

Amplification of array material from model organisms- Yeast

Vishy Iyer, 2001. (Note: this protocol to appear in Molecular Cloning VI)

Saccharomyces cerevisiae (baker's, brewer's, or budding yeast) was the first eukaryote to have its entire genome sequenced [Goffeau, 1996 #16] and fittingly, it was the first organism whose entire genome was represented on a DNA microarray [DeRisi, 1997 #1]. In many respects, the design and manufacture of DNA microarrays for gene expression from this model organism represents an "easy" case. Known and predicted protein coding sequences amount to approximately 6000 genes and the annotation for coding sequences is thought to be reasonably robust. Two properties of the genome influenced the design and manufacture of microarrays from yeast. First, open reading frames (ORFs) in yeast are fairly small (the median size is 1.1 kilobase pairs and 80% of the ORFs are 2 kilobase pairs or less in length). Second, a very small fraction of the transcripts (about 230 of the 6000 genes, or 3.8 %) undergo splicing. These features, combined with a widespread interest in the availability of complete sets of cloned yeast genes, prompted the commercial production of primer pairs for amplifying the full-length sequence for every yeast ORF or gene. Research Genetics (Huntsville AL, <http://www.resgen.com>) synthesized PCR primer pairs specific for each ORF, that could amplify the entire sequence between the known or predicted start and stop codons, using yeast genomic DNA as the template. The primers incorporate restriction enzyme sites to enable cloning of the PCR products into standard plasmid vectors, and also have common 5' ends to permit re-amplification of the initial products with a single pair of flanking primers annealing to the common 5' ends.

Virtually all the published reports to date on yeast gene expression using spotted DNA microarrays have employed these full-length PCR product representations of yeast genes. The basic strategy thus is to use the standard collection of PCR primer pairs to amplify the entire complement of yeast genes. However, several technical and experimental considerations call for a modification or reassessment of this basic strategy.

Given that one is interested in making and using spotted DNA microarrays representing all yeast genes, an important and early decision is whether to represent the genes as PCR fragments or long oligonucleotides (on the order of 70-mers; see the section on production of long oligonucleotides). The use of oligonucleotides of course obviates the need for all the PCR and purification steps described in this protocol, and they have important additional advantages. They can be more easily designed to be complementary to unique or non-conserved sequences within every gene. With a full length PCR representation of the coding sequence, cross-hybridization is a complicating factor when considering the signal from a gene that shares more than 70 to 80 % sequence similarity with another gene.

Studying gene expression at the level of mRNA is just one of the myriad applications of DNA microarrays. For any application that requires the representation of *transcripts* (such as mRNA level measurements, identifying genes associated with distinct polysome fractions or with specific RNA binding

proteins), spotted oligonucleotide arrays are a viable and probably superior alternative. However certain applications call for representing featureless genomic loci on a microarray, and here, oligonucleotides are a less viable alternative. For instance, DNA microarrays can be used to map identical-by-descent loci using the technique of genomic mismatch scanning [Cheung, 1998 #17], to directly compare the genomes of related varieties of *S. cerevisiae* and to map the genome-wide *in vivo* binding distribution of DNA binding proteins on promoters [Reid, 2000 #19; Iyer, 2001 #12; Ren, 2000 #18]. Since the intergenic region upstream of a yeast gene contains its promoter, a PCR product allows it to be easily represented in its entirety on a microarray. The use of oligonucleotides would require either prior knowledge about the location of protein binding sites, or a tiling series that covers the entire region.

Careful advance planning can help make the entire operation proceed smoothly. It is useful to stock all the reagents required for the entire amplification, to avoid potential problems arising from variations in the reagents. The protocols discussed here assume that the entire set of yeast ORFs and/or intergenic regions are being amplified for use in printing microarrays.

Protocol

1.1 Obtaining the template for PCR (yeast genomic DNA).

Yeast genomic DNA is the template for amplifying coding regions and other loci by PCR. The genome that was sequenced was that of the strain called S288c (*MAT alpha mal gal2*). The coordinates of the genomic sets of PCR primer pairs pertain to this strain, and therefore genomic DNA from this strain is typically used as the template.

It is worth noting an interesting consequence of this if microarrays generated from the S288c sequence are used for direct genomic comparisons among related yeast strains. Only loci present in S288c and deleted in another related strain can be identified directly. Loci present in the other strain but deleted in S288c are not represented on the array, and therefore cannot be identified using the same microarrays. This is of little consequence when the arrays are used for most other kinds of experiments, even when the labelled probe is derived from a non-S288c strain (other than the confounding effects of different genetic backgrounds on gene expression).

To maximize the yield from a single round of PCR amplifications, it is recommended that a 100 microlitre PCR reaction be carried out for each gene or microarray element. 40 nanograms of yeast genomic DNA is required as the template for each reaction and for the entire set of yeast genes therefore, one needs approximately 250 micrograms of the template DNA.

1.1.1 Isolation of yeast genomic DNA.

The most straightforward option to obtain the required amount of template is to isolate genomic DNA from yeast. A 500 ml culture of yeast grown nearly to saturation in rich media yields about a milligram of genomic DNA, more than enough for amplifying all genes plus intergenic regions. In addition to the nuclear chromosomal DNA, the resulting DNA also contains mitochondrial and 2 micron plasmid DNA. An alternative to the following simple protocol is to use a commercial kit (e.g. Qiagen) for yeast genomic DNA isolation.

Materials

- § S288c strain. This can be purchased from ATCC, catalog # 26108 or Research Genetics, catalog # 95500.
- § 500 ml YPD (1 % Yeast extract, 2 % peptone, 2 % dextrose)
- § 100 ml lysis buffer (0.9 M sorbitol, 0.1 M Tris-Cl pH 8.0, 0.1M EDTA)
- § Zymolyase
- § 0.28 M beta-mercaptoethanol
- § 50 ml TE buffer (50 mM Tris-Cl pH 8.0, 20 mM EDTA)
- § 10 % sodium dodecyl sulphate (SDS)
- § 5 M potassium acetate
- § 10 mg/ml RNase A
- § 3 M sodium acetate pH 5.2
- § Ethanol

Method

1. Inoculate 500 ml YPD with 1 or 2 ml of an overnight culture of S288c. Grow this culture overnight at 30 °C in a shaking incubator. The culture should be near saturation (O.D. at 600 nm ≈ 5.0)
2. Harvest the cells by centrifugation at 2000 g at room temperature. Discard the supernatant.
3. Resuspend the cells in 25 ml lysis buffer, pipetting up and down with a 10 ml serological pipette and/or vortexing to ensure thorough resuspension. Add 2.5 ml of 0.28 M beta-mercaptoethanol and 2.5 ml of a freshly prepared, 0.5 mg/ml solution of Zymolyase made in lysis buffer. Mix thoroughly by swirling.
4. Divide the suspension equally into two smaller (30 ml) centrifuge tubes. Incubate the tubes at 37 °C for 1 hour, shaking at 200 rpm. This step digests the cell walls to yield spheroplasts. Harvest the spheroplasts by centrifugation at 2000 g at room temperature. Discard the supernatant.
5. Resuspend the spheroplasts in each tube in 10 ml TE buffer, pipetting up and down and/or vortexing to ensure thorough resuspension. Add 1 ml of 10 % SDS to each tube and mix by swirling the tube. Incubate the tubes for 30 minutes in a 65 °C water bath. This step lyses the spheroplasts.
6. Add 4 ml of 5 M potassium acetate to each of the two tubes. Mix by inverting a few times. Place the tubes on ice for 30 minutes to 1 hour.
7. Centrifuge at 5000 g at 4 °C to remove precipitated SDS and protein and cell debris. Collect all the supernatant carefully without disturbing the precipitated material. Divide the supernatant equally (approximately 30 ml total) among 4 centrifuge tubes.
8. To each tube, add 2 volumes (~ 15 ml) of 100 % ethanol. The nucleic acid should precipitate immediately. Let the tubes stand for 5 minutes at room temperature.
9. Centrifuge at 5000 g at 4 °C to pellet the DNA. Resuspend the pellets in 1 ml each of TE buffer. Resuspension might require breaking up the pellet with a 1 ml disposable pipettor tip and pipetting up and down. Consolidate all the solutions into one centrifuge tube (~ 4 ml total)
10. Add 20 microlitre of 10 mg/ml RNase A. Incubate at 37 °C for 30 minutes.

11. Add 400 microlitre of 3 M sodium acetate pH 5.2 and 10 ml 100 % ethanol. Keep tube at room temperature for 5 minutes. The precipitated DNA can also be stored overnight at -20°C at this point (or during step 8) if that is convenient.
12. Centrifuge at 5000 g at 4°C to pellet the DNA. Wash the pellet with an equal volume of 70 % ethanol, dry it, and resuspend in 4 ml of TE (50 mM Tris-Cl pH 8.0, 1 mM EDTA). Quantitate the DNA by spectrophotometry and resuspend at a final concentration of 400 nanogram/microlitre. Analyze a small amount by agarose gel electrophoresis to ensure that the intensity of the DNA band, visualized by ethidium bromide staining, is concordant with the calculated concentration.

1.1.2 Commercial options

S288c genomic DNA can be purchased from a few companies such as Research Genetics (catalog # 40802) and Novagen (catalog # 69240-3).

1.2 PCR primers

1.2.1 Commercially available primer sets

Research Genetics synthesized the first set of PCR primer pairs for amplifying all yeast genes, and they remain the only commercial source of this genomic set of primers. A few options for primer pairs are available from Research Genetics. The primer pairs intended to amplify the entire ORF are called GenePairs™ and the entire set is available at 2 different concentrations. Primer pairs to amplify intergenic regions, likewise are available at 2 different concentrations and are priced the same as the ORF set. A distinct set of primer pairs designated pORF Primers™ are designed to amplify a partial open reading frame. The advantage of these over the standard ORF primer set is that the resulting PCR products are all just under 1 kilobase pair in length. This is intended to alleviate some of the problems associated with full-length amplification of the longer ORFs by PCR. In addition the pORFs can be targeted to the non-conserved regions of ORFs, to alleviate the cross-hybridization seen among gene sequences with high sequence similarity when the full length ORFs are represented. Table 1 below lists the options available from Research Genetics

Catalog number	Description	Quantity	Sufficient for	Price
40603.1P	ORF Genepairs™ in plates	30 microlitres of 10 micromolar mix	3 reactions	\$ 38,000
40603.1T	ORF Genepairs™ in tubes	200 microlitres of 20 micromolar stocks	50 reactions	\$ 68,000
40903.1P	Intergenic Genepairs™ in plates	30 microlitres of 10 micromolar	3 reactions	\$ 38,000

		mix		
40903.1T	Intergenic Genepairs™ in tubes	200 microlitres of 20 micromolar stocks	50 reactions	\$ 68,000
40510.1P	pORF Genepairs™ in plates	50 microlitres of 10 micromolar mix	5 reactions	\$ 21,500
40510.1T	pORF Genepairs™ in tubes	200 microlitres of 20 micromolar stocks	50 reactions	\$ 38,500

1.2.2 Choice of primer pairs

Few labs will have the budget and the freezer space to purchase and store all possible sets of primers. If the only intended use of the primer pairs is to generate microarrays for use in gene expression experiments, the pORF set is probably the better choice. However there is no published data yet on the success rate of PCR using this set and no information on how these results compare to those obtained from the standard full-length representation. Another downside to purchasing this set is that the primers do not have common flanking sequences, making it impossible to re-amplify the PCR products with a common pair of primers, as can be done with the standard ORF set. The standard ORF set is the one to choose, if, in addition to manufacturing microarrays, there is any anticipated need for large scale cloning of yeast genes into plasmid vectors. For genomic mapping and chromatin IP type experiments, one will of course need the intergenic primer set as well. The ORF set and the intergenic set together cover virtually every locus in the yeast genome. Currently, non-ORF features from the nuclear genome, such as rRNAs, tRNAs, snoRNAs, Ty-LTR elements etc., as well as all mitochondrial genes are included in the “intergenic” set (very likely, these can be purchased separately however). A note about annotation: although the annotation of features in the yeast genomic sequence has gone a long way, it is still a work in progress. Scores of ORFs are completely uncharacterized and it is possible that a number of them may not be transcribed. Although these fragments will be included in the ORF set, they should rightfully belong in the intergenic set and this might become apparent only in the future. Likewise, a seemingly featureless intergenic fragment might harbour a cryptic transcript, perhaps one that encodes a small protein or RNA. The researcher therefore has to be able to incorporate new elements into the microarray, or, if using the complete set, be able to change the annotations of existing elements within the archival database.

1.2.3 Custom primer synthesis

Given the availability of complete primer sets from commercial sources, it is unlikely that anyone will want to do a custom synthesis of the entire set. In-house custom synthesis is an option if the lab has access to a high throughput oligonucleotide synthesizer (such as the PolyPlex™ from GeneMachines or the

Mermade™ from BioAutomation), with free synthesis capacity. This can make it possible to generate oligo sets that can amplify multiple, non-overlapping representations of genes. In-house synthesis is cost effective only if fairly large sets of hundreds of oligos are desired. If there is a need to amplify small numbers of elements that are not included in the commercially available sets (such as introns or newly discovered transcripts), it is best to do it through a company.

1.3 Amplification of yeast genomic elements

The genomic primer sets are available in 96-well microtitre plates and all manipulations are done in this format. PCR amplification of such a large set in a reasonable amount of time depends on the availability of some of the infrastructure for doing high-throughput manipulations, something that is becoming increasingly common. Before embarking on this operation, the researcher would be well advised to pilot the entire procedure on a small number (say 4) of 96-well primer plates, to ensure success of every step. PCR and precipitation plates can be labeled in advance to streamline the process. For all pipetting steps that do not involve a multichannel pipettor, use barrier pipette tips to avoid carryover contamination.

Materials (required for steps X.3 and X.4)

- § 96-well plate thermal cycler(s) with a heated lid. The Tetrad™ from MJ Research can accommodate 4 x 96-well plates, permitting high throughput. Another option is the GeneAmp® 9700 from Applied Biosystems
- § 96-well PCR plates and sealing mats suitable for heated lids (e.g. Multiplate™ and Microseal™ A from MJ Research)
- § Adjustable-volume 12-channel pipettors (2.5 microlitre and 200 microlitre capacities)
- § Benchtop/clinical centrifuge with adapters capable of spinning deep and standard 96-well microtitre plates (e.g. Eppendorf 5810R, Allegra 6 from Beckman Coulter or equivalent)
- § Disposable pipettor tips in racks of 8 x 12 to fit pipettors and pipetting basins.
- § Forward and Reverse PCR primer pairs in 96-well plates.
- § Adjustable single channel pipettors and barrier tips
- § 10X PCR buffer II and 25 mM magnesium chloride, catalog # N808-0190 from Applied Biosystems. 10X PCR Buffer II contains 500 mM KCl, 100 mM Tris-HCl, pH 8.3 and has no gelatin. Alternatively; the solutions can be prepared in the lab if adequate care is exercised to prevent DNA contamination.
- § 100mM solution of each of the 4 dNTPs (Amersham Pharmacia catalog # 27-2035-03)
- § Taq DNA polymerase (e.g. Amplitaq Gold™ from Applied Biosystems, 5 units/microlitre)
- § High capacity agarose gel electrophoresis units (e.g. Owl Centipede™)
- § 96-well U bottom polypropylene precipitation plates (e.g. Costar # 3790)
- § 384-well microarray printing plates (catalog no. X5005, Genetix Co., UK)

- § Sorvall RC-3B lowspeed centrifuge with H-6000A swinging bucket rotor plus microplate adapters, or equivalent)
- § Multichannel pipetting robot (e.g. Multimek™ or Biomek FX™ from Beckman Coulter). Although not essential, this is highly desirable.
- § Adhesive sealing lids for 96-well plates.

1.3.1 High throughput 96-well PCR amplification

1. Thaw PCR primer plates from the freezer by placing them at room temperature or in a shallow dish containing warm (~ 45 °C) water. Centrifuge the plates for 2 minutes at 200 g after they have thawed to collect all liquid at the bottom.
2. Carefully peel off the covering from the primer plates. Using a multichannel pipettor, transfer 5 microlitres each of the forward and reverse primers into a 96-well PCR plate that is oriented correctly with respect to the primer plates. While pipetting, mix the primers in the stock plate by pipetting up and down 2 or 3 times before transferring 5 microlitre to the PCR plate.
3. Prepare reaction mix: For **each** 96-well plate, mix:

double distilled water	7.05 ml
10 X PCR buffer II	1 ml
25 mM magnesium chloride	0.8 ml
dNTPs	25 microlitres each of the 4 100 mM stocks
Yeast genomic DNA, 400 nanogram/microlitre stock	10 microlitres

- Add the above components to a disposable multichannel pipettor basin in the order listed. Reaction mix sufficient for four 96-well plates can be mixed in a single basin. Taq polymerase is added just before adding the reaction mix to the primers. For amplifying the entire set, it is convenient to make a large stock of master mix including the above components sufficient for all the plates, and freeze 36 ml aliquots at -20 °C. Each aliquot is sufficient for 4 plates. Thaw the frozen aliquots by microwaving the tubes with loosened caps for 20 seconds. Transfer the thawed mix to a pipetting basin and mix in Taq polymerase just before pipetting.
4. For each 96-well plate, add 40 microlitres of Taq polymerase (5 units/microlitre) to the reaction mix. Dispense the viscous enzyme evenly along the width of the basin.
 5. Using a multichannel pipettor, add 90 microlitres of the reaction mix with Taq polymerase to each row of the 96-well PCR plate containing the primers. During the first pipetting step, mix the components of the reaction mix thoroughly by pipetting up and down several times. When adding the reaction mix to the PCR plate containing the primers, mix by pipetting up and down 2 or 3 times. Use a fresh set of tips for each reaction.
 6. Cover the PCR plate properly with the sealing mat and place in the thermal cycler. The optimal cycling conditions for yeast elements, using

the Research Genetics primers, are

94 °C, 30 seconds

56 °C, 45 seconds

72 °C, 3 minutes 30 seconds

36 cycles

Including an extra 94 °C denaturation step at the beginning and final extension steps at 72 °C and 25 °C do not appear to affect the yield or success.

7. Analyze the reactions by agarose gel electrophoresis immediately or store the products at 4 °C if analyzing them within a few hours. Otherwise, store them at -20 °C.

1.3.2 Agarose gel analysis of PCR reactions

PCR products should be checked by agarose gel. The advantage of doing this immediately after the PCR reactions, rather than after the products have been purified, is that if large numbers of fragments from a plate fail to be amplified, purification of failed products can be avoided. Sporadic failures can occur due to vagaries of the reaction, problems in primer synthesis, problems with the thermal cyclers, and so on. However, for the yeast primer sets, a success rate of 95 % or greater is very realistic and routinely observed. Pay attention to systematic versus sporadic failures. For instance, an entire row missing could be indicative of a pipetting error, while a column that shows several failures could be caused by a malfunctioning position on a multichannel pipettor. A spreadsheet listing the names and sizes of the products should be annotated with the results of the gel analysis. It is worthwhile redoing at least some fraction of the sporadic failed reactions. Sort the list of 96-well plates by the number of failures in each plate. Then pick primers from the plates with the greatest number of failures to make a set of 2 or 3 “failure” plates. These can be reamplified, perhaps tweaking the reaction conditions to favour longer products which are often over represented in the failed set.