[31] Designing Experiments Using Spotted Microarrays to Detect Gene Regulation Differences Within and Among Species

By JEFFREY P. TOWNSEND and JOHN W. TAYLOR

Abstract

Comparative studies of genome-wide gene expression must account for variation not only among species, but also within species. Such studies are necessarily large in scale, because they incorporate experiments on multiple individuals of multiple species in multiple developmental stages in multiple environmental conditions. If the experiments are carefully designed and performed, the data they provide are worth the effort. We describe the utility of spotted microarrays for these studies and highlight experimental design criteria that will maximize inferential and statistical power. We conclude with a discussion of experimental protocols that are designed for investigations of differential gene expression and their pitfalls.

Introduction

DNA microarrays have proved their worth in studies of development and mutation using single experimental strains and will be even more useful for the examination of natural genetic variation by the analysis of multiple individuals from populations and species. DNA microarray technology has particularly strong potential to illuminate studies of the molecular origin of phenotype (Singh, 2003). Where there is variation in phenotype, there will be underlying molecular correlates that will lead to a better understanding of the phenotype. Even complex phenotypes that may involve multiple metabolic and developmental pathways are approachable with appropriate experimental design. Where there is variation within populations or among species, questions about adaptation may be clearly framed. These questions will be best addressed when fitness can be assayed at the same time point for which transcription is profiled. For example, successful spore germination may be assayed among different individuals.

Well-studied model organisms such as yeast or *Escherichia coli* are ideal for this type of research because their genomes are sequenced, microarrays are available, their biochemical metabolism and their cell and molecular biology are well understood, and hypotheses about genes can be tested using molecular genetic experiments. However, non–model

0076-6879/05 $35.00

TABLE I
FEATURES AND DRAWBACKS OF TYPES OF PROBES FOR MICROARRAYS

Type of probe	Genome[a]	Annotated[b]	Sensitivity to divergence	Completeness[c]	Expense
Random genomic clones	No	No	Depends on fragment size	Incomplete	Low
cDNAs from ESTs	No	No	Sensitive	Incomplete	Low
PCR products of ORFs	Yes	Yes	Sensitive	Complete	Laborious, 2 primers/ sequence
70-mers designed for ORFs	Yes	Yes	Very sensitive	Complete, low cross-hybridization	Intermediate
25-mer whole genome tiling arrays	Yes	No	Extremely sensitive	Complete, low cross-hybridization	Expensive

Note: cDNA, complementary DNA; EST, expressed sequence tag; ORF, open reading frame; PCR, polymerase chain reaction.

[a] Requires a known complete sequence for the organismal genome.

[b] Requires a well-annotated genome to design probes.

[c] Results in a complete description of genomic gene expression levels.

systems are not excluded from microarray research. In fact, a useful microarray can be constructed from fragments of the genome of any organism and used for competitive hybridization experiments to identify a pool of potentially interesting genes. The fragments of interest may then be sequenced, and their identification attempted through comparison with known genes. Within fungi, this approach has been used to find genes involved in *Histoplasma capsulatum* pathogenicity (Hwang *et al.*, 2003). There are many ways in which microarrays may be created, and each way has its advantages and disadvantages (Table I). Here, we present guidelines for experimental design of spotted DNA microarray studies, with particular focus on the examination of variation among individuals or species.

Experimental Design

Use of Comparative Hybridizations

Microarray experimental designs are highly influenced by the particular technology employed, and here we emphasize spotted microarray design. The most salient feature of experimental design for spotted microarrays is

their comparative nature. In each use, two samples of messenger RNA (mRNA) are competitively hybridized to the deposited DNA, and the comparison allows for the elimination of many spot-specific confounding factors (Eisen and Brown, 1999; Townsend, 2004). In any experimental use of spotted DNA microarrays, the first question to consider is ''how should samples of interest be compared?'' Experimental design in this context may be approached by the use of universal reference samples (Novoradovskaya *et al*., 2004; Puskás *et al*., 2004), pooled samples, or circuits of comparisons. For most purposes, circuits of comparisons have the greatest statistical power to detect differences in gene expression among all samples examined (Kerr and Churchill, 2001a; Townsend, 2003; Townsend and Hartl, 2002; Wolfinger *et al*., 2001; Yang and Speed, 2002); thus, this method of experimental design is described here.

When comparisons are made in circuits, complex experiments can be designed to accommodate three major axes of biological variation. The first axis is that of genotype and is typically manifested in multiple individuals, $G_1 \ldots G_i$, where i is the number of individuals and each individual is a natural or mutational variant. The second axis includes any investigated aspects of the environment, $E_1 \ldots E_j$, where j is the number of environments, typically comprising several experimental treatments, microhabitats, or geographic locations. Lastly, a third axis along which samples are frequently compared is an axis of developmental states, typically time points in a developmental course or cycle, for example, $T_1 \ldots T_k$, where k is the number of time points. Experimental designs incorporating one, two, or three of these dimensions may be most clearly planned and depicted using design graphs.

Graph Theoretical Depiction of Experimental Designs

Spotted microarray experimental designs are commonly depicted using directed multigraphs (e.g., Fig. 1, a simple design incorporating all three major axes of variation; see also Figs. 2A–C and 3A–C). In these graphs, circular nodes represent samples of mRNA that have been harvested from a particular genotype, environment, and developmental state. Arcs between nodes represent competitive hybridizations between the mRNAs. This depiction is a *multi*graph because more than one arc can connect any two nodes. The number of these replicated hybridizations, or arcs, is vital to the degree of statistical significance for both large and small differences in expression level. Simultaneous comparison of two mRNAs on a single microarray is made possible by labeling the samples with two fluorescent dyes. The direction of the arc in the diagram conventionally indicates which sample is labeled with which fluorophore, (e.g., the pointed end

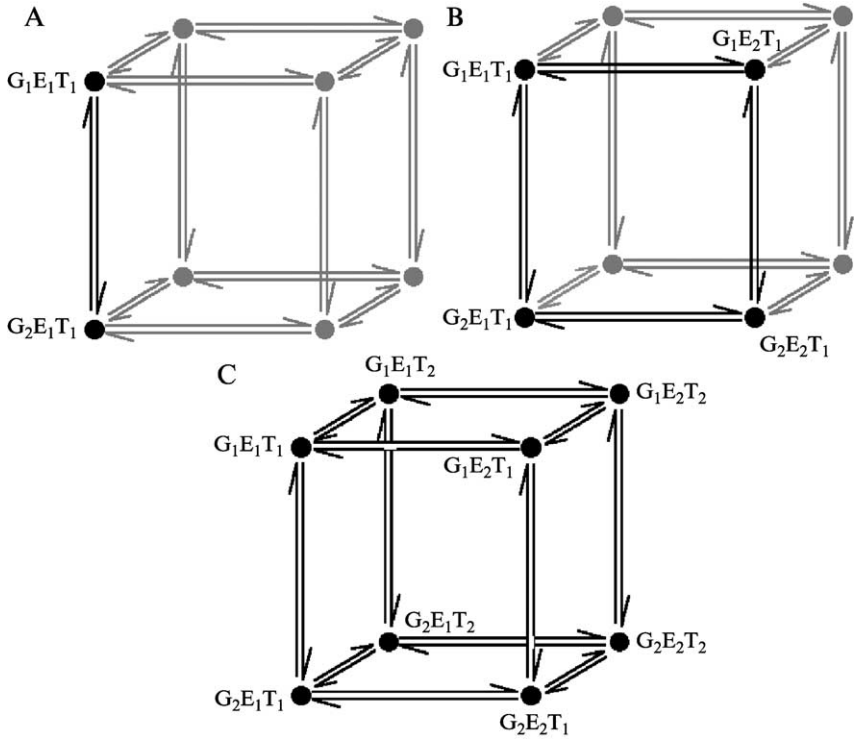

FIG. 1. A general example of spotted complementary DNA (cDNA) microarray experimental design that incorporates interrogation of eight messenger RNA (mRNA) samples from (A) two genotypes (G_1 and G_2), (B), two environments (E_1 and E_2) and (C) two developmental states (T_1 and T_2). Nodes represent the mRNA samples, and arcs represent competitive hybridizations. The direction of the arc indicates which mRNA sample was labeled with which fluorophore (e.g., the blunt end indicates the node labeled with Cy5, and the pointed end indicates the node labeled with Cy3). With eight mRNA samples, 24 competitive hybridizations are proposed. An additional 24 competitive hybridizations could be made symmetrically by including direct comparisons across the diagonals of the cube graph faces (i.e., $G_1E_1T_1$ to $G_2E_2T_1$, $G_2E_1T_1$ to $G_1E_2T_1$, etc.).

indicates the complementary DNA [cDNA] labeled with Cy5, and the blunt end indicates the cDNA labeled with Cy3). Between the fluorophore and the DNA spotted, there is a small interaction effect on the hybridization intensity, so it is recommended to perform two hybridizations with ''flipped fluorophores'' between any two samples that are to be directly compared. This pair of replicates is indicated in a directed multigraph by a pair of oppositely directed arcs (Fig. 1A).

For a given mRNA sample, or node, the number of adjacent nodes is the unweighted valence and the number of arcs is the weighted valence. In Fig. 1A, each mRNA is to be competitively hybridized two times against each of three other mRNAs, for a total of six hybridizations. Each node has an unweighted valence of three and a weighted valence of six. Thus, the weighting of the valence of a node in microarray experimental design is the degree of experimental replication applied to that particular mRNA sample. To increase replication in the experimental design in Fig. 1, for example, flipped-fluorophore cross-comparisons could be added to span from corner to corner of each face of the cubic graph. These added comparisons would increase the weighted valence of each node from 6 to 12 and allow one to resolve differences in gene expression as low as 1:1.2 instead of 1:1.4 (see discussion of statistical significance).

If variation along an axis of interest is continuous or ordinal (e.g., a quantitative genetic trait, a temperature gradient, a variable reagent concentration, or a time course), the experiment should be designed to compare each node with its nearest neighbors. This guideline ensures that the greatest power is applied to the detection of differences between the most similar samples. It might be argued that differences between the most similar samples are not of as much interest; in that case, the prudent action would be to pare down the number of nodes in the design until all samples are of interest.

If variation along an axis of interest is discrete (e.g., wild type vs. mutant, presence vs. absence of environmental contaminants, or a series of developmentally plastic phenotypes), it is useful to compare each node with the others in as symmetrical a fashion as possible. This guideline ensures that there will be equivalent power to detect differences among all nodes in your design. Of course, if one experimental node is of greater interest, then it is sensible for your experimental design to feature a higher valence for that node.

Generally, it is easy to design a microarray experiment that examines many factors for their influence on gene expression in organisms. The difficult part is carefully performing such an experiment. Our ability to conceive of interesting variables to manipulate far exceeds our ability to examine them. Figure 2 depicts a more complex example of an experiment along the lines of Fig. 1, but with greater inferential power. However, this experiment would require more than four times as many microarray hybridizations as the experiment depicted in Fig. 1. Figure 3 depicts an example of an experiment to jointly examine population and species variation in gene expression, and this design would need almost as many microarrays as in the experiment depicted in Fig. 2. The technical requirements of these fairly simple-to-conceive designs make clear how rapidly the multiplicative examination of various influences on gene expression can

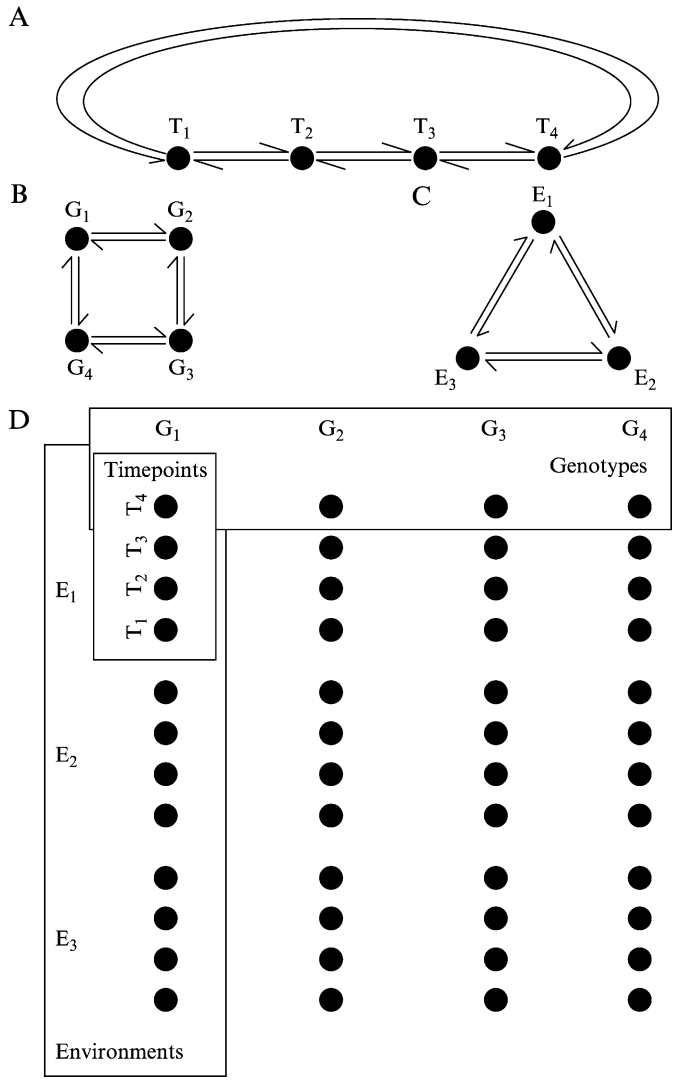

FIG. 2. A general example of an extensive spotted cDNA microarray experimental design, incorporating factors including developmental state ($T_1 \ldots T_4$), genotype ($G_1 \ldots G_4$), and environment ($E_1 \ldots E_3$). Nodes (filled circles) represent 48 mRNA samples, and arcs represent competitive hybridizations. The direction of the arc indicates which mRNA sample was labeled with which fluorophore (e.g., the blunt end indicates the node labeled with Cy5, and the pointed end indicates the node labeled with Cy3). (A) Circuit design for a time course. Each ordinal time point has been compared to its neighbor, and additionally the last time point has been compared to the initial time point. This design comprises eight microarrays and

make experimental designs costly and time consuming. Fortunately, spotted microarrays are not as expensive as other high-throughput expression analyses, and large numbers of arrays can be used. Completing all parts of the experiments diagrammed in Fig. 1A–C or Fig. 2A–C as depicted in Fig. 1D or Fig. 2D results in every experimental node possessing a weighted valence of 12, as well as resolution of the significance of most differences in gene expression of a factor of 1.2 or more. A key step after brainstorming an interesting set of experiments lies in paring down those experiments so they are feasible in both time and expense while retaining the essential comparisons to answer the questions of importance in your system. This stage of the design cannot be prescribed but will always belong to the experimentalist, who must apply his or her expertise and insight to narrow down the biological factors examined to those that are truly relevant.

Replication and Resolution of Differences in Gene Expression

Replication of microarray hybridizations is essential to precise inference of gene expression level (Lee *et al*., 2000; Pan *et al*., 2002). Flipping fluorophores is one important method of replication. However, just two hybridizations between a pair of nodes frequently are insufficient for the resolution of small differences in gene expression. Closed circuit designs like those in Figs. 1–3 help to increase the resolution of differences in gene expression, because data from alternate longer paths in the design also provide moderate inferential power (Townsend, 2003), Performing additional hybridizations beyond the minimum of two between two nodes is another option. Additional replicate comparisons increase the probability that the experiment performed will resolve large and small differences in gene expression as statistically significant. A useful summary statistic of the power of an experiment to resolve the statistical significance of small

should detect as significant most differences greater than about 1.6-fold. Cross-comparisons from T_2 to T_4 and from T_1 to T_3 could be added to increase resolution. (B) Circuit design for comparison of four genotypes; note that the graph is topologically the same as in panel A. (C) Circuit design for comparison of three environments. This design comprises six microarrays and should detect as significant most differences greater than about 1.6-fold. (D) Diagram showing all 48 mRNA samples that would have to be collected to examine all of these factors. Boxes indicate sets of mRNA samples that correspond to nodes in graphs A–C. An examination of all of these factors in one experiment could be conducted by performing the comparisons in panel A for all combinations of the four genotypes and three environments, by performing the comparisons in panel B for all combinations of the four time points and three environments, and by performing the comparisons in panel C for all combinations of the four time points and the four genotypes. Combined into a single dataset and analyzed together, these 288 microarrays should yield a net resolution of the statistical significance of most differences greater than about 1.2-fold.

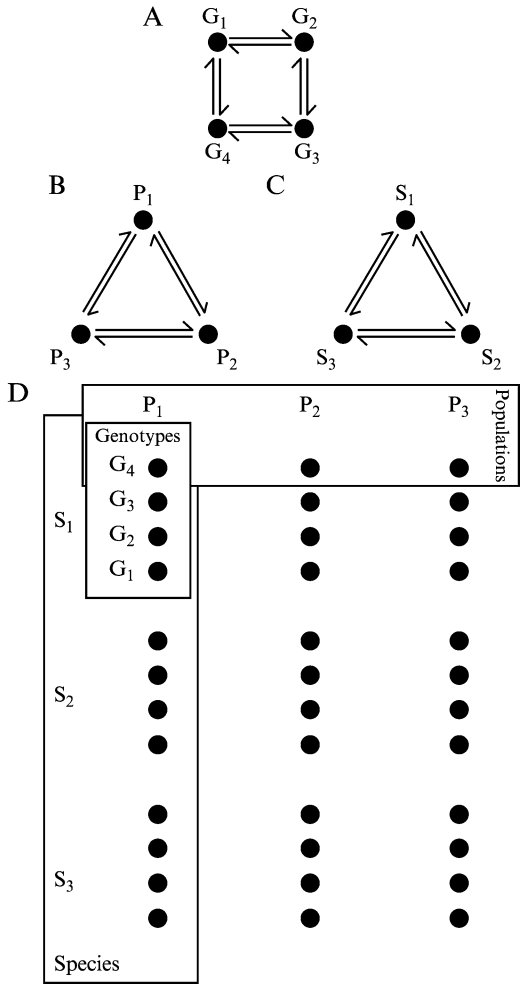

FIG. 3. A general example of the application of spotted cDNA microarray experimental design applied to the study of variation within populations and among species. In this figure, factors examined include genotype ($G_1 \ldots G_4$), population ($P_1 \ldots P_3$), and species ($S_1 \ldots S_3$). Nodes represent mRNA samples, and arcs represent competitive hybridizations. The direction of the arc indicates which mRNA sample was labeled with which fluorophore (e.g., the blunt end indicates the node labeled with Cy5, the pointed end indicates the node labeled with Cy3). (A) Circuit design comparing genotypes G_1 through G_4. This design comprises eight microarrays and should detect as significant most differences greater than about 1.6-fold. Cross-comparisons from G_2 to G_4 and from G_1 to G_3 could be added to increase resolution. (B) Circuit design comparing localities a, b, and c. This design comprises six microarrays and should detect as significant most differences greater than about 1.6-fold. (C) Circuit design comparing species S_1, S_2, and S_3, graph theoretically the same as in panel B.

differences in gene expression is the gene expression level (GEL) at which there is an empirical 50% probability of a statistically significant call (GEL_{50}). For any pair of samples, this statistic may be calculated across all genes present on a microarray by logistic regression of the statistical significance call on estimated GEL (Fig. 4). Additional replicates increase the precision of estimates of GEL, decrease the GEL_{50} for an experiment, and thus increase the power to resolve small differences in gene expression.

Statistical Significance and Its Importance

Early transcriptional profiling experiments identified genes as differentially expressed by a ''twofold threshold,'' so genes whose expression level was greater or lesser by a factor of 2 in a comparison of an experimental to a reference sample were considered differentially expressed (Alexandre *et al*., 2001; DeRisi *et al*., 1997; Lyons *et al*., 2000; Sudarsanam *et al*., 2000). The twofold threshold has no theoretical basis and often, inappropriately, serves double duty as a signifier of both statistical and biological significance.

Disentangling statistical and biological significance is essential to understanding the power of a study to reveal biological differences among samples (Townsend and Hartl, 2002; Tseng *et al*., 2001; Wolfinger *et al*., 2001). The finer resolution of smaller and smaller differences in gene expression with increased replication demonstrated in several empirical studies (Townsend, 2004) and in simulations (see Fig. 4) shows that a particular transcriptional profiling experiment cannot reveal all of the differential expression among treatments or genotypes. Rather, a transcriptional profiling experiment reveals those genes whose differences in expression are sufficiently large and sufficiently consistent in measurement to be statistically different. The number of such genes is strongly influenced by the replication present in the experimental design.

(D) Diagram showing all 36 mRNA samples that would have to be collected to examine all of these factors. Each mRNA sample is represented by a filled circle. Boxes indicate sets of mRNA samples that comprise comparison diagrams A–C. An examination of all of these factors in one experiment could be conducted by performing the comparisons in panel A for all combinations of the three localities and three species, by performing the comparisons in panel B for all combinations of the four genotypes and three species, and by performing the comparisons in panel C for all combinations of the four genotypes and the three localities. Combined into a single dataset and analyzed together, these 216 microarrays should yield resolution of the statistical significance of most differences greater than about 1.2-fold.

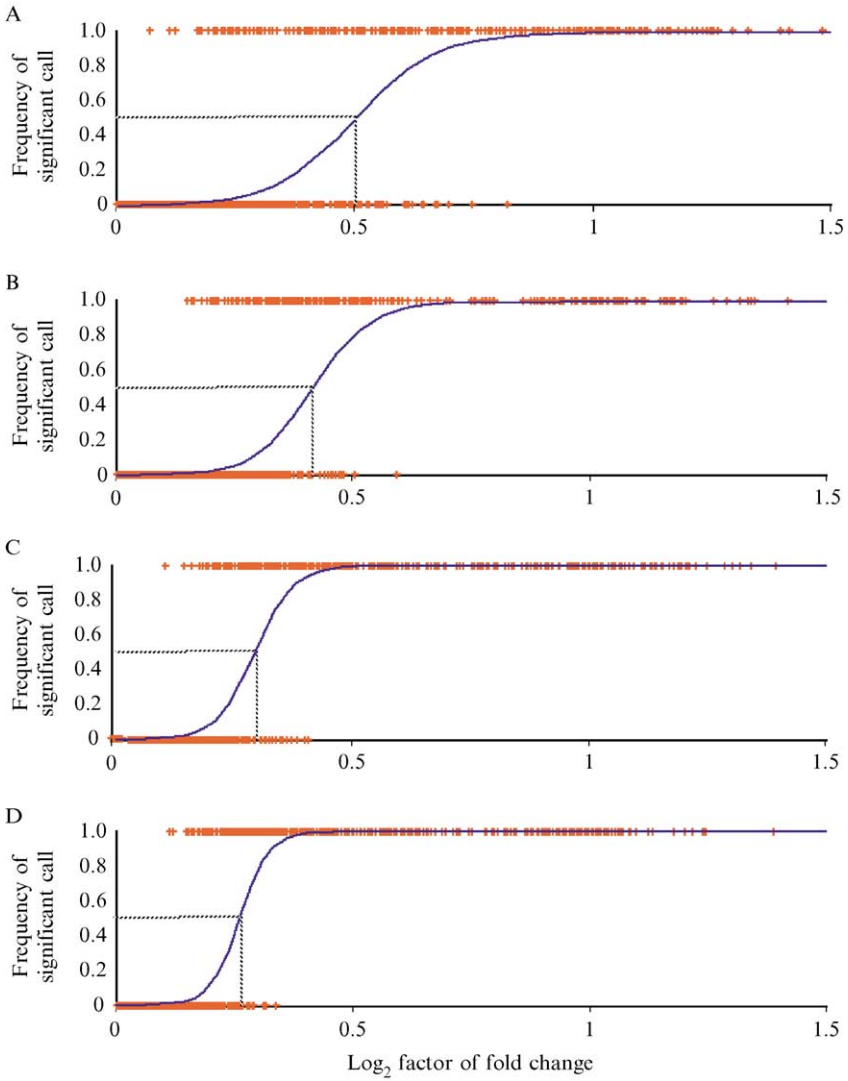

FIG. 4. Logistic regressions of the frequency of affirmative significance call over $\log_2$ factor of difference in gene expression. The logistic model plotted is that $\log_e p/(1-p) = mx + b$, where x is the $\log_2$ factor of difference in gene expression. Cross symbols represent estimated expression levels from simulated data using a Bayesian analysis of gene expression level with additive small error terms. Each cross is placed on the abscissa at the estimated expression level, either at the top of the plot (significant, S) or at the bottom (not significant, NS). Ratio data were simulated using the probability distribution of (Fieller, 1932) assuming a constant coefficient of variation across samples. Logistic regressions are on the factors of difference

Furthermore, restricting analysis to only those genes showing statistically significant changes yields much more meaningful results. Genes whose expression level measurements were highly variable in a given study should be filtered from large-scale comparisons because of their lack of statistical significance (i.e., large credible or confidence intervals for expression level). The filtered subset of expression level measurements demonstrate vastly increased biological correlation (Townsend, 2003; Townsend and Hartl, 2002; Townsend *et al.*, 2003), because the well-measured genes are not swamped by a morass of poorly measured genes. For instance, large numbers of poorly resolved genes will by chance be clustered within small clusters of genes that are well measured and have true biological association, obscuring otherwise clear functional groupings. The best verification of the results of a DNA microarray study, in the end, is the concordance of expression data with known biology, in particular with molecular biological data on transcription factor recruitment, metabolic pathways, and protein–protein interaction. For DNA microarray studies on population and species differences in expression, it should be kept in mind that these molecular data are well known only for model organisms. Thus, using prior knowledge of organismal biology to troubleshoot the development of DNA microarray technology and protocols is much easier when model organisms and their close relatives are the initial organisms of study.

estimated from simulated data comparing two samples to each other with different numbers of replicates. (A) An example with four replicates. The model has a highly significant fit ($\chi^2 = 1593.1$, $P < .0001$). The estimated intercept for the log odds, b, of a significant call versus no significant call is -5.4 (significant, $P < .0001$), and the estimated slope with $\log_2$ factor of difference in gene expression, m, is 11.5 (significant, $P < .0001$). The factor of gene expression at which 50% of estimated differences were identified as significant (GEL_{50}) was 1.4-fold. (B) An example with six replicates. The model has a highly significant fit ($\chi^2 = 1186.8$, $P < .0001$). The estimated intercept for the log odds, b, of a significant call versus no significant call is -6.1 (significant, $P < .0001$), and the estimated slope with $\log_2$ factor of difference in gene expression, m, is 16.5 (significant, $P < .0001$). The factor of gene expression at which 50% of estimated differences were identified as significant (GEL_{50}) was 1.29-fold. (C) An example with eight replicates. The model has a highly significant fit ($\chi^2 = 1186.8$, $P < .0001$). The estimated intercept for the log odds, b, of a significant call versus no significant call is -7.3 (significant, $P < .0001$), and the estimated slope with $\log_2$ factor of difference in gene expression, m, is 26.7 (significant, $P < .0001$). The factor of gene expression at which 50% of estimated differences were identified as significant (GEL_{50}) was 1.21-fold. (D) An example with 10 replicates. The model has a highly significant fit ($\chi^2 = 1307.2$, $P < .0001$). The estimated intercept for the log odds, b, of a significant call versus no significant call is -8.6 (significant, $P < .0001$), and the estimated slope with $\log_2$ factor of difference in gene expression, m, is 35.0 (significant, $P < .0001$). The factor of gene expression at which 50% of estimated differences were identified as significant (GEL_{50}) was 1.18-fold.

Comparing within and among Species

Comparisons among different individuals are possible within and among species. Interpretation of such comparisons will be most convincing if information is already available on the circumscription of populations and species, preferably by phylogenetic and population genetic studies using nucleic acid variation and appropriate methods of statistical analysis. Typically, arrays of any kind will be designed based on sequence of one individual in one species. To incorporate natural variation into microarray studies, it will be necessary to use arrays designed from the sequence of one individual for competitive hybridizations among genetically different individuals or species. The utility of the array will decrease as the genetic distance increases between the ''design'' individual and the ''experimental'' individuals. Detailed knowledge of populations and species will allow researchers to select individuals with increasing genetic distances to test the range of a microarray to aid experimental design. Fragmented DNA can be used as the probe in these experiments.

In terms of design, and in keeping with the aforementioned guidelines, it is ideal to make direct comparisons among the most closely related populations and/or species and then work toward more divergent comparisons. This approach is needed because sequence divergence will increase with phylogenetic distance and confound the hybridizations at each spot on a microarray (Bozdech *et al*., 2003; Letowski *et al*., 2004; Nagpal *et al*., 2004). In principle, a consistent low level of divergence across genes should not present a serious problem, because any decrease in hybridization due to that divergence will then be approximately constant across genes in that sample and a global normalization will compensate appropriately. However, there is considerable variation in the rate of divergence of genes (Graybeal, 1994). Thus, experimentalists must be wary of interpreting differential hybridization as representative of differential expression when mRNA is derived from divergent organisms.

Although divergence in DNA sequence creates a difficulty for the study of gene expression in closely to moderately diverged species, it is an advantage for the study of gene expression in very distantly related species that live in close association. For instance, many of the socially important and evolutionarily interesting questions concerning fungi involve mutualisms with other organisms, ranging from pathogenesis to symbiosis. Microarrays containing neighboring spots of DNA sequence for genes of both species can be used to simultaneously measure gene expression in the two organisms with little concern for cross-hybridization. Thus, single arrays composed of sequences homologous to both partners may be used to identify pools of candidate genes that are coordinately regulated in both

partners in the mutualism (Johansson *et al.*, 2004). Microarrays are designed to measure mRNA level and profile transcription throughout the genome, but they also can be used to discriminate between alleles at single nucleotide polymorphism (SNP) loci throughout the genome (Steinmetz *et al.*, 2002). Therefore, experiments can be designed to combine QTL analysis and transcription profiling in progeny from parents with phenotypes important to adaptation: pathogenicity, industrial production, and so on.

Disturbance Involved in Harvesting mRNA

A key element to good competitive hybridization experiments is to minimize disturbance of the organisms in the process of harvesting RNA, to avoid accidentally studying the effects of the disturbance. Moreover, it is absolutely vital to confine differences among experimental treatments to those being tested in the experiment; any other differences will confound analysis and interpretation. Consider with extreme care how samples will be handled immediately before RNA harvest. An incautiously designed microarray experiment may easily reveal only the manifold effects of centrifugation, filtering, or other lab manipulations. Of course, the same experimental protocols cannot be applied to all organisms, for example, contrast the swift and uniform arresting of bacterial cell activity at -80° or with ethanol (EtOH)–phenol treatment (Zimmer *et al.*, 2000) with the quite variable treatments of human and other primate organs before RNA extraction (Enard *et al.*, 2002). Even routine experimental manipulations can have a strong effect on transcription, as exemplified by a study of gravitropism in plants that showed significant changes in transcriptional profile due solely to the minimal handling of plants before RNA extraction (Moseyko *et al.*, 2002).

Technical Protocols

DNA Microarray Construction

DNA microarrays are multiplexed Southern hybridizations (Fig. 5). Using terminology that mirrors the functional roles of the nucleic acids in Southern hybridizations, the ''probe'' in a microarray experiment is unlabeled and affixed to a solid substrate (the coated glass slide). The ''target,'' counterintuitively, is washed over the multiple affixed probes and consists of fluorescently labeled cDNA made from mRNA harvested from the organism. Spotted microarrays can be constructed by deposition of polymerase chain reaction (PCR) products for each and every open reading

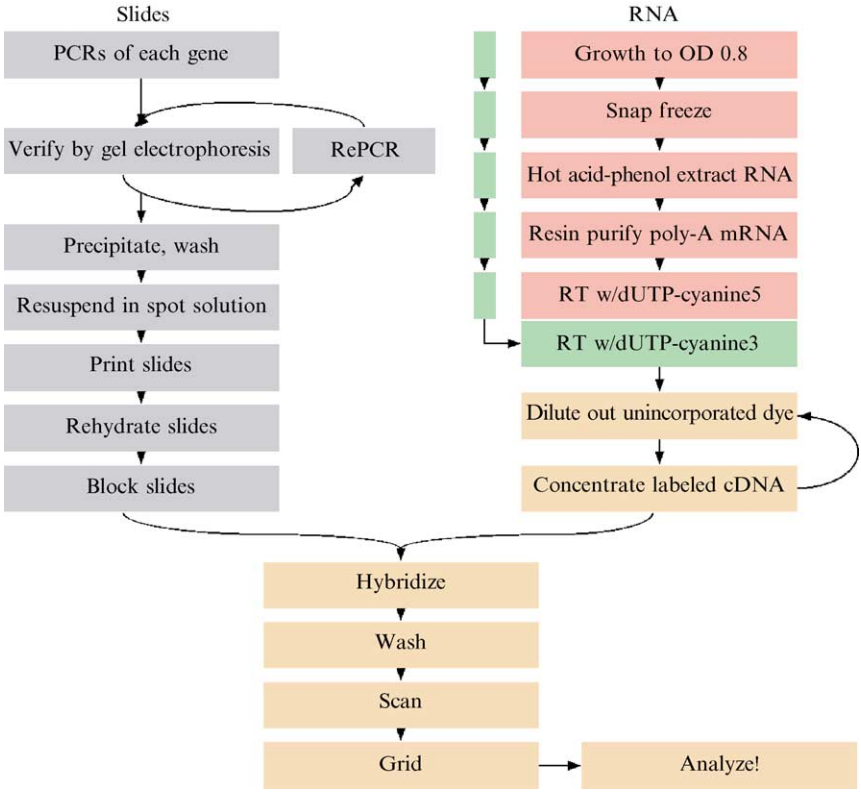

FIG. 5. Schematic diagram of the steps employed to create spotted microarrays from polymerase chain reaction products and to perform a comparative microarray hybridization.

frame (ORF) (Eisen and Brown, 1999), by deposition of oligomers (70 or 50 nucleotides) designed from predicted ORF sequences, or by deposition of random clones from a cDNA library or even from random clones of genomic DNA fragments. Oligomer arrays and cDNA arrays appear to have similar sensitivities for differential gene expression (Lee *et al.*, 2004), but there are advantages and disadvantages to each method (Table I). For example, it is easier to design oligomer probes that minimize cross-hybridization, but 10% sequence divergence randomly located within a 70-mer long oligo may reduce hybridization intensity by 64% (Bozdech *et al.*, 2003). Conversely, whole-ORF PCR products do cross-hybridize but appear not to be very sensitive to SNPs (Ranz *et al.*, 2003).

ORFs may be amplified from clones or from genomic DNA by means of the PCR. Longer ORFs may be amplified with any of a variety of special polymerase kits that have high processivity. Each amplified product should be confirmed for correct length by agarose gel electrophoresis. Amplified DNA may be precipitated in 96-well format with isopropanol, washed with 70% EtOH, and resuspended in a salt spotting solution such as 3× SSC. DNA is commonly spotted on polylysine-coated glass slides (Eisen and Brown, 1999) or γ-aminopolysilane (GAPS)–coated glass slides (Corning, Corning, NY), using a microarraying robot (e.g., http://cmgm.stanford.edu/pbrown/mguide/index.html). GAPS slides are more expensive but have a better shelf life and tend to be less variable in quality.

Extraction of RNA and Reverse Transcription

The key reagent in transcriptional profiling is the RNA harvested from an organism. This RNA is used as a template for reverse transcription to make cDNA for competitive hybridization against the affixed probes on the microarray. RNA is best extracted from flash-frozen pellets of tissue or culture grown in meticulously maintained common garden conditions. The flash-frozen matter should have its RNA extracted in a manner that will not result in mRNA degradation (i.e., performed rapidly upon or while thawing). Nucleic acids may be EtOH precipitated, washed, dried, and redissolved in TE buffer. Yield ranges from organism to organism, but a spectrophotometric ratio of absorption (260 nm/280 nm) of about 2.0 indicates a clean preparation without much protein contaminant. mRNA may be purified easily using a Qiagen Extraction Kit (Valencia, CA), which contains columns that retain poly-A RNA and allow much of the tRNA and rRNA to pass through. For eukaryotes, reverse transcription of eluted mRNA may be performed with oligo-dT primers of an appropriate length to bind with the poly-A tails of mRNA from the organism of study. For both eukaryotes and prokaryotes, the reactions may be primed with random hexamer primers, supplied with deoxyribonucleic triphosphates (dNTPs), performed by a reverse transcriptase such as Superscript II. To provide a ligand for dye labeling, amino-allyl-dUTP is incorporated into the cDNA along with the dNTPs. After at least 2 h of reverse transcription, the approximately 20-μl reaction should be stopped with 10 μl of 1 *M* NaOH and 10 μl of 0.5 *M* of ethylenediaminetetraacetic acid (EDTA), and the mix is incubated at 65° for 15 min. Then 25 μl of 1 *M* HEPES pH 7.5 is added to stabilize the solution.

Both total RNA and purified mRNA have been successfully used as templates for the production of labeled cDNA for microarray hybridization.

Also, both oligo-dT alone and a mixture of oligo-dT and oligo-dN primer have been used as primers to reverse transcribe mRNA. For the most part, both purifying the mRNA and using oligo-dT alone help decrease noise caused by errant cDNAs from tRNA and rRNA. However, using total RNA is cheaper and faster, and using oligo-dN primers can dramatically increase signal. The tradeoff must be examined in each particular microarray experimental context, because these factors interact with other aspects of experimental setup such as spotted DNA fragment size, slide chemistry, deposition solution, and hybridization conditions. Figure 6 may be used as a guide for optimization.

Cyanine Dye Coupling

To retain the cDNA and discard unincorporated nucleotides, the products of reverse transcription are diluted and filtered in a Microcon-30 microconcentrator, which retains the long polymers of cDNA but not unincorporated nucleotides. Typically, an initial 10-fold dilution of the reverse transcription reaction product is followed by a 20-fold concentration. Two more rounds of 20-fold dilution and concentration complete the cleanup. NHS dye may be bound to cDNA via amino-allyl-dUTP residues by raising the pH. To 10–13 μl of purified concentrate, 0.8 μl of 1 *M* $NaHCO_3$ pH 9 can be added, with an appropriate NHS-cyanine dye aliquot. This coupling reaction is incubated in the dark at 25° for 75 min and then stored in the dark at 4° and used in less than 24 h. The labeled cDNA may then be purified with a QIAquick column. This elution of about 55 μl of purified cyanine-labeled cDNA may also be stored at 4° and should be used in less than 24 h.

	TTT... primers	TTT... primers and NNN... primers
Total RNA	Low signal Intermediate noise	High signal High noise
Poly-A RNA	Low signal Low noise	High signal Intermediate noise

FIG. 6. Diagram relating signal and noise characteristics for microarray hybridizations conducted using messenger RNA (mRNA) or total RNA and poly-T or polyT + polyN primers used in the reverse-transcriptase reaction for production of the labeled mRNA target to be washed over the microarray of fixed probe spots.

Hybridization

For each competitive hybridization, the labeled target cDNAs from two samples are used, i.e., one labeled with cyanine-3 and one labeled with cyanine-5. The labeled cDNA is concentrated to 20 μl in Microcon-30 microconcentrators, combining appropriate cyanine-3– and cyanine-5–labeled paired samples. Note that 1.5 μl of poly-dA oligomers of appropriate length for the organism of study may be added to block poly-T tails of the cDNA. Next, 3 μl of 20× SSC and 0.5 μl of 1 *M* HEPES pH 7.0 are added. The mix can then be filtered of any dust or residues with a wetted (10 μl ddH_2O) Millipore-0.45 μm filter; 10% sodium dodecyl sulfate (SDS) is added, and the mix is then boiled for 2 min to denature the nucleic acids. It should then be cooled at 27° for 10 min. Hybridizations using labeled target at temperatures above room temperature can result in extremely high background fluorescence. A microarray slide is set in a hybridization chamber. To keep the slide stable within the chamber, deposit drops of 3× SSC on the underside of the slide, allowing them to adsorb to the slide corners and the chamber bottom. To prevent dehydration of the labeled cDNA solution from beneath the coverslips, 3× SSC is added to the hybridization chamber wells. A coverslip (LifterSlips are very convenient for this purpose) should be cleaned with EtOH and then placed over the printed microarray. The labeled cDNA mix is then injected at the corners of the coverslip, and the chamber is sealed and then placed level in a 60° waterbath, to be incubated at 60–63° for 12–15 h to reach equilibrium (Sartor *et al*., 2004).

Array Wash

Hybridized microarray slides can be washed by repeated plunging in a solution of 387 ml of purified water, 12 ml of 20× SSC, and 1 ml of 10% SDS, and rinsed by repeated plunging in a solution of 399 ml of purified water and 1 ml of 20× SSC. The array should be scanned as soon as possible; if there must be a delay, the array may be stored in the dark, but for no more than 2 h.

Data Acquisition and Analysis

Fluorescent DNA bound to the microarray may be detected with a GenePix 4000 microarray scanner (Axon Instruments, Foster City, CA), using the GenePix 4000 software package to locate spots in the microarray. Other scanners are available, as are alternative, open source, and freely downloadable scanning software, such as TIGR's Spotfinder (http://www.tigr.org/software).

Normalization

Fluorescence intensity values are commonly adjusted in each channel by subtracting the average background intensity observed surrounding the spot from the observed foreground intensity of the spot. Typical foreground intensities for quality microarray hybridizations should be 10 or more times as intense as background. However, some spots do not achieve this ideal even in excellent hybridizations. To eliminate signals that are most prone to estimation error, any spot can be excluded from analysis if both the Cy3 and the Cy5 fluorescence intensities for that spot are within 3 standard deviations of the distribution of intensities of the background pixels for that spot. This procedure avoids artificially inflated measurements of relative expression for the competing mRNA samples to that spot due to near-zero background-subtracted intensity values in one fluorescence channel.

Relative expression levels for the two competing dye-labeled samples may be normalized by linear scaling of the cyanine-5 values so the mean cyanine-5 and cyanine-3 background-corrected intensity values of nonexcluded spots are equal when hybridizations are of uniformly high quality. This straightforward method should then yield a linear log-log cyanine-3–cyanine-5 intensity and no further normalization will be necessary. If the relationship is not linear because of systemic technical problems, it may be useful to use LOWESS smoothing or other statistical approaches. However, curved or abnormally shaped log-log regressions generally indicate poor and therefore misleading data, regardless of applied statistical sophistication. It should be considered by the experimentalist whether extra analytical efforts to glean the most information from poor experiments are more fruitful than repeating hybridizations and refining experimental technique.

Data Analysis

Multifactorial experimental designs, as described earlier, should be analyzed statistically. Two major methods for analysis of such designs are analysis of variance (ANOVA) methods (Kerr and Churchill, 2001a,b; Wolfinger *et al*., 2001) and Bayesian methods (Townsend, 2004; Townsend and Hartl, 2002). The details of these approaches are adequately covered in the primary literature. The two approaches yield consistent results when used to analyze the same dataset (Whitfield *et al*., 2003), so the choice of method may come down to a combination of philosophical preference and practicality. ANOVA presents a powerful method that can, with a fair degree of statistical savvy on the part of the user, be tailored tightly to incorporate normalization and downstream analysis (Jin *et al*., 2001) into a single cohesive whole. A powerful and flexible method for the Bayesian analysis of gene expression levels (BAGEL), from multifactorial

experiments, has been implemented to analyze normalized data on multiple platforms in a straightforward and freely available software package (http://web.uconn.edu/townsend/software.html).

Acknowledgments

We thank the Miller Institute for Basic Research and NIH grant GM068087 for supporting our research.

References

Alexandre, H., Ansanay-Galeote, V., Dequin, S., and Blondin, B. (2001). Global gene expression during short-term ethanol stress in *Saccharomyces cerevisiae*. *FEBS Lett.* **498,** 98–103.

Bozdech, Z., Zhu, J., Joachimiak, M. P., Cohen, F. E., Pulliam, B., and DeRisi, J. L. (2003). Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biol.* **4,** R9.

DeRisi, J. L., Iyer, V. R., and Brown, P. O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278,** 680–686.

Eisen, M. B., and Brown, P. O. (1999). DNA arrays for analysis of gene expression. *Methods Enzymol.* **303,** 179–205.

Enard, W., Khaitovich, P., Klose, J., Zollner, S., Heissig, F., Giavalisco, P., Nieselt-Struwe, K., Muchmore, E., Varki, A., Ravid, R., Doxiadis, G. M., Bontrop, R. E., and Paabo, S. (2002). Intra- and interspecific variation in primate gene expression patterns. *Science* **296,** 340–343.

Fieller, E. C. (1932). The distribution of the index in a normal bivariate population. *Biometrika* **24,** 428–440.

Graybeal, A. (1994). Evaluating the phylogenetic utility of genes: A search for genes informative about deep divergences among vertebrates. *Syst. Biol.* **43,** 174–193.

Hwang, L., Hocking-Murray, D., Bahrami, A. K., Andersson, M., Rine, J., and Sil, A. (2003). Identifying phase-specific genes in the fungal pathogen *Histoplasma capsulatum* using a genomic shotgun microarray. *Mol. Biol. Cell* **14,** 2314–2326.

Jin, W., Riley, R. M., Wolfinger, R. D., White, K. P., Passador-Gurgel, G., and Gibson, G. (2001). The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nat. Genet.* **29,** 389–395.

Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R., Lundeberg, J., Uhlen, M., and Tunlid, A. (2004). Transcriptional responses of Paxillus involutus and Betula pendula during formation of ectomycorrhizal root tissue. *Mol. Plant Microbe Interact.* **17,** 202–215.

Kerr, M. K., and Churchill, G. A. (2001a). Experimental design for gene expression microarrays. *Biostatistics* **2,** 183–201.

Kerr, M. K., and Churchill, G. A. (2001b). Statistical design and the analysis of gene expression microarray data. *Genet. Res.* **77,** 123–128.

Lee, H. S., Wang, J. L., Tian, L., Jiang, H. M., Black, M. A., Madlung, A., Watson, B., Lukens, L., Pires, J. C., Wang, J. J., Comai, L., Osborn, T. C., Doerge, R. W., and Chen, Z. J. (2004). Sensitivity of 70-mer oligonucleotides and cDNAs for microarray analysis ofgene expression in Arabidopsis and its related species. *Plant Biotechnol. J.* **2,** 45–57.

Lee, M.-L. T., Kuo, F. C., Whitmore, G. A., and Sklar, J. (2000). Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations. *Proc. Natl. Acad. Sci. USA* **97,** 9834–9839.

Letowski, J., Brousseau, R., and Masson, L. (2004). Designing better probes: Effect of probe size, mismatch position and number on hybridization in DNA oligonucleotide arrays. *J. Microbiol. Methods.* **57,** 269–278.

Lyons, T. J., Gasch, A. P., Gaither, L. A., Botstein, D., Brown, P., and Eide, D. (2000). Genome-wide characterization of the Zap1p zinc-responsive regulon in yeast. *Proc. Natl. Acad. Sci.* **97,** 7957–7962.

Moseyko, N., Zhu, T., Chang, H. S., Wang, X., and Feldman, L. J. (2002). Transcription profiling of the early gravitropic response in Arabidopsis using high-density oligonucleotide probe microarrays. *Plant Physiol.* **130,** 720–728.

Nagpal, S., Karaman, M. W., Timmerman, M. M., Ho, V. V., Pike, B. L., and Hacia, J. G. (2004). Improving the sensitivity and specificity of gen expression analysis in highly related organisms through the use of electronic masks. *Nucleic Acids Res.* **32,** e51.

Novoradovskaya, N., Whitfield, M. L., Basehore, L. S., Novoradovsky, A., Pesich, R., Usary, J., Karaca, M., Wong, W. K., Aprelikova, O., Fero, M., Perou, C. M., Botstein , D., and Braman, J. (2004). Universal reference RNA as a standard for microarray experiments. *BMC Genomics* **5,** 20.

Pan, W., Lin, J., and Le, C. T. (2002). How many replicates are required to detect gene expression changes in microarray experiments? A mixture model approach. *Genome Biol.* **3,** research 0022.1–0022.10.

Puskás, L. G., Zvara, Á., Hackler, L., Micsik, T., and van Hummelen, P. (2004). Production of bulk amounts of universal RNA for DNA microarrays. *Biotechniques* **33,** 898–904.

Ranz, J. M., Castillo-Davis, C. I., Meiklejohn, C. D., and Hartl, D. L. (2003). Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* **300,** 1742–1745.

Sartor, M., Schwanekamp, J., Halbleib, D., Mohamed, I., Karyala, S., Medvedovic, M., and Tomlinson, C. R. (2004). Microarray results improve significantly as hybridization approaches equilibrium. *Biotechniques* **36,** 790–796.

Singh, R. S. (2003). Darwin to DNA, molecules to morphology: The end of classical population genetics and the road ahead. *Genome* **46,** 938–942.

Steinmetz, L. M., Sinha, H., Richards, D. R., Spiegelman, J. I., Oefner, P. J., McCusker, J. H., and Davis, R. W. (2002). Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416,** 326–330.

Sudarsanam, P., Iyer, V. R., Brown, P. O., and Winston, F. (2000). Whole-genome expression analysis of *snf/swi* mutants of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **97,** 3364–3369.

Townsend, J. P. (2003). Multifactorial experimental design and the transitivity of ratios with spotted DNA microarrays. *BMC Genomics* **4,** 41.

Townsend, J. P. (2004). Resolution of large and small differences in gene expression using models for the Bayesian analysis of gene expression levels and spotted DNA microarrays. *BMC Bioinformatics* **5,** 54.

Townsend, J. P., Cavalieri, D., and Hartl, D. L. (2003). Population genetic variation in genome-wide gene expression. *Mol. Biol. Evol.* **20,** 955–963.

Townsend, J. P., and Hartl, D. L. (2002). Bayesian analysis of gene expression levels: Statistical quantification of relative mRNA level across multiple treatments or samples. *Genome Biol.* **3,** research 0071.1–0071.16.

Tseng, G. C., Oh, M.-K., Rohlin, L., Liao, J. C., and Wong, W. H. (2001). Issues in cDNA microarray analysis: Quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res.* **29,** 2549–2557.

Whitfield, C. W., Cziko, A.-M., and Robinson, G. E. (2003). Gene expression profiles in the brain predict behavior in individual honey bees. *Science* **302,** 296–299.

Wolfinger, R. D., Gibson, G., Wolfinger, E. D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C., and Paules, R. S. (2001). Assessing gene significance from cDNA microarray expression data via mixed models. *J. Comput. Biol.* **8,** 625–637.

Yang, Y. H., and Speed, T. (2002). Design issues for cDNA microarray experiments. *Nat. Rev.* **3,** 579–588.

Zimmer, D. P., Soupene, E., Lee, H. L., Wendisch, V. F., Khodursky, A. B., Peter, B. J., Bender, R. A., and Kustu, S. (2000). Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: Scavenging as a defense against nitrogen limitation. *Proc. Natl. Acad. Sci. USA* **97,** 14674–14679.

[32] Methods for Studying the Evolution of Plant Reproductive Structures: Comparative Gene Expression Techniques

By ELENA M. KRAMER

Abstract

A major component of evolutionary developmental (evo-devo) genetics is the analysis of gene expression patterns in nonmodel species. This comparative approach can take many forms, including reverse-transcriptase polymerase chain reaction, Northern blot hybridization, and *in situ* hybridization. The choice of technique depends on several issues such as the availability of fresh tissue, as well as the expected expression level and pattern of the candidate gene in question. Although the protocols for these procedures are fairly standard, optimization is often required because of the specific characteristics of the species under analysis. This chapter describes several methods commonly used to determine gene expression patterns in angiosperms, particularly in floral tissues. Suggestions for adapting basic protocols for diverse taxa and troubleshooting are also extensively discussed.

General Considerations for Working with RNA

RNA is, by nature, a less stable molecule than DNA and there are many sources of RNase in the environment, which can lead to its rapid degradation. These facts often lead to trepidation on the part of researchers who are not experienced with RNA work. However, such concerns are largely unnecessary because simple precautions can effectively prevent

 0076-6879/05 $35.00