Temporal analysis of *Candida albicans* gene expression during biofilm development

Kathleen M. Yeater,¹ Jyotsna Chandra,² Georgina Cheng,¹ Pranab K. Mukherjee,² Xiaomin Zhao,¹ Sandra L. Rodriguez-Zas,³ Kurt E. Kwast,⁴ Mahmoud A. Ghannoum² and Lois L. Hoyer¹

¹Department of Pathobiology, University of Illinois, Urbana, IL 61802, USA
²Center for Medical Mycology, Department of Dermatology, University Hospitals of Cleveland and Case Western Reserve University, Cleveland, OH 44106, USA
³Department of Animal Sciences, University of Illinois, Urbana, IL 61801, USA
⁴Department of Molecular and Integrative Physiology, University of Illinois, Urbana, IL 61801, USA

Microarrays were used to identify changes in gene expression associated with *Candida albicans* biofilm development. Two biofilm substrates (denture and catheter), and two *C. albicans* strains for each substrate, were tested to remove model- and strain-dependent variability from the overall dataset. Three biofilm developmental phases were examined: early (6 h), intermediate (12 h), and mature (48 h). Planktonic specimens were collected at the same time points. Data analysis focused primarily on gene expression changes over the time-course of biofilm development. Glycolytic and non-glycolytic carbohydrate assimilation, amino acid metabolism, and intracellular transport mechanisms were important during the early phase of biofilm formation. These early events increase intracellular pools of pyruvate, pentoses and amino acids, which prepare the biofilm for the large biomass increase that begins around 12 h of development. This developmental stage demands energy and utilizes specific transporters for amino acids, sugars, ions, oligopeptides and lactate/pyruvate. At mature phase (48 h), few genes were differentially expressed compared with the 12 h time point, suggesting a relative lack of initiation of new metabolic activity. Data analysis to assess biofilm model-specific gene expression showed more dynamic changes in the denture model than in the catheter model. Data analysis to identify gene expression changes that are associated with each strain/substrate combination identified the same types of genes that were identified in the analysis of the entire dataset. Collectively, these data suggest that genes belonging to different, but interconnected, functional categories regulate the morphology and phenotype of *C. albicans* biofilm.

INTRODUCTION

Micro-organisms attach to surfaces in many natural, industrial and medical environments and can develop into biofilms. Biofilms are a functional association of microbial cells that are enveloped within extracellular polymer matrices and associated with surfaces (Costerton *et al.*, 1987; Donlan, 2002; Donlan & Costerton, 2002). Phenotypically, biofilm cells are different from planktonic (free-floating) cells (Costerton *et al.*, 1995), with one of the most important manifestations of these differences being the significantly decreased susceptibility of microbial biofilms to antimicrobial agents (Rupp, 2005; Schwank *et al.*, 1998; Stewart *et al.*, 2004; Wilson, 1996). Although bacterial biofilms have been the focus of many studies, detailed investigations into the biology and pathogenesis of fungal biofilms have only recently been initiated (reviewed by Douglas, 2003; Ghannoum & O’Toole, 2004). *Candida albicans* is the most common fungal pathogen associated with colonization and biofilm formation on the surfaces of indwelling medical devices (IMDs) such as dentures and intravenous catheters (Kojic & Darouiche, 2004). Biofilm formation by *C. albicans* can promote superficial or systemic disease. For example, denture stomatitis, a superficial candidiasis, occurs in 65% of denture-wearing...
biofilms do not have distinct layers and are typically material) influences biofilm architecture and morphology. Candida which (mature phase; Chandra phase); and maturation into an ECM-encased community of extracellular matrix (ECM) production (intermediate cellular aggregation, microcolony formation, and initiation adhesion to substrate and cell proliferation (early phase); Biofilm formation proceeds in three developmental phases: adhesion to substrate and cell proliferation (early phase); cellular aggregation, microcolony formation, and initiation of extracellular matrix (ECM) production (intermediate phase); and maturation into an ECM-encased community (mature phase; Chandra et al., 2001). The substrate on which Candida biofilms are grown (e.g. denture or catheter material) influences biofilm architecture and morphology. Model denture biofilms, grown on saliva-coated poly-methylmethacrylate strips, primarily consist of yeast forms. These biofilms do not have distinct layers and are typically between 20 and 30 μm thick (Chandra et al., 2001). In contrast, C. albicans biofilms formed on serum-coated silicone elastomer catheter discs consist of a basal layer of yeast cells covered by a layer of hyphae in the mature phase of growth. Model catheter biofilms are typically 200–300 μm thick (Chandra et al., 2001; Kuhn et al., 2002). Regardless of the architecture, C. albicans biofilm grown on either substrate is more resistant than planktonic cells to antifungals such as fluconazole, nystatin, amphotericin B and chlorhexidine (Baillie & Douglas, 1999; Chandra et al., 2001; Hawser & Douglas, 1995; Kuhn et al., 2002). Recent completion and annotation of the C. albicans genome sequence (Braun et al., 2005; Jones et al., 2004), and the availability of C. albicans microarrays, have made it possible to perform whole-genome transcriptional profiling of biofilms formed by these fungi. Three previous studies assessed C. albicans biofilm transcriptional profiles to identify genes that are involved in this process (Cao et al., 2005; Garcia-Sanchez et al., 2004; Murillo et al., 2005). However a study that follows gene expression from an early stage of biofilm growth to biofilm maturity has not previously been reported. The goal of the present study was to use microarray analysis to identify genes that are differentially expressed during biofilm development, indicating their potential importance in biofilm formation, antifungal resistance and persistence strategies in the host environment. Biofilm specimens and planktonic control cultures were collected at 6, 12 and 48 h, representing the early, intermediate and mature phases of biofilm development respectively. Analysis of two biofilm model systems and two C. albicans strains within each system allowed us to remove model- and strain-dependent variability from the experiment. Gene expression profiles were analysed (a) as a function of time in the biofilm and planktonic cultures separately, and (b) between the biofilm and planktonic forms of C. albicans at each time point. These analyses led to a description of the temporal changes in gene expression that suggests cellular mechanisms responsible for the phenotypic changes that occur during biofilm development.

METHODS

C. albicans microarrays. Construction of the microarray used in these studies was described previously (Zhao et al., 2005). Briefly, the microarray included 6928 spots representing C. albicans ORFs and/or gene sequences and controls, which were printed in duplicate on the array. Of the 6928 spots, 6715 represented C. albicans ORFs from Assembly 6 of the genome sequence (http://www-sequenc.stanford.edu/group/candida) and are likely to include redundancies among genes because the number of ORFs in the C. albicans genome is estimated to be 6354 (Braun et al., 2005). The ORFs were printed on Arrayl brand SuperAmine Substrate microarray slides (Telechem) at Microarrays, Inc. (Nashville, TN, USA). Controls included Arabidopsis thaliana Spot Report genes and mammalian DNA fragments (Stratagene). C. albicans genes of known expression profiles, and buffer blanks. Additional information about the microarrays is available at http://candida.cvm.uiuc.edu.

Growth of model biofilms and planktonic cell cultures. C. albicans strains SC5314 (Gillum et al., 1984), GDH2346 and M61 were used. Strain GDH2346 was isolated from a patient with denture stomatitis and provided by Dr Julia Douglas (University of Glasgow, UK). Strain M61 was obtained from a patient with catheter-associated candidiasis (intravascular line) at the University Hospitals of Cleveland (USA) (Kuhn et al., 2002). Model biofilms of strains GDH2346 and SC5314 were grown on denture acrylic coated with saliva (Chandra et al., 2001); model biofilms of strains M61 and SC5314 were formed on serum-coated silicone elastomer catheter material (Kuhn et al., 2002) as described previously. All biofilms were grown in yeast nitrogen base supplemented with 50 mM glucose (Chandra et al., 2001; Kuhn et al., 2002). Samples were grown for 6, 12 and 48 h. Planktonic cultures of each strain were grown and collected for each substrate and time combination. Planktonic growth conditions were identical to those used for biofilm formation except that the bioprosthetic material was not present in the culture well. Planktonic cells were aspirated from the sample well with a pipette. Samples were flash-frozen in liquid nitrogen and stored at −80 °C until used. Two independent replicates were used in the microarray hybridization analysis. The experimental design of the array experiments was a four-block (one for each strain and substrate combination), simple loop design with a dye swap (Fig. 1). Loop designs, in which samples are compared to each other in a daisy-chain fashion, are efficient and amenable to statistical analysis (Churchill, 2002). Interwoven-loop designs, such as the one used in our study, are efficient and robust because they create multiple links between the samples.

Microarray experiments. Total RNA was obtained from cells using a hot phenol extraction method (Collart & Oliviero, 1993). Total RNA was further purified using an RNeasy kit (Qiagen) according to the manufacturer’s suggested protocol (Mackenzie et al., 1997). The methods for labelling probes and hybridization to microarrays were adapted from protocols developed by The Institute for Genomic
Phase-dependent gene expression in C. albicans biofilms

Using GenePix Pro version 4.0, each hybridization produces a pair of red (Cy5) and green (Cy3) fluorescence intensities that serve as the inputs for all statistical analyses. For statistical analysis, the background-corrected median fluorescence intensity of each individual DNA spot was used. To determine an appropriate normalization method for the data, MA-plots (Dudoit et al., 2002), were used to initially analyse the fluorescence (R [Cy5], G [Cy3]) data. These are plots of the intensity log-ratio $M = \log_{2} R/G$ vs the mean log intensity $A = \log_{2} (R+G)$. A global normalization model of $\log_{2}$ transformed fluorescent data was used (Wolfgang et al., 2001). The residuals from this model, which are regarded as a crude indicator of relative expression level and are referred to as ‘normalized expression levels’, were then analysed in individual gene-specific models. The normalization and gene models were fit using the MIXED procedure of SAS (SAS Institute, 2000). SAS code for this analysis was based on code available in the Introductory Manual for Mixed Model Analysis of Microarray Data (MANNMADA, 2002).

To assess the magnitude and significance of the effects of the strain, substrate, time, and their interactions on the normalized expression level of each gene, least square means were calculated using the ESTIMATE option of SAS (SAS Institute, 2000) on 6928 ORFs on the array, as described in the MANNMADA (2002). The ESTIMATE option calculates the mean value for each target category, with adjustment for the other terms in the model. For each pair-wise comparison, the difference between the two normalized expression levels and their standard errors are calculated along with a t-test for statistical significance (Littell et al., 1996). Incorporation of two interconnected mixed linear models accounts for variability across and within genes. This statistical approach is superior to using raw fold-change ratios, which results in common type I and type II errors (Tanaka et al., 2000). Our statistical approach allows more direct control over the percentage of false positives and false negatives (Wolfgang et al., 2001). Information about all hybridizations that were performed and the complete set of data and image files formatted using MIAME criteria, can be downloaded from http://candida.cvm.uiuc.edu. Data in MIAME format can also be found in the Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under the GEO accession number GSE7825. Genes were assigned to functional categories using information from the Candida Genome Database (http://www.candidagenome.org). Although some genes could be assigned to multiple categories, a single assignment was made based on available information. Hierarchical clustering was performed using Genespring 7.1 (Silicon Graphics) using the gene tree clustering tool with a distance similarity measure.

Verification of differential gene expression using real-time RT-PCR. The following genes were selected to verify results of array hybridization comparisons: ERG10, ADE1, FBA1, Ccw12, PDC11, GDH3, ENO1 and ICL1. These genes were differentially expressed between biofilm and planktonic cells at the 6 h time point. Because of the interwoven-loop experimental design, validation of data at this time point validates data from other time points as well. PCR primers were designed using PrimerExpress (Applied Biosystems) with a melting temperature ($T_m$) between 59 and 60 °C and an amplicon size between 50 and 100 bp (Table 1). A pool of total biofilm RNA was made by combining 1.5 µg RNA from each 6 h strain/substrate biofilm sample for a total RNA pool of 6 µg. A 6 µg pool of total planktonic RNA was also made by combining 1.5 µg RNA from each 6 h strain/substrate planktonic sample. Total RNA was DNase treated and purified with the RNaseasy kit (Qiagen). The cDNA was made using the SuperScript First-Strand Synthesis System (Invitrogen) and 1 µg of pooled RNA. The cDNA was diluted 1:5 with RNase-free water (Ambion).

PCR reactions contained 100 nM of each primer, 1 × Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen), 1 × ROX Reference dye

Fig. 1. Schematic of the microarray experimental design. The experimental design of the microarray experiments was a four-block (one for each strain and substrate combination), simple loop design with a dye swap. This schematic shows one block in the design. Arrows represent hybridizations between the RNA samples and the microarray. The sample at the head of the arrow was labelled with green (Cy3) dye. For example, 6 h biofilm, labelled red, and 6 h planktonic, labelled green, were hybridized to one array, then 6 h biofilm, labelled green, and 6 h planktonic, labelled red, were hybridized to another. This pattern continued around the loop, and the hybridizations were completed for a second independent replication. This experimental block was completed for each strain/substrate combination using a total of 96 microarrays.

Research (TIGR, http://pga.tigr.org/protocols.html). cDNA was synthesized using Superscript II RT (Invitrogen Life Technologies; 100 units of reverse transcriptase) in reaction mixtures containing 0.5 mM (total) deoxynucleoside triphosphates (4:1). After cDNA synthesis overnight at 42 °C, the RNA was hydrolysed with 1 M NaOH and 0.5 M EDTA. The reaction was neutralized with 1 M HCl. The cDNA was purified using the Qiagen QIAquick PCR purification kit and modified protocol (TIGR). The cDNA was recovered and dried in a SpeedVac (ThermoLabs).

The cDNA was coupled with fluorescent dyes by resuspension in 0.05 M sodium bicarbonate buffer, pH 9.0, and incubation with either Cy3 or Cy5 dye (GE Healthcare) for 1 h at room temperature in the dark. Unincorporated dye was removed using a Qiagen QIAquick PCR purification kit and 0.05 M sodium bicarbonate buffer, pH 9.0, and incubation with a dye swap. This schematic shows one block in the design. Arrows represent hybridizations between the RNA samples and the microarray. The sample at the head of the arrow was labelled with green (Cy3) dye. For example, 6 h biofilm, labelled red, and 6 h planktonic, labelled green, were hybridized to one array, then 6 h biofilm, labelled green, and 6 h planktonic, labelled red, were hybridized to another. This pattern continued around the loop, and the hybridizations were completed for a second independent replication. This experimental block was completed for each strain/substrate combination using a total of 96 microarrays.

**Data analysis and statistical inference.** Arrays were scanned on a GenePix 4000B scanner (Axon Instruments) and data were quantified using GenePix Pro version 4.0. Each hybridization produces a pair of red (Cy5) and green (Cy3) fluorescence intensities that serve as the inputs for all statistical analyses. For statistical analysis, the background-corrected median fluorescence intensity of each individual DNA spot was used. To determine an appropriate normalization method for the data, MA-plots (Dudoit et al., 2002), were used to initially analyse the fluorescence (R [Cy5], G [Cy3]) data. These are plots of the intensity log-ratio $M = \log_{2} R/G$ vs the mean log intensity $A = \log_{2} (R+G)$. A global normalization model of $\log_{2}$ transformed fluorescent data was used (Wolfgang et al., 2001). The residuals from this model, which are regarded as a crude indicator of relative expression level and are referred to as ‘normalized expression levels’, were then analysed in individual gene-specific models. The normalization and gene models were fit using the MIXED procedure of SAS (SAS Institute, 2000). SAS code for this analysis was based on code available in the Introductory Manual for Mixed Model Analysis of Microarray Data (MANNMADA, 2002).

To assess the magnitude and significance of the effects of the strain, substrate, time, and their interactions on the normalized expression level of each gene, least square means were calculated using the ESTIMATE option of SAS (SAS Institute, 2000) on 6928 ORFs on the array, as described in the MANNMADA (2002). The ESTIMATE option calculates the mean value for each target category, with adjustment for the other terms in the model. For each pair-wise comparison, the difference between the two normalized expression levels and their standard errors are calculated along with a t-test for statistical significance (Littell et al., 1996). Incorporation of two interconnected mixed linear models accounts for variability across and within genes. This statistical approach is superior to using raw fold-change ratios, which results in common type I and type II errors (Tanaka et al., 2000). Our statistical approach allows more direct control over the percentage of false positives and false negatives (Wolfgang et al., 2001). Information about all hybridizations that were performed and the complete set of data and image files formatted using MIAME criteria, can be downloaded from http://candida.cvm.uiuc.edu. Data in MIAME format can also be found in the Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under the GEO accession number GSE7825. Genes were assigned to functional categories using information from the Candida Genome Database (http://www.candidagenome.org). Although some genes could be assigned to multiple categories, a single assignment was made based on available information. Hierarchical clustering was performed using Genespring 7.1 (Silicon Graphics) using the gene tree clustering tool with a distance similarity measure.

**Verification of differential gene expression using real-time RT-PCR.** The following genes were selected to verify results of array hybridization comparisons: ERG10, ADE1, FBA1, Ccw12, PDC11, GDH3, ENO1 and ICL1. These genes were differentially expressed between biofilm and planktonic cells at the 6 h time point. Because of the interwoven-loop experimental design, validation of data at this time point validates data from other time points as well. PCR primers were designed using PrimerExpress (Applied Biosystems) with a melting temperature ($T_m$) between 59 and 60 °C and an amplicon size between 50 and 100 bp (Table 1). A pool of total biofilm RNA was made by combining 1.5 µg RNA from each 6 h strain/substrate biofilm sample for a total RNA pool of 6 µg. A 6 µg pool of total planktonic RNA was also made by combining 1.5 µg RNA from each 6 h strain/substrate planktonic sample. Total RNA was DNase treated and purified with the RNaseasy kit (Qiagen). The cDNA was made using the SuperScript First-Strand Synthesis System (Invitrogen) and 1 µg of pooled RNA. The cDNA was diluted 1:5 with RNase-free water (Ambion).

PCR reactions contained 100 nM of each primer, 1 × Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen), 1 × ROX Reference dye

http://mic.sgmjournals.org
Table 1. Primers used for real-time RT-PCR validation of microarray results

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’–3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADE1RTF</td>
<td>GCC ACC GAC AGA ATT TCT GC</td>
</tr>
<tr>
<td>ADE1RTR</td>
<td>AGG GAT ACC ATT GGC CAT AAT AAC</td>
</tr>
<tr>
<td>CCW12RTF</td>
<td>CGC ACA ACA AAT GTG CTA CTG T</td>
</tr>
<tr>
<td>CCW12RTR</td>
<td>TGT TGA CTG ATT GAA CAG CTG GA</td>
</tr>
<tr>
<td>ERG10RTF</td>
<td>TGG CTC ATC TTC TTG GTC TG C</td>
</tr>
<tr>
<td>ERG10RTR</td>
<td>ACA ATA AAG TGA CAA CAA TGC TTC TG CAC</td>
</tr>
<tr>
<td>ENO1RTF</td>
<td>TCC AGC CAA GAT ACG TCT AGC</td>
</tr>
<tr>
<td>ENO1RTR</td>
<td>CAA CTT CAA CGG TTG GGT ACA C</td>
</tr>
<tr>
<td>FBA1RTF</td>
<td>CAT TGT TCT CAT CCC ACA GTG TG</td>
</tr>
<tr>
<td>FBA1RTR</td>
<td>TTT CGT CAT CGG TTT CTT CAG A</td>
</tr>
<tr>
<td>GDH3RTF</td>
<td>TGT GTC CAG ATA CCG ATG TTC C</td>
</tr>
<tr>
<td>GDH3RTR</td>
<td>CTC TAC CAC CAA CAC CAA TAT ACA C</td>
</tr>
<tr>
<td>ICL1RTF</td>
<td>CAA TCC AGC TTG CTC AAG TGG</td>
</tr>
<tr>
<td>ICL1RTR</td>
<td>CGG AAA CGT AAA TGG GGT CCA A</td>
</tr>
<tr>
<td>PDC11RTF</td>
<td>CCA AAT GGT TTA TCA GCT TTA GTC AG</td>
</tr>
<tr>
<td>PDC11RTR</td>
<td>AGC AAT GGC ATT AGT CAA ACT CAA T</td>
</tr>
</tbody>
</table>

(Invitrogen), and 5 μl diluted cDNA in a final volume of 25 μl. PCR reactions were run on the ABI PRISM 7000 Sequence Detection System (SDS) with a 50 °C uracil-DNA glycosylase (UDG) incubation step for 2 min, 95 °C for 2 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Primer specificity was verified with a dissociation profile of each amplicon.

A standard curve for each primer set was performed with 1:10, 1:25, 1:50, 1:100, 1:250 and 1:500 dilutions of the cDNA. The slopes of the standard curves were within 10% of 100% efficiency. Crossing threshold (Ct) values were determined using the Autoanalyse feature of the SDS software. Relative gene expression was quantified using the 2–ΔΔCt method (Livak & Schmittgen, 2001), where 2−ΔΔCt = (Ct, Biofilm – Ct, TEF) – (Ct, Planktonic – Ct, TEF). The fold change in the target gene, normalized to TEF1, was calculated for each sample using 2−ΔCt. Two separate pooled RNA preparations (independent samples) were made for each target and carried through the analysis. Each RNA replication was treated separately and the results were averaged after the 2−ΔΔCt calculation of each PCR run. The mean, SD and CV were then determined from two independent replications with two individual runs of triplicate samples per pooled RNA preparation.

RESULTS AND DISCUSSION

Datasets and data comparisons

The main goal of this study was to identify genes that were differentially expressed in biofilm-grown cells compared to planktonic cells, without regard to C. albicans strain or to the substrate on which the biofilm was grown. The experimental design included two different biofilm substrates: polymethylmethacrylate (denture; Chandra et al., 2001) and silicone elastomer (catheter; Kuhn et al., 2002). Two different C. albicans strains were grown in each model (GDH2346 and SC5314 in the denture model; M61 and SC5314 in the catheter model). Three different time points were examined, representing important developmental periods during biofilm formation. Biofilms were studied at early phase (6 h), when antifungal drug resistance is first observed, intermediate phase (12 h), when the ECM begins to form, and mature phase (48 h), when the biofilm has reached its mature structure (Chandra et al., 2001). Planktonic control cultures were grown in the same manner as the biofilm specimens, except that the substrate (denture or catheter) was omitted from the culture well. Planktonic control samples were collected at the same time points as used for the biofilm specimens. Biofilm and planktonic data were analyzed separately across the experimental time-course. This method was used to identify genes that were differentially expressed during biofilm development and also as planktonic cells multiplied in the culture wells. Analysis of different biofilm models and different C. albicans strains within each model allowed us to remove model- and strain-dependent variability from the final dataset. Genes that were differentially expressed (≥2-fold change, P<0.05) were considered significant. Due to the asynchronous nature of biofilm formation, twofold increases in gene expression are likely to highlight significant cellular processes. Supplementary Table S1 (available with the online version of this paper) lists the 243 differentially expressed genes grouped by functional category, and identifies the analysis from which they were derived. Comparisons were also made between the biofilm and planktonic specimens at each time point to provide a relative magnitude of gene expression differences between biofilm and planktonic C. albicans cells. This set of 191 differentially expressed genes is shown in Supplementary Table S2. Together, these data tables demonstrate that only a small percentage of the C. albicans genome was differentially expressed as a function of biofilm development phase or between biofilm and planktonic cells at any of the time points. Microarray data were validated using real-time RT-PCR analysis as described in Methods. Microarray and real-time RT-PCR estimates of gene expression differences were consistent, and supported the integrity of the microarray experiments (Table 2).

Genes differentially expressed in this study were compared to genes identified by two other global analyses of C. albicans gene expression during biofilm formation. In the first study, Garcia-Sanchez et al. (2004) used a macroarray that included approximately one-third of the genome and compared transcriptional profiles between biofilms grown under a variety of conditions and several different planktonic control cultures. These growth conditions utilized different types of vessels, and variations in flow of growth medium, percentage glucose, aeration, time and temperature. For each gene on the macroarray, a statistical measure of central tendency (the mean is cited in one place in the manuscript while another states that the median was used) was derived from the normalized gene expression values for all of the biofilm samples; the same process was applied to all of the planktonic samples. The derived, single statistical measure for the biofilm specimens was divided by the corresponding measure for the planktonic specimens to yield a ratio for each gene on the macroarray. This method produced a set of 325 genes that are differentially
Table 2. Comparison of fold-induction values of real-time RT-PCR to microarray experiments

<table>
<thead>
<tr>
<th>Gene</th>
<th>orf19 ID</th>
<th>Relative gene expression</th>
<th>Microarray fold change</th>
<th>RT-PCR mean fold change</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADE13</td>
<td>19.3870</td>
<td>-2.40*</td>
<td>0.52 ± 0.06*</td>
<td></td>
</tr>
<tr>
<td>CCW12/PGA6</td>
<td>19.4765</td>
<td>3.46</td>
<td>12.83 ± 1.93</td>
<td></td>
</tr>
<tr>
<td>ENO1</td>
<td>19.395</td>
<td>2.00</td>
<td>1.69 ± 0.08</td>
<td></td>
</tr>
<tr>
<td>ERG10</td>
<td>19.1591</td>
<td>2.50</td>
<td>1.81 ± 0.04</td>
<td></td>
</tr>
<tr>
<td>FBA1</td>
<td>19.4618</td>
<td>2.04</td>
<td>2.16 ± 0.26</td>
<td></td>
</tr>
<tr>
<td>GDH3</td>
<td>19.4716</td>
<td>3.04</td>
<td>13.91 ± 1.94</td>
<td></td>
</tr>
<tr>
<td>ICL1</td>
<td>19.6844</td>
<td>-4.40*</td>
<td>0.01 ± 0.00*</td>
<td></td>
</tr>
<tr>
<td>PDC11</td>
<td>19.2877</td>
<td>3.02</td>
<td>3.28 ± 0.31</td>
<td></td>
</tr>
</tbody>
</table>

*Genes with a negative fold change in the microarray analysis were downregulated in biofilm. Fold-change values less than 1.0 for real-time RT-PCR also indicate downregulation of genes; values closer to zero indicate greater downregulation.

Phase-dependent gene expression in C. albicans biofilms and planktonic conditions. In the same study (Garcia-Sanchez et al., 2004), comparisons of gene expression profiles between strain CAI4, which grows in wild-type morphology, and a hypha-negative strain that lacks functional Cph1p and Efg1p identified 317 genes that are specific for biofilm development, independent of hypha formation. Genes from these datasets that match genes identified in our current study are noted in our supplementary tables with a ‘U’ to indicate upregulation in the published work (Garcia-Sanchez et al., 2004), or a ‘D’ to indicate downregulation. In the second global study of C. albicans gene expression during biofilm formation, Murillo et al. (2005) focused on the early phase of biofilm formation (to approximately 6 h) on a polystyrene surface. Data analysis identified genes differentially expressed in biofilm and those differentially expressed in planktonic cells. A common gene set (CGS) including genes similarly regulated in biofilm and planktonic cultures was also compiled. In addition, gene expression was compared in C. albicans cells that were adherent to the polystyrene compared to those that were non-adherent. Overlaps between these datasets and our analysis are noted in the supplementary tables available with the online version of this paper. Genes that are differentially expressed in more than one microarray analysis can be identified readily by examining the right-hand columns of each supplementary table.

The gene lists in Supplementary Tables S1 and S2 illustrate how conclusions about gene expression patterns are influenced by the comparison that is made. In Supplementary Table S1, comparisons are made between two time points during biofilm formation (6 vs 12 h, or 12 vs 48 h) or between two time points during the growth of the planktonic culture. These comparisons provide a longitudinal view of processes that occur in the specimens. In contrast, comparisons in Supplementary Table S2 are between biofilm and planktonic specimens at each time point (6, 12 or 48 h). The analyses displayed in Supplementary Tables S1 and S2 can be combined to produce a visual summary of the gene expression data for the biofilm and planktonic specimens, without consideration of C. albicans strain or biofilm model studied (Fig. 2).

For example, expression of RPS4, a gene encoding a ribosomal protein, is more than twofold different between 12 and 48 h in biofilm (Supplementary Table S1). At the 12 h time point, RPS4 expression is over twofold higher in biofilm than in planktonic cells (Supplementary Table S2; Fig. 2). RPS6A1 shows a similar expression pattern (Fig. 2), but is more than twofold different between 6 and 12 h in planktonic cells, between 12 and 48 h in biofilm, and between biofilm and planktonic cells at 6 and 12 h (Supplementary Tables S1 and S2). Expression patterns of two histone-encoding genes (HHF1 and HTA1) are similar to each other and show more than twofold change between 6 and 12 h in both biofilm and planktonic cells (Supplementary Table S2) but are not mentioned on the list of genes differentially expressed between biofilm and planktonic specimens (Supplementary Table S1) since the magnitude of gene expression differences does not exceed twofold at each time point. ALS1 shows a different expression pattern from any of the other examples, with the gene apparently downregulated in biofilms compared to planktonic cells at 6 h (Supplementary Table S2), upregulated during biofilm development at 12 h relative to 6 or 48 h (Supplementary Table S1) and upregulated in planktonic cells at 12 vs 48 h. Although the analysis presented in Supplementary Table S2 can be used to better understand the relationships between gene expression in biofilm and planktonic cells, the nature of the comparisons makes these data more susceptible to influence by growth differences between the biofilm and planktonic specimens at each time point. For this reason, emphasis was placed on interpreting gene expression patterns over time as presented in Supplementary Table S1.

Phase-dependent gene expression in C. albicans biofilm and planktonic cultures

Of the 243 genes that are differentially expressed over the experimental time-course in either biofilm or planktonic cells (Supplementary Table S1), the vast majority (191 genes) are differentially expressed only during biofilm development. Fig. 3 summarizes differential gene expression data over the time-course of biofilm development by listing functional categories of genes that are important at each developmental stage studied. Details regarding differential gene expression during biofilm development are provided in Supplementary Table S1. Five of the 14 differentially expressed amino acid biosynthetic genes are
involved in sulfur amino acid metabolism. A similar emphasis on sulfur amino acid biosynthetic pathways was observed in previous studies that transcriptionally profiled C. albicans biofilm (García-Sánchez et al., 2004; Murillo et al., 2005). Also represented among genes in this functional category were those involved in metabolism of branched-chain amino acids, glutamate and glycine. Genes in the glycolytic pathway showed the largest fold change in expression among those in Supplementary Table S1. Glycolytic genes were similarly upregulated at 6 vs 12 h in planktonic cells. Among the other carbohydrate and carbon compound metabolism genes differentially expressed during biofilm development were those from the tricarboxylic acid cycle, pyruvate metabolism, gluconeogenesis, the pentose phosphate pathway and the glyoxylate cycle.

Upregulation of histone-encoding genes (HHF1, HTA1 and HTB1) at 6 vs 12 h was a common feature of biofilm development and planktonic culture growth (Supplementary Table S1). In biofilms, these genes are part of a larger group of upregulated genes that functions in chromatin assembly and chromosome function. A cytoskeletal gene (TUB2), as well as genes that encode subunits of the ribonucleotide reductase complex (RNRI and RNRR2), were also upregulated at 6 vs 12 h in biofilm. The ribonucleotide reductase complex catalyses the rate-limiting step in deoxyribonucleotide synthesis and is regulated by DNA replication and DNA damage checkpoint pathways (Elledge et al., 1993). Many genes involved in cellular transport processes were differentially regulated during biofilm development, and tended to not be included in the list of differentially expressed genes in planktonic cells (Supplementary Table S1). Genes involved in intracellular transport were upregulated mostly at 6 vs 12 h. Genes upregulated at the 12 vs 6 h or 12 vs 48 h time point tended to be involved in uptake of amino acids, ions, oligopeptides, nucleosides and lactate/pyruvate that can feed the metabolic needs of the biofilm as it increases its biomass. The upregulation of genes encoding high-affinity, broad-specificity hexose transporters (HGT6, HGT7, HGT8 and HGT19) underscores the importance of carbohydrate metabolic pathways in biofilm formation.

Genes encoding cell wall proteins were upregulated at both the 6 and 12 h time points. Among the upregulated genes

Fig. 2. Graphical representation of fold-change data for expression of five different C. albicans genes in biofilm and planktonic specimens over the experimental time-course. Data from Supplementary Tables S1 (data analysis over time for either biofilm or planktonic specimens) and S2 (analysis at each time point comparing biofilm and planktonic specimens) can be combined to generate the graphical representations of fold change in gene expression shown here. Genes encoding ribosomal proteins (RPS4 and RPS6A) are shown in (a) and (b). Genes encoding histones (HHF1 and HTA1) are shown in (c) and (d). The expression pattern for ALS1 is shown in (e). Trends over the experimental time-course for each type of specimen (biofilm or planktonic) are evident, as well as the relationship between the gene expression values for biofilm (diagonal striped bars) and planktonic (light grey bars) specimens at each developmental time point. Construction of the graphs using the log of the mean fluorescence intensity units produced histograms (data not shown) for biofilm and planktonic cells with the same basic relationships as those shown here.
are many that encode GPI-linked cell wall proteins and also genes that are involved in synthesis of cell wall components such as $\beta$-1,3-glucan, $\beta$-1,6-glucan and mannoproteins. Upregulation of ALS genes was also noted in our studies. Our comparisons between $ALS1$ expression in planktonic cells and biofilm at various time points (Supplementary Table S2; Fig. 2) showed a relative downregulation of $ALS1$ at 6 h in biofilm and no differential expression between the two growth conditions at the other time points. During biofilm development (Supplementary Table S1; Fig. 2), $ALS1$ expression surges between 6 and 12 h and then declines in mature biofilm. These data are consistent with previous in vitro investigations that showed that $ALS1$ expression peaks shortly after $C. albicans$ cells from a saturated culture are placed into fresh culture medium and then steadily decreases over time (Green et al., 2005). $ALS1$ expression also peaks shortly after $C. albicans$ cells are placed into growth medium that promotes germ tube formation (Green et al., 2005; Zhao et al., 2004). Collectively, these data demonstrate that the $ALS1$ expression level change is influenced by the growth stage of a culture and that conclusions about whether $ALS1$ is up- or downregulated in biofilm depend on the comparison that is made.

Garcia-Sanchez et al. (2004) also observed $ALS1$ to be the most differentially expressed gene in their analysis of biofilm formation. This result led O’Connor et al. (2005) to use quantitative reverse transcriptase PCR to test a set of clinical isolates for $ALS1$ expression in flask-grown planktonic cells (37 °C, 24 h) and 48 h catheter biofilm. Upregulation of $ALS1$ in biofilm was observed for each $C. albicans$ isolate. Nailis et al. (2006) observed $ALS1$ upregulation in catheter biofilm grown in a bioreactor with continuous growth medium replacement compared to a flask-grown planktonic control culture. $ALS1$ was one of the cell-surface proteins regulated by Bcr1p that was proposed to play an adhesive role during biofilm formation (Nobile et al., 2006a) although Als1p was not required for biofilm formation in vitro. Murillo et al. (2005) demonstrated that $C. albicans$ cells that adhered to the polystyrene substrate had threefold less $ALS1$ transcription than cells that were free-floating in the culture well. However, as biofilm development progressed, $ALS1$ expression increased. It is possible that this $ALS1$ expression increase is due, at least in part, to formation of germ tubes that occurs in the developing biofilm. Although $ALS1$ may react differently depending on the biofilm substrate, current evidence points to a role for $ALS1$ later in biofilm formation rather than in initial attachment to the biofilm substrate.

### Table 1: Gene Expression Changes in $C. albicans$ Biofilm Development

<table>
<thead>
<tr>
<th>Time (h)</th>
<th>Gene Categories</th>
</tr>
</thead>
</table>
| 6        | - Amino acid metabolism  
|          | - Glycolysis  
|          | - Chromatin assembly  
|          | - DNA synthesis  
|          | - Intracellular transport  
|          | - Cell wall synthesis  
|          | - Lipid biosynthesis  
|          | - Stress response  |
| 12       | - Amino acid metabolism  
|          | - Pyruvate metabolism  
|          | - Uptake of amino acids, ions, oligopeptides, nucleosides, lactate/pyruvate  
|          | - Hexose transport  
|          | - Cell wall synthesis  
|          | - Energy production  
|          | - Nucleotide biosynthesis and salvage  
|          | - Protein synthesis/translation  
|          | - Stress response  
|          | - Transcription  |
| 48       | - GIT1  
|          | - UGT51C1  
|          | - HSP104  
|          | - TEC1  
|          | - orf19.7065  |

**Fig. 3.** Summary of the cellular processes that are associated with genes upregulated at the different time points during $C. albicans$ biofilm development. The time-course of biofilm development is shown, highlighting the time points (6, 12 and 48 h) studied by microarray analysis. Descriptions of cellular processes are summarized from the detailed data presented in Supplementary Table S1. Categories of genes upregulated at 6 vs 12 h are summarized under the 6 h heading. Data from both the 12 vs 6 h and 12 vs 48 h comparisons are placed under the 12 h heading. Few genes are upregulated at 48 h compared to 12 h, suggesting that initiation of new metabolic activity is relatively low in the mature biofilm. The individual genes upregulated at 48 vs 12 h are listed.
Genes involved in energy production were upregulated at 12 h of biofilm development, primarily compared to the 48 h time point (Supplementary Table S1). These genes are associated with a variety of energy-related processes including structural constituents of the H⁺-transporting ATP synthase (ATP2, ATP4 and ATP5), mitochondrial electron transport (orf19.1710, CYT1, ABCI and RPI1), respiratory chain complexes (orf19.4016), ATP:ADP antiporter activity (PET9), and linking of haem to cytochrome c (CYC3). Only one gene encoding a structural constituent of the H⁺-transporting ATP synthase (PMA1) was upregulated solely in planktonic cells at 6 vs 12 h. In contrast to the energy-related genes, genes encoding proteins involved in fatty acid, lipid and isoprenoid metabolism were upregulated at each biofilm developmental time point assayed (Supplementary Table S1). These genes are primarily involved in synthesis of ergosterol and other membrane lipids, and in peroxisomal fatty acid degradation. Upregulation of ERGI0 during the early phase of biofilm development is consistent with previous observations that showed high levels of ergosterol at the early phase, with approximately a 50% reduction by the intermediate and mature phases of growth (Mukherjee et al., 2003).

Multiple genes involved in biosynthesis or salvage of purines or pyrimidines are upregulated in biofilm at 12 h relative to 48 h (Supplementary Table S1). This pattern of expression is similar to that observed for genes involved in protein synthesis/translation and corresponds to the time point at which the biofilm is accumulating the most biomass (Kuhn et al., 2002). Most of the genes in the protein synthesis/translation category encode ribosomal components, although several translation initiation and elongation factors are also included. Only one of these genes, RPS6A, is upregulated during planktonic growth, but at the 6 vs 12 h time point, rather than the 12 vs 48 h point most commonly observed for biofilm-associated genes.

Genes that respond to environmental conditions are differentially expressed at various time points over the course of biofilm development. This list includes genes involved in response to oxidative stress (CAT1, GCS1, TSA1, NCE103 and orf19.5180), osmotic stress (RHR2), and thermal stress (HSP90 and HSP104). Upregulation of NCE103, which encodes carbonate dehydratase, and GCS1, which encodes the rate-limiting enzyme in glutathione biosynthesis, were observed during the very early stages of biofilm development in the study by Murillo et al. (2005). NCE103 expression is upregulated immediately upon attachment of C. albicans to a polystyrene surface while GCS1 upregulation is associated with biofilm development after a second basal layer of yeast cells has formed. The relationships between these enzymes and sulfur metabolism were noted (Murillo et al., 2005).

Genes associated with transcriptional processes are also differentially expressed at various time points over the course of biofilm development. Featured in this group are several genes encoding RNA polymerase and transcription factors. TEC1, which regulates hypha-specific genes, was one of a few genes upregulated at 48 vs 12 h. TEC1 is one of a very few C. albicans genes that when disrupted results in a strain with a severe defect in biofilm formation on a silicone elastomer surface (Nobile & Mitchell, 2005). The list of genes differentially expressed in biofilm and planktonic specimens also includes 38 genes (16%) that encode unnamed, hypothetical proteins (Supplementary Table S1). Although most of these genes are differentially expressed only in biofilm, expression of six genes is unique to planktonic cells in our experiments. Expression patterns represented within this gene set are variable, and span the entire experimental time-course.

Fifty-two genes were differentially expressed as a function of time during planktonic culture growth (Supplementary Table S1). Differential expression of 18 genes, involved in diverse cellular processes, was unique to planktonic cells. Of the 52 genes, 34 were also differentially expressed during biofilm development. Over half of these genes (18/34) showed a similar temporal pattern of expression in both planktonic cells and developing biofilms. Examples of similarly regulated genes included those encoding histones (HHFI, HTAI and HTBI) and glycolytic enzymes (ENO1, FBA1, PDC11, PGK1, GDN1 and TPI1). Our criteria for differential gene expression (≥2-fold change) applied to far fewer genes in the planktonic cells compared to those in biofilm (Supplementary Table S1). These results create the impression of a very dynamic biofilm development process that is marked by larger changes in gene expression levels, and a comparatively static planktonic culture in which changes in gene expression levels are of smaller magnitude. To determine whether planktonic growth was characterized by upregulation <2-fold or whether gene expression changes did not occur at all, we examined expression data for genes involved in energy generation (Supplementary Table S1). This functional category was represented strongly in our data, and generally lacked genes that were upregulated in planktonic cells. Moreover, energy production clearly is important for growth of the planktonic culture. We found that expression levels of key genes were upregulated during planktonic growth, but at a magnitude <2-fold. For example, RIP1, which is involved in mitochondrial electron transport, was upregulated at 12 vs 6 h with a fold-change value of 1.98 that just missed our ≥2.0 cut-off value (Supplementary Table S1). Similar to its expression pattern in biofilm, ATP2, which encodes a subunit of the H⁺-transporting ATP synthase, is also upregulated in planktonic cells at 12 vs 48 h, with a fold-change value of 1.6.

Another possible explanation for the appearance of a comparatively static planktonic culture is that the planktonic culture has progressed further in its growth curve than the biofilm and that the dynamic processes associated with exponential growth have already occurred by the 6 h time point. For example, the large increase in expression of
genes encoding ribosomal subunits that is observed in biofilm apparently did not occur in the planktonic culture (Supplementary Table S1). Further examination of the data showed upregulation of expression of these genes at 6 vs 12 h in the planktonic culture, although with values that were between 1.0 and 2.0 and, therefore, not mentioned in the data tables. The relationship between expression in planktonic cells and biofilm of two representative ribosomal subunit-encoding genes (RPS4 and RPS6A) showed expression levels apparently rising toward 12 h in biofilm, but static by 12 h in planktonic cells (Fig. 2). These data suggest a difference in growth phase between planktonic culture and biofilm at each of the time points. However, the expression patterns of other genes do not support this conclusion. For example, the expression of histone-encoding genes was similar in both planktonic culture and biofilm at each of the time points. However, the overall conclusion suggested by these analyses is that cells committed to biofilm growth share many similar processes with cells growing in suspension, but that there are numerous differences between the two growth forms.

**Biofilm model-specific and C. albicans strain-specific gene expression patterns**

One feature that makes this study unique is the parallel comparisons between different biofilm models and C. albicans strains within those models. Analysis presented in Supplementary Tables S1 and S2 utilized the biofilm and planktonic datasets as a whole, regardless of biofilm model or C. albicans strain. Supplementary Table S3 compares gene expression results for the denture and catheter biofilm models. This analysis shows that growth of C. albicans in the denture model is marked by more dynamic changes in gene expression than C. albicans growing in the catheter biofilm model (Supplementary Table S3). These differences are most notable in the early (6 h) and mature (48 h) phases of growth. The differences between the biofilm models that were highlighted in Supplementary Table S3 are expanded upon in Supplementary Table S4, in which biofilm gene expression data from the four strain/substrate combinations (GDH2346/denture, SC5314/denture, M61/catheter and SC5314/catheter) are analysed separately. Supplementary Table S5 shows analysis of the planktonic gene expression data from the four strain/substrate combinations. Overall, there was greater variability among the biofilm datasets (Supplementary Table S4) than among the planktonic datasets (Supplementary Table S5). Among the biofilm datasets (Supplementary Table S4), there was greater variability between the denture biofilm datasets (strains GDH2346 and SC5314) than between the catheter biofilm datasets (strains M61 and SC5314; Supplementary Table S4). When comparing all four biofilm datasets at the 6 h time point, far more genes were unique to the SC5314/denture biofilm dataset than the other strain/substrate combinations examined. At the 12 h time point, however, the GDH2346/denture biofilm dataset had the highest number of unique genes. These results suggested differences in the SC5314/denture specimens at the early time point that were not apparent as the biofilm matured. In contrast, the GDH2346/denture specimens diverged from the group at 12 h. By 48 h, there were few differentially expressed genes in any of the four datasets, making the datasets for the different strain/substrate combinations appear fairly consistent compared to earlier time points when gene lists were more extensive.

Analysing data from each strain/substrate combination individually increased the number of genes differentially expressed in biofilm by nearly threefold (comparing Supplementary Tables S1 to S4). The protein modification, transcription, unknown protein function, and amino acid metabolism functional categories showed the largest increases in gene number. By contrast, very few extra genes involved in energy generation were added by the strain/substrate-specific analysis. Genes that were differentially expressed in the strain/substrate-specific analysis (Supplementary Table S4) tended to fill in the same metabolic pathways and be associated with the same basic cellular processes that were represented in the overall analysis (Supplementary Table S1). An example is the presence of one gene involved in iron transport (FET35) in the overall analysis, compared to five genes in the strain/substrate-specific analysis, which were particularly associated with the GDH2346/denture dataset. Genes involved in metabolism of sulfur amino acids were also more numerous in the strain/substrate-specific analysis than in the overall analysis (Supplementary Tables S1 and S4). Analysis of the planktonic datasets showed the same trends (Supplementary Table S5). In this case, over fourfold more genes were identified as differentially expressed in planktonic cells when the strain/substrate datasets were analysed individually (Supplementary Table S5) rather than together (Supplementary Table S1). Many of the genes identified in the strain/substrate-specific analysis of the planktonic cell datasets (Supplementary Table S5) were the same as those that were differentially expressed in the biofilm strain/substrate-specific analysis (Supplementary Table S4). However, 38 genes that were differentially expressed in planktonic cells (shaded pink in Supplementary Table S5) were not included in any lists of genes differentially expressed in biofilms. These genes were particularly abundant in the cell wall (ALS7, PIR1, orf19.3895, CHT3, PAGA58, WSC1 and RBR3) and unknown protein function (nine genes) categories.

**Gene expression changes associated with time-point-related biofilm phenotypic changes**

The time points utilized in this study were selected because of their association with phenotypic changes that are important to biofilm development and antifungal drug resistance. For example, the 6 h time point was selected because it is the point at which antifungal drug resistance becomes measurable in the model systems studied.
(Chandra et al., 2001). The 12 h time point was selected since it marks the start of the tremendous increase in biomass that occurs during biofilm development.

Because drug resistance assays often compare biofilm to planktonic cells at the same time points (Chandra et al., 2001), the analysis in Supplementary Table S2 is appropriate to examine for differentially expressed genes that might explain biofilm-related changes in drug resistance. These microarray data are consistent with previous phenotypic observations. For example, upregulation of ERG10 at the 6 h time point (Supplementary Table S2) is in agreement with the findings of Mukherjee et al. (2003), who demonstrated that azole resistance in C. albicans biofilms is associated with phase-dependent alterations in membrane sterol levels and composition. The strain/substrate-based analysis of biofilm gene expression data also highlighted differential expression of many genes that might contribute to biofilm antifungal drug resistance (Supplementary Table S4). Additional genes in the ergosterol biosynthetic pathway (ERG4, ERG11 and ERG24) were differentially expressed, as were genes encoding efflux pumps and transporters (QDR1, MDR97, CDR4, ATM1 and orf19.2350). In most cases, differential expression was associated with specific strain/substrate combinations (Supplementary Table S4). This strain/substrate specificity of gene expression parallels data from planktonic cells that show strain-dependent differences in the mechanisms underlying drug resistance (Ghannoum & Rice, 1999; White, 1997). Consistent with previous findings, our overall microarray analysis suggests that multi-component, phase-specific mechanisms are operative in the antifungal resistance of C. albicans biofilms in a strain- and/or substrate-dependent manner.

Gene expression was analysed at the 12 h time point to understand the mechanisms at work during the large increase in biomass that is initiated at this time point. Much of the biomass increase is due to accumulation of ECM (Kuhn et al., 2002). More genes are upregulated in biofilm at the 12 h time point than at the other time points assayed (Supplementary Tables S1 and S2). However, we did not observe upregulation of genes that would implicate a novel biochemical process in the formation of the ECM. Instead, the upregulated genes function in the same basic processes that are required for C. albicans growth. This observation is consistent with the conclusion that C. albicans biofilm ECM is composed of the same types of molecules that are found in planktonic cells. Recently Al-Fattani & Douglas (2006) demonstrated that glucose is the major component of the C. albicans biofilm ECM. The increased expression of hexose transporters at the 12 h time point could also function to support ECM production.

**Progress toward defining biofilm-associated gene expression**

The temporal changes in gene expression that we observed in biofilm provide a mechanistic understanding of the phenotypic changes during biofilm development. Glycolytic and non-glycolytic carbohydrate assimilation, amino acid metabolism, and upregulation of intracellular transport mechanisms are important during the early phase of biofilm development, where C. albicans cells have adhered to the IMD surface and are dividing to form microcolonies. Early events result in an increased intracellular pools of pyruvate, pentoses and amino acids that prepare the biofilm for the large increase in biomass that is, in part, due to the accumulation of large amounts of carbohydrate-rich ECM that begins at approximately 12 h of development. More genes are upregulated in biofilm at the 12 h time point than at the other points assayed. The cellular processes occurring in the intermediate phase of biofilm development demand energy, consistent with the observed upregulation of genes involved in ATP synthesis at this time point. Upregulation of genes encoding specific transporters for amino acids, sugars, ions, oligopeptides and lactate/pyruvate suggests an increased requirement or decreased synthesis of these key building blocks, which are needed to support the large increase in biofilm biomass. As the biofilm reaches mature phase, expression of few genes is differentially upregulated compared to 12 h suggesting a relative lack of initiation of new metabolic activity. Reduced metabolic activity may conserve energy and nutrients and avoid stress inducers, and may also provide an explanation for the drug-resistant phenotype of C. albicans biofilm. Our extensive microarray analyses indicate that many of the processes that contribute to biofilm development are the same as those that are associated with planktonic growth.

Despite differences in experimental design and focus, our data compare favourably with those from two other investigations of biofilm gene expression (Garcia-Sanchez et al., 2004; Murillo et al., 2005). Far fewer genes were differentially expressed at 48 h in our study than reported by Garcia-Sanchez et al. (2004), who analysed a diverse biofilm dataset, composed of a statistical summary of data points from biofilms grown under many conditions. Key among these growth conditions was biofilm formation in a microfermenter that supplied fresh medium to the biofilms over the course of 48–72 h and growth of planktonic cultures primarily for 48 h. These experimental conditions and the subsequent data comparisons most closely resemble those in our Supplementary Table S2. Overlap between the published gene sets (Garcia-Sanchez et al., 2004) and ours occurred primarily for genes identified at our 12 h time point, suggesting that the 48 h biofilms analysed by Garcia-Sanchez et al. (2004) were still undergoing active growth. Although overlap between lists of differentially expressed genes was somewhat limited, the cellular processes identified in the two studies are similar and include protein synthesis/translation, and amino acid and nucleotide metabolism. Differential expression of genes involved in transport processes was not as marked in the work of Garcia-Sanchez et al. (2004) as in our current study. Our microarray analysis is perhaps less comparable to the work of Murillo et al. (2005) who
utilized a polystyrene substrate and studied closely spaced time points within approximately the first 6 h of biofilm development. Most prominently featured among genes upregulated within the first 30 min of biofilm development are those involved in sulfur and phosphate metabolism and iron assimilation and homologues of genes in the *Saccharomyces cerevisiae* Ribi regulon, which is involved in ribosome biogenesis. Parallels between our dataset and that of Murillo et al. (2005) were perhaps most notable in expression of genes involved in sulfur metabolism and oxidative metabolism as discussed above. The fact that parallels can be found between disparate datasets suggests that processes fundamental to biofilm development are conserved across various models.

Based on their microarray results, which showed the importance of amino acid biosynthesis genes in biofilm formation, Garcia-Sanchez et al. (2004) tested and demonstrated the importance of the Gcn4p transcriptional activator for normal biofilm formation. However, transcriptional profiling of biofilm is not the only method that has been used to gain insight into the genes that are associated with *C. albicans* biofilm development. Assay of transcription factor mutants showed that Efg1p is required for wild-type biofilm formation on plastic surfaces, suggesting a requirement for Efg1p-mediated filamentation or for the product of an Efg1p-regulated gene for biofilm growth (Ramage et al., 2002). Screening of a set of transcription factor mutants revealed that strains lacking Bcr1p could not form wild-type biofilms (Nobile & Mitchell, 2005). Further work demonstrated the contributions of Bcr1p targets ALS3, ALS1, HWI1 and ECE1 to biofilm formation (Nobile & Mitchell, 2005; Nobile et al., 2006a, b). Proteomic screening of early phase *C. albicans* biofilms grown in the catheter model showed decreased production of alcohol dehydrogenase (Mukherjee et al., 2006). Evaluation of sets of genes or gene families indicated that loss of wild-type *ALS3* (Zhao et al., 2005) or ALS3 (Zhao et al., 2006) decreases biofilm mass in the catheter model and that deletion of genes encoding the CFEM (Common in several Fungal Extracellular Membrane) domain results in abnormal biofilm formation on a polystyrene surface (Perez et al., 2006). Likewise, several of the *PMT* genes, which encode protein mannosyltransferases responsible for O-glycosylation of mannoproteins, are required for accumulation of wild-type biofilm mass when assayed on a polystyrene surface (Peltroche-Llacsaahuanga et al., 2006). Screening of a set of *C. albicans* mutant strains showed that disruption of *NUP85*, *MDS3*, *KEM1* or *SUV3* blocks biofilm development and suggests the importance of hypha formation in this process (Richard et al., 2005). Evaluation of the cellular phenotype of *C. albicans* mutant strains showed that *NOT4* (Krueger et al., 2004), *ACE2* (Kelly et al., 2004) and *MKC1* (Kumamoto, 2005) were required for wild-type biofilm formation. Deletion of the yeast form-specific gene *YWIP1* resulted in *C. albicans* cells with increased adhesiveness and biofilm formation (Granger et al., 2005). Many, varied experimental approaches have contributed to our current understanding of gene expression associated with biofilm development. Our microarray analyses presented here suggest that *C. albicans* biofilms exhibit changes in gene expression to adjust to varying requirements during development, with the expression of specific subsets of genes at different developmental phases. Understanding changes in gene expression that are associated with biofilm formation on denture and catheter surfaces increases the utility of these models for testing of potential inhibitors of biofilm formation and for testing of *C. albicans* mutant strains to understand the effects of candidate genes on biofilm development.

**ACKNOWLEDGEMENTS**

We thank Bill Schnitzlein for contributing his real-time RT-PCR expertise to this study, M. A. G. and L. L. H. thank Mira Edgerton and their fellow members of the Consortium for Candida DNA Microarray facilities (J. Berman, P. Orlean, P. T. Magee and P. Sundstrom) for their efforts in constructing the *C. albicans* microarray. Judith Berman, Suzanne Grindle and John Crow contributed bioinformatics analysis that was essential for microarray construction. The skilled and highly accurate efforts of Lynne Herron were instrumental in microarray construction. We thank Joel Peek at Microarrays, Inc. for technical guidance in designing and printing the custom PCR-based microarrays. We thank Emma Hoyer for devoting considerable effort to completing the tables for this manuscript. This project was funded by United States Public Health Service grant R01 DE13932 from the National Institute of Dental and Craniofacial Research (NIDCR) and the Bristol Myers Squibb Freedom to Discover Award (to M. A. G.), and by grant R01GM59826 (to K. E. K.). Microarray construction was funded by grant R01 DE10641-01 (to Mira Edgerton) from NIDCR. Sequence data for *Candida albicans* were generated at the Stanford DNA Sequencing and Technology Center with the support of NIDCR and the Burroughs Wellcome Fund. Funding was also provided by Scientist Development Grant 0335313N (to P. K. M.) from the American Heart Association and a Pilot and Feasibility Study (to P. K. M.) from the Skin Diseases Research Center (Department of Dermatology, Case Western Reserve University). This investigation was conducted in a facility constructed with support from Research Facilities Improvement Program Grant Number C06 RR16515-01 from the National Center for Research Resources, National Institutes of Health.

**REFERENCES**


