin resistance), then four Km^r colonies were chosen for colony-PCR to screen for positive clones (Fig 9b). After being screened, the colonies were miniprepped, and the plasmid DNA was used as the entry vector for the LR gateway recombination reaction (Fig 9c-d). Constructs *FIlin-32*::*HIS3* and *FIlin-32*:: *LacZ* need to be linearized so they can enter the yeast YM4271 genome, so a digest was performed for each construct which was then run on a 1% agarose gel against its uncut LR miniprep DNA, to determine a successful digest (Fig 9e). To determine whether both *FIlin-32*::*HIS3* and *lin-32*::*LacZ* were integrated into the genome, the yeast was grown on a Sc-U-H-T plate and then screened by Yeast PCR (Fig 9f-g). The *HIS3* construct has histidine, the *LacZ* construct contains URA3 on its plasmid, and the yeast plasmid contains thymidine. Yeast PCR is performed with Primers M13 F and His232R (adds 410 bp) for *FIlin-32*::*HIS3,* and primers P1H1F and Lac595R (adds 960bp) for *FIlin-32*:: *LacZ.*

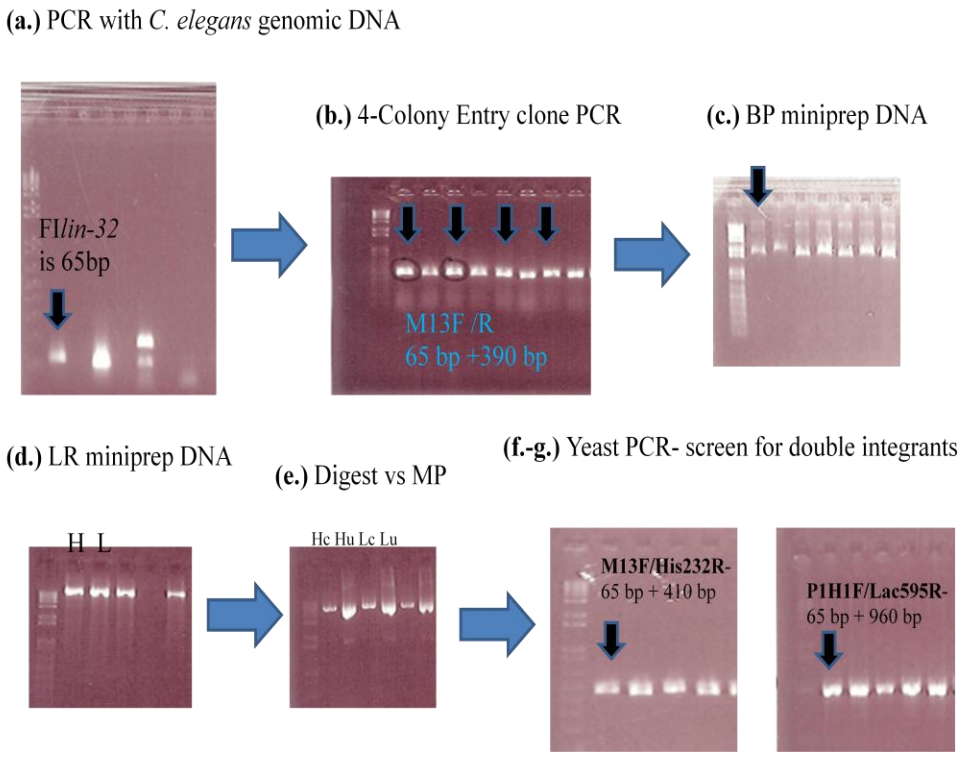


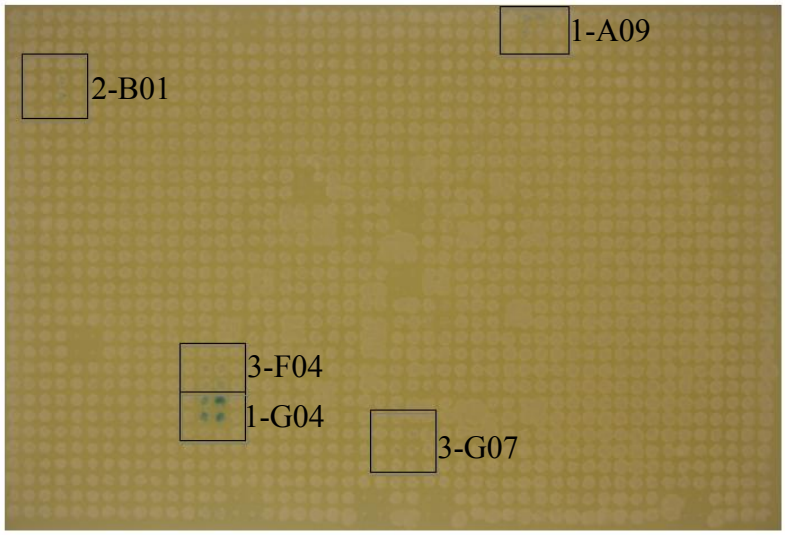
Figure 9. Results for Bait *FIlin-32.* (**a.**) The *lin-32* first introns were amplified by PCR. *FIlin-32* is 65 bp. (**b.**) Colony PCR was performed to verify Gateway BP recombination. Primers M13F and M13R add 390 bp to the intron (black arrows denote the four colonies analyzed). (**c.**) Miniprep DNA of BP Entry clone. (**d.**) Miniprep DNA of Desintaion clones His (H) and Lac (L). (**e.**) The cut *FIlin-32::His3* construct (Hc) and *FIlin-32::LacZ* construct (Lc) were run on a gel against the uncut constructs (Hu, Lu) to verify linearization. (**f.**) Yeast PCR was performed to verify double integrants. Primers M13F and His232R add 410 bp to the *FIlin-32*::*His 3* construct and (**g.**) Primers P1H1F and Lac595R add 960 bp to the *FIlin-32*::*Lac Z* construct.

Y1H ASSAY

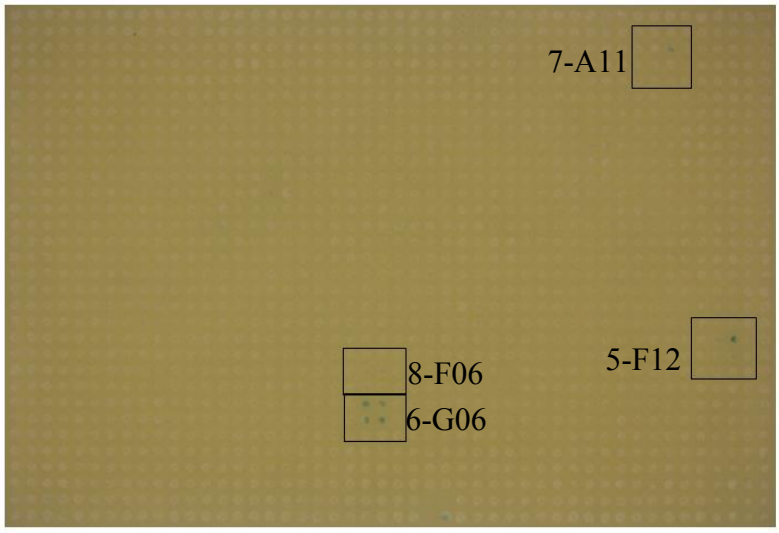
After the introns were integrated into the YM4271 yeast genome, a Singer RoToR HDA robot was used for the mating assay. An array of 800 *C. elegans* TFs, previously cloned into the AD fusion vector, and transformed into the Yα1867 yeast strain, was generated in 1536 format with each TF present in "quads" on three agar plates. The 1536 array was mated with each "bait" strain, and transferred to agar plates that detect the expression of the HIS3 (Sc-U-H-T +3mM, 5mM, and 10mM 3AT) and LacZ (Sc-U-H-T +5mM 3AT+X-gal) reporters that occurs when a TF interacts with the DNA bait.

Pictures were taken on days 3, 5 and 7 to observe the reporter expression. The three plates below (Figure 10) for *FIlin-32* are on day 7 on the Sc-U-H-T 5mMAT + Xgal. Plate 1 had 5 interactors (shown as boxes) (Fig 10a), plate 2 had 4 interactors (Fig 10b) and plate 3 had 4 interactors (Fig 10c). Interactor identity was confirmed by PCR from the AD vector (Fig 11b) within each positive yeast, along with PCR from the baits using primers M13F and His232R followed by sequencing of the PCR product (Fig 11a).

a.



b.



c.

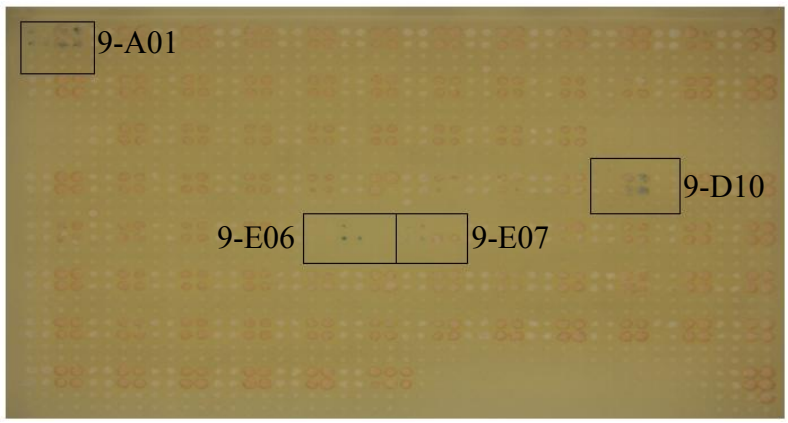


Figure 10. Example Mating Assay in 1536 Arrays for Day 7 of 5mM 3AT + Xgal plate of bait *FIlin-32*. (**a.**) Array plate 1 shows 5 interactions (boxed). (**b.**) Array plate 2 shows 4 interactions. (**c**) Array plate 3 shows 4 interactions. 14 first introns were screened by the Y1H assay and 8 had interactors.

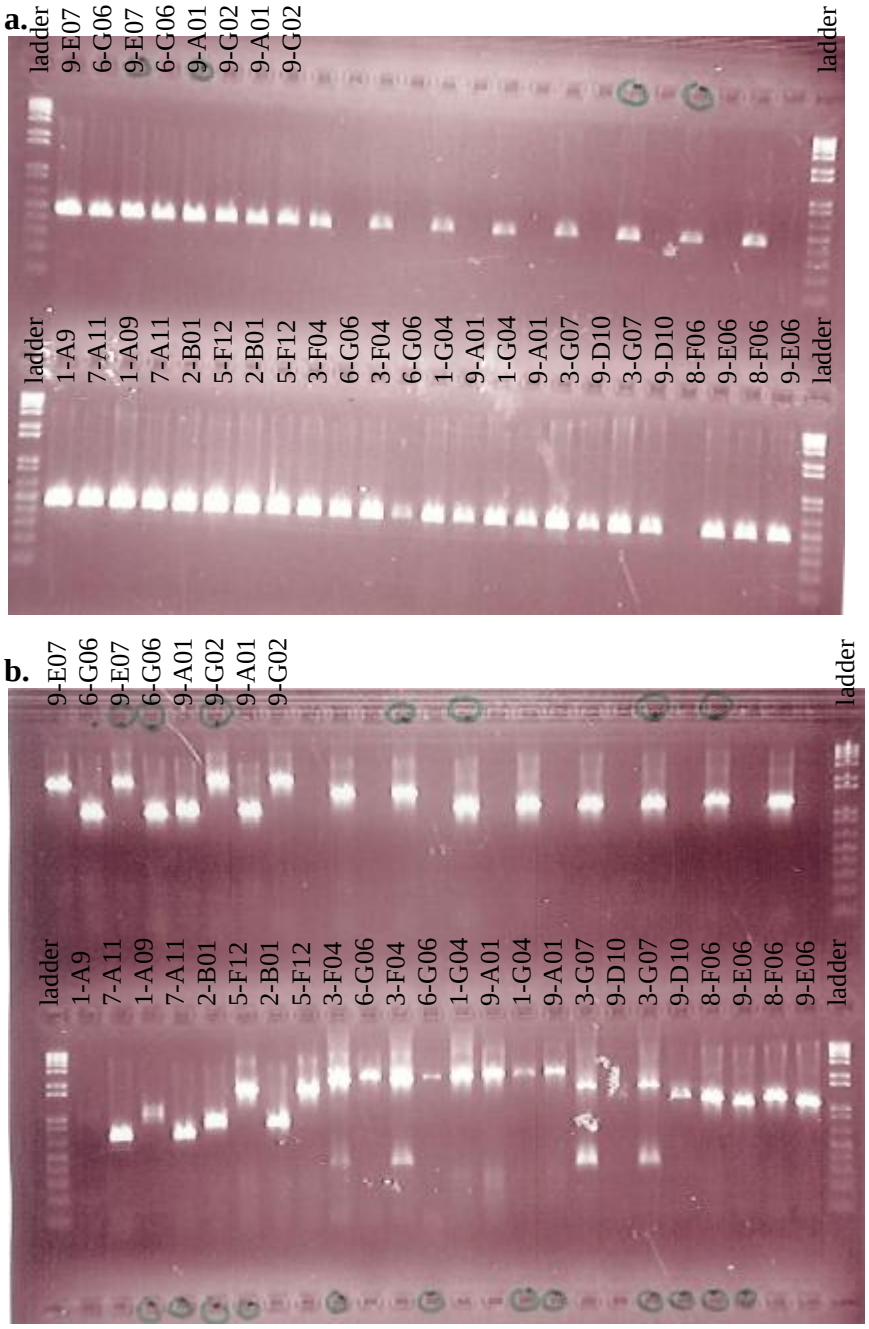


Figure 11. Yeast PCR to Screen the Success of the Integration. Yeast colony PCR was performed to verify that the TF is bound to the DNA bait. Primers (**a**) M13F and His232R were used to verify the DNA bait and (**b**) AD and TERM were used to verify the TF binding. 12 interactors are present in this gel for *FIlin-32.*

RESULTS SUMMARY

Overall, 14 first introns and 11 promoters were screened by the Y1H assay. All the interactors are shown in Figure 12. 8 first introns have interactors but 6 do not have any. Two TFs that interact with *C49C3.5* also interact with the promoter (highlighted in red). A closeup of each gene in this figure is presented in Figure 13.

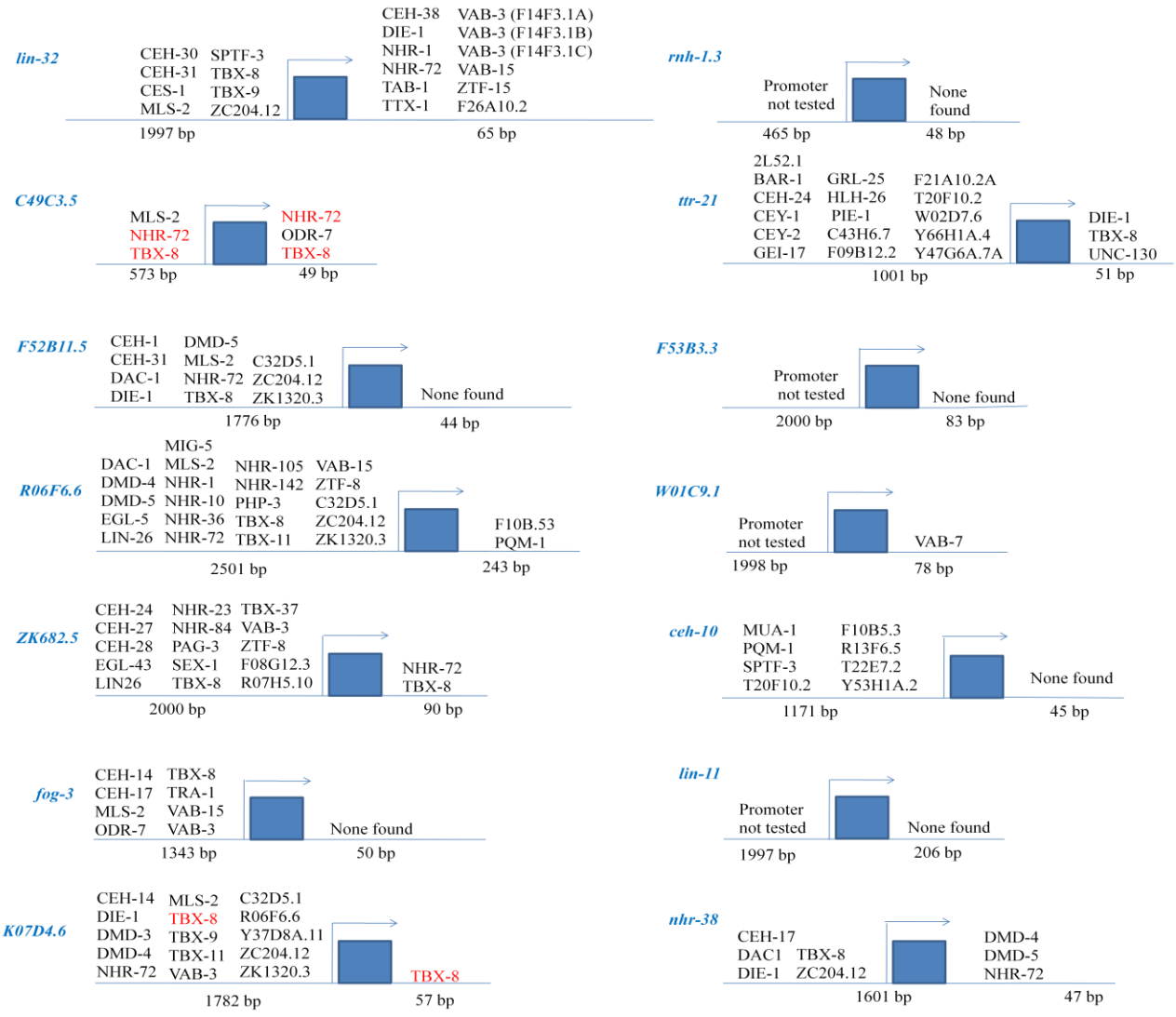


Figure 12. Results Summary. 14 first introns and 11 promoters were screened by the Y1H assay. A cartoon of each gene locus shows the promoter (upstream from arrow), first exon (blue box), and first intron (downstream from blue box). The size of each promoter fragment and intron is labeled below; the interactors are listed above. TFs labeled in red are found in both the first intron and promoter.

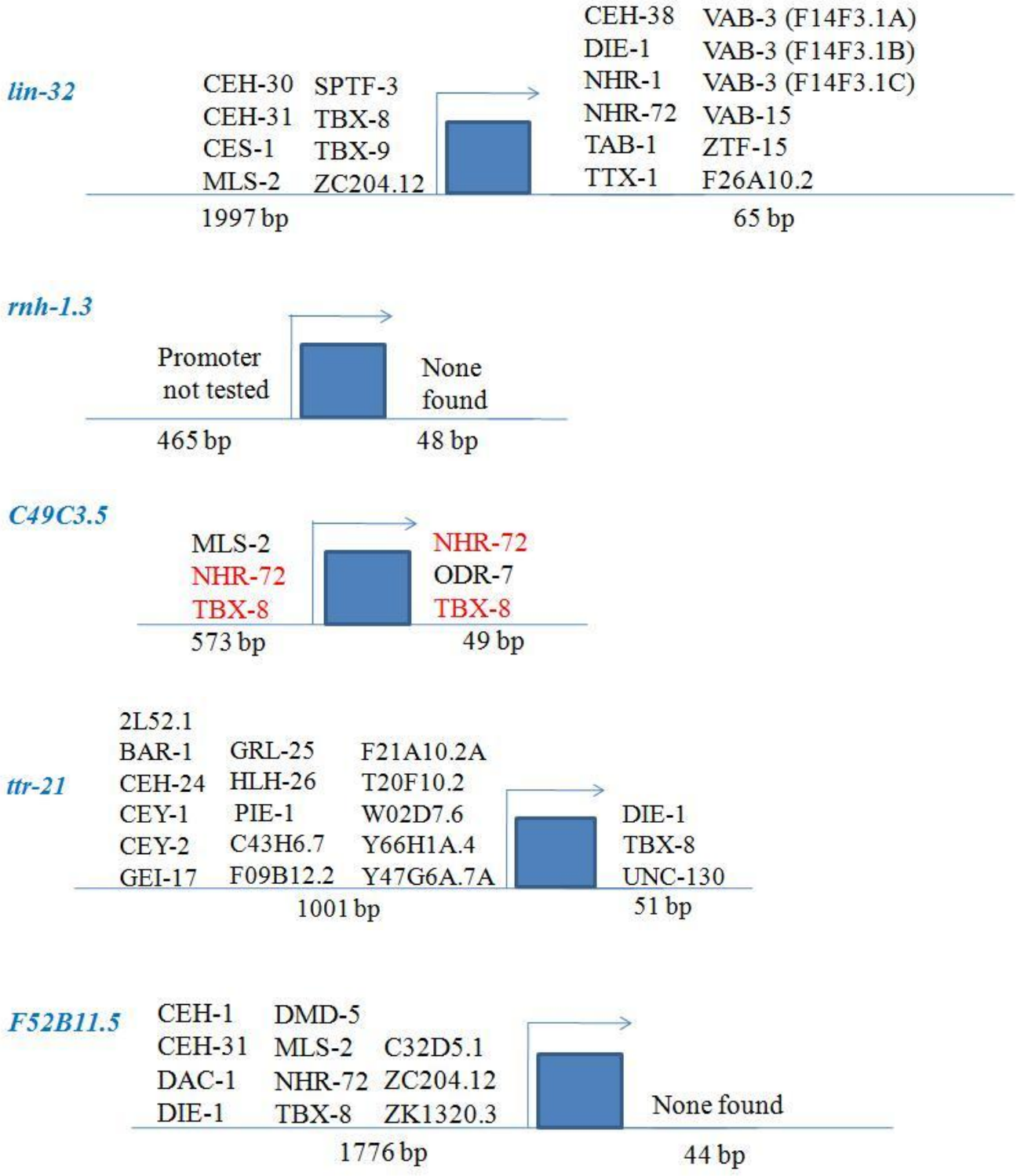


Figure 13a. Closeup Diagrams from the Previous Figure.

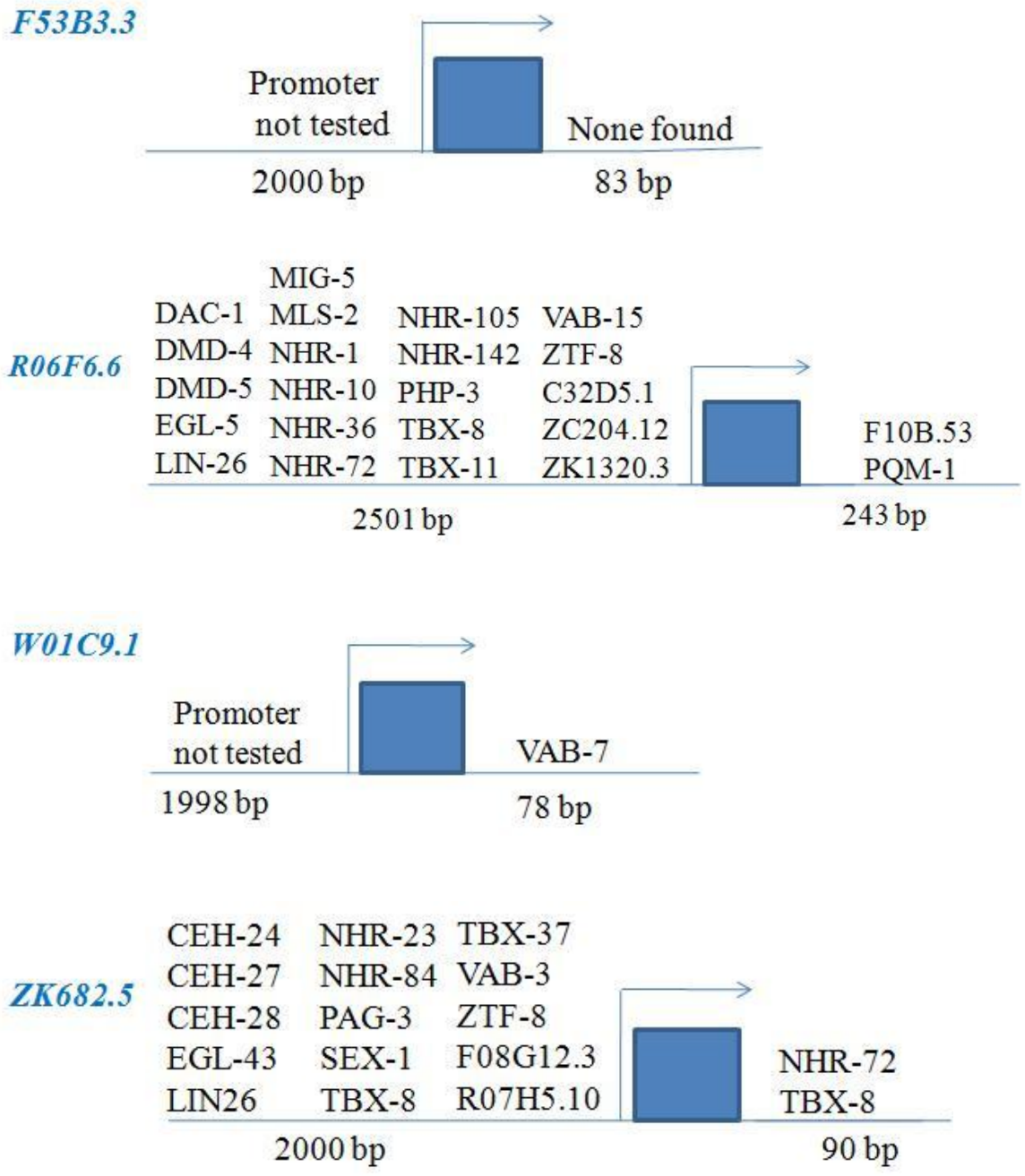


Figure 13b. Closeup Diagrams from the Previous Figure.

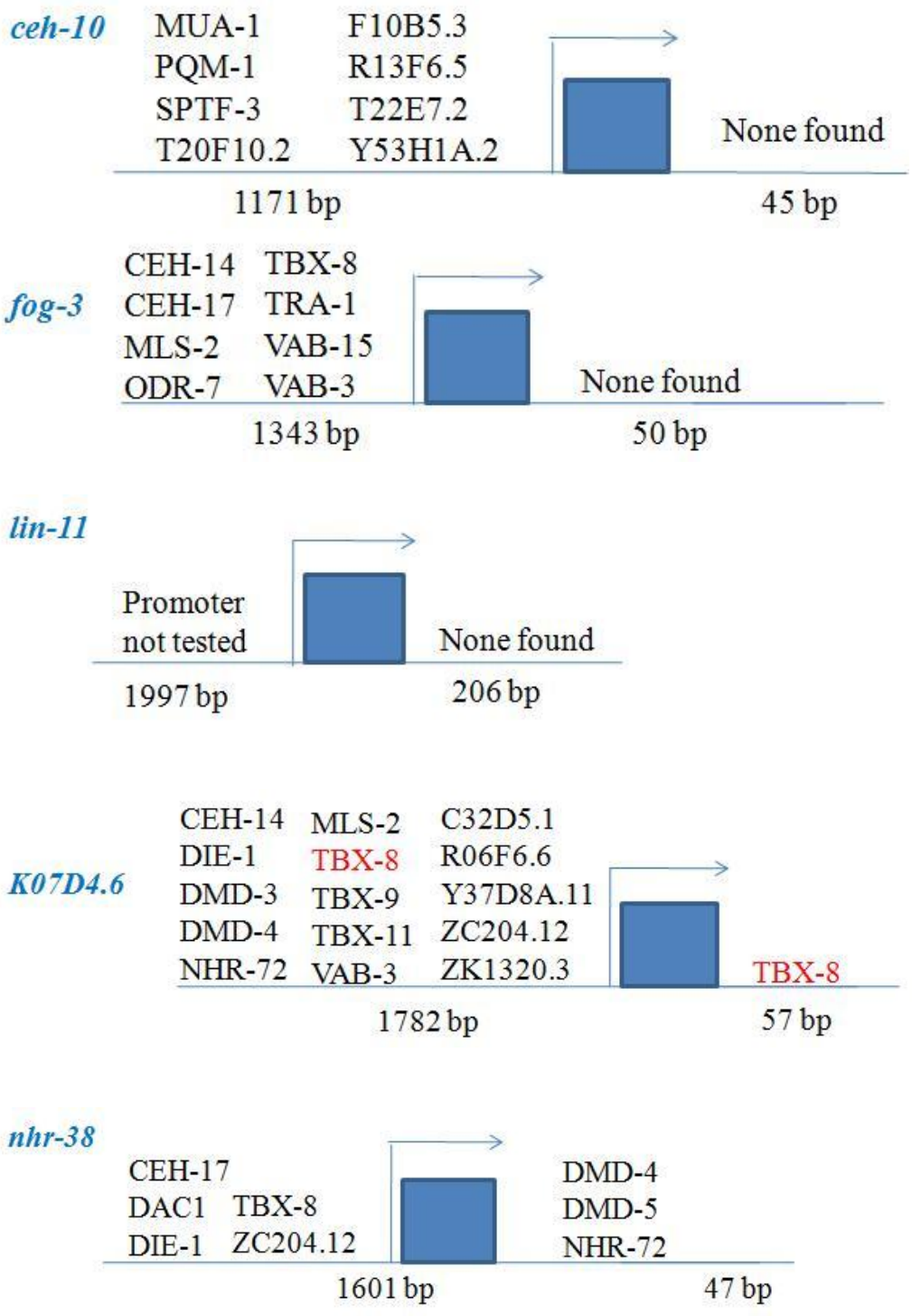


Figure 13c. Closeup Diagrams for the Previous Figure.

DISCUSSION

The first intron of 14 *C. elegans* genes, and most of the promoters of these same genes, were assayed in the Y1H system. 8 of the first introns had interactors and 6 did not have any. The small first intron (65 bp) of *lin-32* has 12 interactors, while its larger promoter (1997 bp) had only has 8 interactors. It is interesting that such a small DNA region can have so many interactors. Conversely, *R06F6.6* has 21 interactors in its promoter (2501), but its first intron (243 bp) only has 2 interactors. Another interesting result is that two of the three TFs that bind in first intron of *C49C3.5*, also bind in the promoter. Promoters are traditionally considered the main binding sites for TFs so the polymerase can be recruited and the transcription start site can be activated. It would be interesting to see if transcription is affected if this intron is removed, since it shares TFs with the promoter or vice versa.

From this data it is clear that TFs do bind the first intron of some genes. While the majority of these first introns show TF binding, the results from only 14 genes are not enough to make any general conclusions about all first introns having a role in gene regulation. But this data does show that the first intron is an element that needs to be further studied.

Due to time constraints only 14 introns were assayed in this study, but 34 more introns have been cloned and stored in glycerol stocks and will soon be processed in the Y1H assay. This increased data set will help give a better idea whether TF binding to first introns is uncommon as suggested by previously published DNase I hypersensitivity data (Shi et al. 2009) or is actually more important than previously thought.

To convincingly show that the first introns are important for gene regulation in *C. elegans, in vivo* experiments must be performed to validate the TF binding events found by Y1H assays. A typical experiment would be to insert a vector in the worm that has the gene promoter and first intron driving expression of GFP, and to observe the effects of mutating the TF or deleting the TF binding site in the first intron. If the GFP expression is reduced when the binding site is removed then the TF was using that site in the first intron to activate gene expression. Such *in vivo* tests would be similar to the findings for TFs binding in the first intron of the *C. elegans* genes *hlh-1* (Lei and Liu, 2009) and *lim-7* (Voutev and Keating, 2008).

In summary, this study provides an indication that some of the first introns in *C.elegans* are important for gene regulation, but more first introns need to be studied, and *in vivo* experiments need to be performed.

BIBLIOGRAPHY

Bradnam, K.R. and Korf, I. Longer first introns are a general property of eukaryotic gene structure. PLoS ONE 2008 Aug 29;3(8):e3093.

Brenowitz, M., Senear, D.F., Kingston, R.E. DNase I footprint analysis of protein-DNA binding. *Curr Protoc Mol Biol*. 2001. May; Chapter 12:Unit 12.4.

Deplancke, B., Dupuy, D., Vidal, M., and Walhout, A.J.M. 2004. A Gateway-compatible yeast one-hybrid system. *Genome Res.***14:** 2093–2101.

Deplancke, B., Mukhopadhyay, A., Ao, W., Elewa, A.M., Grove, C.A., Martinez, N.J., Sequerra, R., Doucette-Stamm, L., Reece-Hoyes, J.S., Hope, I.A., Tissenbaum, H.A., Mango, S.E., Walhout, A.J. A gene-centered C. elegans protein-DNA interaction network. *Cell* 2006 Jun 16; **125**(6): 1193-1205.

Durand-Dubief M, Ekwall K. Chromatin immunoprecipitation using microarrays. *Methods Mol Biol* 2009; **529**: 1-17.

Fields, S. and Song, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature* **340:** 245-246.

Flames, N., Hobert, O. Gene regulatory logic of dopamine neuron differentiation. *Nature* 2009 Apr 16; **458**(7240): 885-889.

Hartwell, L.H., et al, Genetics from genes to genomes. New York: McGraw-Hill, 2004.

International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature 2004 Oct 21; **431**(7011): 931-945.

King, M.W. "RNA Synthesis" . The Medical Biochemisty Page. 2009. IU School Of Medicine. 3 Mar. 2009 <http://themedicalbiochemistrypage.org/rna.html>.

Lee, T.I. and Young, R.A. 2000. Transcription of eukaryotic protein-coding genes. *Annu. Rev. genet.* **34:** 77-137.

Lei, H., Liu., J. Caudal-like PAL-1 directly activates the bodywall muscle module regulator hlh-1 in *c.elegans* to iniatate the embryonic muscle gene regulatory network. 9 Feb 2009: ePress

Martinez NJ, Ow MC, Barrasa MI, Hammell M, Sequerra R, Doucette-Stamm L, Roth FP, Ambros VR, Walhout AJ. A C. elegans genome-scale microRNA network contains composite feedback motifs with high flux capacity. *Genes Dev* 2008 Sep 15; **22**(18): 2535-2549.

McEwen, D.G., Ornitz, D.M. Regulation of the fibroblast growth factor receptor 3 promoter and intron 1 enhancer by Sp1 family transcription factor. *J Biol Chem* 1998. **273:** 5349-5357.

Okkema, P.G, Harrison, S.W., Plunger, V., Aryana,A., Fire, A. Sequence Requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. *Genetics* 1993. **135:** 385-404.

Polakowska, R.R., Graf, B.A., Falciano, V., LaCelle, P. Transcription regulatory elements of the first intron control human transglutaminase type 1 gene expression in epidermal keratinocytes. *J Cell Biochem* 1999. **73:** 355-369.

Reece-Hoyes, J.S., Deplancke, B., Shingles, J., Grove, C.A., Hope, I.A., and Walhout, A.J.M. 2005. A compendium of *C. elegans* regulatory transcription factors: A resource for mapping transcription regulatory networks. *Genome Biol.* **6:** R110.

Sabo, P. J. *et al.* Genome-scale mapping of DNase I sensitivity *in vivo* using tiling DNA microarrays. *Nature Methods* 2006. **3:** 511–518.

Shah, R.N., Ibbitt, J.C., Alitato, K., et al. FGFR4 over expression in pancreatic cancer is mediated by a intronic enhancer activated by HNF1α. *Oncogene* 2002. **21:** 8251-8261.

Shi, B., Guo, X., Wu, T., Sheng, S., Wang, J., Skogerbo, G., Zhu, X., Chen, R. Genome-scale identification of caenorhabditis elegans regulatory elements by tiling-array mapping of DNase I hypersensitive sites. BMC Genomics 2009. **10:** 92.

Taylor, J.D., Ackroyd, A.J., Halford, S.E. The Gel Shift Assay for Analysis of DNA-Protein interactions. *Springer Protocols* 1994. **30:** 263-279.

The *C. elegans* Sequencing Consortium, Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 1998. **282**: 2012-2018.

Vermeirssen V, Barrasa MI, Hidalgo CA, Babon JA, Sequerra R, Doucette-Stamm L, Barabasi AL, Walhout AJ. Transcription factor modularity in a gene-centered C. elegans core neuronal protein-DNA interaction network. *Genome Res* 2007 **7:** 1061-1071.

Voutev R. and Keating. R. Characterization of celegans islet LIM-homodomain ortholog, *lim-7*. *FEBS Letters* 2009.

Walhout, A.J.M. Unraveling transcription regulatory networks by protein-DNA and protein-protein interaction mapping. *Genome Res* 2006.**16:** 1445-1454.

Wilson MA, Chrysogelos SA. Identification and characterization of a negative regulatory element within the epidermal growth factor receptors gene first intron in

hormone-dependent breast cancer cells. *J Cell Biochem* 2002. **85:** 601-614.

APPENDIX

The following table includes all primers for the first introns which were designed to have gateway ends (attB4 tails for forward primers and attB1R for reverse primers compatible with pDONR P4-P1R vector).

Gateway 5' end
GGGGACAACTTTGTATAGAAAAGTTG

Gateway 3' end
GGGGACTGCTTTTTTGTACAAACTTGTC

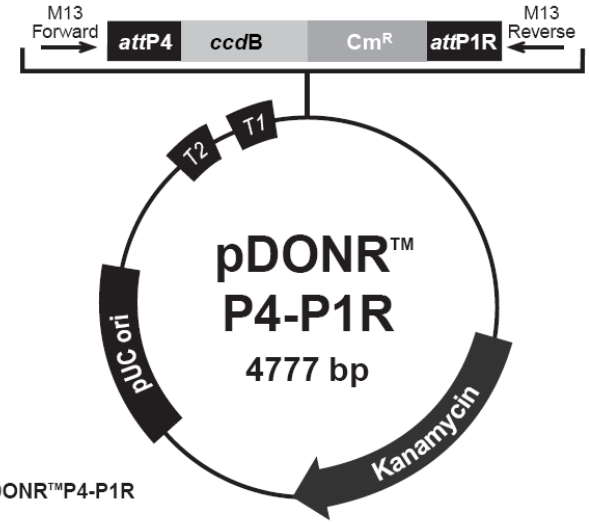
FORWARD PRIMERS

Plate	Row	Column	Primer Name	Sequence
PLATE01F	A	1	FI_lin-32_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagtttaaccatatcatctttttttca
PLATE01F	A	2	FI_rnh-1.3_F1	GGGGACAACTTTGTATAGAAAAGTTGtacggttttttgtgcaattaattgt
PLATE01F	A	3	FI_C18A11.1_F1	GGGGACAACTTTGTATAGAAAAGTTGtgagtagacaaacttttttttatcga
PLATE01F	A	4	FI_tag-97_F1	GGGGACAACTTTGTATAGAAAAGTTGtgaggaagcttgaagttttgt
PLATE01F	A	5	FI_C49C3.5_F1	GGGGACAACTTTGTATAGAAAAGTTGtgactttggaactgatgcat
PLATE01F	A	6	FI_ttr-21_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagttatttttttcctgaattttaaaagtaa
PLATE01F	A	7	FI_F26A10.2_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagtcacaagggggactct
PLATE01F	A	8	FI_F52B11.5_F1	GGGGACAACTTTGTATAGAAAAGTTGtgagttggttttcagccatt
PLATE01F	A	9	FI_F53B3.3_F1	GGGGACAACTTTGTATAGAAAAGTTGtaaacatatatccatgacaagacgg
PLATE01F	A	10	FI_R03C1.1_F1	GGGGACAACTTTGTATAGAAAAGTTGtgagtgtctaaaacttgattcccc
PLATE01F	A	11	FI_R06F6.6_F1	GGGGACAACTTTGTATAGAAAAGTTGtggctcacttcttttgcat
PLATE01F	A	12	FI_T23G5.3_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagtctaaaaaaaaagttgtaacttacaaa
PLATE01F	B	1	FI_W01C9.1_F1	GGGGACAACTTTGTATAGAAAAGTTGtatttctatatttcatctgcattataata
PLATE01F	B	2	FI_Y49E10.4_F1	GGGGACAACTTTGTATAGAAAAGTTGtgagtttttgcctgatcttca
PLATE01F	B	3	FI_ZK287.1_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagaacaatgactataatctaaaaactttagt
PLATE01F	B	4	FI_ZK682.5_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagtttgaaactgaaaatattcgtttcc
PLATE01F	B	5	FI_cdr-4_F1	GGGGACAACTTTGTATAGAAAAGTTGttagtttcctcttacgacattattga
PLATE01F	B	6	FI_ceh-10_F1	GGGGACAACTTTGTATAGAAAAGTTGtaatatcattcgtattccgaagtgc
PLATE01F	B	7	FI_ceh-13_F1	GGGGACAACTTTGTATAGAAAAGTTGttagttttcaagagttttaattgttaaa
PLATE01F	B	8	FI_ceh-36_F1	GGGGACAACTTTGTATAGAAAAGTTGtagttaatcaggggggatcaaaaaa
PLATE01F	B	9	FI_ceh-43_F1	GGGGACAACTTTGTATAGAAAAGTTGtatgaaactcggaattgaaaaagc
PLATE01F	B	10	FI_ceh-5_F1	GGGGACAACTTTGTATAGAAAAGTTGtaaggcagaaaatgaaactgttg
PLATE01F	B	11	FI_cyp-33C8_F1	GGGGACAACTTTGTATAGAAAAGTTGtatgtagtctatgttttccttgaagtc
PLATE01F	B	12	FI_elt-4_F1	GGGGACAACTTTGTATAGAAAAGTTGtgagtttgcaacttttttttcaaatc
PLATE01F	C	1	FI_emb-5_F1	GGGGACAACTTTGTATAGAAAAGTTGtaattataattccggggtatttctat
PLATE01F	C	2	FI_F38B6.1_F1	GGGGACAACTTTGTATAGAAAAGTTGtttcaataattaaaaaaaaattaagaataatca
PLATE01F	C	3	FI_fog-3_F1	GGGGACAACTTTGTATAGAAAAGTTGtaatttgaaacaatttttttgatggt
PLATE01F	C	4	FI_inx-6_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagtttattgtgctttttttccgag
PLATE01F	C	5	FI_lin-11_F1	GGGGACAACTTTGTATAGAAAAGTTGtatcggtccctcacctttttt
PLATE01F	C	6	FI_lin-49_F1	GGGGACAACTTTGTATAGAAAAGTTGtaatatacttctacgtttgagaaacattt
PLATE01F	C	7	FI_nhr-68_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagaatttatatcatgtttagccca
PLATE01F	C	8	FI_nhr-79_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagtgctcaaaatgttggga
PLATE01F	C	9	FI_pop-1_F1	GGGGACAACTTTGTATAGAAAAGTTGtgagttttccaataaatttgatcg
PLATE01F	C	10	FI_pos-1_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagtaatcgttgatcctacaaattt
PLATE01F	C	11	FI_pqn-26_F1	GGGGACAACTTTGTATAGAAAAGTTGtaatttactatttttttaaatttgttgaattt
PLATE01F	C	12	FI_pqn-60_F1	GGGGACAACTTTGTATAGAAAAGTTGtgagttctctgaaattttgaattttt
PLATE01F	D	1	FI_unc-130_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagttgggaacttgggaaaa
PLATE01F	D	2	FI_unc-42_F1	GGGGACAACTTTGTATAGAAAAGTTGtaatttgctaactaattctagtctccg
PLATE01F	D	3	FI_zag-1_F1	GGGGACAACTTTGTATAGAAAAGTTGtaggtcctagtgctcgccgagt
PLATE01F	D	4	FI_gpa-10_F1	GGGGACAACTTTGTATAGAAAAGTTGtcagtttttgagttacatgtttcataa
PLATE01F	D	5	FI_ceh-23_F1	GGGGACAACTTTGTATAGAAAAGTTGtaacatcaaatccattttgaaatgtg
PLATE01F	D	6	FI_nhr-83_F1	GGGGACAACTTTGTATAGAAAAGTTGtgggttttcttttttactggcc
PLATE01F	D	7	FI_sma-3_F1	GGGGACAACTTTGTATAGAAAAGTTGttagtcaaatatatttttttaaaggaatttataa
PLATE01F	D	8	FI_K07D4.6_F1	GGGGACAACTTTGTATAGAAAAGTTGtatgttaataaatacattttatgacgttt
PLATE01F	D	9	FI_lin-48_F1	GGGGACAACTTTGTATAGAAAAGTTGtatttgagtcttgatttcaaaaaaaaaac
PLATE01F	D	10	FI_ttr-54_F1	GGGGACAACTTTGTATAGAAAAGTTGtaagtttttttattttttgaaatgtcgt
PLATE01F	D	11	FI_hlh-6_F1	GGGGACAACTTTGTATAGAAAAGTTGtgagtgaatgttttatattgcacc
PLATE01F	D	12	FI_nhr-38_F1	GGGGACAACTTTGTATAGAAAAGTTGtaatctaaagttgaacaaaatataaatgtaaca

REVERSE PRIMERS

Plate	Row	Column	Primer Name	Sequence
PLATE01R	A	1	FI_lin-32_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgcaacaagaaatcaattagaaatc
PLATE01R	A	2	FI_rnh-1.3_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaatatgttcaaataaggaacaattaattg
PLATE01R	A	3	FI_C18A11.1_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtggaaattatgaaaactttttttttttta
PLATE01R	A	4	FI_tag-97_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaatcatcaaatttgtattgtaaatagg
PLATE01R	A	5	FI_C49C3.5_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaattgtattctatgctgtctaaaatg
PLATE01R	A	6	FI_ttr-21_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgtaaattttaattttgtttacttttaaaattc
PLATE01R	A	7	FI_F26A10.2_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtaaaaacgcattatatattagttatttttgtt
PLATE01R	A	8	FI_F52B11.5_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtggaatcggggcatttattg
PLATE01R	A	9	FI_F53B3.3_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaaaagtgcctatttggtg
PLATE01R	A	10	FI_R03C1.1_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaagattcttgttagttcaactcc
PLATE01R	A	11	FI_R06F6.6_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtggaatttttttttgaaaaaaaagg
PLATE01R	A	12	FI_T23G5.3_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaataagcattcaaattttcc
PLATE01R	B	1	FI_W01C9.1_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgtgagaattagaaaaagcattg
PLATE01R	B	2	FI_Y49E10.4_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaattatgtataaattttctgaagatc
PLATE01R	B	3	FI_ZK287.1_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaaattataaactgaaaaaaaaaact
PLATE01R	B	4	FI_ZK682.5_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaatatttaacatcaaaagattaata
PLATE01R	B	5	FI_cdr-4_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgcaaataaatctgtcatattcaataatgt
PLATE01R	B	6	FI_ceh-10_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtaaaaattactgatttgtgcacttcg
PLATE01R	B	7	FI_ceh-13_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaatattttctattgatttgagatg
PLATE01R	B	8	FI_ceh-36_R1	GGGGACTGCTTTTTTGTACAAACTTGTCttaaaaaaatcgttaaagattttaatttca
PLATE01R	B	9	FI_ceh-43_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgcaaacttttttttataataacccct
PLATE01R	B	10	FI_ceh-5_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaaatgtgattctcaaaaacac
PLATE01R	B	11	FI_cyp-33C8_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtataaaaattgcaaaaaattctcaagtt
PLATE01R	B	12	FI_elt-4_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgagaaataaaccattagatttgaaa
PLATE01R	C	1	FI_emb-5_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtaaaaatgaacaaatgaaaatagaaatac
PLATE01R	C	2	FI_F38B6.1_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaagcgcatgttatttg
PLATE01R	C	3	FI_fog-3_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtggaaaatataaattaactctaaaccatc
PLATE01R	C	4	FI_inx-6_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaattcaagttaaaaattcgaga
PLATE01R	C	5	FI_lin-11_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaagtaagtgaaagtgtcggc
PLATE01R	C	6	FI_lin-49_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaagtaacttttttttaaaatgtttc
PLATE01R	C	7	FI_nhr-68_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtaaattagaatatttgtgggctaaac
PLATE01R	C	8	FI_nhr-79_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtggaaatattgggtctcgaaa
PLATE01R	C	9	FI_pop-1_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaatttgataattattcggttag
PLATE01R	C	10	FI_pos-1_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaattccagtcataaaataagaa
PLATE01R	C	11	FI_pqn-26_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaaatcgttcaagttgttg
PLATE01R	C	12	FI_pqn-60_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtggaaaattattgtgatttcagattt
PLATE01R	D	1	FI_unc-130_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgcaaaattatagagttacagttgttt
PLATE01R	D	2	FI_unc-42_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaatatagatcatgatgtaggttc
PLATE01R	D	3	FI_zag-1_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgcaaaaaataaagtaaaaacattaatttatc
PLATE01R	D	4	FI_gpa-10_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaatattaacgacaattatgaaaca
PLATE01R	D	5	FI_ceh-23_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaattcataaaatgatatgtcagc
PLATE01R	D	6	FI_nhr-83_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaaaattgtgcaaaataattgag
PLATE01R	D	7	FI_sma-3_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtggaatttagaatttgaaatgga
PLATE01R	D	8	FI_K07D4.6_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtggaaacagtgctcatttgaa
PLATE01R	D	9	FI_lin-48_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgaattcacactttcaagttttttt
PLATE01R	D	10	FI_ttr-54_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgtaatttaatttttttttcagttcaaca
PLATE01R	D	11	FI_hlh-6_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtgtaatatattcaaaaaatgatataagtacatac
PLATE01R	D	12	FI_nhr-38_R1	GGGGACTGCTTTTTTGTACAAACTTGTCtttaattctttttgttacatttatattttgtt

The following figure shows the pDONR P4-P1R vector.



Comments for pDONR™P4-P1R
4777 nucleotides

*rrn*B T2 transcription termination sequence: bases 268-295 (c)
*rrn*B T1 transcription termination sequence: bases 427-470 (c)
M13 Forward (-20) priming site: bases 537-552
*att*P4 recombination site: bases 593-824 (c)
*ccd*B gene: bases 1181-1486 (c)
Chloramphenicol resistance gene: bases 1828-2487 (c)
*att*P1R recombination site: bases 2748-2979 (c)
M13 Reverse priming site: bases 3042-3058
Kanamycin resistance gene: bases 3171-3980
pUC origin: bases 4101-4774

(c) = complementary strand

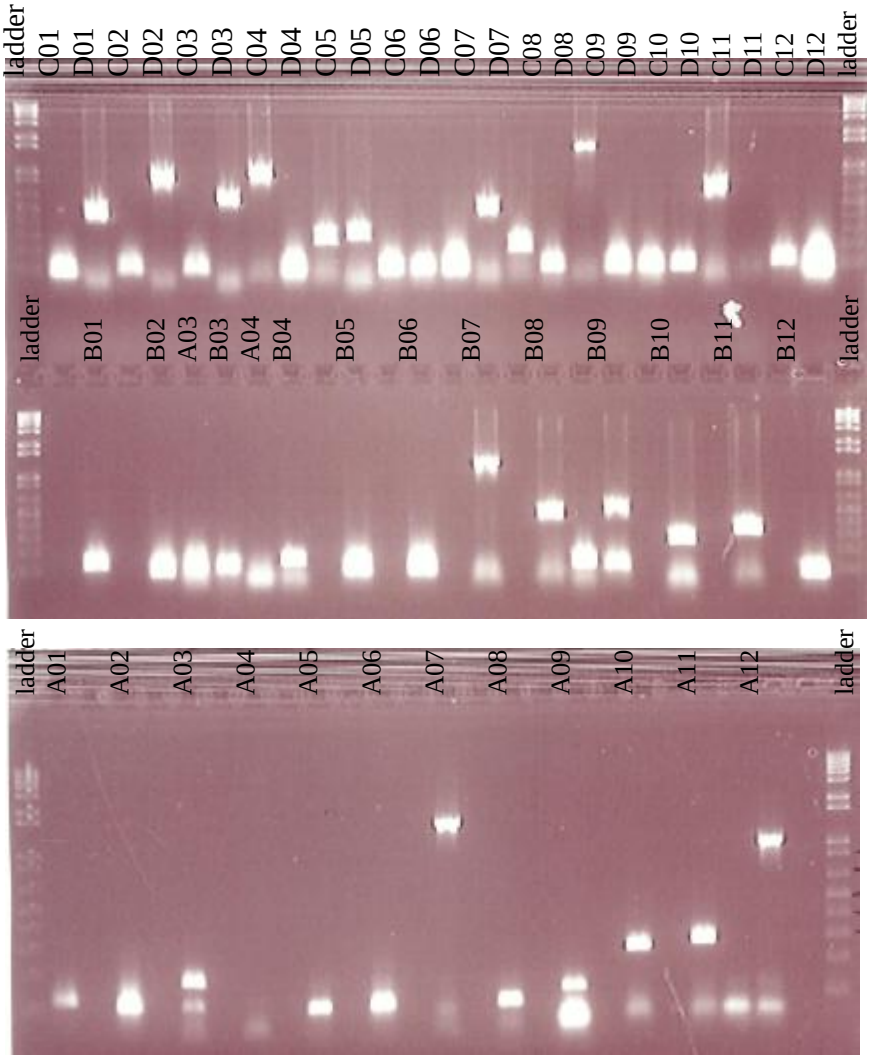
The following is a list of primers used throughout the experiment.
M13RV, 5'- CAGGAAACAGCTATGACC-3'
M13FW, 5'-GTAAAACGACGGCCAGT-3'
HIS293RV, 5'-GGGACCACCCTTTAAAGAGA-3'
P1H1FW, 5'-GTTCGGAGATTACCGAATCAA-3'
LacZ592RV, 5'-ATGCGCTCAGGTCAAATTCAGA-3'
AD, 5'-CGCGTTTGGAATCACTACAGGG-3'
TERM, 5'-GGAGACTTGACCAAACCTCTGGCG-3'

The following table includes all the data collected regarding intron number, sizes etc. used for the genes selected in the experiment. The coordinate in the 96-well plate of each intron is listed in the table. Genes highlighted in purple are the introns that went through the Y1H assay, all the others are still in glycerol stocks and need to be integrated into the YM4271 yeast genome.

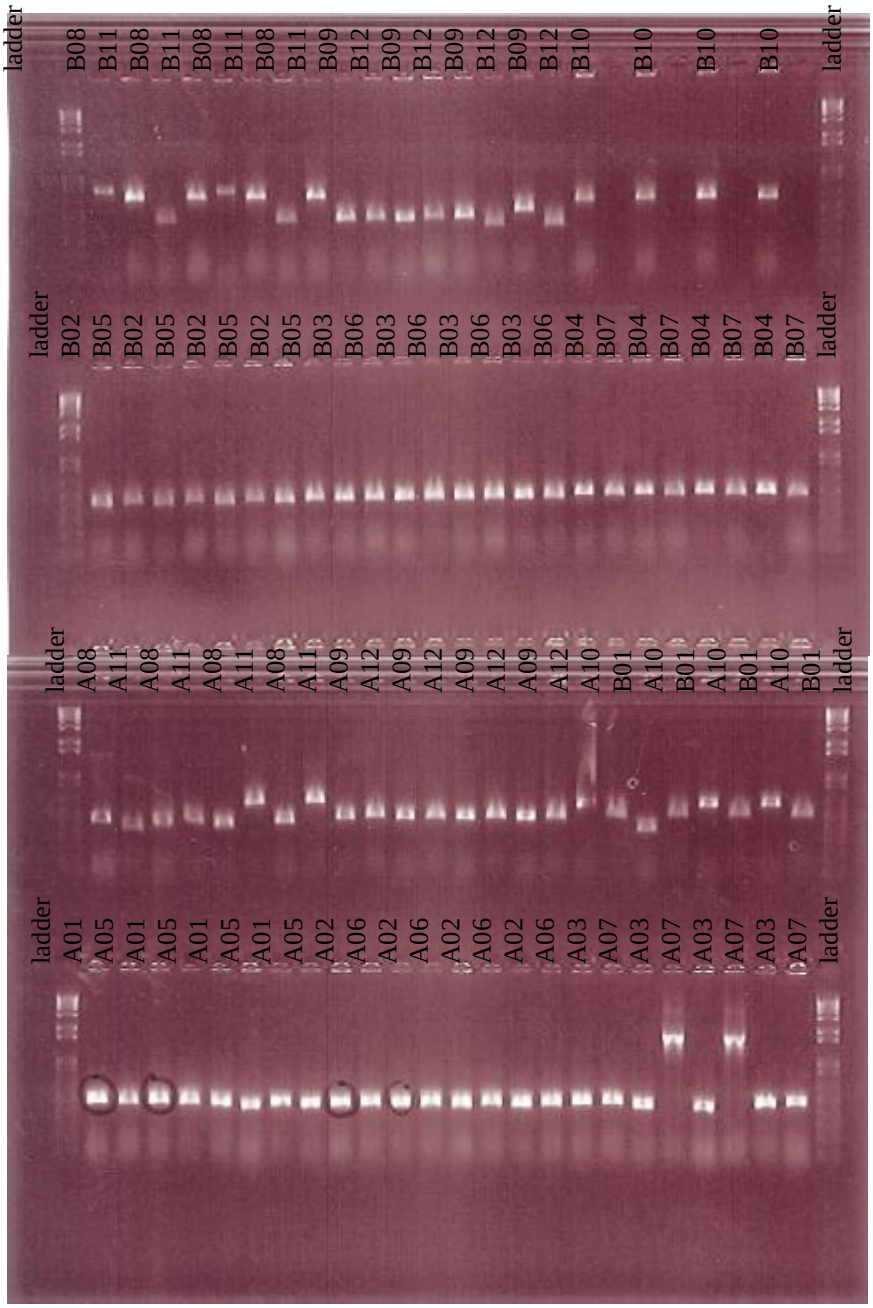
Promoter	Main Name	Sequence Name	Well	Comment	# of TFs	Promoter Length	# of Introns	Length of First Intron
Plin-32	lin-32	T14F9.5	A01	no EST	5	1997	2	65
PC04F12.9	rnh-1.3	C04F12.9	A02	begins with UTR	7	465	3	48
PC18A11.1	C18A11.1	C18A11.1	A03	begins and ends with UTR	6	1635	2	137
PC33A11.4	tag-97	C33A11.4	A04	very long intron	5	2000	9	4567
PC49C3.5	C49C3.5	C49C3.5	A05	first 3 no EST	1	573	5	49
PF09F3.6	ttr-21	F09F3.6	A06	1 intron begin and end UTR	1	1001	1	51
PF26A10.2	F26A10.2	F26A10.2	A07	long intron?, 1 EST	1	2000	8	1488
PF52B11.5	F52B11.5	F52B11.5	A08	begin and end with UTR	1	1776	2	44
PF53B3.3	F53B3.3	F53B3.3	A09	1/2 EST confirmed	4	2000	10	83
PR03C1.1	R03C1.1	R03C1.1	A10	1/2 EST confirmed	1	2000	7	214
PR06F6.6	R06F6.6	R06F6.6	A11	4 EST confimed	4	2501	7	243
PT23G5.3	T23G5.3	T23G5.3	A12		7	1743	4	1012
PW01C9.1	W01C9.1	W01C9.1	B01	only 2 EST	2	1998	4	78
PY49E10.4	Y49E10.4	Y49E10.4	B02	begin and end with UTR	2	1873	4	44
PZK287.1	ZK287.1	ZK287.1	B03	begin and end with UTR	1	2000	5	175
PZK682.5	ZK682.5	ZK682.5	B04	missing 3 EST	3	2000	9	90
Pcdr-4	cdr-4	K01D12.11	B05	only 2 EST confirmed	1	848	5	48
Pceh-10	ceh-10	W03A3.1	B06	missing only 1 EST	4	1171	4	45
Pceh-13	ceh-13	R13A5.5	B07	end with large Utr, EST where no exon	10	2000	2	1235
Pceh-36	ceh-36	C37E2.4	B08	begin with UTR, missing 1 EST	3	2003	5	412
Pceh-43	ceh-43	C28A5.4	B09		2	382	4	479
Pceh-5	ceh-5	C16C2.1	B10	begin and end with UTR	2	1801	2	222
Pcyp-33C8	cyp-33C8	R08F11.3	B11		1	536	3	290
Pelt-4	elt-4	C39B10.6	B12	only 1 intron	14	2501	1	45
Pemb-5	emb-5	T04A8.14	C01		4	1548	10	47
PF38B6.1	F38B6.1	F38B6.1	C02	no EST	4	301	3	53
Pfog-3	fog-3	C03C11.2	C03	missing 3 EST	4	1343	7	50
Pinx-6	inx-6	C36H8.2	C04	long intron?	14	1880	7	914
Plin-11	lin-11	ZC247.3	C05	long intron?	3	1997	9	206
Plin-49	lin-49	F42A9.2	C06		4	1984	9	50
Pnhr-68	nhr-68	H12C20.3	C07		6	494	7	44
Pnhr-79	nhr-79	T26H2.9	C08	long intron?, missing 4 EST	42	1798	10	155
Ppop-1	pop-1	W10C8.2	C09	long intron?, missing 1 EST	4	2501	3	1520
Ppos-1	pos-1	F52E1.1	C10	1 intron	8	620	1	52
Ppqn-26	pqn-26	DY3.5	C11	long intron?, no EST	2	2469	9	656
Ppqn-60	pqn-60	R11G11.7	C12	1 intron	5	817	1	69
Punc-130	unc-130	C47G2.2	D01	2 intron UTR begin and end	12	1997	2	434
Punc-42	unc-42	F58E6.10	D02	missing 2 EST	13	2000	6	869
Pzag-1	zag-1	F28F9.1	D03	missing 2 EST	5	2000	6	562
Pgpa-10	gpa-10	C55H1.2	D04	no EST	37	2003	8	46
Pceh-23	ceh-23	ZK652.5	D05	no EST	23	2000	3	240
Pnhr-83	nhr-83	F48G7.3	D06	no EST	14	1674	11	47
Psma-3	sma-3	R13F6.9	D07	UTR at end, no EST	13	1552	11	458
PK07D4.6	K07D4.6	K07D4.6	D08	no EST	10	1782	5	57
Plin-48	lin-48	F34D10.5	D09	no EST	8	2000	5	48
PT14G10.4	ttr-54	T14G10.4	D10	no EST	7	774	2	60
Phlh-6	hlh-6	T15H9.3	D11	no EST	6	1192	3	141
Pnhr-38	nhr-38	K01H12.3	D12	no EST	6	1601	7	47

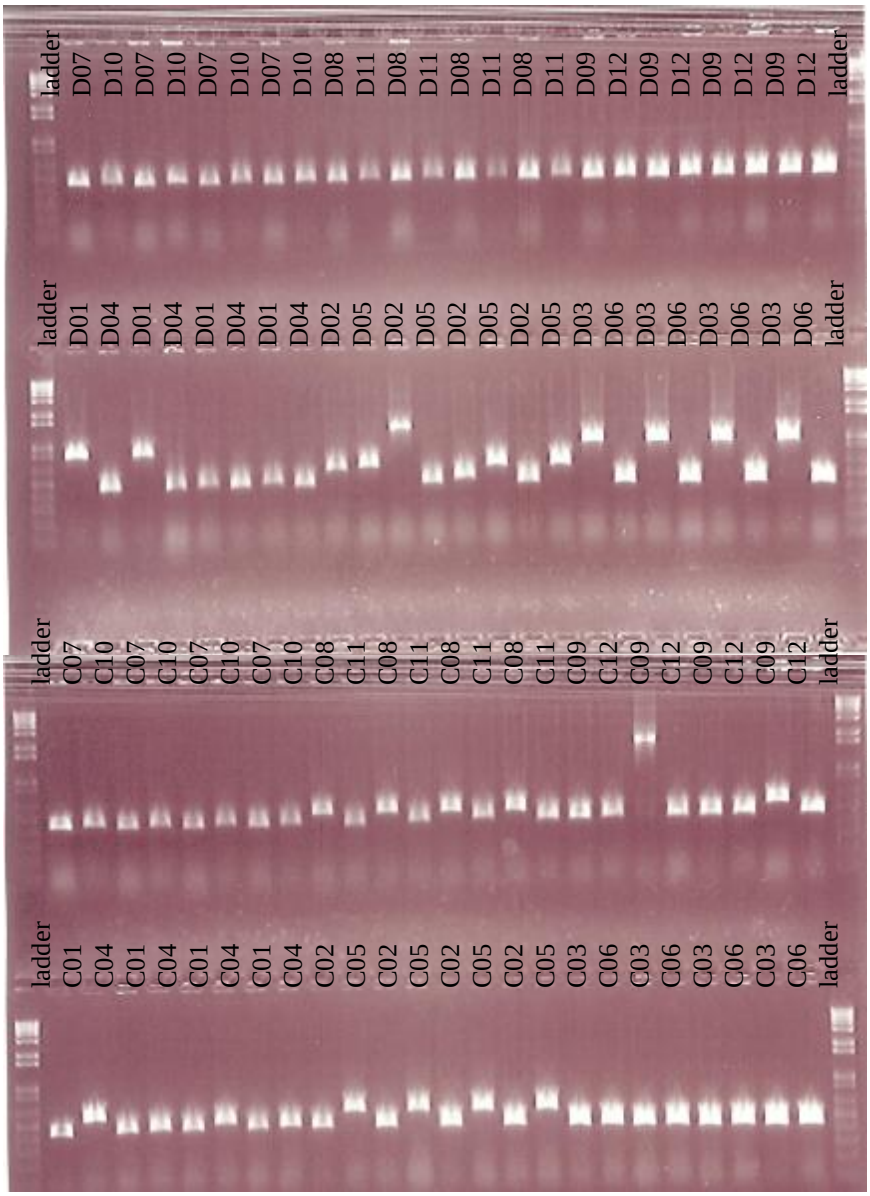
The following are the remainder of gel pictures for the first intron baits.

FIRST INTRON PCR AMPILFICATION

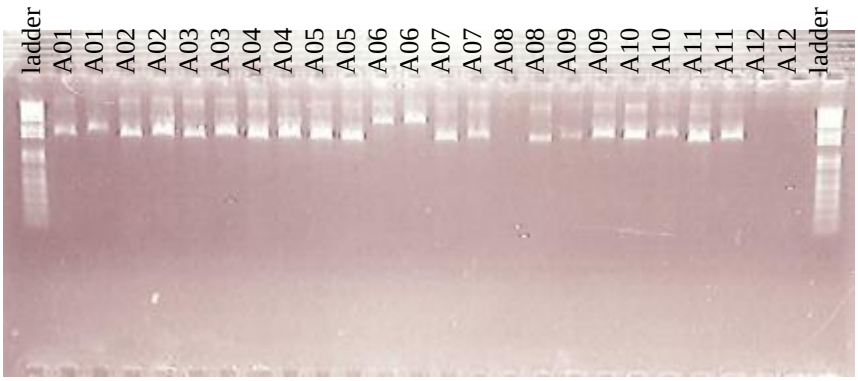


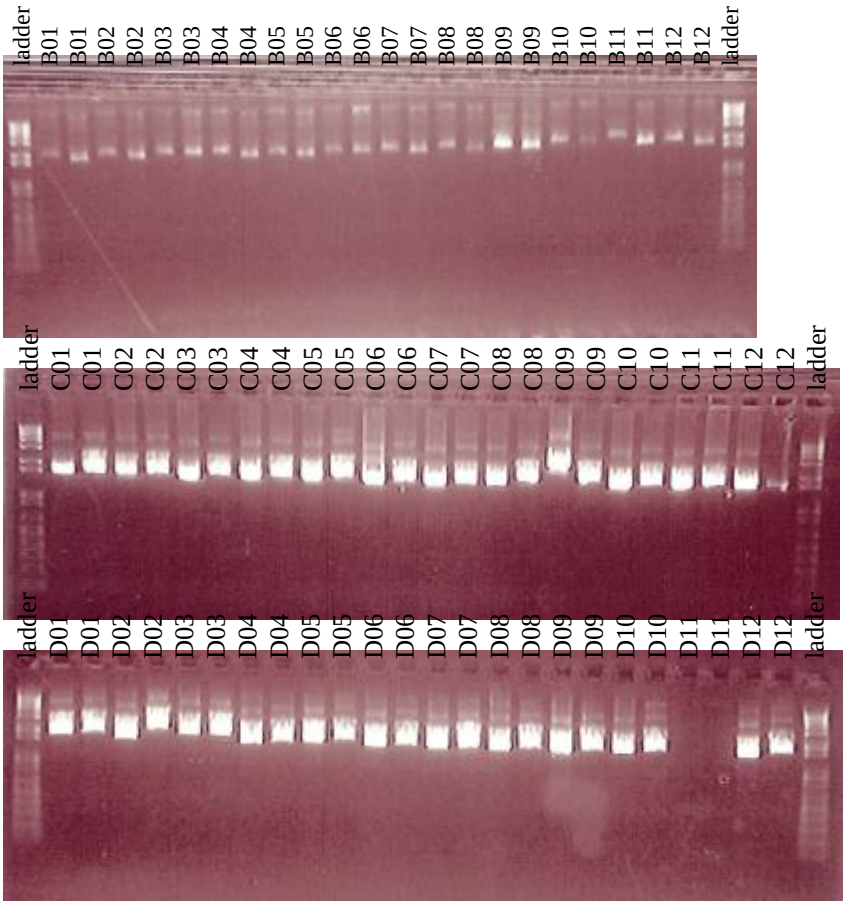
BP 4-COLONY PCR




ladder C01 C04 C01 C04 C01 C04 C01 C04 C02 C05 C02 C05 C02 C05 C02 C05 C03 C06 C03 C06 C03 C06 C03 C06 ladder
ladder C07 C10 C07 C10 C07 C10 C07 C10 C08 C11 C08 C11 C08 C11 C08 C11 C09 C12 C09 C12 C09 C12 C09 C12 ladder
ladder D01 D04 D01 D04 D01 D04 D01 D04 D02 D05 D02 D05 D02 D05 D02 D05 D03 D06 D03 D06 D03 D06 D03 D06 ladder
ladder D07 D10 D07 D10 D07 D10 D07 D10 D08 D11 D08 D11 D08 D11 D08 D11 D09 D12 D09 D12 D09 D12 D09 D12 ladder

BP MINIPREP

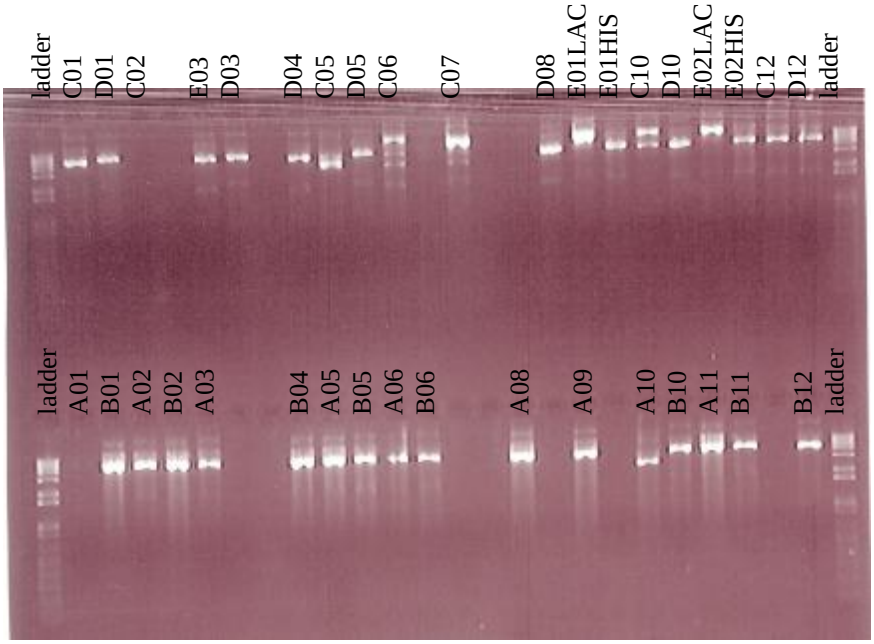

ladder A01 A01 A02 A02 A03 A03 A04 A04 A05 A05 A06 A06 A07 A07 A08 A08 A09 A09 A10 A10 A11 A11 A12 A12 ladder



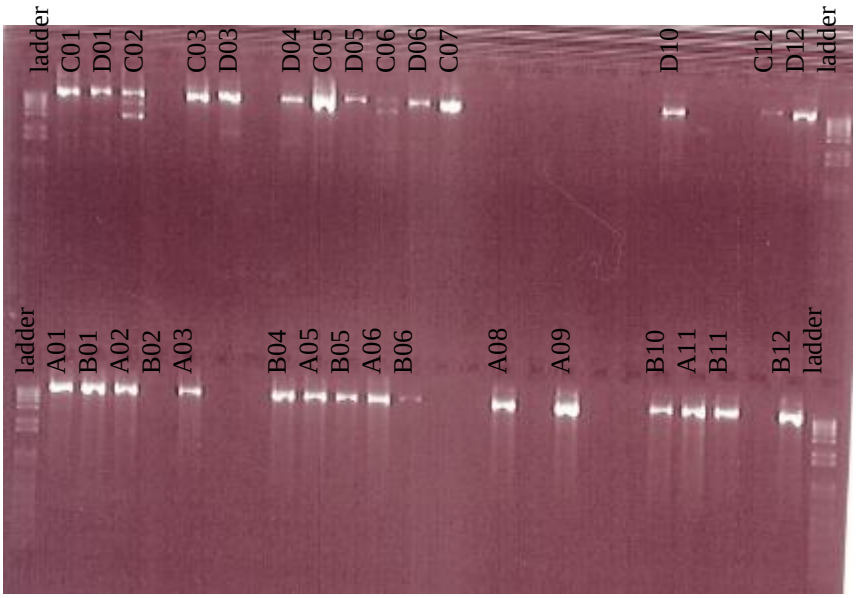
LR MINIPREP

01-DEC

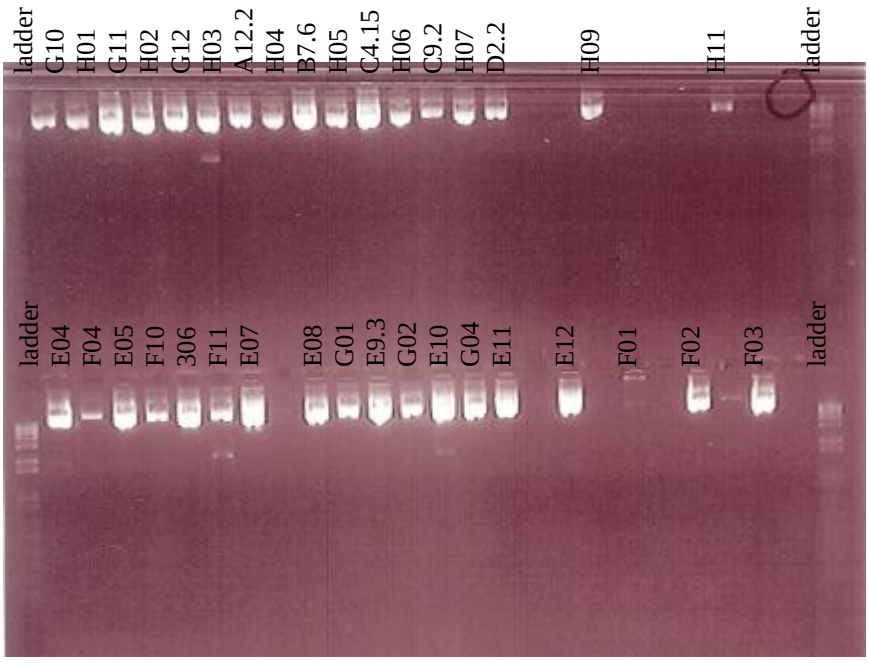
HIS



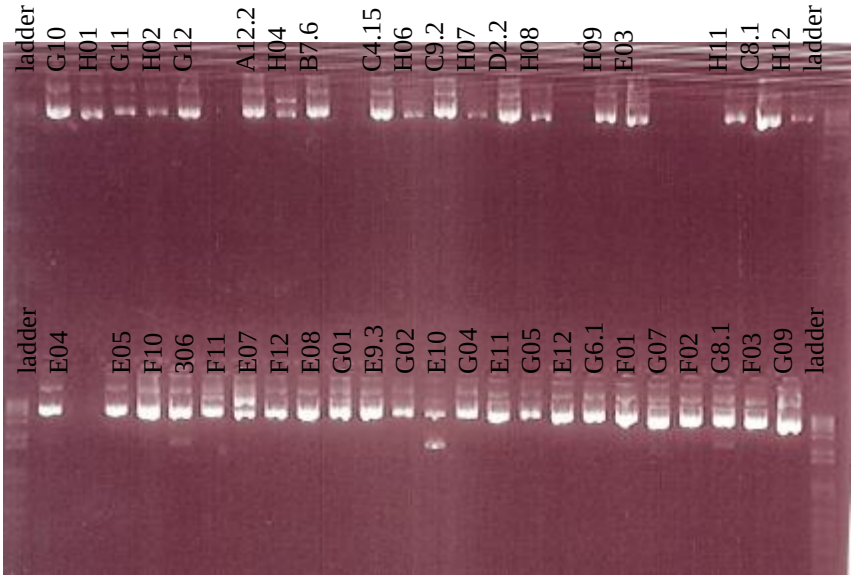
LAC



11-DEC
LAC

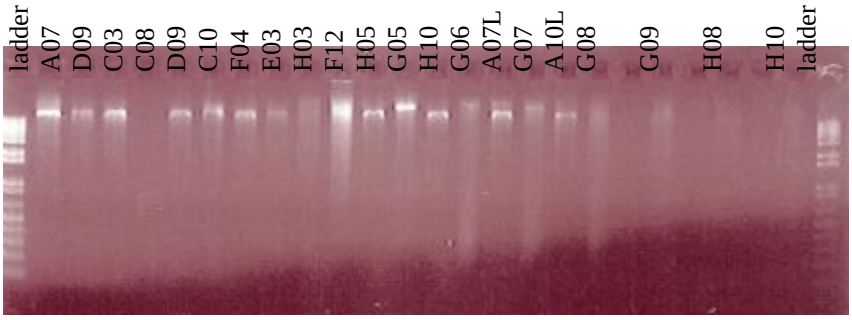


HIS

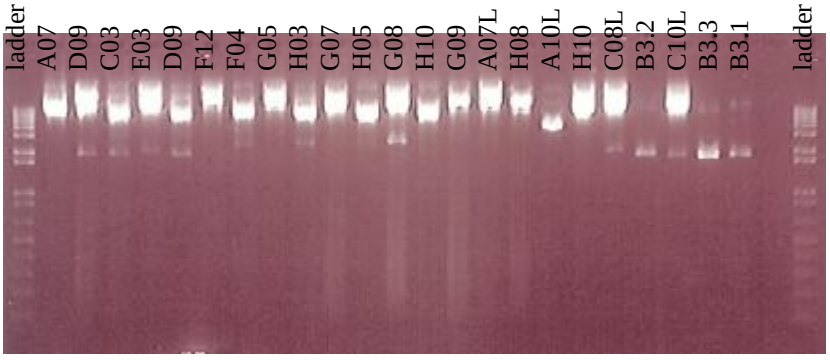


LR GLYCEROL STOCK TO MP

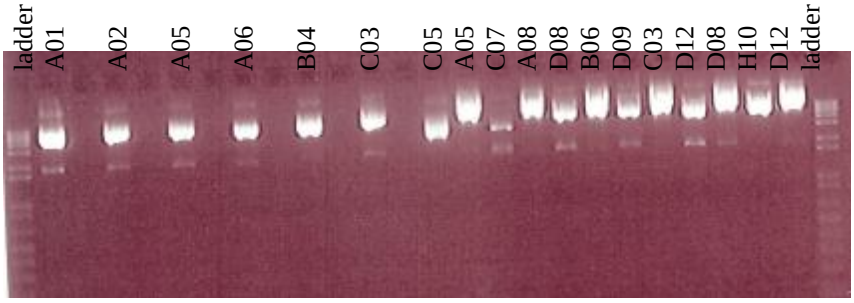
8-JAN
ROW A=HIS B=LAC



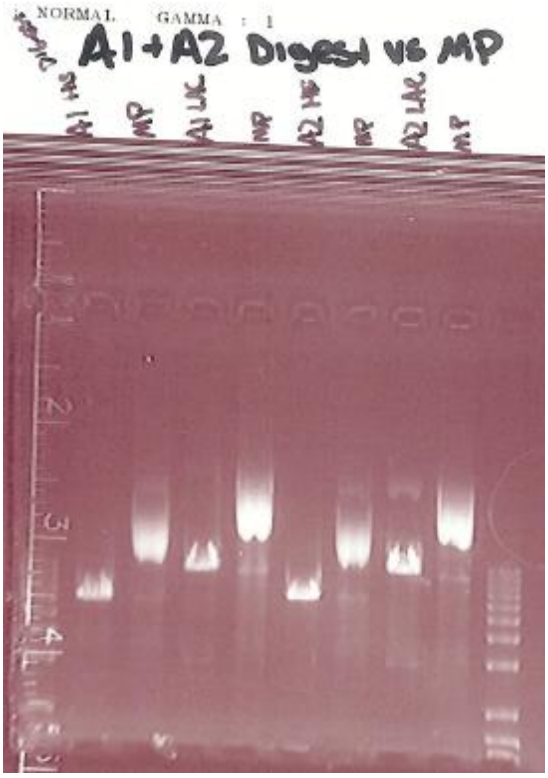
10-JAN
ROW A=HIS B=LAC

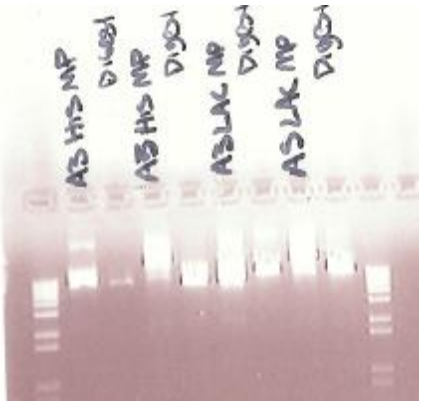


27-JAN
ROW A=HIS B=LAC

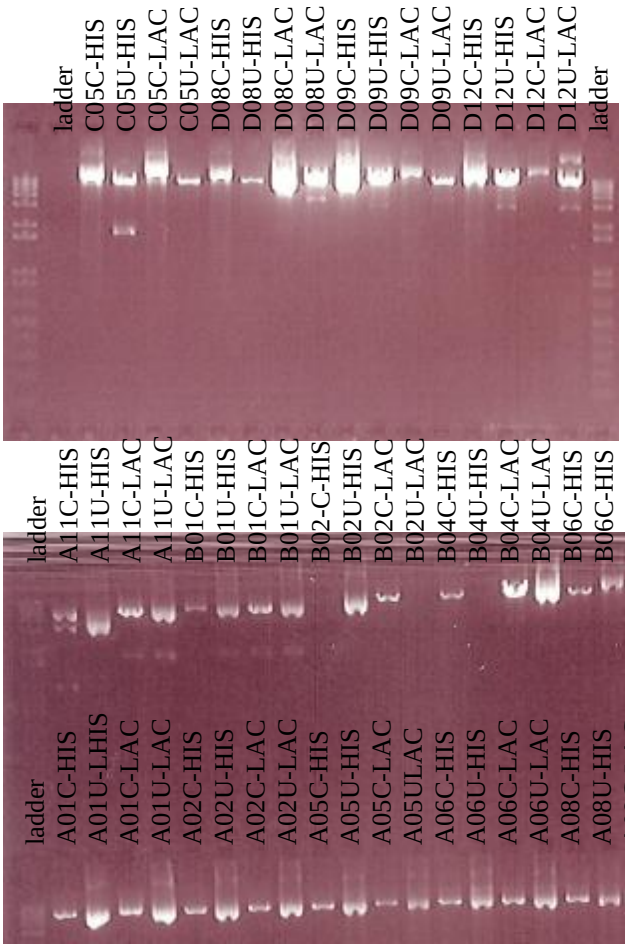


DIGESTS VS MP



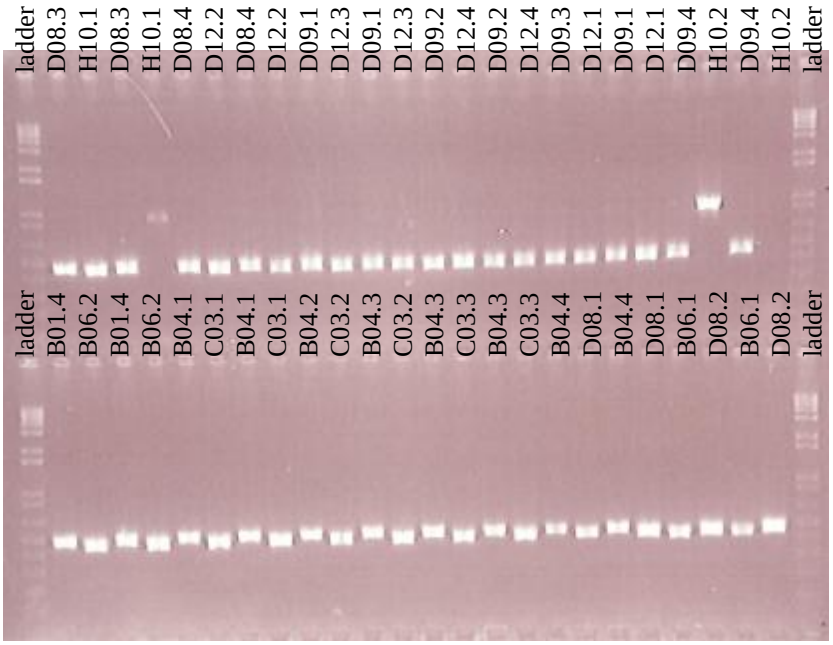


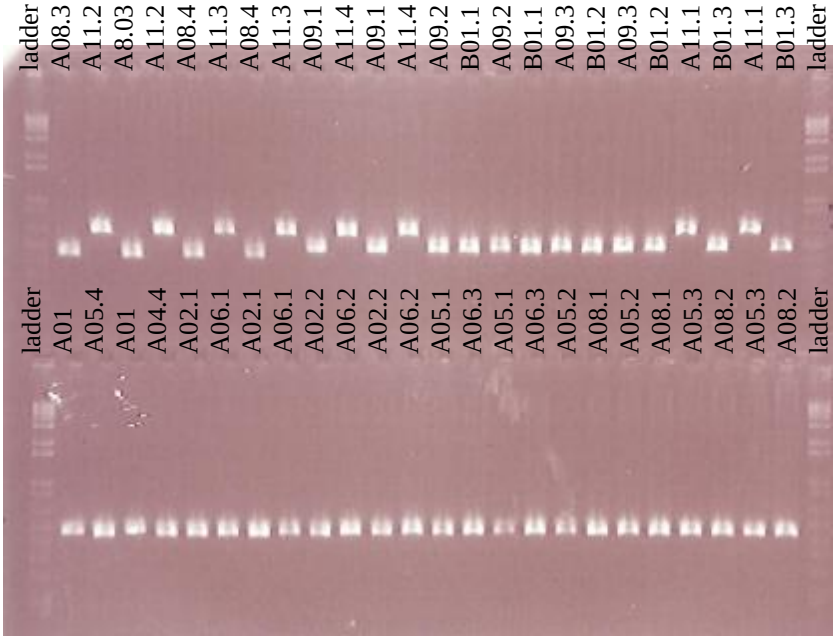
30-JAN


ladder
C05C-HIS
C05U-HIS
C05C-LAC
C05U-LAC
D08C-HIS
D08U-HIS
D08C-LAC
D08U-LAC
D09C-HIS
D09U-HIS
D09C-LAC
D09U-LAC
D12C-HIS
D12U-HIS
D12C-LAC
D12U-LAC
ladder
ladder
A11C-HIS
A11U-HIS
A11C-LAC
A11U-LAC
B01C-HIS
B01U-HIS
B01C-LAC
B01U-LAC
B02-C-HIS
B02U-HIS
B02C-LAC
B02U-LAC
B04C-HIS
B04U-HIS
B04C-LAC
B04U-LAC
B06C-HIS
B06C-HIS
B06C-LAC
B06U-LAC
C03C-HIS
C03U-HIS
CO3C-LAC
C03U-LAC
ladder
ladder
A01C-HIS
A01U-LHIS
A01C-LAC
A01U-LAC
A02C-HIS
A02U-HIS
A02C-LAC
A02U-LAC
A05C-HIS
A05U-HIS
A05C-LAC
A05ULAC
A06C-HIS
A06U-HIS
A06C-LAC
A06U-LAC
A08C-HIS
A08U-HIS
A08C-LAC
A08U-LAC
A09C-HIS
A09U-HIS
A09C-LAC
A09ULAC
ladder

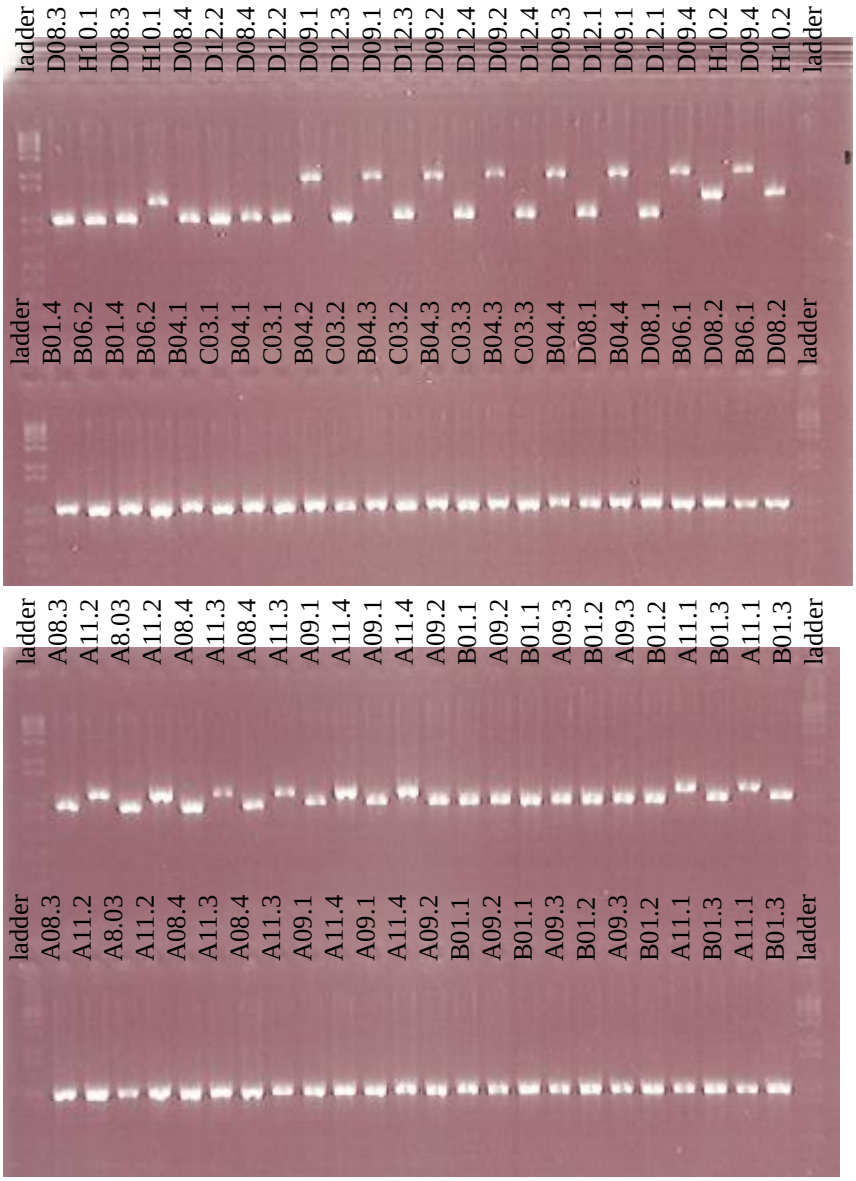
YEAST PCR TO SCREEN INTEGRATIONS

HIS





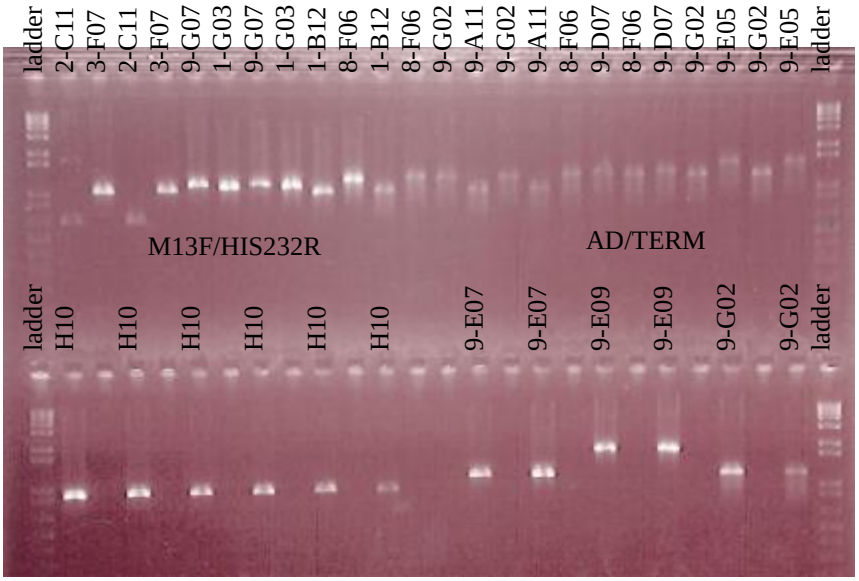
LAC


ladder D08.3 H10.1 D08.3 H10.1 D08.4 D12.2 D08.4 D12.2 D09.1 D12.3 D09.1 D12.3 D09.2 D12.4 D09.2 D12.4 D09.3 D12.1 D09.1 D12.1 D09.4 H10.2 D09.4 H10.2 ladder
ladder B01.4 B06.2 B01.4 B06.2 B04.1 C03.1 B04.1 C03.1 B04.2 C03.2 B04.3 C03.2 B04.3 C03.3 B04.3 C03.3 B04.4 D08.1 B04.4 D08.1 B06.1 D08.2 B06.1 D08.2 ladder
ladder A08.3 A11.2 A8.03 A11.2 A08.4 A11.3 A08.4 A11.3 A09.1 A11.4 A09.1 A11.4 A09.2 B01.1 A09.2 B01.1 A09.3 B01.2 A09.3 B01.2 A11.1 B01.3 A11.1 B01.3 ladder
ladder A08.3 A11.2 A8.03 A11.2 A08.4 A11.3 A08.4 A11.3 A09.1 A11.4 A09.1 A11.4 A09.2 B01.1 A09.2 B01.1 A09.3 B01.2 A09.3 B01.2 A11.1 B01.3 A11.1 B01.3 ladder

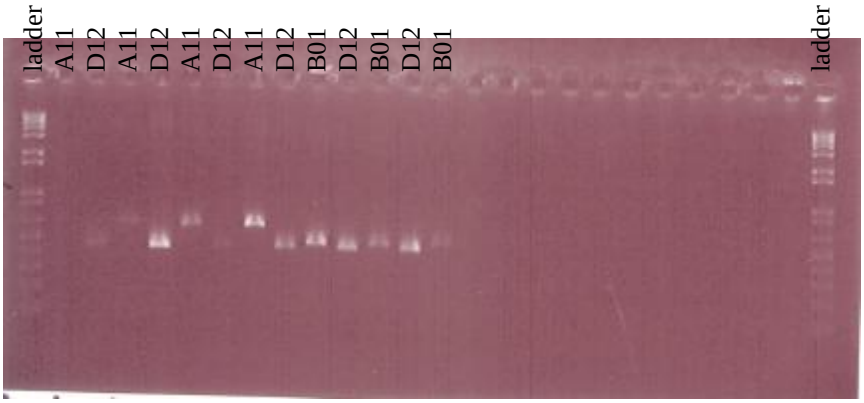
YEAST PCR SCREEN FOR POSITIVES

1-APRIL

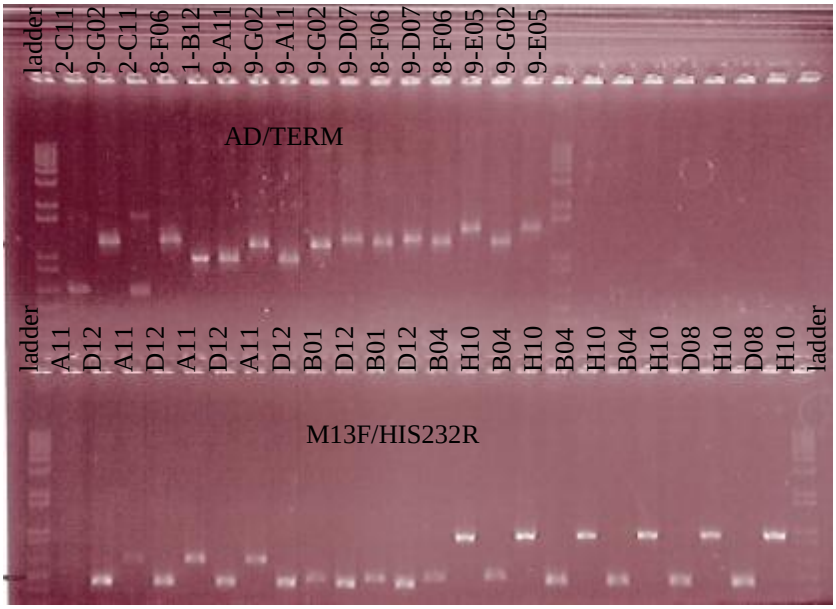
AD/TERM



M13F/HIS232R

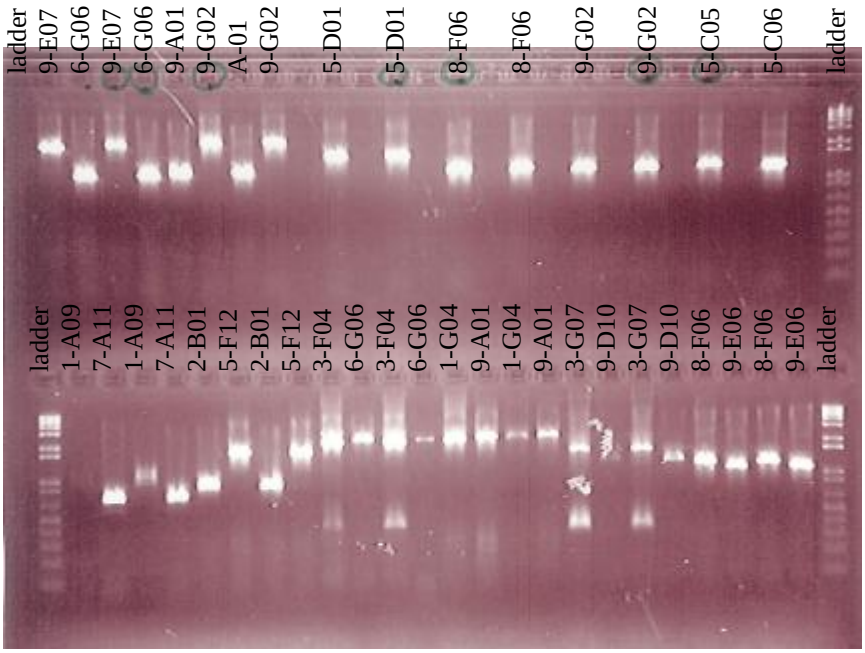


2-APRIL

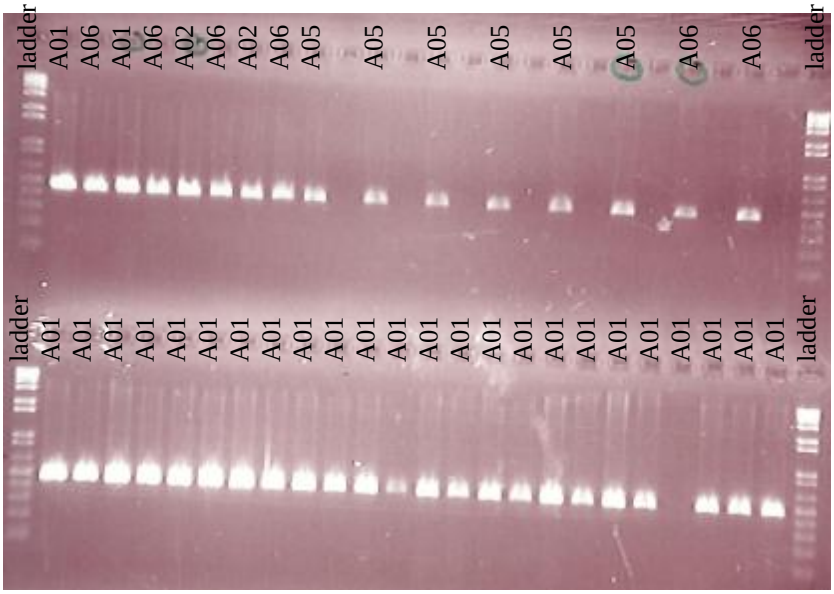


20-MARCH

AD/TERM



M13F/HIS232R



The following figures are of the mating assay in 1536 arrays for Day 7 of 5mM 3AT + Xgal plates of the first intron and TF preys.

FIrnh-1.3
Plate 1.

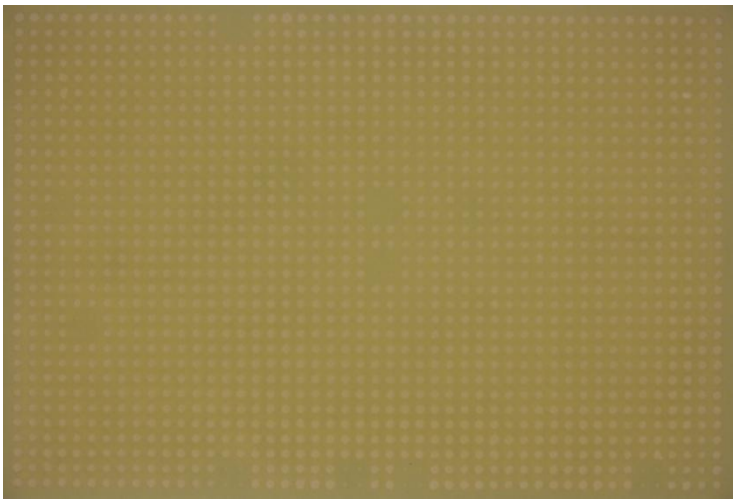


Plate 2

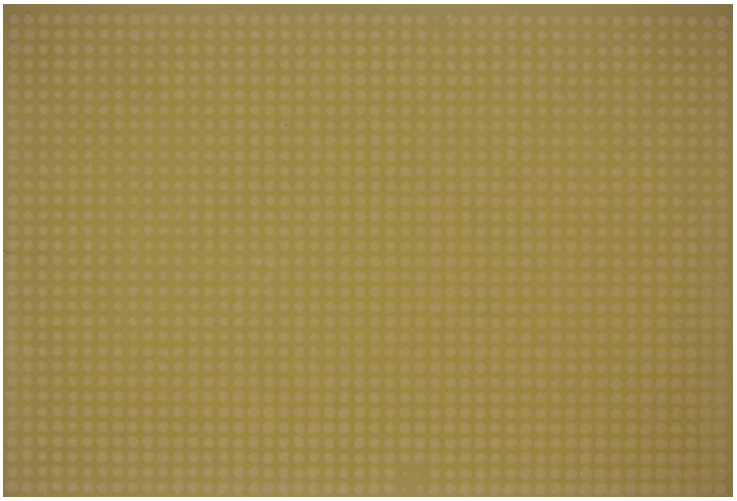
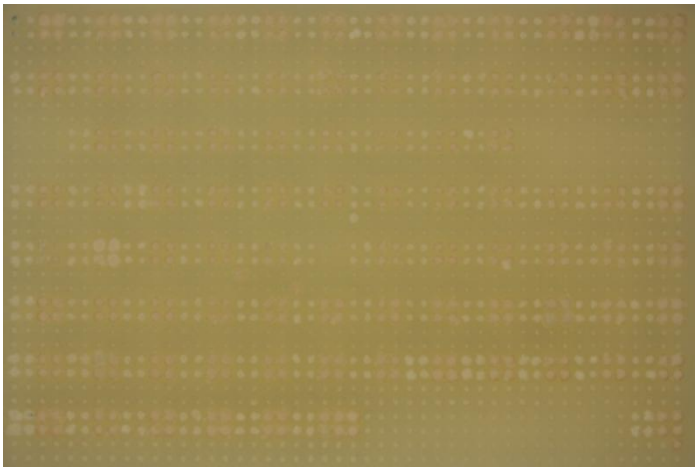


Plate 3



FIC49C3.5

Plate 1

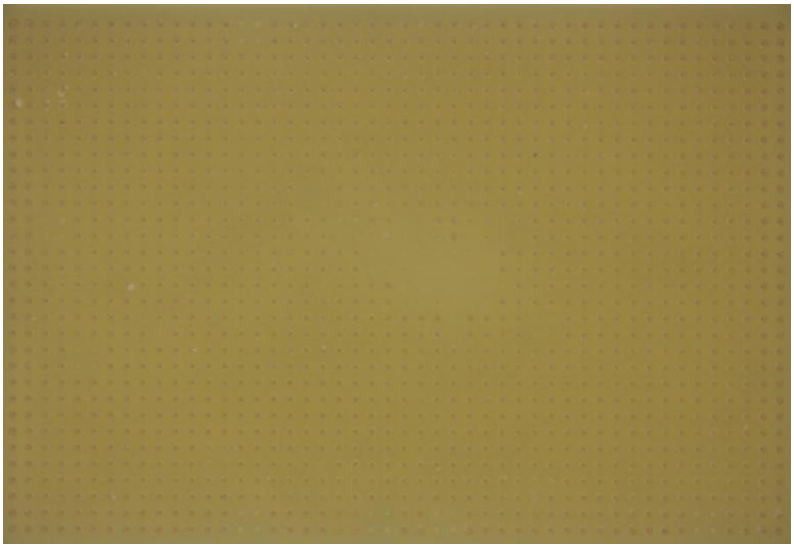


Plate 2

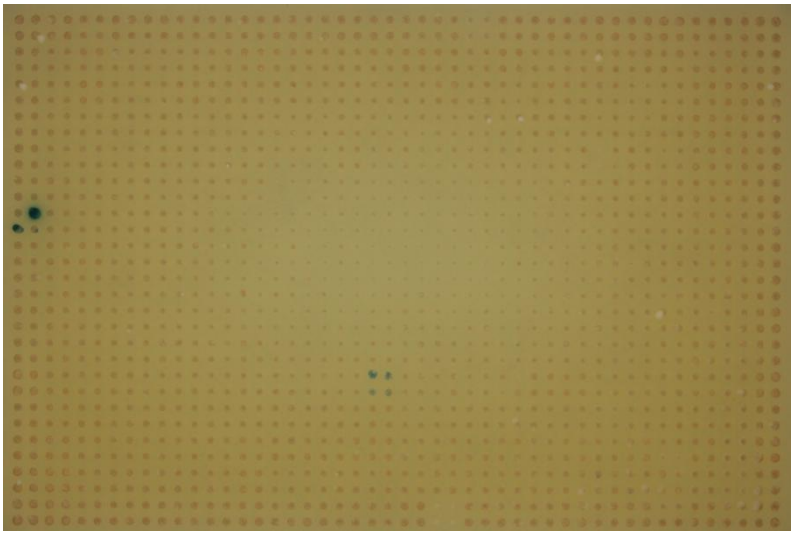
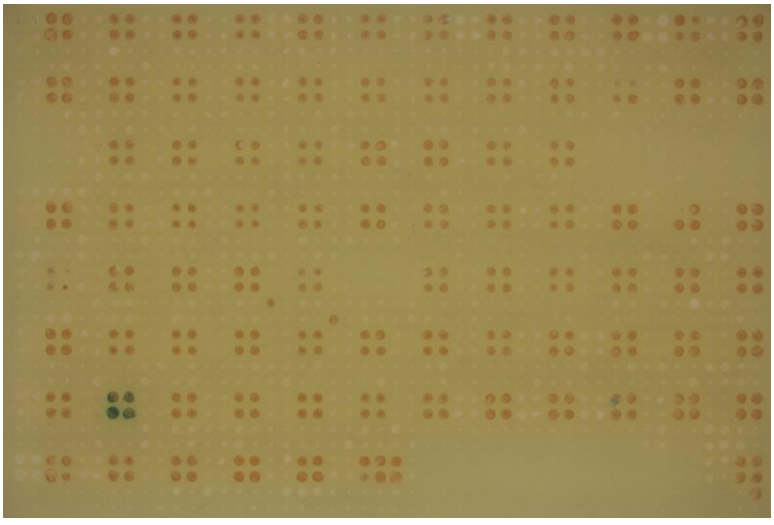


Plate 3



FIttr-21

Plate 1

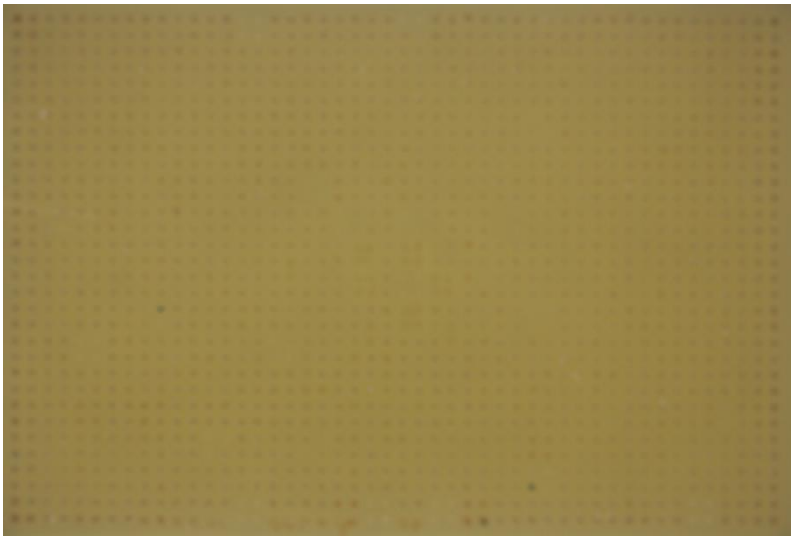


Plate 2

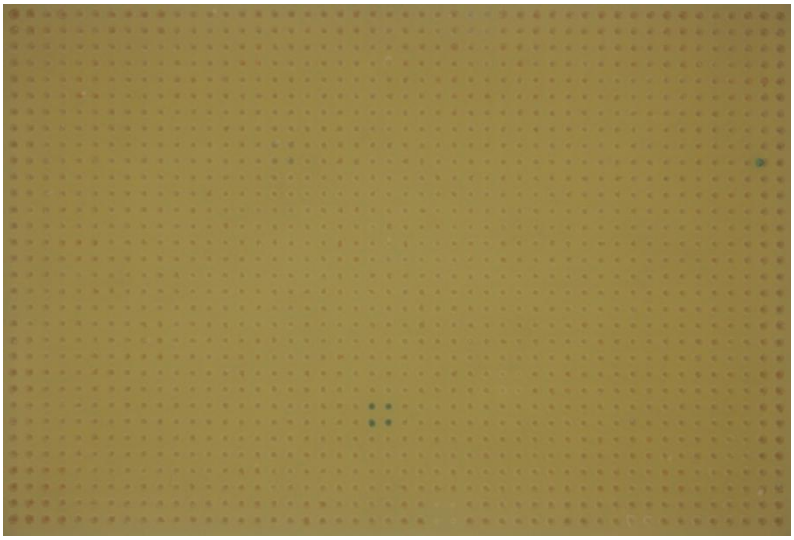
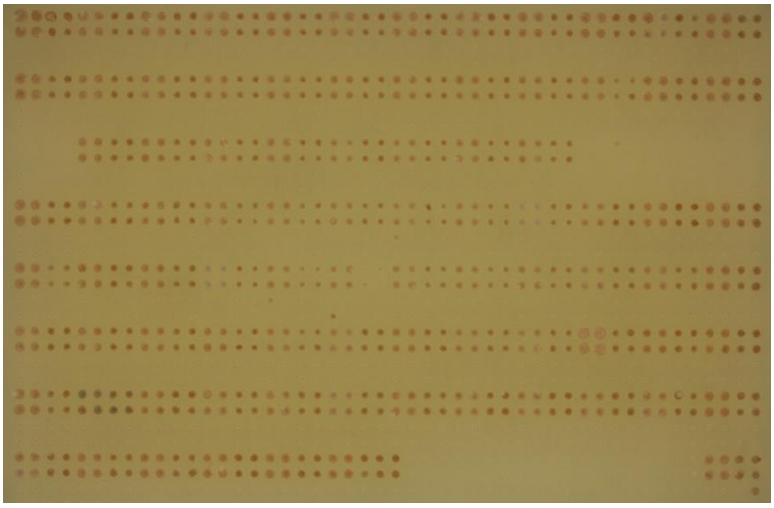


Plate 3



FIF52B11.5

Plate 1

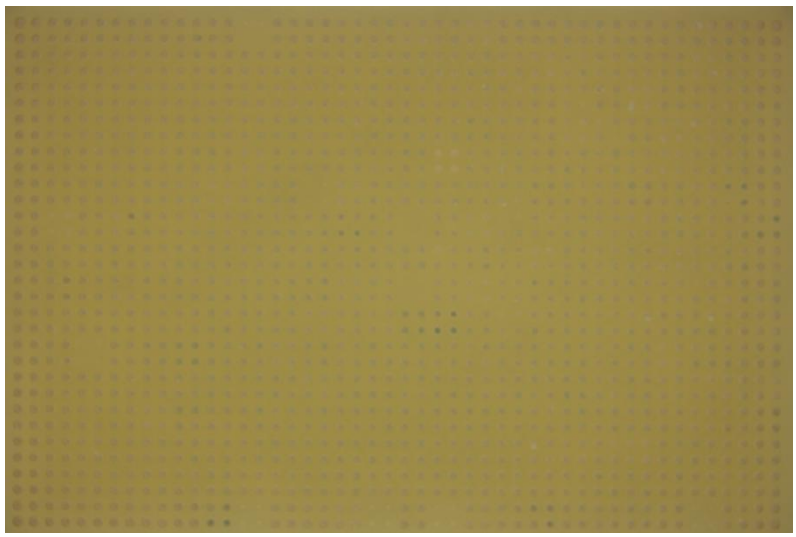


Plate 2

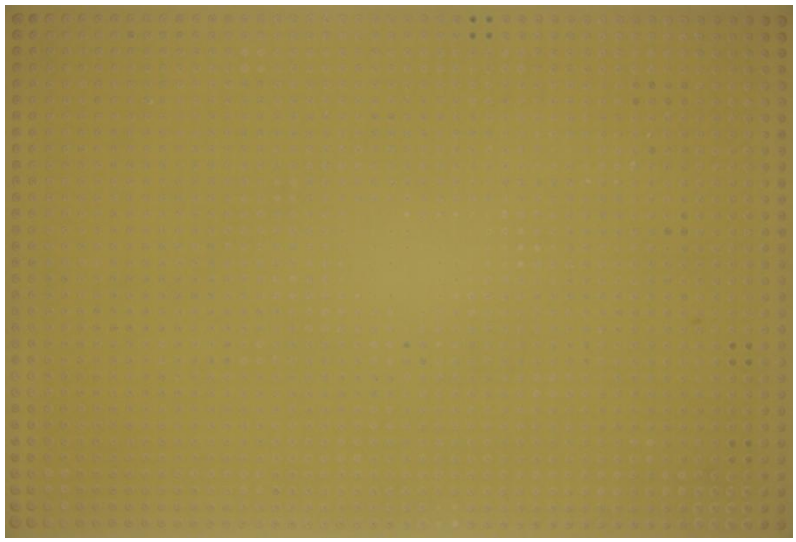
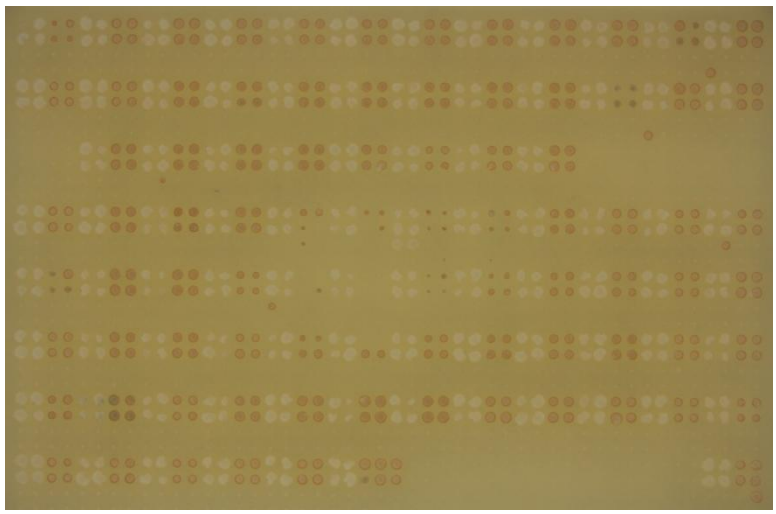


Plate 3



FIF53B3.3

Plate 1

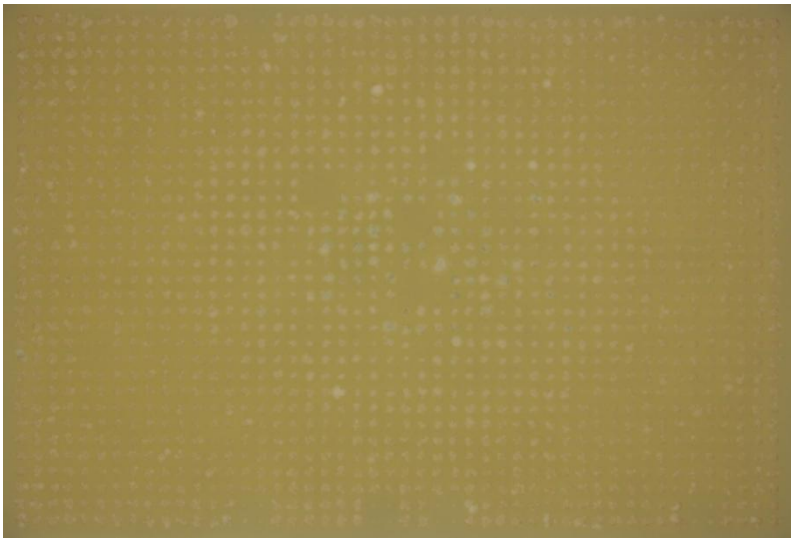


Plate 2

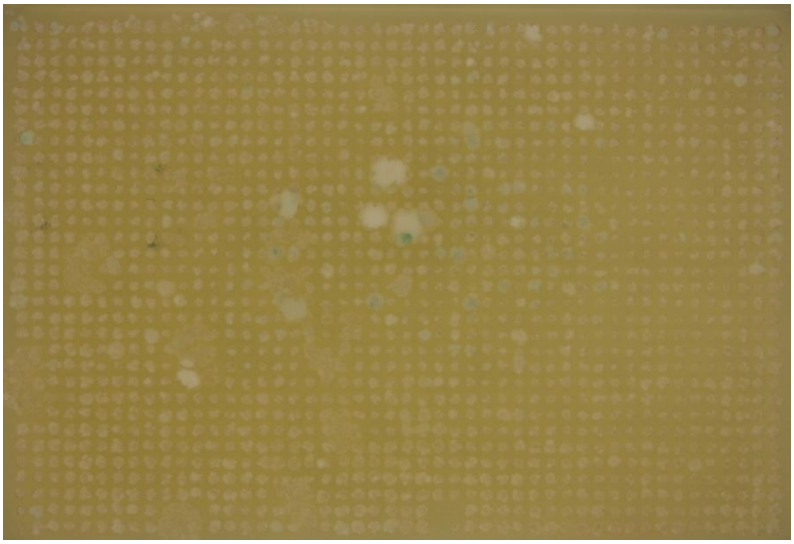
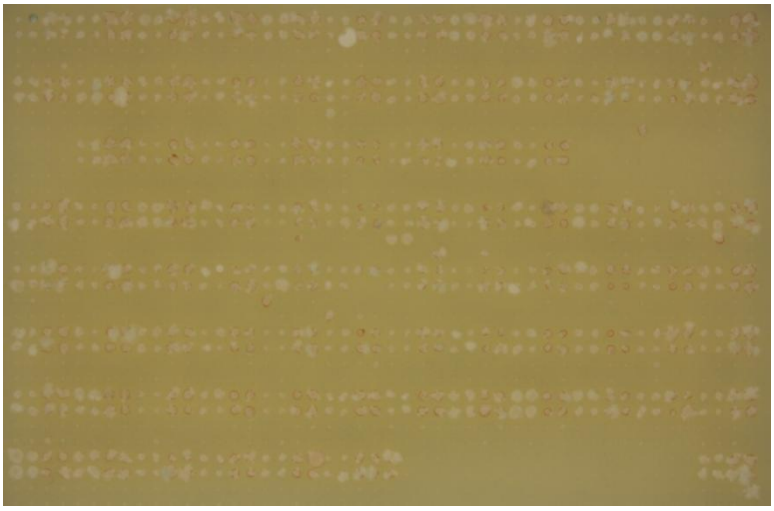


Plate 3



FIR06F6.6

Plate 1

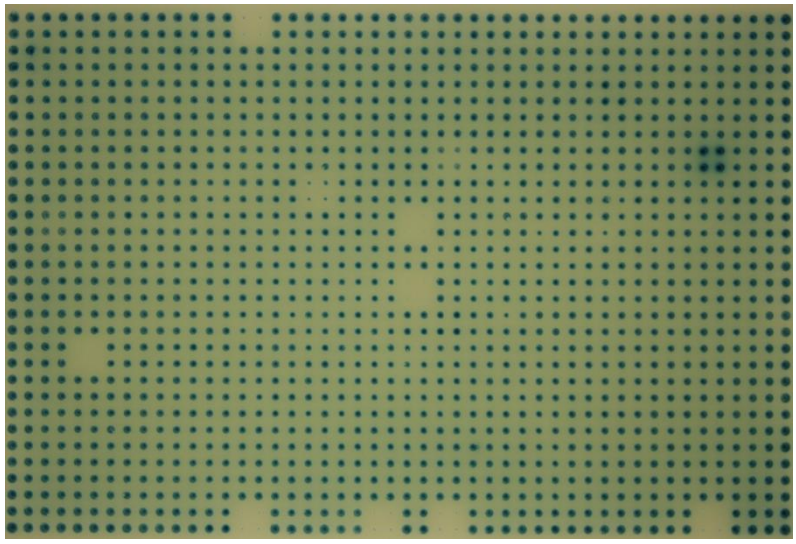


Plate 2

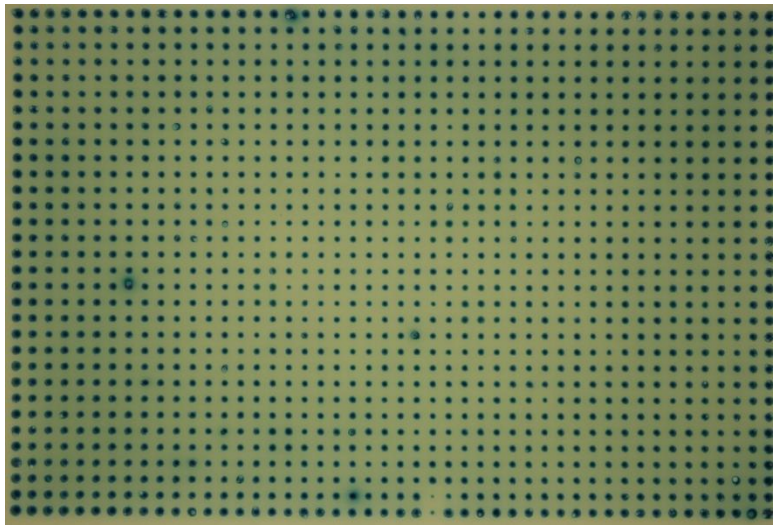
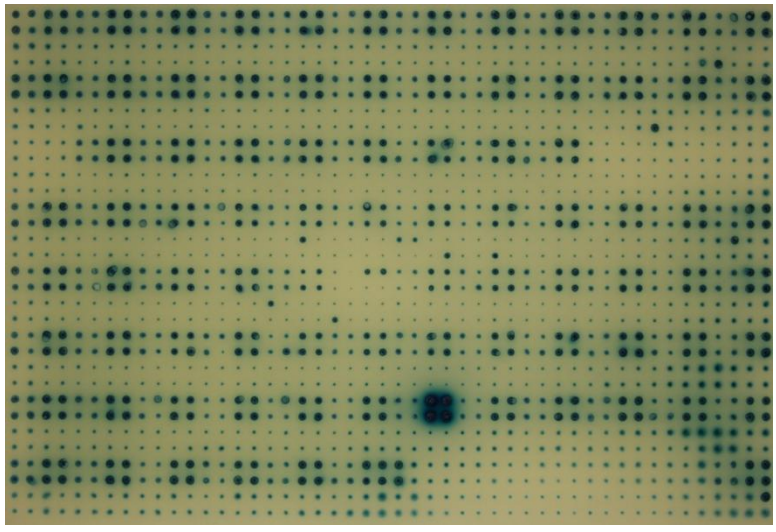


Plate 3



FIW01C9.1

Plate 1

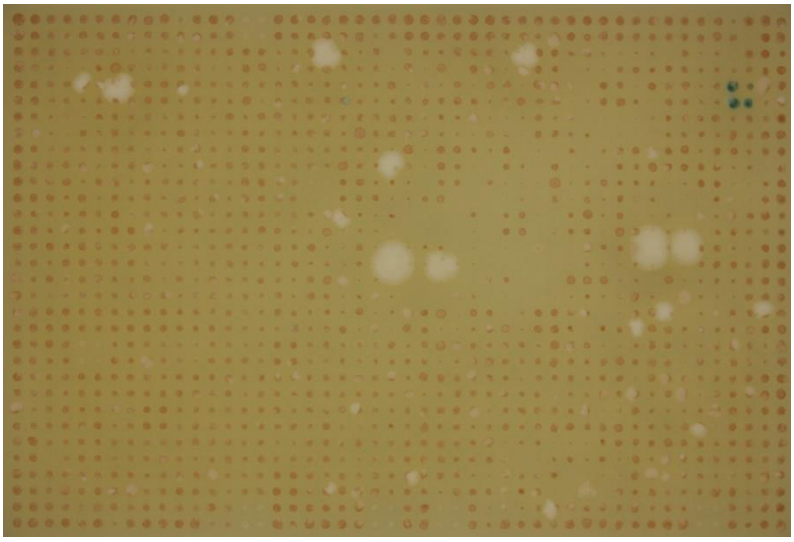


Plate 2

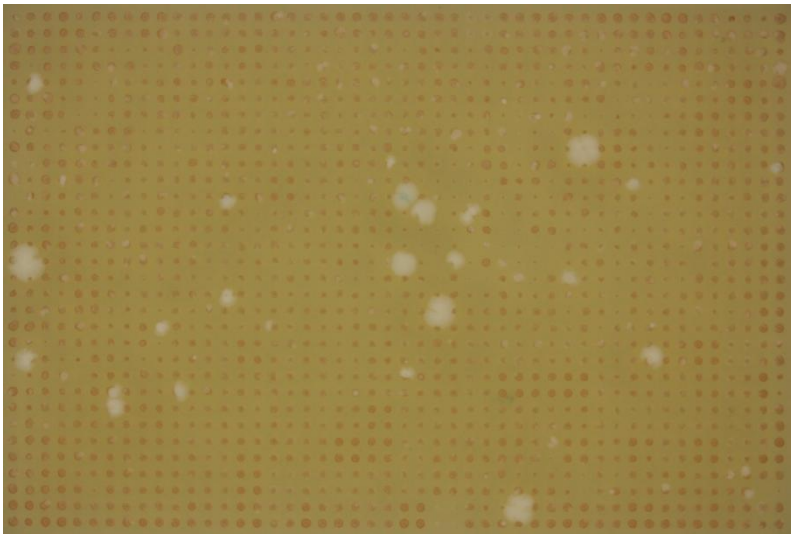
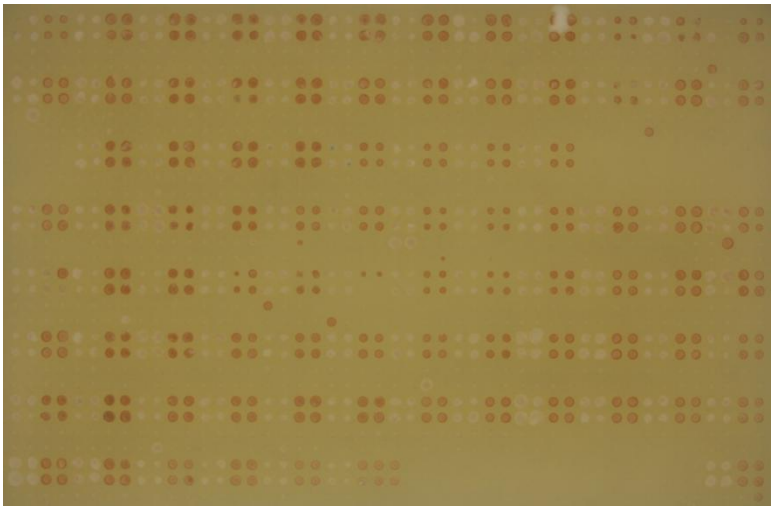


Plate 3



FIZK682.5

Plate 1

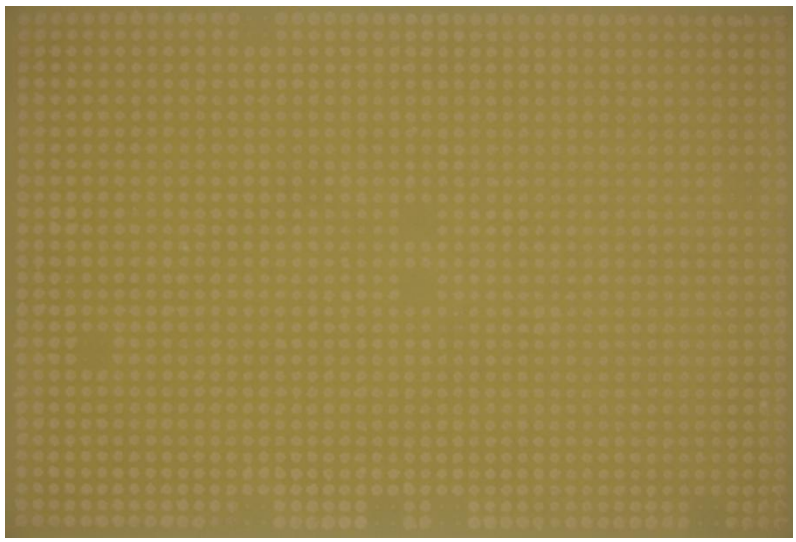


Plate 2

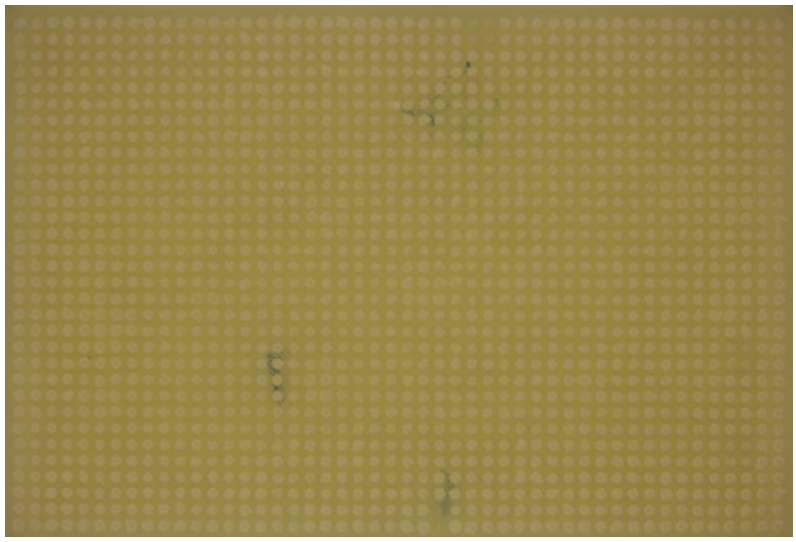
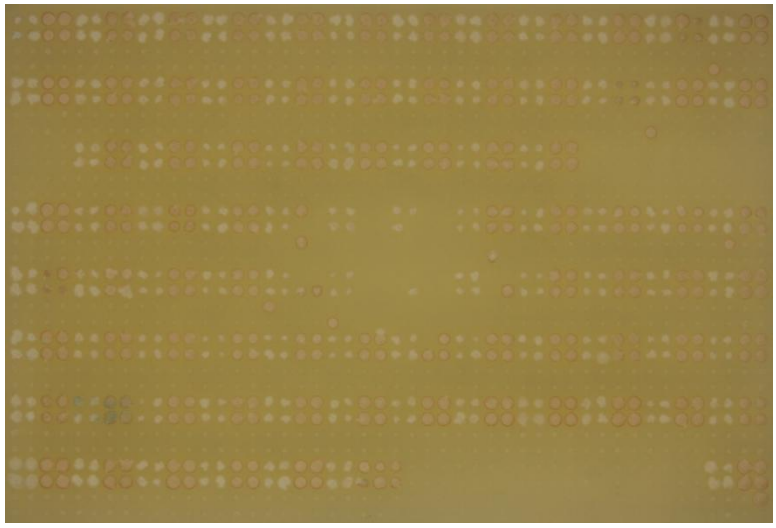


Plate 3



FIceh-10

Plate 1

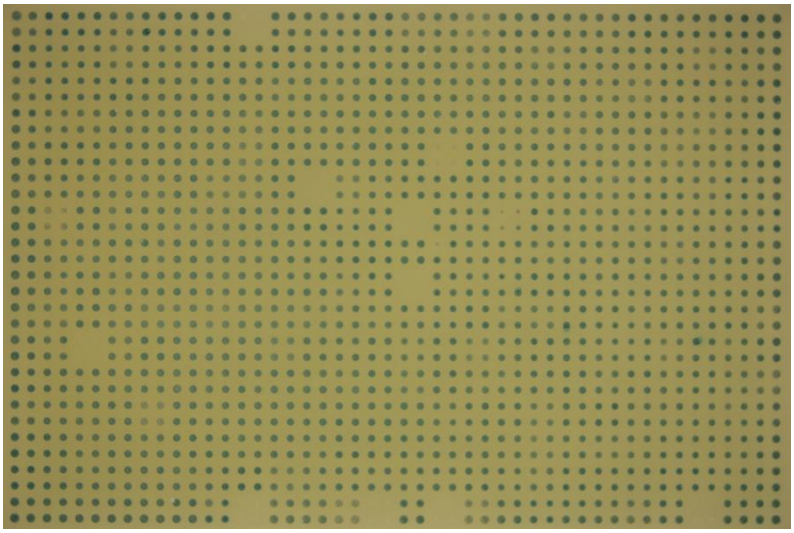


Plate 2

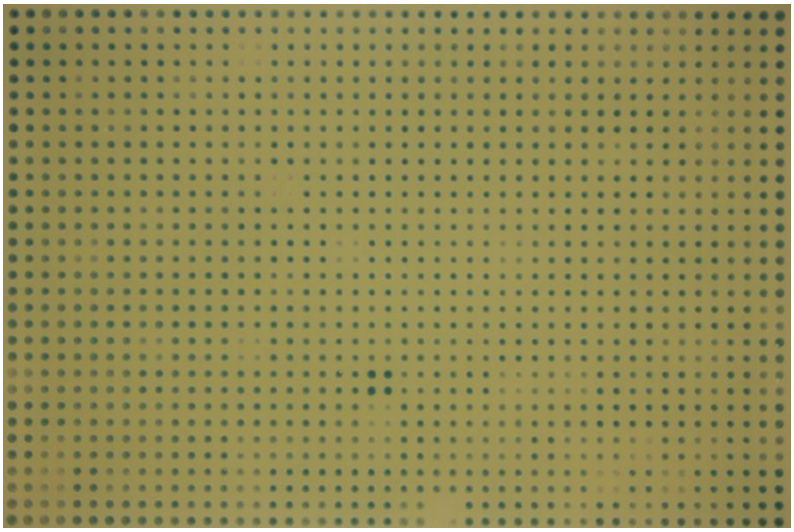
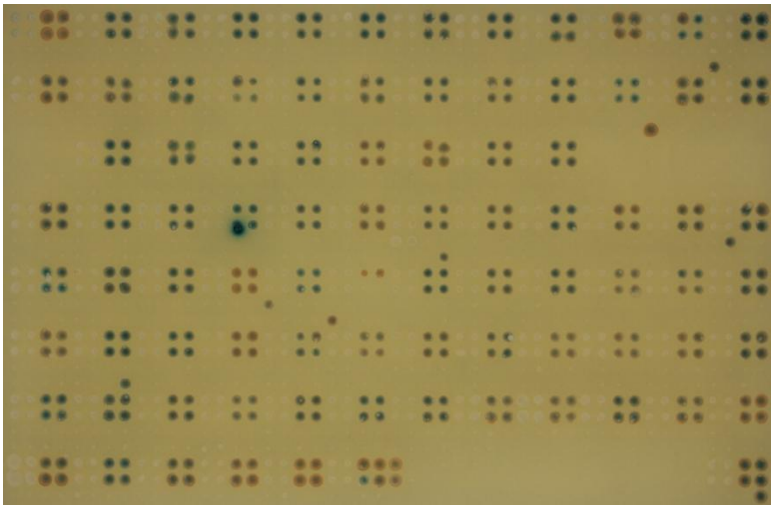


Plate 3



FIfog-3

Plate 1

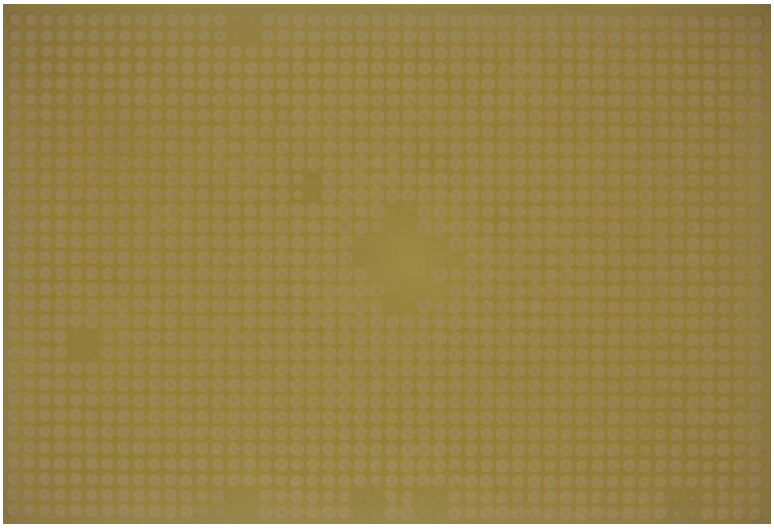


Plate 2

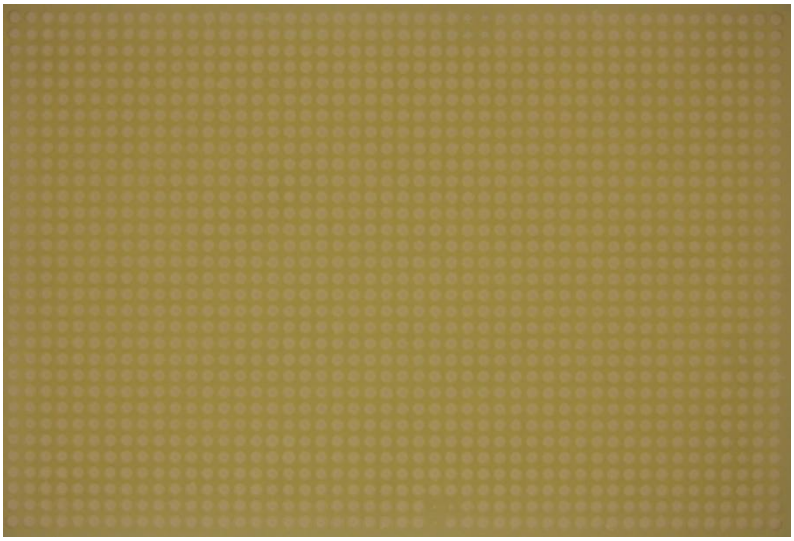
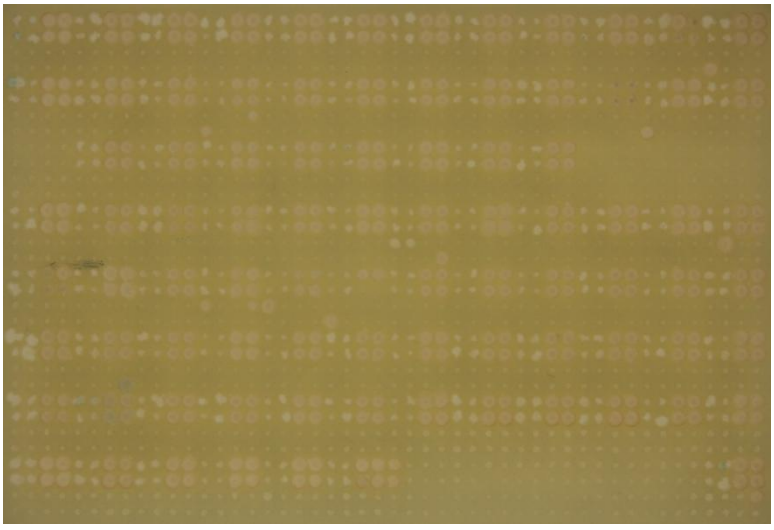


Plate 3



FIK07D4.6

Plate 1

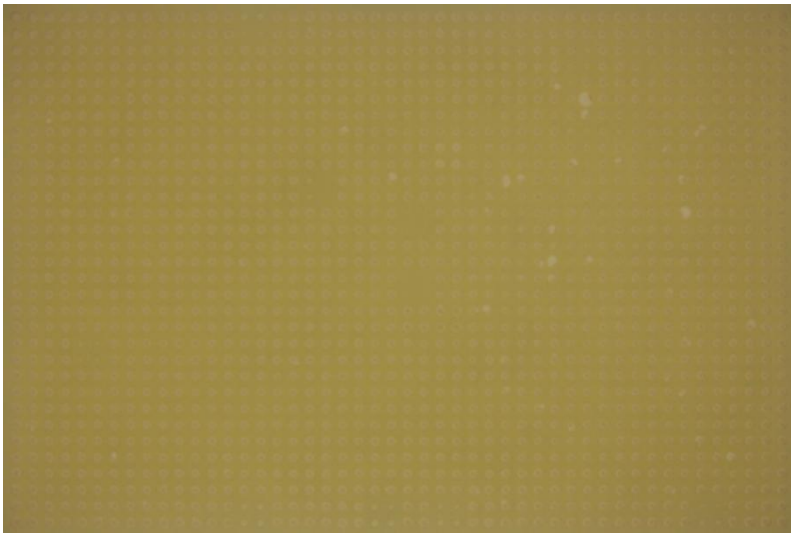


Plate 2

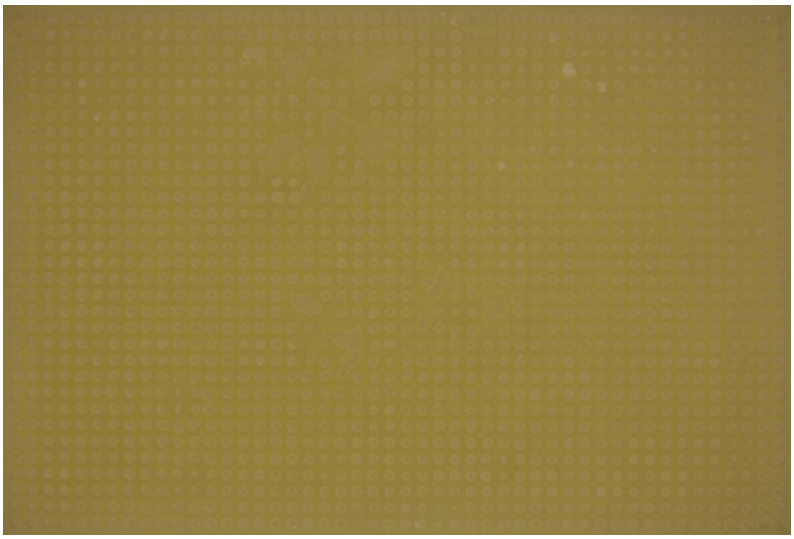
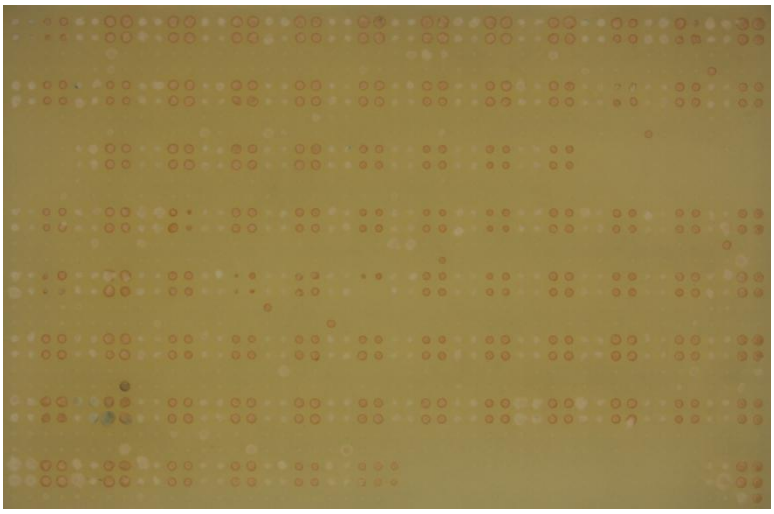


Plate 3



FInhr-38

Plate 1

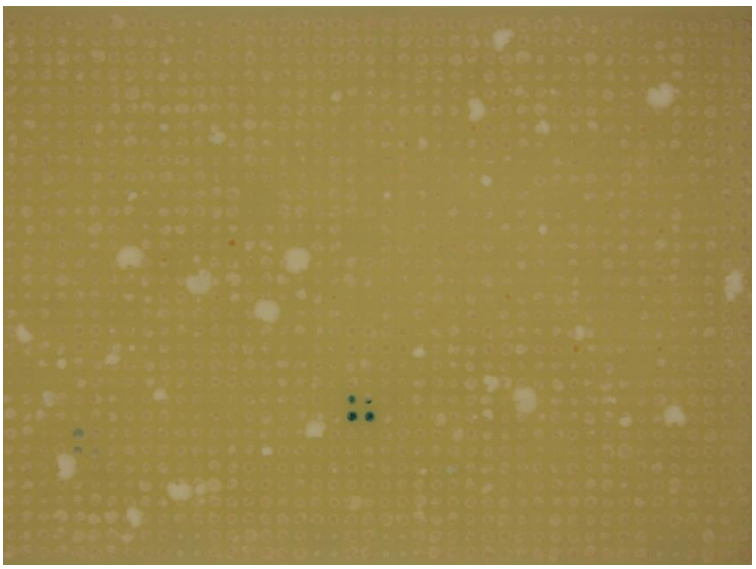


Plate 2

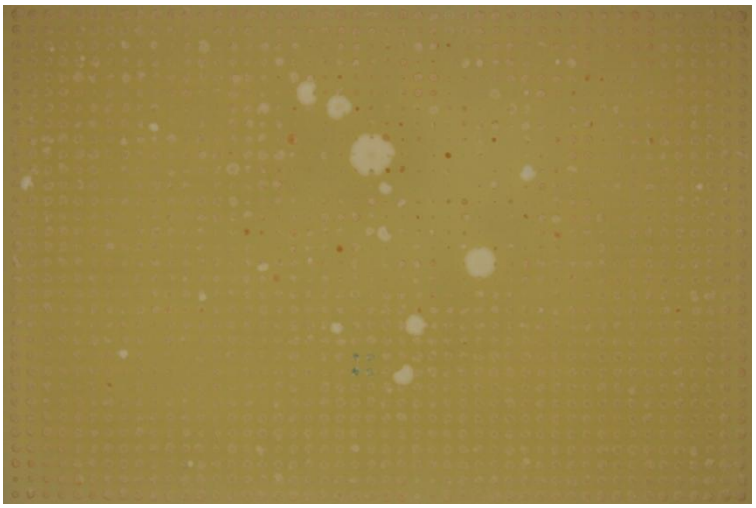
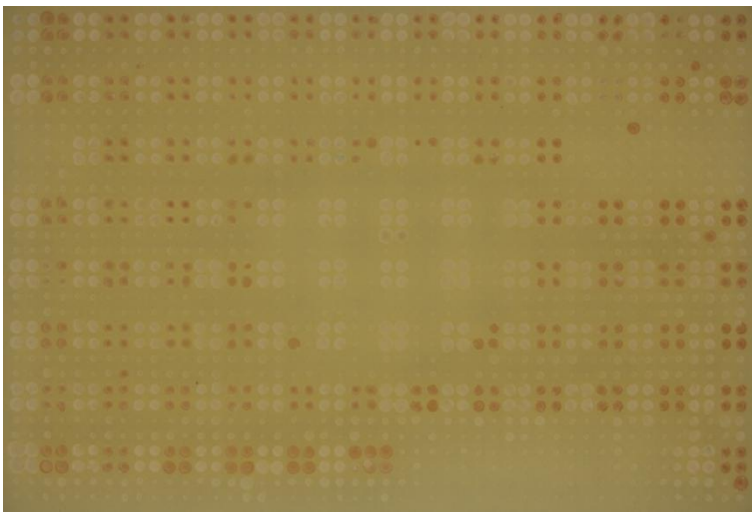


Plate 3



The following figures are of the mating assay in 1536 arrays for Day 7 of 5mM 3AT + Xgal plates of the promoter and TF preys.

Plin-32

Plate 1

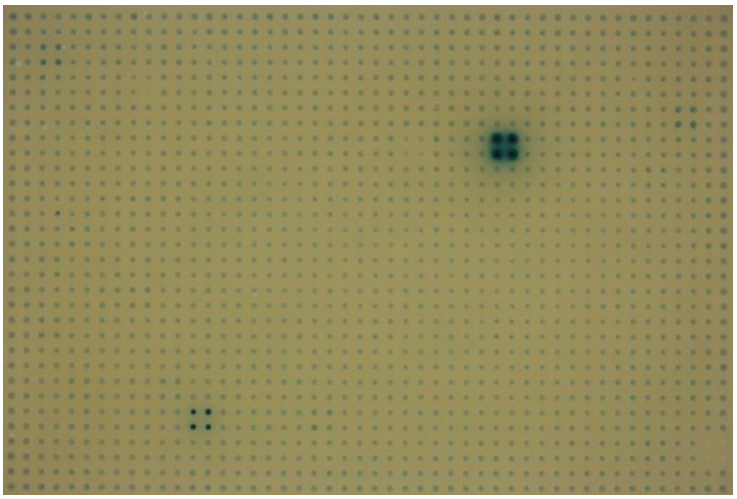


Plate 2

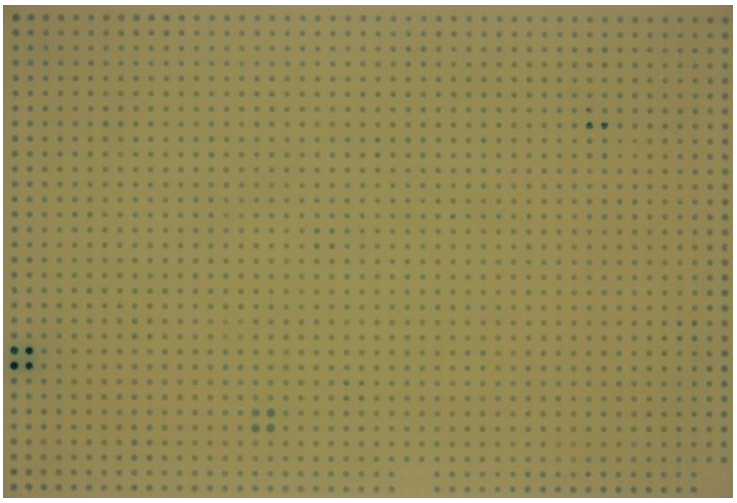
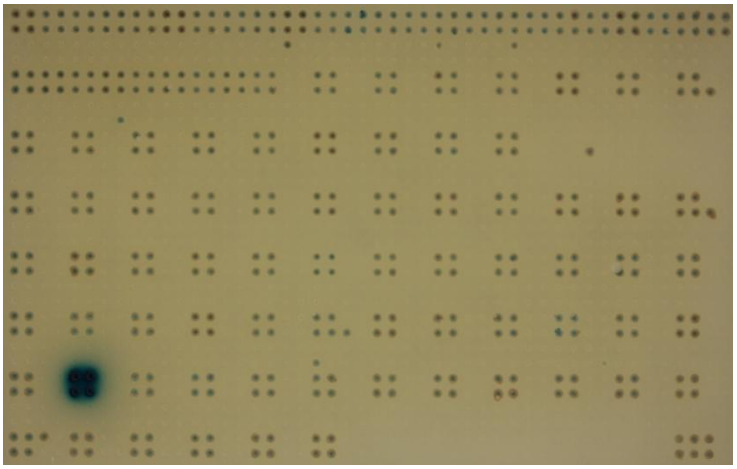


Plate 3



PC49C3.5

Plate 1

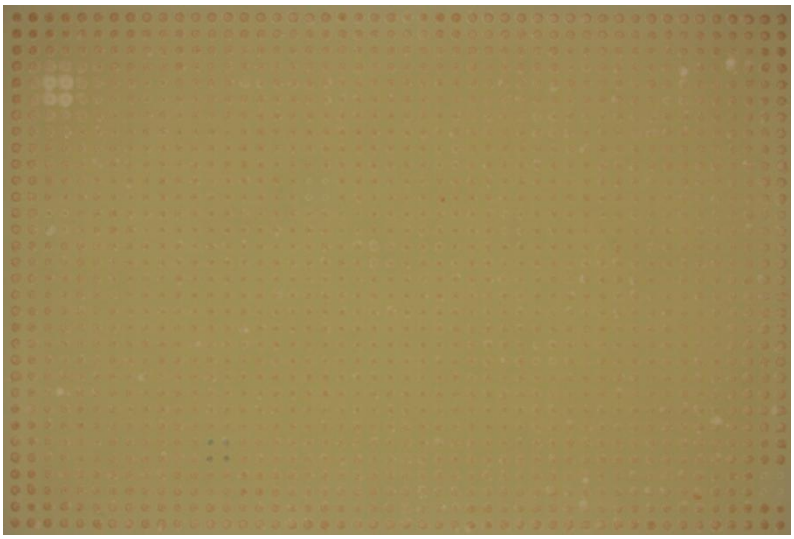


Plate 2

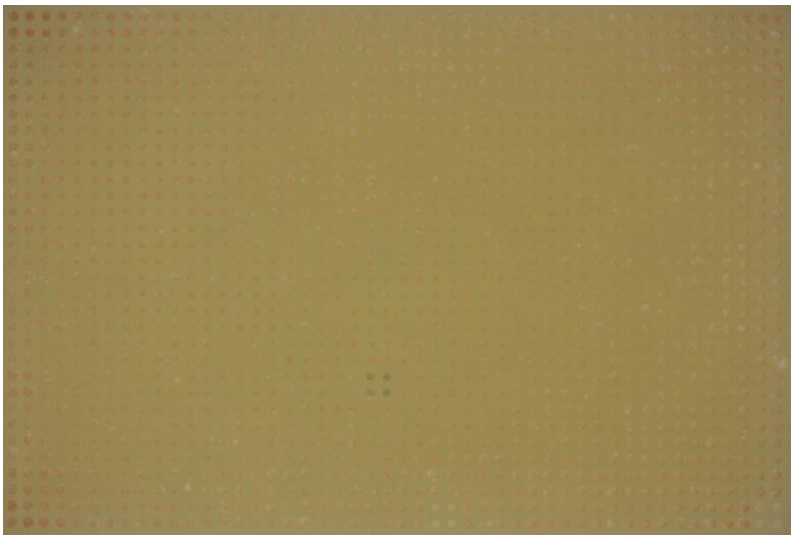
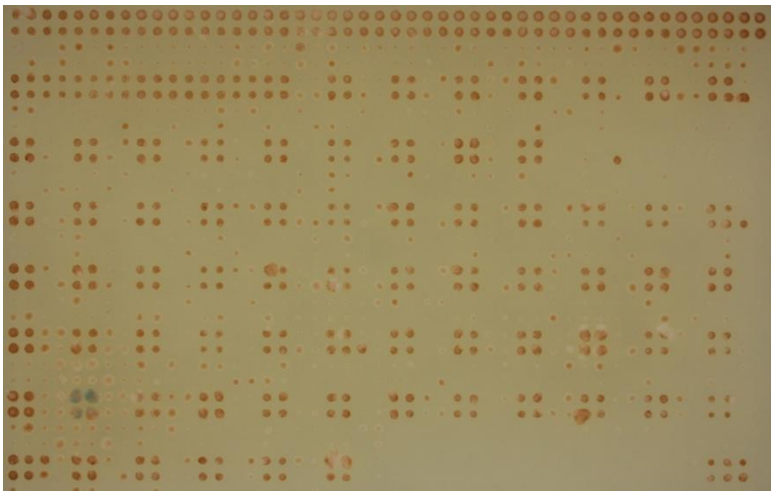


Plate 3



Pttr-21

Plate 1

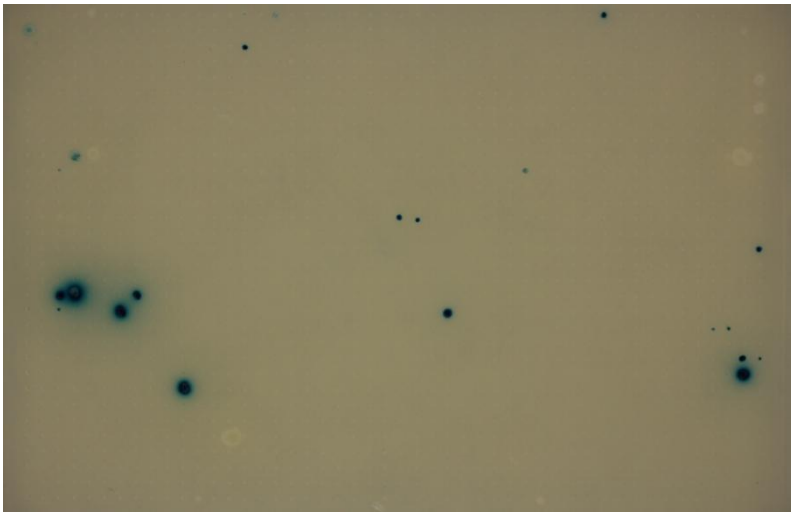


Plate 2

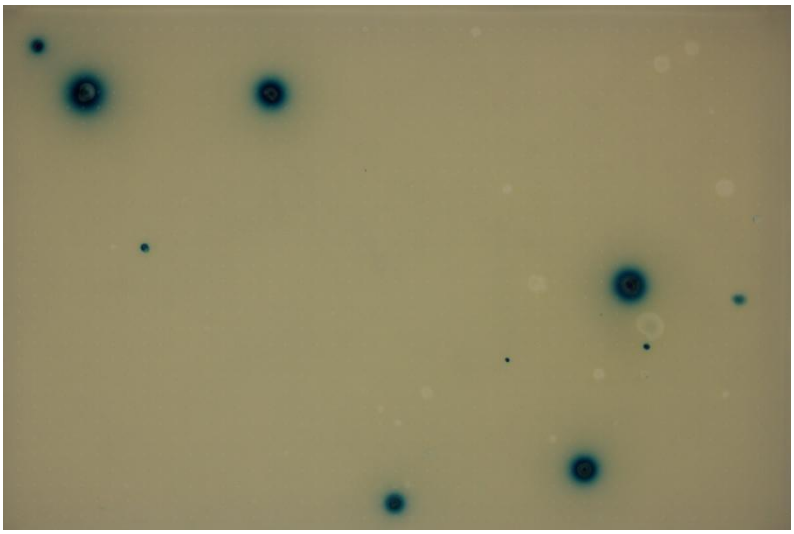
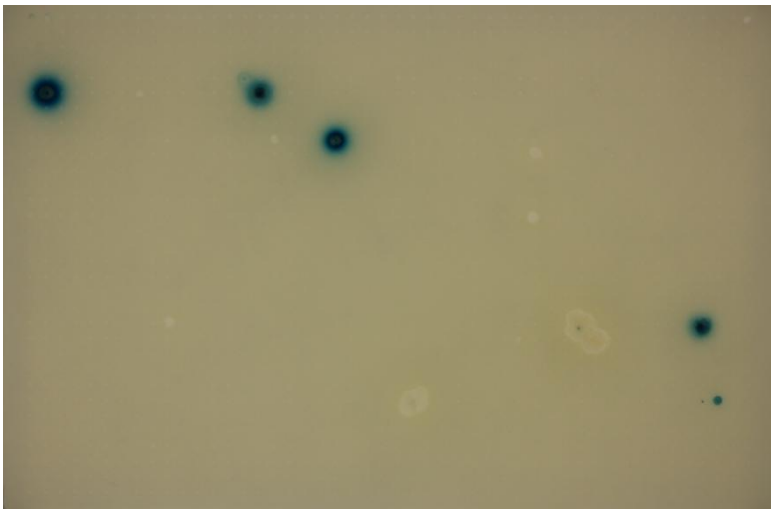


Plate 3



PF52B11.5

Plate 1

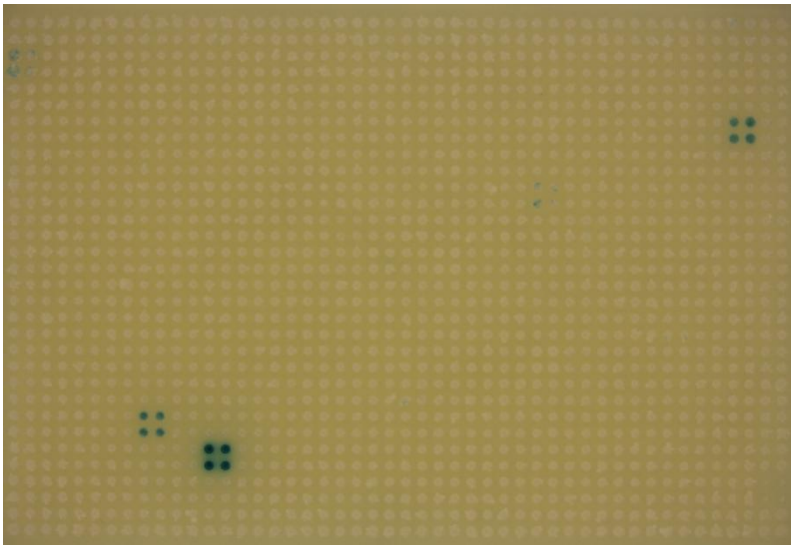


Plate 2

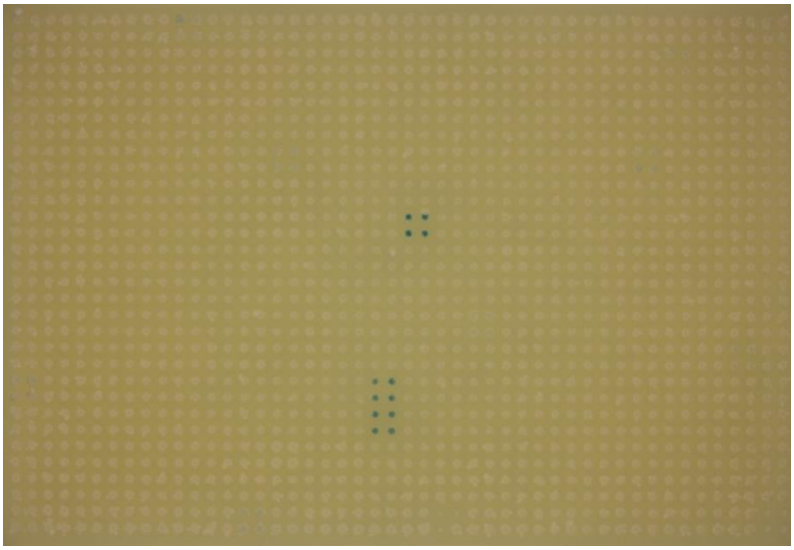
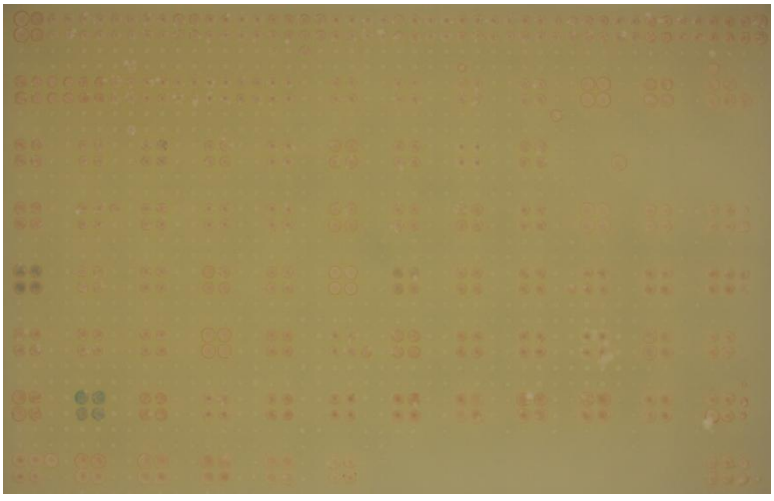


Plate 3



PR06F6.6

Plate 1

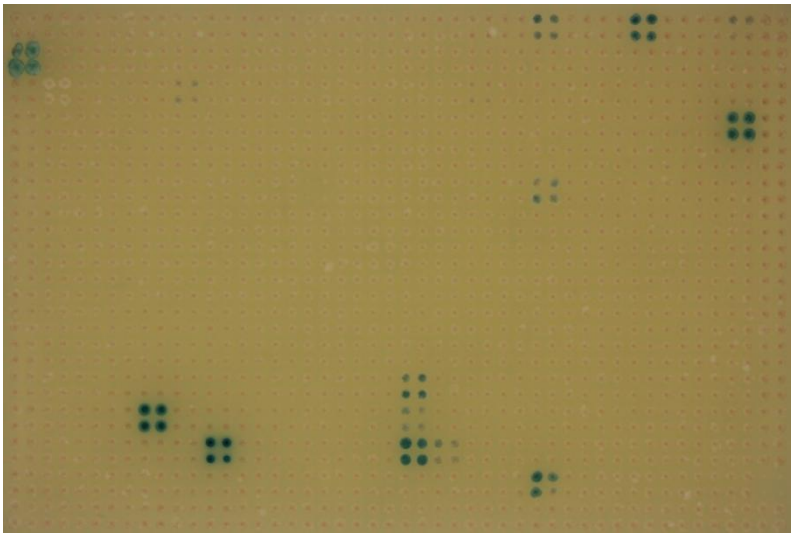


Plate 2

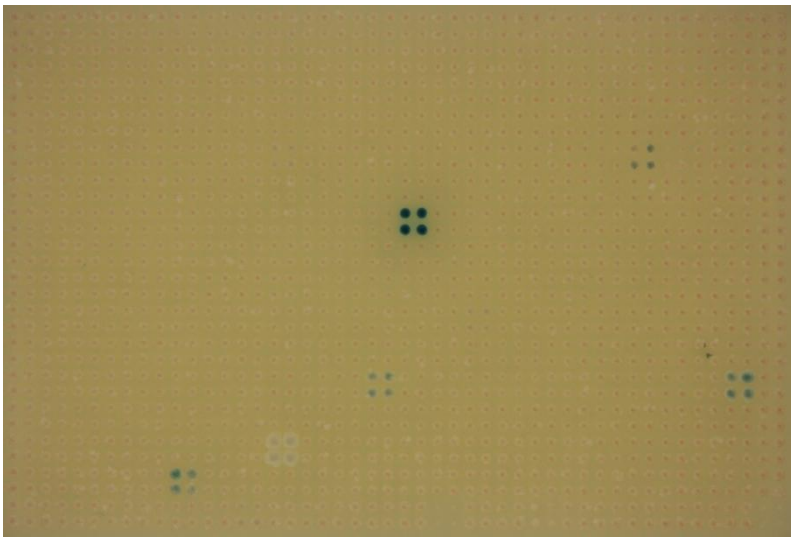
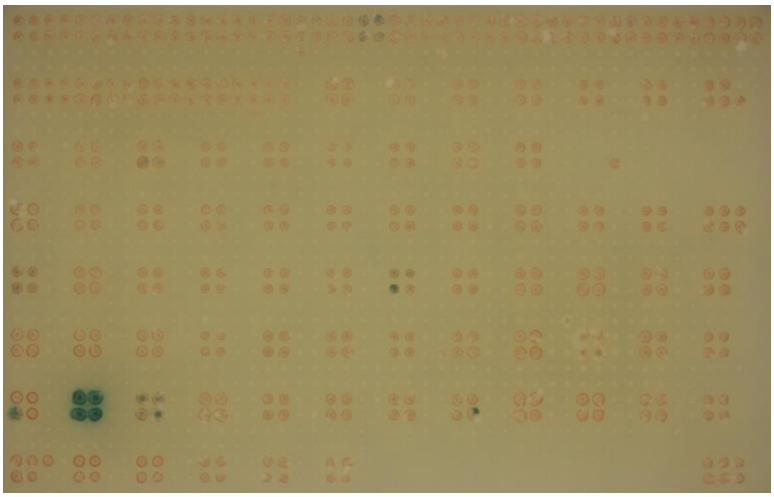


Plate 3



PZK682.5

Plate 1

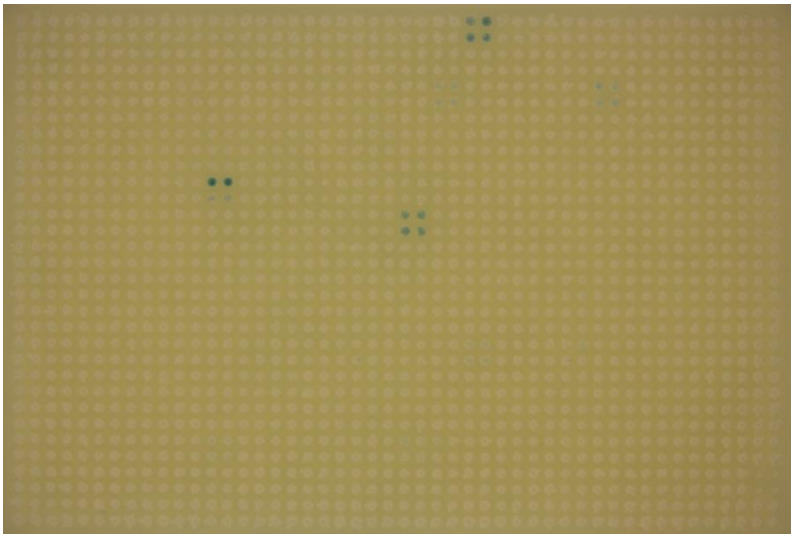


Plate 2

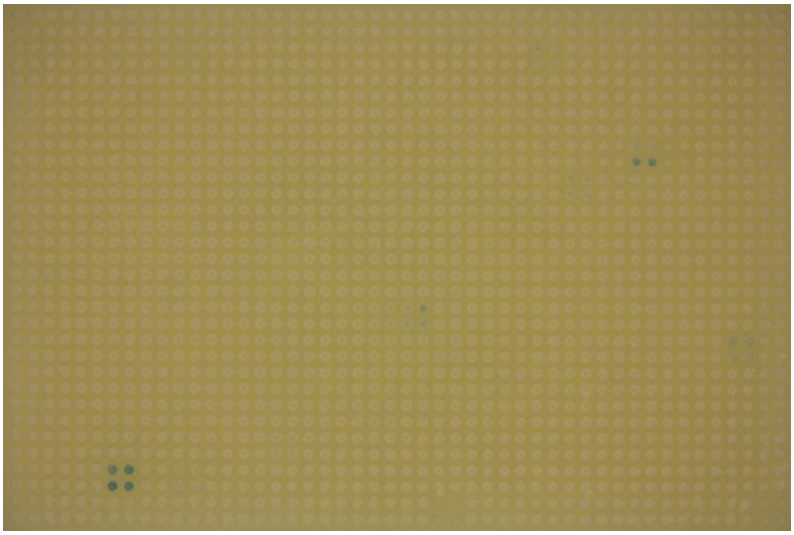
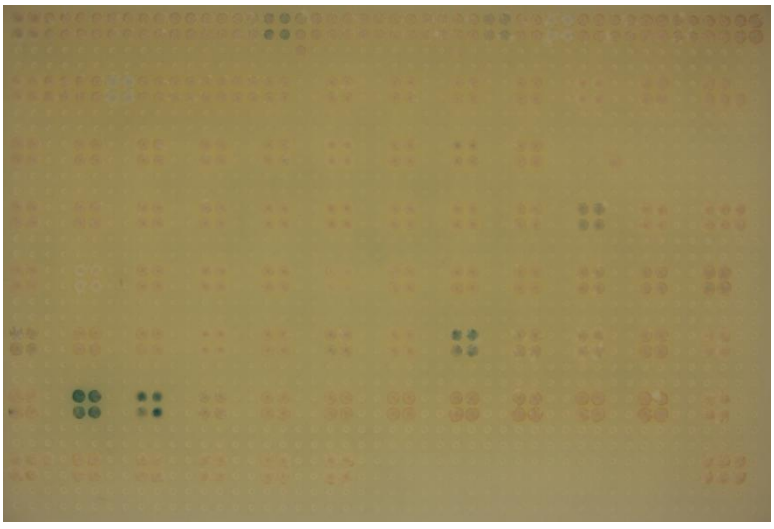


Plate 3



Pceh-10

Plate 1

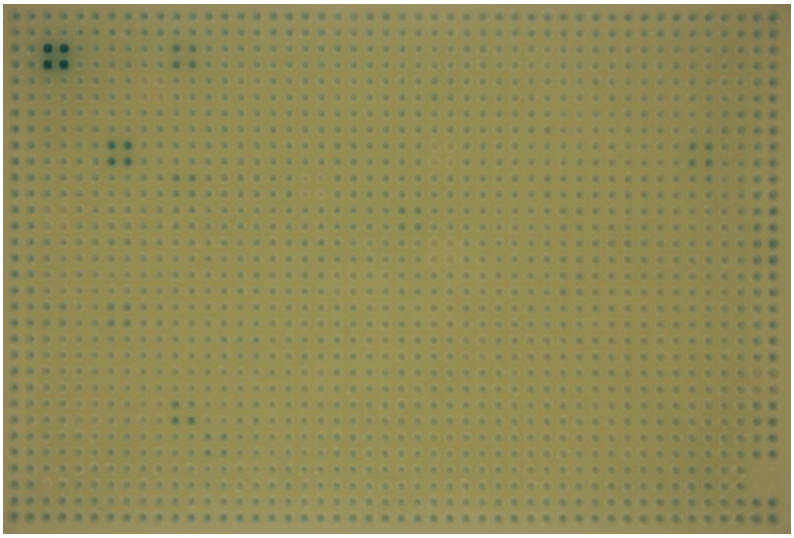


Plate 2

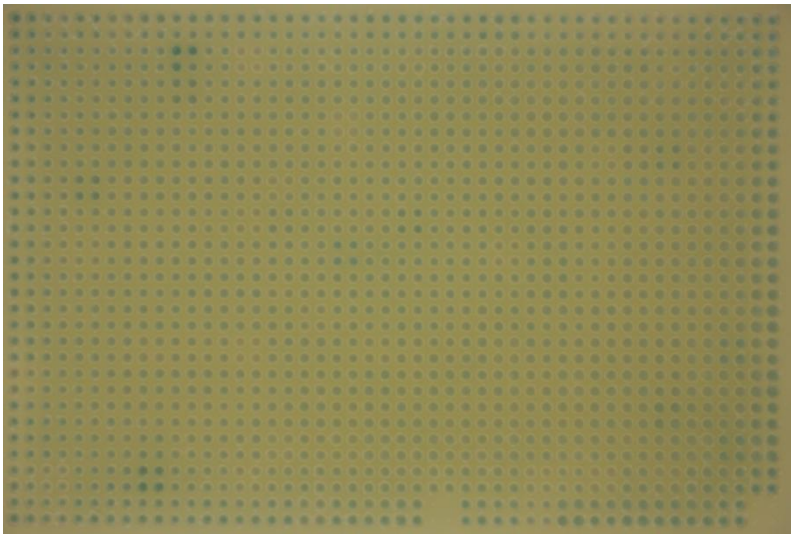
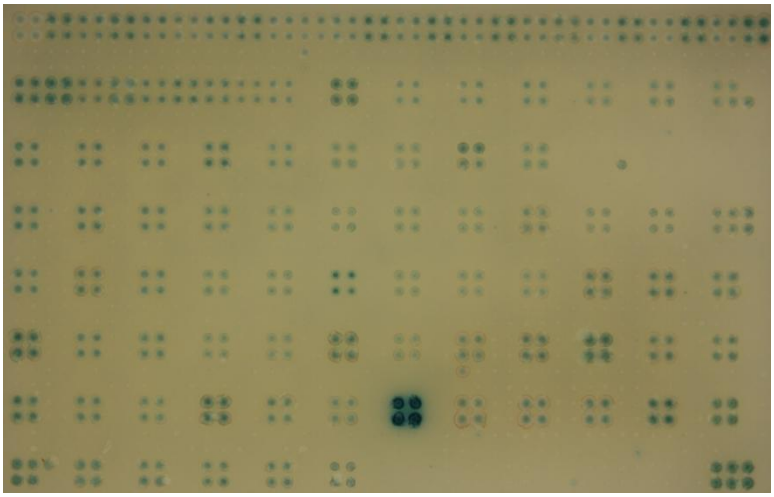


Plate 3



PK07D4.6

Plate 1

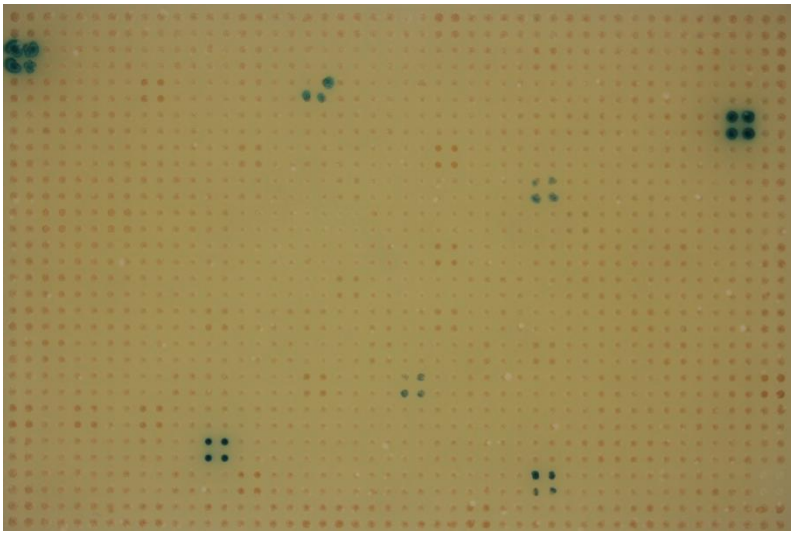


Plate 2

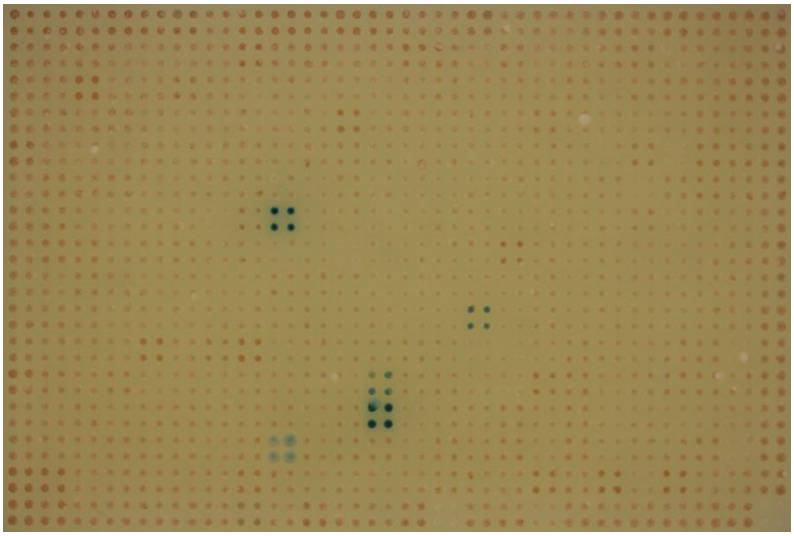
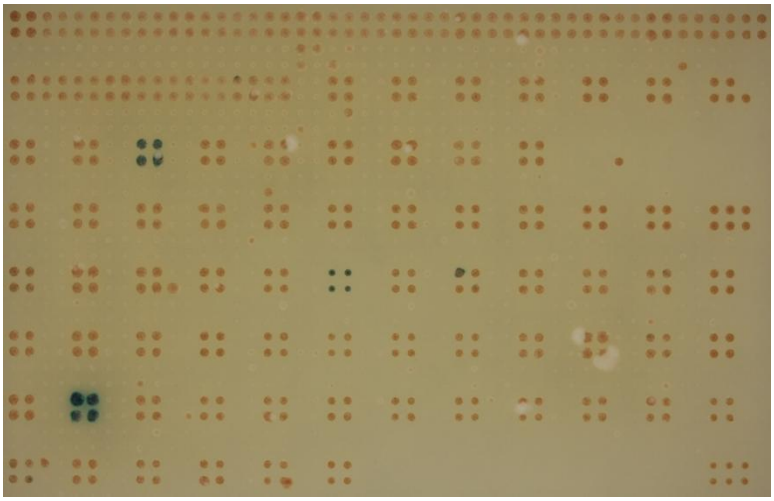


Plate 3



Pnhr-38

Plate 1

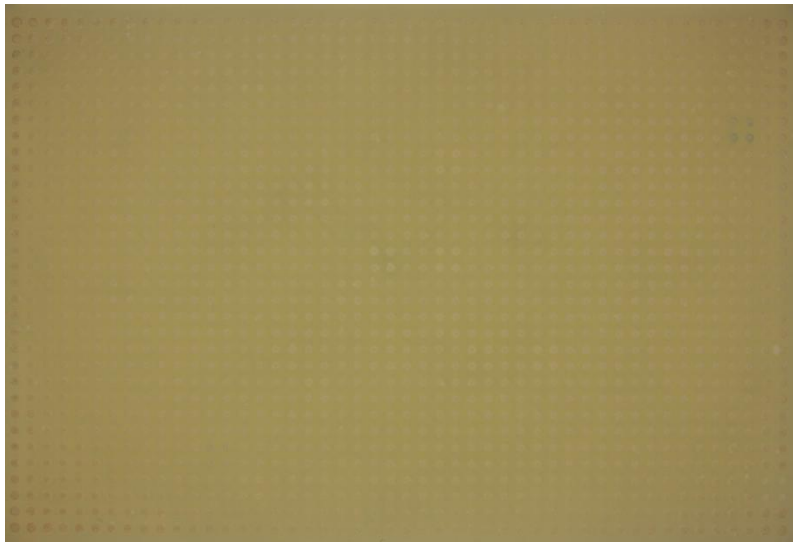


Plate 2

Plate 3