4-2009

Gene expression of Saccharomyces cerevisiae exposed to commercial wood preservatives by DNA Microarray Analysis and RT-PCR

Madison M. Stevens
Longwood University

Gene expression of *Saccharomyces cerevisiae* exposed to commercial wood preservatives by DNA Microarray Analysis and RT-PCR

A Senior Honors Thesis

Submitted to the Faculty of Longwood University

in fulfillment of the requirements for the degree of

Bachelor of Science with Honors

Biological and Environmental Sciences

Longwood University

April 2009

Madison M. Stevens

with Assistance from Advisor

Dr. Consuelo J. Alvarez

Table of Contents

Abstract 3

Introduction 4

Materials and Methods 7

Cell Preparation 8

Total RNA Extraction 9

cDNA Preparation and Hybridization 9

Yeast DNA Microarray chips: Source and Scanning 10

Dye-Swap DNA Microarray Chips 11

Data Analysis 11

MAGIC Tool 11

Statistical Analysis Using MAGIC Tool and Microsoft Excel 12

Gene Expression Validation by RT-PCR 14

BLASTn and BLASTp of Significant Genes 14

Results 15

Statistical Analysis 15

RT-PCR Validation 17

MAGIC Tool Clustering and BLASTn and BLASTp Analysis 19

Discussion and Conclusions 25

Acknowledgements 28

References 29

Appendices 31

ABSTRACT

Creosote and pentachlorophenol (PtCP) are two commercial wood preservatives that are regulated by the EPA because of their toxicity to wildlife and humans. Creosote and PtCP have been suspected of causing cancer in humans, but that claim has not been proven. To observe changes in gene expression in organisms exposed to these compounds, a model system such as *Saccharomyces cerevisiae* (baker's yeast) is used. *S. cerevisiae* cells were exposed to a creosote concentration of 50ng/ml and to a PtCP concentration of 50μM. Since creosote and PtCP were suspended in methylene chloride and ethanol, respectively, yeast cells were also exposed separately to the solvents as controls. cDNA was prepared from a total RNA extraction of exposed and non-exposed *S. cerevisiae* cells and was hybridized onto microarray chips containing the entire yeast genome using Genisphere Array Kit procedures. A total of twenty microarray chips were tested (seven creosote chips, seven PtCP chips, three methylene chloride chips, and three ethanol chips) for this study. Analysis of the microarray data was done using MAGIC Tool software and Microsoft Excel to find statistical significance in gene expression. Genes showing significant changes in genes expression underwent real-time polymerase chain reaction (RT-PCR) to validate that their change in gene expression was correctly measured in the microarray experiment and that it is due to genetic regulation. Because creosote and PtCP have been indirectly linked to causing cancer in humans, clustering analysis in MAGIC Tool and BLAST analysis on the National Center for Biotechnology Information (NCBI) website compared genes with significant changes in gene expression to other genes in the *S. cerevisiae* genome and genes within the human genome. In both experimental treatments, creosote and PtCP, genes with roles in cell cycle regulation,

drug transport, transcription regulation, and response to stress had significant changes in gene expression. RT-PCR analysis verified that the changes in gene expression could be validated. Clustering analysis in MAGIC Tool revealed highly correlated gene expression in genes associated with mitotic controls. BLASTn and BLASTp analysis confirmed that some genes with significant changes in gene expression had homology to human nucleotide and protein sequences. Overall, the results of this DNA microarray study of *S. cerevisiae* cells exposed to wood preservatives are a sign of the necessity for more studies to be done by the EPA and workers' health associations in order to establish job/health regulations. The results of this study could be a starting point for R-1 institutions that concentrate in cancer studies.

INTRODUCTION

With the ever increasing abundance of genomic information, technologies such as DNA microarray allow researchers to move away from traditional genetics research where one gene is studied at a time. Instead, an entire genome can be evaluated simultaneously and potential interactions among genes can be investigated. To explore potential effects of chemical stressors on gene expression, *Saccharomyces cerevisiae* may be used as a model genetic system due to its simple structure and regulatory gene functions (Pollack and Iyer, 2002). Gene expression is the process of transferring information from DNA to mRNA and from mRNA to proteins; so, the measure of changes in mRNA levels will be an indication of differential gene expression in an organism. DNA microarray analysis coupled with bioinformatics and statistical analysis represents a powerful tool for the study of variations in gene expression.

Two compounds, creosote and pentachlorophenol (PtCP) were selected as the experimental treatments, and DNA microarray analysis was applied to provide relevant information about their carcinogenic properties (ATSDR, 2001 and 2002). Creosote and PtCP are commercial wood preservatives and are known to be toxic to humans and wildlife. However, they are still in use today protecting railroad ties, utility poles, and wharf pilings from insect and water damage. They are slowly being phased out of use by commercial companies because of their carcinogenic properties. Human exposure to these wood preservatives is thought to cause cancer; however, there is no solid evidence to directly link the two. The EPA has released risk assessments of creosote and PtCP. The excerpt below from the EPA risk assessment of PtCP:

> "Some studies have found an increase in cancer risk in workers exposed to high levels of technical grade pentachlorophenol for a long time, but other studies have not found this. Increases in liver, adrenal gland, and nasal tumors have been found in laboratory animals exposed to high doses of pentachlorophenol. The EPA has determined that pentachlorophenol is a probable human carcinogen and the International Agency for Cancer Research (IARC) considers it possibly carcinogenic to humans (ATSDR, 2001)."

Creosote (Figure 1A in Appendix 1) is a mixture of approximately 300 different constituents, the most prevalent coming from coal tar (ATSDR, 2002). It is used as a fungicide, insecticide, sporicide and wood preservative (EPA, 2008). Creosote is used in some medicines to treat psoriasis (LOSH, 2003). However, creosote has also been linked to increased cancer risks. In one case, a 50-year-old railroad worker developed squamous cell carcinoma on his knee after 30 years of handling creosote soaked wood. Sun-related

carcinogenesis was ruled out because the affected area was clothed, thus increasing the likelihood that the cause of the cancer was linked to repeated creosote exposure (Carlsten *et al.,* 2005). Creosote contamination has been noted to increase the levels of a cancer biomarker in the liver and gills of rainbow trout (McClain *et al.,* 2003).

PtCP (Figure 1B in Appendix 1) is a chemical that, upon human exposure, causes cells in the body to produce excess heat, and thus, damages body tissues (ATSDR, 2001). Rats treated with increasing doses of PtCP (from 200 to 3,200 ppm) showed fluctuations in liver weight, thinning hair, and loss of appetite compared to non-treated rats (U.S. DHH 1999). Studies by Kitagawa *et al.* (2003) suggest that PtCP produces the highest differential expression in genes involved in cellular transport, protein distribution, and metabolism.

The same study by Kitagawa *et al.* (2003) reports that exposing *S. cerevisiae* to PtCP does result in changes in gene expression, but the study was limited by grouping the result only by general cell function defects. In an effort to gain deeper understanding of how compounds like PtCP and creosote affect cellular function, one focus of this project was to identify specific subsets of genes with similar expression that also have a correlation between their biological functions, molecular functions, and cellular components.

Concentrations for creosote and PtCP were chosen based on published reports that showed changes in gene expression in yeast and fish (Roling *et al,* 2004; McClain *et al.*, 2003; Kitagawa *et al.*, 2003). A creosote concentration of 50ng/mL was employed in experiments that exposed mummichogs (small fish also known as a killifish) to pyrenes (poly aromatic hydrocarbons which have similar properties as creosote) and also showed

gene expression changes (Roling *et al.,* 2004). A 50μM concentration of PtCP, was previously reported to cause changes in gene expression in *Saccharomyces cerevisiae* (Kitagawa *et al.*, 2003).

Studying the gene expression of *S. cerevisiae* when exposed to the wood preservatives is a way to see how genes are affected when exposed to carcinogens. Data analysis to find genes with a significant change in gene expression was performed with MicroArray Genome Imaging and Clustering Tool (MAGIC Tool) (Heyer *et al.*, 2005) and Microsoft Excel. In addition, by using the real time Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) those genes with significant changes in expression were validated. It is important to first understand the trends in genes expression of *S. cerevisiae* before comparing the data to any other organism. After a thorough examination of the trends in the significant genes, websites such as the National Center for Biotechnology Information (NCBI) allow for comparisons between the significant *S. cerevisiae* genes found in the experiment and human genes associated with cancer.

MATERIALS AND METHODS

All compounds were obtained from commercial sources. The pentachlorophenol (PtCP) 86% was obtained from Acros Organics a company within Fisher Scientific, Pittsburg, PA, catalog # 161120010; creosote in methylene chloride (organic standard of 1000 ug/mL) was obtained from Fisher Scientific, Pittsburg, PA, catalog # S-3847; ethanol (ethyl alcohol 200 proof) was obtained from Pharmaco Products Inc, Brookfield, CT, catalog # 111HPL200CS4L; and the methylene chloride 99% (solvent for creosote) was obtained from Fisher Scientific, Pittsburg, PA, catalog # D-143-4. Creosote was obtained from the distributor without further purification. However, pentachlorophenol is

a solid that is insoluble in the standard yeast growth media used, so 15.5 mg of the PtCP was suspended in 1mL of ethanol to obtain a 1000X stock solution (50mM). The cMasterPure^{TM} Yeast RNA Purification Kit was obtained from Epicentre, Madison, WI, catalog # MPY03100. The 3DNA 350 kit was obtained from Genisphere, Hatfield, PA, catalog # W300180.

Cell Preparation

The *Saccharomyces cerevisiae* yeast strain X2180-1A was provided by Dr. Kirk Anders from the Department of Genetics at Stanford University School of Medicine. A full loop of cells (from a sample stored at –80°C) was inoculated in a 5.0 mL aliquot of sterile liquid media and allowed to grow at 30°C in a shaking water-bath. The liquid media is made up of 10.0 g of yeast extract, 20.0 g of yeast peptone, and 20.0 g of dextrose for a 1.00 L volume (YPD media). Once a cell culture reached an optical density of 2.0 at 600 nm, the sample was diluted to a total volume of 5.0 mL with fresh YPD media and incubated overnight to obtain a cell culture with an optical density of approximately 0.2-0.3. For each compound investigated, two separate samples were utilized: one labeled as wild-type (no exposure to the compound) and the second labeled as exposed. To the "exposed" PtCP samples, 5 uL of 50 mM stock PtCP solution was used to obtain a concentration of 50uM (Kitagawa *et al.*, 2003). For the creosote, 12.5 uL of the purchased solution was used to obtain a concentration of 50ng/mL (Roling *et al.*, 2004). The wild type sample was used to evaluate overall cell growth, and all samples were allowed to grow until the wild type sample obtained an optical density at 600 nm of 1.1. This required five hours of growth. Cell growth occurred at 30°C in a

shaking, water bath. All cell culture samples were then harvested in preparation for RNA extraction (as described in Walker *et al.*, 2008).

Total RNA Extraction

Total RNA was isolated by using a modified version of the MasterPure^{TM} Yeast RNA Purification Kit from Epicentre. The modifications are as follows. For part A (RNA purification), a total of 2.0 ml of cell culture was used (instead of the suggested 1.0–1.5 mL). The incubation at 70°C was extended from 10–15 minutes to 30 minutes total time. For part B (Removal of contaminating DNA from RNA preparations), the DNase I treatment was always used. The incubation time at 37°C was increased from 10 minutes to 30 minutes total time. A second centrifugation step was added before step 9 to obtain a completely clear supernatant. Finally, at the end of step 13, the cleaned RNA preparation was allowed to precipitate in ethanol at –20°C and was stored until the cDNA preparation was initiated. The afternoon before cDNA preparation began, RNA was recovered from the ethanol suspension and was suspended in 12 uL of RNase free water, and RNA sample integrity was assessed by running 0.3-0.5 uL of each sample in a 0.8% agarose gel prepared with Tris-acetate-EDTA (TAE) buffer and using λ Hind III DNA molecular weight markers as size markers. The remaining aqueous RNA suspension was stored at –20°C until use the next day (as described in Walker *et al.*, 2008).

cDNA Preparation and Hybridization

The Genisphere 3DNA 350 kit protocol was employed without modification. The option of using vial 7 (2X formamide-based hybridization buffer) was always chosen. The complete protocol can be found on the Genomic Consortium for Active Teaching (GCAT) website (GCAT, 2003).

Yeast DNA Microarray chips: Source and Scanning

Microarray chips were purchased from GCAT. The chips used were produced by Washington University at Saint Louis, MI. The identification numbers of the chips from Washington University are: A41, 535, 536, 538, 539, 984 and 985 for creosote; A45, 537, and 540 for methylene chloride; A40, A42, 726, 729, and 737, 738, and 983 for PtCP; and A44, 486, and 740 for ethanol. All the chips contain two duplicate metagrids where each metagrid contains 16 grids and each individual grid contains 22 columns and 21 rows for a total of 462 potential gene locations per grid. Therefore a metagrid from WUSL contains 7,392 potential gene locations (locations for all of the controls are included) (Figure 2A in Appendix 1).

All microarray chips are 70mer oligos printed on epoxy slides. Gene lists for individual chips are provided by the manufactures, can be retrieved from the GCAT website, and must be linked with the data obtained by scanning the microarray chips (GCAT, Gene Information, 2008). After completing the hybridization procedure, scans of the DNA chips were performed at the Biology Department of Davidson College at Davidson, NC using the scanner model ArrayWoRx manufactured by Applied Instruments. Chips were sent by overnight shipment, covered with foil, and were scanned within 48 hours after completion of the exposure experiment. The scanned files are retrieved from the Institute for System Biology (ISB) server by using a file transfer program (ftp). Each chip used produces two images, one for the Cy3 fluorescent green dye scanned at 595 nm and the second one for the Cy5 fluorescent red dye scanned at 685 nm (as described in Walker *et al.*, 2008).

Dye-Swap DNA Microarray Chips

Dye-swap DNA microarray chips were paired with experimental microarray chips to assess any dye bias that may have occurred during implementation of the hybridization protocol. On the non-dye-swap microarray chips, the Cy3 green fluorescent dye was always used with the wild type RNA, and the Cy5 red fluorescent dye was always used for the stressed sample (sample exposed to PtCP, ethanol, creosote, or methylene chloride) (Figure 2B in Appendix 1).

On the dye-swap microarray chips the Cy3 green fluorescent dye was always used with the stressed sample (sample exposed to PtCP, ethanol, creosote, or methylene chloride), and the Cy5 red fluorescent dye was used for the wild type RNA (Figure 2B). The non-dye-swap chips include A41, 535, 536, 984, and 985 of creosote; A45 and 537 of methylene chloride; A40, A42, 726, 738, and 983 of PtCP; and A44 and 740 of ethanol. The dye-swap chips include: 538 and 539 of creosote; 540 of methylene chloride; 729 and 737 of PtCP; and 486 of ethanol. The non-dye-swap and dye-swap chip numbers for each treatment are summarized in the Table 1 of Appendix 2.

DATA ANALYSIS

MAGIC Tool

The computer program MAGIC Tool was used to obtain the data from the scanned chips (Heyer *et al.*, 2005 and GCAT, MAGIC Tool v2.1, 2008). MAGIC Tool retrieves the scanned images and the gene list that contains the order in which the genes are printed on a chip. The data is converted into a color image that ranges from green at one end of the spectrum, passing through yellow, to red at the other end of the spectrum. Because the Cy3 green fluorescent dye was always used with the wild type RNA, and the

Cy5 red fluorescent dye was always used for the stressed sample (sample exposed to creosote, methylene chloride, PtCP and ethanol) on non-dye-swap chips, for any given gene location on the grid, a green color is a visual indication that repression may have occurred, and a red image is a visual indication that induction may have occurred. A yellow color is an indication that genes in both samples, wild type and stressed, were affected in a similar manner (Figures 3A and 3B in Appendix 1).

"Addressing" was accomplished by assigning numbers to each of grids on the DNA microarray chip and the number of rows and columns in each grid. Additionally, the position of gene number one relative to gene number two must be specified. For all chips used, gene number one was located in the upper left corner and gene number two lies horizontally one space to the right. "Gridding" was accomplished following the steps outlined by the MAGIC Tool tutorial. On the "Segmentation" screen, the default segmentation method, fixed circle, with a radius of seven pixels was used to obtain an expression file. The expression file may be saved either as an "exp" expression file or as a "raw" expression file. The "exp" files are useful within the framework of the MAGIC Tool program. The "raw" expression files can be used for manual evaluation within the Microsoft Excel program and provide readily available information related to signal intensity that are not easily accessible with the "exp" files.

Statistical Analysis Using MAGIC Tool and Microsoft Excel

After gridding and addressing each chip, MA plots of each chip were made (from the "Segmentation" screen) to assess the Log_2 expression ratios vs. the total dye intensity values. Thus, for each of the 14,000 spots on the chip, Log_2 (R/G) was on the y-axis and (1/2)Log_2 (R*G) was on the x-axis (Figures 4A and 4B in Appendix 1). Because the

majority of the genes should have no change in expression, the points on the MA plot should be vertically centered at zero (since the $\log_2 1 = 0$). Therefore, based on the MA plots, normalization of each chip by standardization was necessary. On each chip the mean was subtracted from each expression ratio and then divided by the standard deviation of the chip. This set the mean of each chip to zero and the standard deviation to 1, allowing chips of the same experimental treatment to be compared using normalized values.

Each chip contains two spots of data or replicates per gene, one replicate in the first metagrid and the other in the second metagrid, accounting for nearly 14,000 spots per chip. The expression ratios (red/green) of the two replicates were averaged and transformed into Log_2. Then, the Log_2 expression ratios were standardized on each chip by subtracting the mean and dividing by the standard deviation. The dye-swap chip data were converted into non-dye-swap data using the "Dye-Swap" option under the Expression and Manipulate Data Tab in Magic Tool.

The creosote and methylene chloride chips were merged into one expression file (ten chips total). The PtCP and ethanol chips were also merged into one expression file (ten chips total). Using Microsoft Excel, the average and standard deviation of the expression of each gene across the 7 experimental chips (creosote and PtCP separately) was calculated.

A one sample t-test was performed, comparing the average gene expression of each gene across the 7 experimental chips to zero (μ_0: $\mu = 0$). Any genes that were significant at the $\alpha \leq 0.05$ level were considered to have significant gene expression, and the null was rejected. To make sure expression was not due to the solvent, a two sample

t-test was performed, comparing the mean gene expression of each gene in the experimental chips to the mean gene expression of each gene in the solvent chips (μ_0: $\mu_{experimental} = \mu_{solvent}$). Any genes that were significant at the $\alpha \leq 0.05$ level were considered to have significant gene expression, and the null was rejected. Genes that were significant at the $\alpha \leq 0.05$ level in both the one and two sample t-tests were considered to have significant gene expression and were deemed appropriate for gene expression validation by RT-PCR. The one and two sample t-tests were done for creosote and PtCP separately.

Gene Expression Validation by RT-PCR

From the list of genes with significant gene expression, determined by the statistical analysis of the one and two sample t-tests, 28 genes were chosen for gene expression validation by RT-PCR. Genes were chosen based on their biological and molecular roles in cell cycle regulation, drug transport, transcription regulation, and response to stress. All of these functions are commonly altered by chemical stressors. Oligos were ordered from Integrated DNA Technologies Inc. (IDT®). AffinityScript™ Multiple Temperature cDNA Synthesis Kit from STRATAGENE® was used for the cDNA synthesis of the total RNA. The Brillant® SYBR® Green QPCR Master Mix and the Mx3000P® protocol from STRATAGENE® were used for the RT-PCR reactions. The SYBR® Green was used as the fluorescent reagent, and ROX was used as a reference dye. The RT-PCR machine used was Mx3000P® from STRATAGENE®.

BLASTn and BLASTp of Significant Genes

All significant genes were analyzed for their homology to human nucleotide sequences and protein sequences using the National Center for Biotechnology

Information (NCBI) blast tools (NCBI, BLAST tools, 2008). Nucleotide blasts and protein blasts were performed for each gene. Accession numbers were acquired from the NCBI website by searching the gene database on the NCBI homepage using the *S. cerevisiae* Systematic Names.

RESULTS

The DNA microarray analysis of *S. cerevisiae* cells grown under the exposure of wood preservatives, creosote and PtCP, resulted in changes of gene expression (Figures 5A and 5B in Appendix 1). The dye-swap chips did not indicate any die bias, and by normalizing the gene expression across all the chips, any small bias that may have been present was minimized.

Overall, there was equal induction and repression in creosote with 3,167 genes induced and 3,167 repressed, totaling 6,334 genes. In the solvent, methylene chloride, there was more induction than repression with 3,226 genes induced and 3,108 genes repressed, totaling 6,334 genes. In PtCP there was less induction than repression with 2,988 genes induced and 3,113 genes repressed, totaling 6,101 genes. In ethanol there was less induction than repression with 2,975 genes induced and 3,126 genes repressed, totaling 6,101 genes. The total number of induced genes and repressed genes for each treatment is summarized in Table 2 in Appendix 2.

Statistical Analysis

In the one sample t-test of creosote, 445 genes had significant expression at the $\alpha \leq 0.05$ level, which is approximately the top 5% of the 6,334 genes present on the chip. However, when the gene expression of creosote was compared to the solvent, methylene chloride with a two sample t-test, there were 27 genes with significant gene expression at

the $\alpha \leq 0.05$ level. In the one sample t-test for PtCP, 3,047 genes had significant expression at the $\alpha \leq 0.05$ level. When compared to the solvent, ethanol, in the two sample t-test, there were 180 genes with significant gene expression at the $\alpha \leq 0.05$ level. Both solvents, methylene chloride and ethanol, had an effect on the gene expression of *S. cerevisiae*, which limited the number of genes with significant changes in gene expression that could be attributed to creosote and PtCP.

There was only one gene in common between the two treatments, PAU12 (YGR294W), which was induced in both treatments. PAU12 has not been characterized. In creosote, 12 of the 27 genes were induced (shown in red), and 15 were repressed (shown in green) (Table 3A in Appendix 2). In PtCP, 105 of the 180 genes were induced, and 75 genes were repressed (Table 3B in Appendix 2).

A file containing gene ontology information (biological function, molecular function, and cellular component) of all genes present on the DNA microarray chip was added to the MAGIC Tool expression files. Using the gene ontology information, many genes with biological functions dealing with mitotic regulation were found to have significant gene expression. In creosote, transcription and chromosome segregation were repressed; there are a limited number of genes with significant expression in creosote (Table 4A in Appendix 2). In PtCP there were 18 genes found to be involved in various biological processes in the nucleus: chromatin silencing, chromatin remodeling, regulation of transcription, response to stress, and ubiquitin-dependent processes (Table 4B in Appendix 2). Chromatin silencing was both induced and repressed, while chromatin remodeling was repressed. Genes with other functions such as regulation of transcription, initiation of transcription, and response to stress were induced. Protein

biosynthesis in the cytosol of the PtCP treated *S. cerevisiae* was repressed, but amino acid production in the cytosol was induced.

RT-PCR Validation

Of the significant genes, fifteen were chosen for further validation. Five genes were chosen for creosote, and ten genes were chosen for PtCP. TDH1 (glyceraldehyde 3-phosphate dehydrogenase -- GAPDH) was used as a control for normalization. A standard curve, using TDH1 primers, was used to quantify the amplified cDNA product (Figure 6 in Appendix 1). Relative quantities to the standard were used to calculate the gene expression fold ratios. To compare RT-PCR data to the MAGIC Tool data, fold ratios of the RT-PCR were transformed into Log_2.

PAU12 was chosen because it was the only gene significantly expressed in both creosote and PtCP. PAU12 was induced in both treatments and has unknown biological function, molecular function, and cellular component. PAU genes are in the family of seripauperins, which is the largest multigene family in yeast (Rachidi *et al.*, 2000). For creosote, BNA3 (YJL060W), Kynurenine aminotransferase involved in biosynthesis of nicotinic acid (BNA) was induced, and it is a potential cell division cycle protein (Cdc28p) substrate (YGD 2008 and NCBI 2008). PAT1 (YCR077C), Topoisomerase II-associated deadenylation-dependent mRNA-decapping factor, was repressed and is required for faithful chromosome transmission, maintenance of rDNA locus stability, and protection of mRNA 3'-UTRs from trimming (YGD, 2008 and NCBI, 2008). HAP2 (YGL237C), subunit of the heme-activated, glucose-repressed Hap2p/3p/4p/5p CCAAT-binding complex, was repressed and is a transcriptional activator and global regulator of respiratory gene expression (YGD, 2008 and NCBI, 2008). MYO3 (YKL129C), one of

two type I myosins, was induced and localizes to actin cortical patches (YGD, 2008 and NCBI, 2008).

For PtCP, BIT61 (YJL058C), subunit of TORC2, was induced and is a membrane-associated complex that regulates cell cycle-dependent actin cytoskeletal dynamics during polarized growth and cell wall integrity (YGD, 2008 and NCBI, 2008). WSC3 (YOL105C), sensor-transducer of the stress-activated PKC1-MPK1 signaling pathway involved in maintenance of cell wall integrity, was induced and is involved in the response to heat shock and other stressors (YGD, 2008 and NCBI, 2008). MDG1 (YNL173C) is a plasma membrane protein involved in G-protein mediated pheromone signaling pathway and was repressed (YGD, 2008 and NCBI, 2008). ITC1 (YGL133W) is a component of the ATP-dependent Isw2p-Itc1p chromatin remodeling complex required for repression of early meiotic genes during mitotic growth, and repression of INO1(1, 2, 3) (YGD, 2008 and NCBI, 2008). ITC1 was induced in the experiment. RSC30 (YHR056C), component of the RSC chromatin remodeling complex, was induced and is a non-essential gene required for regulation of ribosomal protein genes and the cell wall/stress response (YGD, 2008 and NCBI, 2008). GAT1 (YFL021W), transcriptional activator of genes involved in nitrogen catabolite repression, was induced. RPG1 (YBR079C), subunit of the core complex of translation initiation factor 3(eIF3), was repressed and is part of a subcomplex that stimulates binding of mRNA and tRNA(i)Met to ribosomes (YGD, 2008 and NCBI, 2008). RPN4 (YDL020C), transcription factor that stimulates expression of proteasome genes, was highly induced and is regulated by various stress responses. DBF2 (YGR092W), Serine/Threonine kinase involved in transcription and stress response, was repressed and functions as part of a network of

genes in exit from mitosis (GCAT, Gene Information, 2008). Table 5 in Appendix 2 summarizes the results of the RT-PCR validation.

Although most of the genes followed the same pattern of either induction or repression when comparing the DNA microarray results to the RT-PCR results, only four genes had significant changes in expression in the RT-PCR: BNA3 ($\alpha<0.10$) and HAP2 ($\alpha<0.05$) in creosote and RSC30 ($\alpha<0.10$), and GAT1($\alpha<0.01$) in PtCP. The RT-PCR results of DBF2, MDG1, and RPG1 did not follow the same pattern of expression in the DNA microarray results. DBF2 was not expressed in the wild type RNA or the PtCP RNA.

MAGIC Tool Clustering and BLASTn and BLASTp Analysis

In MAGIC Tool, dendrograms of the 7 treatment chips of creosote and PtCP were made, and hierarchical clustering was performed. Dissimilarities of Log_2 transformed data were calculated for each gene using MAGIC Tool, using the default setting, *distance = 1 - correlation*. A correlation of 1 between two genes corresponds to equal gene expression (distance of 0), and a correlation of -1 between genes indicates opposite gene expression (distance of 2). The dissimilarity calculations compare the gene expression of each gene on the chip to every other gene. In clustering, the genes with the smallest distance, corresponding to a correlation of ~1, are grouped together first and then the genes with the next smallest distance to the first two are grouped, etc. Dendrograms display the genes in a tree-like arrangement connecting genes with the most similar gene expression across the seven chips being connected by nodes. Genes are positioned to the right of the branches, which are anchored by the nodes and arranged in a tree (Figure 7 in Appendix 1).

There were many genes in both creosote and PtCP that had highly correlated gene expression and similar biological and molecular functions as other genes. Genes with significant expression were often highly correlated as well. Some genes with significant expression also have homologous DNA and protein sequences in *Homo sapiens*; however, no orthologs were found between the subset of significant genes in *S. cerevisiae* with *H. sapiens*. Blast results for all genes with significant changes in gene expression in creosote and PtCP are listed in Table 6A and 6B in Appendix 2.

In PtCP, WAR1 (YML076C) is a transcription factor that binds to a weak acid response (WAR) element (GCAT, Gene Information, 2008). WAR1 is induced and highly correlated with 36 genes: NOG1 (YPL093W), PKH2 (YOL100W), MCM2 (YBL023C), RRN3 (YKL125W) (transcription factor independent of DNA template; involved in recruitment of RNA polymerase I to rDNA), DPB2 (YPR175W) (required for normal yeast chromosomal replication, expression peaks at the G1/S phase boundary, and potential Cdc28p substrate), SLD5 (YDR489W) (assembly of DNA replication machinery), SQS1 (YNL224C), SKT5 (YBL061C), RPC40 (YPR110C), CUE3 (YGL110C), TRP3 (YKL211C), SEC12 (YNR026C), YGL081W (no other alias), YOR019W (no other alias), YOL159C (no other alias), YJR079W (no other alias), TBS1 (YBR150C), INP53 (YOR109W), ITC1 (chromatin remodeling), TRK1 (YJL129C), PPT1 (YGR123C) (Protein serine/threonine phosphatase with similarity to human phosphatase PP5), YHR131C (no other alias), RSM23 (YGL129C) (similarity to mammalian apoptosis mediator proteins and null mutation prevents induction of apoptosis), LAG2 (YOL025W) (protein involved in determination of longevity and overexpression extends the mean and maximum life span of cells), YJL185C (no other

alias), FPK1 (YNR047W) (protein kinase proposed to regulate the putative phospholipid translocases Lem3p-Dnf1p/Dnf2p; overexpression interferes with pheromone-induced growth arrest), GDE1 (YPL110C), DUG1 (YFR044C), NCA2 (YPR155C), OST4 (YDL232W), YOR389W (no other alias), SYT1 (YPR095C), FMP45 (YDL222C), MTQ2 (YDR140W), and LDB17 (YDL146W) (YGD, 2008 and NCBI, 2008) (Figure 8 in Appendix 1) .

In the highlighted branches of the dendrogram, four of the genes were significantly induced in PtCP according to the statistical analysis performed: WAR1, RRN3, ITC1, and YJL185C (no other alias). In BLASTp analysis, WAR1 is homologous to the *H. sapiens* RAS found in rat sarcoma and EF-hand domain (RASEF). RASEF proteins belong to the small Guanosine TriPhosphatase (GTPase) superfamily and bind Guanosine DiPhosphate (GDP) and Guanosine TriPhosphate (GTP). They regulate the growth of cells, and if knocked out in yeast, cells cannot multiply at all (NCBI, BLAST tools, 2008).

YBR259W (no other alias) is significantly induced in PtCP, has unknown biological function, and is highly correlated with seven other genes: POP2 (YNR052C) (regulation of transcription), SIP1 (YDR422C) (protein kinase activity), PAM18 (YLR008C), BPL1 (YDL141W) (protein modification), SEC62 (YPL094C) (protein binding activity), MST28 (YAR033W), IRC3 (YDR332W) (helicase activity) (GCAT, Gene Information, 2008). YBR259W is homologous to Homo sapiens tumor protein p53 binding protein 1 (TP53BP1) (NCBI, BLAST tools, 2008).

ABP1 (YCR088W) is involved in the establishment of cell polarity and is significantly repressed in PtCP. ABP1 is highly correlated to 16 genes: COX23

(YHR116W) (aerobic respiration), BCY1 (YIL033C) (cAMP-dependent protein kinase), JSN1 (YJR091C) (mRNA binding activity), SNC1 (YAL030W) (endocytosis), RSM18 (YER050C) (protein biosynthesis), YAR068W (no other alias), KRE9 (YJL174W) (cell wall organization and null mutation leads to severe growth defects), YDL199C (no other alias), SFP1 (YLR403W) (transcription factor activity and regulation of cell size), ADI1 (YMR009W), LSB1 (YGR136W), MDY2 (YOL111C), RCN2 (YOR220W), ARC40 (YBR234C) (actin filament organization), RFA3 (YJL173C) (DNA recombination and DNA replication), and MDG1 (YNL173C) (signal transduction during conjugation with cellular fusion) (GCAT, Gene Information, 2008). ABP1, MDY2, and MDG1 are all significantly repressed in PtCP. However, the RT-PCR validation of MDG1 expression was not statistically significant. ABP1 is homologous to *H. sapiens* serpin peptidase inhibitor (SERPING1) (NCBI, BLAST tools, 2008).

SHM2 (YLR058C) is significantly repressed in PtCP and is involved in one-carbon compound metabolism. It is homologous to *H. sapiens* endothelial cell adhesion molecule (ESAM) (NCBI, BLAST tools, 2008). It is highly correlated with six other genes with important biological and molecular functions: ADE5,7 (YGL234W) (purine base metabolism), RRP15 (YPR143W), HMS2 (YJR147W) (transcription factor activity and pseudohyphal growth, ADE4 (YMR300C) (purine base metabolism), PHO3 (YBR092C) (acid phosphatase activity), and AGA1 (YNR044W) (cell adhesion receptor activity) (GCAT, Gene Information, 2008). SHM2 and AGA1 are both significantly repressed in PtCP.

OYE3 (YPL171C) is involved in Nicotinamide Dinucleotide Phosphate (NADPH) dehydrogenase activity, and it is significantly induced in PtCP. It has highly

correlated expression with eight other genes: ERG25 (YGR060W) (C-4 methyl sterol oxidase activity), GAC1 (YOR178C) (protein phosphatase type 1 activity), HTS1 (YPR033C) (histidine-tRNA ligase activity), GRE2 (YOL151W) (oxidoreductase activity and response to stress), YKR075C (no other alias), CYT2 (YKL087C) (holodytochrome c synthase activity), HYM1 (YKL189W) (transcriptional repressor activity), and NUR1 (YDL089W) (GCAT, Gene Information, 2008). OYE3, GRE2, ERG25 are all significantly induced in PtCP.

BUD6 (YLR319C) was identified as a bipolar budding mutant and a potential Cdc28p substrate. BUD6 is repressed in creosote, and it is involved in cytoskeletal regulatory protein binding activity (GCAT, Gene Information, 2008). BUD6 is homologous to *H. sapiens* poly(A) binding protein, cytoplasmic 1 (PABPC1) and has a conserved protein domain, AIP3 superfamily: Actin interacting protein (NCBI, BLAST tools, 2008). It is highly correlated with YOR343C-A, NTH2 (YBR001C) (response to stress), YBR285W (no other alias), VPS20 (YMR077C) (late endosome to vacuole transport), PEX15 (YOL044W) (peroxisome organization and biogenesis), YCT1 (YLL055W), YJR142W (no other alias), LDB19 (YOR322C), PSR1 (YLL010C) (protein phosphatase activity and response to stress), COQ1 (YBR003W) (trans-hexaprenyltranstransferase activity), YNL144C (no other alias), RSF2 (YJR127C) (transcription factor activity), AFR1 (YDR085C) (receptor signaling protein activity), YGR012W, YFL023W, YKR070W, YJR049C (NAD+ kinase activity and iron ion homeostasis), YBR075W (no other alias), and SSQ1 (YLR369W) (chaperone activity and DNA dependent DNA replication) (GCAT, Gene Information, 2008). Although

BUD6 was the only gene with significant expression, the genes have highly correlated expression across the seven chips and similar biological and molecular functions.

HAP2, whose gene expression was validated in the RT-PCR, is a gene responsible for transcriptional activator activity in the nucleus, significantly repressed in creosote, and highly correlated with other genes that are also significantly expressed with similar biological and molecular functions (GCAT, Gene Information, 2008). It is highly correlated with six other genes: PAU3 (YCR104W), YPT31 (YER031C), YKL077W (no other alias), RTF1 (YGL244W), RUB1 (YDR139C), and ADE5,7. PAU3 has unknown biological function, significant gene expression, and is highly correlated with HAP2, which is an interesting result considering that there are only 27 genes in creosote with statistically significant gene expression. HAP2 is homologous to *H. sapiens* hemoglobin, gamma G (HBG2) (NCBI, BLAST tools, 2008).

There are two genes in creosote that have interesting *H. sapiens* homologs, but are not highly correlated with other genes having significant changes gene expression: MFB1 (YDR219C) and BNA3, whose gene expression was also validated in the RT-PCR. MFB1 is involved in protein binding activity and is repressed in creosote (GCAT, Gene Information, 2008). It is homologous to *H. sapiens* fibroblast growth factor receptor 2 (FGFR2), which is involved in craniofacial dysostosis 1, Crouzon syndrome, Pfeiffer syndrome, and Jackson-Weiss syndrome (NCBI, BLAST tools, 2008). BNA3, which is responsible for nicotinamide adenine dinucleotide biosynthesis and arylformamidase activity, is significantly induced in creosote (GCAT, Gene Information, 2008). It is homologous to *H. sapiens* myeloid/lymphoid or mixed-lineage leukemia 3 (MLL3) (NCBI, BLAST tools, 2008).

DISCUSSION AND CONCLUSIONS

Yeast is used as a model organism to study and understand how human cell processes are regulated. When applying the yeast model to human cancer, it is important to look at genes that regulate mitotic controls (Wassmann and Benezra, 2001). In both experimental treatments, creosote and PtCP, genes with roles in cell cycle regulation, drug transport, transcription regulation, and response to stress had significant changes in gene expression. RT-PCR analysis verified that the changes in gene expression could be validated. Clustering analysis in MAGIC Tool revealed highly correlated gene expression in genes associated with mitotic controls. BLASTn and BLASTp analysis confirmed that some genes with significant changes in gene expression had homology to human nucleotide and protein sequences.

The highly correlated expression and function of genes in response to creosote and PtCP treatment show parallels with gene expression that can be associated with cancerous growth. There are certain properties that allow cells to escape programmed cell death and promote cancerous growth: disregard for signals that regulate cell proliferation, avoidance of apoptosis and senescence, and genetic instability (Alberts *et al.,* 2004*)*. WAR1, the significantly induced transcription factor mentioned in the PtCP treatment, is one example of a gene with expression that can be associated with tumorigenesis. It was highly correlated with genes that have roles in cell cycle regulation: RRN3 (significantly induced transcription factor independent of DNA template and involved in recruitment of RNA polymerase I to rDNA), DPB2 (required for normal yeast chromosomal replication, expression peaks at the G1/S phase boundary, and potential Cdc28p substrate), SLD5 (assembly of DNA replication machinery), ITC1

(significantly induced gene associated with chromatin remodeling), RSM23 (similarity to mammalian apoptosis mediator proteins and null mutation prevents induction of apoptosis), and LAG2 (protein involved in determination of longevity and overexpression extends the mean and maximum life span of cells) (YGD, 2008 and NCBI, 2008). Genes that are highly correlated in their expression and also have biological functions responsible for cell cycle regulation, such as the genes correlated with WAR1, will provide a possible target for further investigation.

Genes associated with important cell cycle checkpoints also had significant changes in gene expression and are potential substrates of Cdc28, the equivalent to Cdc2 in fission yeast, *Schizosaccharomyces pombe*. Activation of Cdc28-substrate complexes is one of the final signals that activate replication fork movement in the M and G_1 phase of mitosis (Murakami and Nurse, 2000). Overexpression of Cdc28 leads to cell cycle arrest, while low expression levels of Cdc28 allows the cell cycle to proceed normally (Murakami and Nurse, 2000). BNA3, BUD6, XBP1 (YIL101C), SIT1 (YEL065W), and KEL2 (YGR238C) are all genes with significant changes in gene expression and potential substrates of Cdc28, which may also lead to further research.

WAR1, BNA3, and YBR259 had homologous sequences with genes in humans that are important in regulated the cell cycle. All three genes were induced in the experiment. WAR1 was homologous with RASEF, which regulates the growth of cells and controls cell division in yeast (NCBI, BLAST tools, 2008). BNA3 is homologous to *H. sapiens* myeloid/lymphoid or mixed-lineage leukemia 3 (MLL3) (NCBI, BLAST tools, 2008). MLL3 regulates the gene expression of certain genes, for example HOX genes (homeobox genes that regulate morphogenesis in organisms), and is a target site for

translocations in leukemia (NCBI, 2008). YBR259W has unknown biological function; however, YBR259W is homologous to *Homo sapiens* tumor protein p53 binding protein 1 (TP53BP1), a well documented gene involved in tumorigenesis (NCBI, BLAST tools, 2008). The homology of these three genes to important genes in cell cycle control and their significant induction in this experiment may lead to investigation of their role in tumorigenesis and possibly the characterization of YBR259W.

Creosote and PtCP are two compounds that are possibly carcinogenic and therefore are potentially dangerous to humans and wildlife. The EPA has assessed the risk and regulated the use of creosote and PtCP when used in the presence of workers and the environment. Their presence causes changes in the gene expression of certain genes and can effect DNA replication, cell growth, and cell death. The gene expression trends in this experiment do not exclusively indicate carcinogenic effects. The changes in gene expression may be due to the toxicity of the chemicals. The toxicity of the chemicals is most evident in the highly induced expression of genes dealing with response to stress and response to drug including: XBP1 (response to stress), GRE2 (YOL151W) (response to stress), PDR5 (YOR153W) (response to drug), SNG1 (YGR197C) (response to drug), and SNQ2 (YDR011W) (response to drug) (GCAT, Gene Information, 2008). The concentration and exposure time to creosote, PtCP and their solvents may need to be adjusted to decrease the toxicity of the chemicals. Perhaps with further study, the EPA will be able to change their statement in the Toxicology Profile for Pentachlorophenol from "a probable human carcinogen" and "possibly carcinogenic" (ATSDR, 2001) to "is a human carcinogen" and "is carcinogenic."

This study involved a combination of biotechnology, genetics, and mathematics. Without the cooperation and collaboration of different university departments, comprehensive research using DNA microarrays cannot be accomplished. This DNA microarray study was educational as well as scientifically worthwhile. Overall, the results of this DNA microarray study of *S. cerevisiae* cells exposed to wood preservatives are a sign of the necessity for more studies to be done by the EPA and workers' health associations in order to establish job/health regulations. The results of this study could be a starting point for R-1 institutions that concentrate in cancer studies.

ACKNOWLEDGEMENTS

I would like to thank Dr. Consuelo Alvarez, without whom I would not have been exposed to this research experience. Her countless hours spent teaching me and helping me with this project are a testament to her dedication to her students. I would also like to the Dr. Leigh Lunsford for her time and effort in working with Dr. Alvarez and I on several DNA microarray projects. I would like to thank the Genome Consortium for Active Teaching (GCAT) for supplying and scanning the DNA microarray chips. I would like thank the Honors Program and Dr. Geoffrey Orth for giving me the opportunity to pursue undergraduate research. I would also like to acknowledge the Cormier Scholarship Funds as well as the Dean's Funds for financial support to carry out the experiments.

REFERENCES

Agency for Toxic Substances and Disease Registry (ATSDR). 2001. Toxicology Profile for Pentachlorophenol. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service.

Agency for Toxic Substances and Disease Registry (ATSDR). 2002. Potential for human exposure. In: Toxicological Profile for Wood creosote, Coat Tar creosote, Coal Tar, Coal Tar Pitch, and Coal Tar Pitch Volatiles. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service. Available: http://www.atsdr.cdc.gov/toxprofiles/tp85-c1.pdf [accessed 11 Feb 2007].

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, and Walter P. 2002. 4th ed. Molecular Biology of the Cell. Garland Science, New York. pp. 1463.

Carlsten C, Hunt SC, and Kaufman JD. 2005. Squamous Cell Carcinoma of the Skin and Coal Tar creosote Exposure in a Railroad Worker. Environmental Health Perspectives 113(1): 96-7.

Environmental Protection Agency (EPA). 2008. Reregistration Eligibility Decision for Creosote (Case 0139). pp. 1-91.

Genome Consortium for Active Teaching (GCAT). 2008. Gene Lists and Gene Information. http://www.bio.davidson.edu/projects/MAGIC/MAGIC.html. [20 Aug 2008].

Genome Consortium for Active Teaching (GCAT). 2008. MicroArray Genome Imaging and Clustering Tool: MAGIC Tool v2.1. http://www.bio.davidson.edu/projects/MAGIC/MAGIC.html [20 Aug 2008].

Genome Consortium for Active Teaching (GCAT). 2003. Report of Labeling Micrarrays with Genisphere Array 350 Kit. http://www.bio.davidson.edu/projects/gcat/protocols/3DNAmethod.doc. [20 Oct 2008].

Heyer LJ, Moskowitz DZ, Abele JA, Karnik P, Choi D, Campbell AM, Oldham EE, and Akin BK. 2005. Gene Expression. MAGIC Tool: integrated microarray data analysis. Bioinformatics 21: 2114-2115.

Kitagawa E, Momose Y, Iwahashi H. 2003. Correlation of the structures of agricultural fungicides to gene expression in *Saccharomyces cerevisiae* upon exposure to toxic doses. Environmental Science & Technology 37: 2788-2793.

Labor Occupational Safety & Health Program (LOSH), UCLA. 2003. Creosote: What You Need to Know. Los Angeles: University of California. Available: www.losh.ucla.edu/catalog/factsheets/creosote_english.pdf [accessed 11 Feb 2007].

McClain JS, Oris JT, Burton GA, and Lattier D. 2003. Laboratory and Field Validation of Multiple Molecular Biomarkers of Contaminant Exposure in Rainbow Trout (Oncorhynchus Mykiss). Environmental Toxicology and Chemistry 22(2): 361-370.

Murakami H, and Nurse P. 2000. DNA replication and damage checkpoints and meiotic cell cycle controls in the fission and budding yeasts. Biochemical Journal. 349: 1-12.

National Center for Biotechnology Information (NCBI). 2008. BLASTn and BLASTp tools. http://blast.ncbi.nlm.nih.gov/Blast.cgi. [8 Mar 2009].

National Center for Biotechnology Information (NCBI). 2008. Gene Information Database. [1 Apr 2009].

Pollack JR, and Iyer VR. 2002. Characterizing the physical genome. Nature Genetics 32: 515-521.

Rachidi N, Marie-JoseÂe M, Barre P, and Blondin B. 2000. *Saccharomyces cerevisiae* PAU genes are induced by anaerobiosis. Molecular Microbiology 35(6), 1421-1430.

Roling JA, Bain LJ, and Baldwin WS. 2004. Differential gene expression in mummichogs (Fundulus heteroclitus) following treatment with pyrene: comparison to a creosote contaminated site. Marine Environmental Research 57: 377-395.

U. S. Department of Health and Human Services (U.S. DHH). 1999. Toxicology and carcinogenesis studies of pentachlorophenol in F344/N rats. NTP Technical Report No. 99-3973: 5-182.

Walker DE, Lutz GP, and Alvarez CJ. 2008. Development of a Cross-Disciplinary Investigative Model for the Introduction of Microarray Techniques at Non-R1 Undergraduate Institutions. CBE: Life Sciences Education 7(1): 118-31.

Wassmann K, and Benezra R. 2001. Mitotic checkpoints: from yeast to cancer. Current Opinion in Genetics & Development 11: 83–90.

Yeast Genome Database (YGD). 2008. Scientific database of the molecular biology and genetics of the yeast *Saccharomyces cerevisiae*: *S. cerevisiae Gene Information.* http://www.yeastgenome.org/. [1 Apr 2009].

Appendices Table of Contents

Appendix 1 32

Figure 1A and Figure 1B 32

Figure 2A and Figure 2B 33

Figure 3A and Figure 3B 34

Figure 4A and Figure 4B 35

Figure 5A and Figure 5B 36

Figure 6 37

Figure 7 38

Figure 8 39

Appendix 2 40

Table 1, Table 2, Table 3A 40

Table 3B 41

Table 4A and Table 4B 43

Table 5 44

Table 6A 45

Table 6B 46

APPENDIX 1

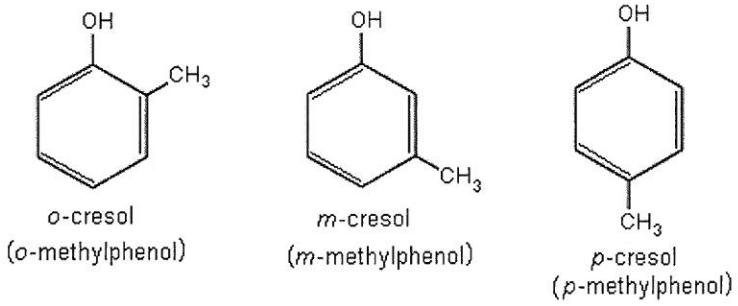


Figure 1A. Some Phenolic Constituents of Creosote borrowed from http://www.angelo.edu/faculty/kboudrea/molecule_gallery/06_phenols/00_phenols.htm.

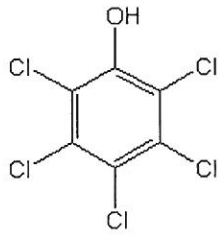


Figure 1B. Pentachlorophenol (PtCP) borrowed from http://toxipedia.org/wiki/display/toxipedia/Pentachlorophenol.

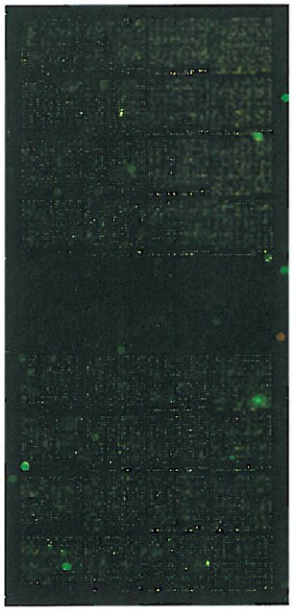


Figure 2A. The scanned PtCP DNA microarray chip shows the 1st metagrid (top half of chip) and 2nd metagrid (bottom half of chip). In each metagrid there are 16 grids total for a total of 32 grids per chip.

Figure 2B. Non-Dye-Swap chip 726 of PtCP (left) and Dye-Swap chip 729 of PtCP (right).

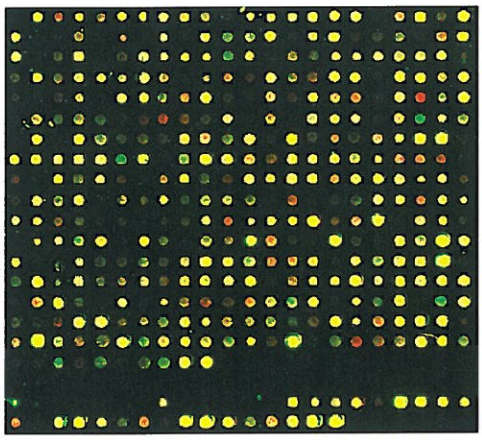


Figure 3A. Close-up of Scanned Chip A40 of PtCP. Shows induction (red), repression (green), and equal (yellow) expression in one grid of the chip. One grid contains 22 columns and 21 rows of spots. The black spots are left as controls.

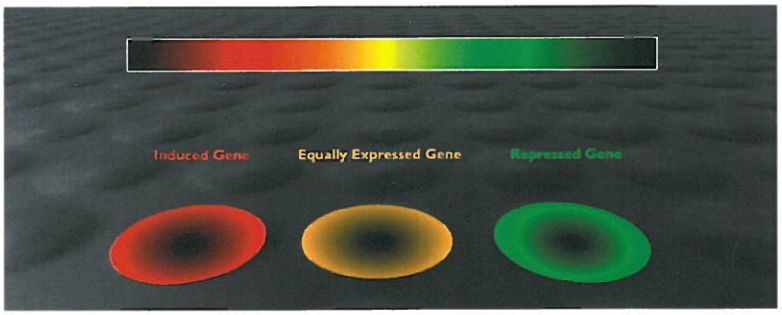


Figure 3B. Gene Expression Color Scale

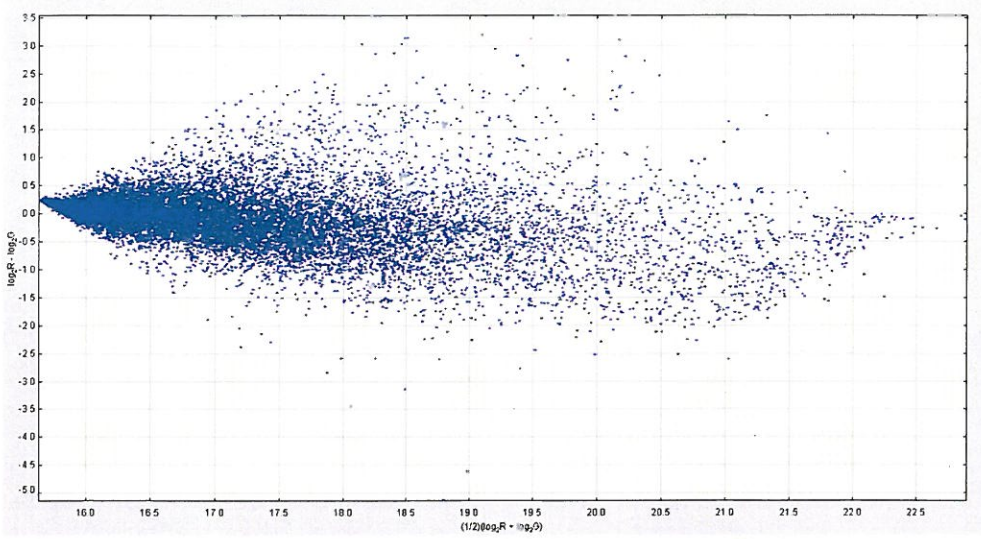


Figure 4A. MA Plot of PtCP Chip A40.

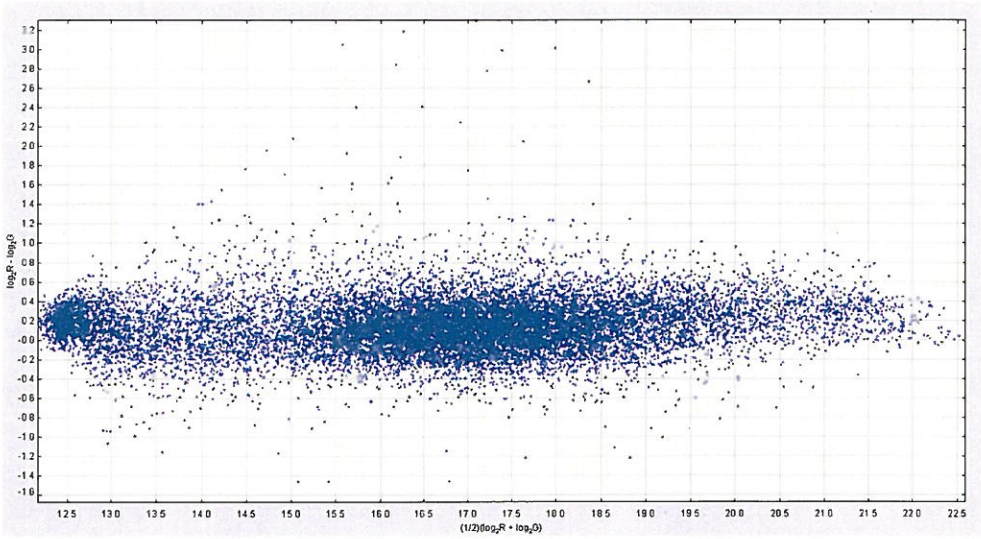


Figure 4B. MA Plot of Creosote Chip A41.

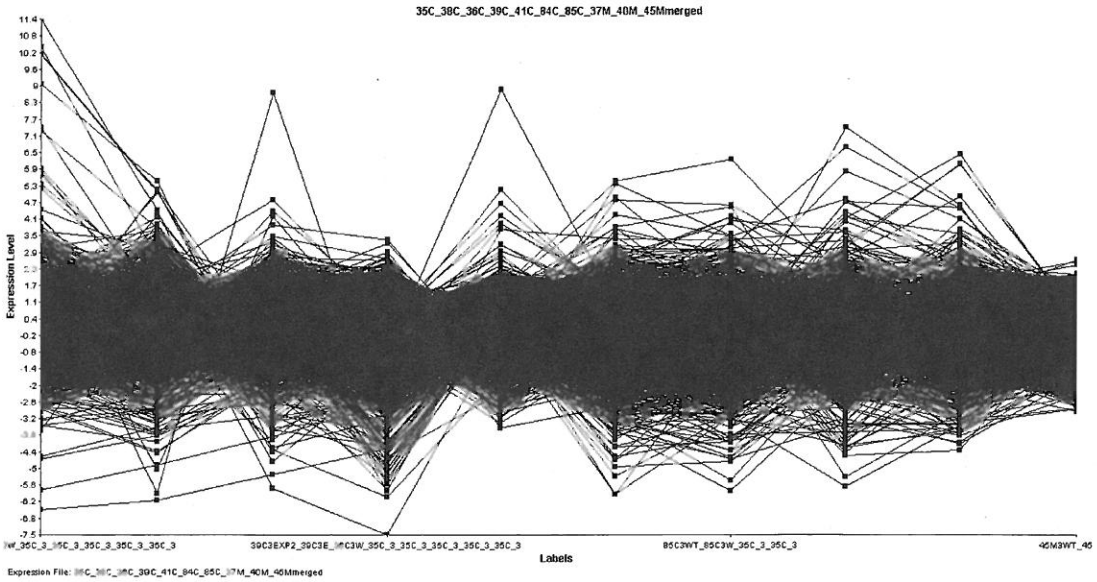


Figure 5A. Gene Expression Across the 7 chips of creosote and the 3 chips of methylene chloride (in order from left to right: 535, 538, 536, 539, A41, 984, 985, 537, 540, and A45).

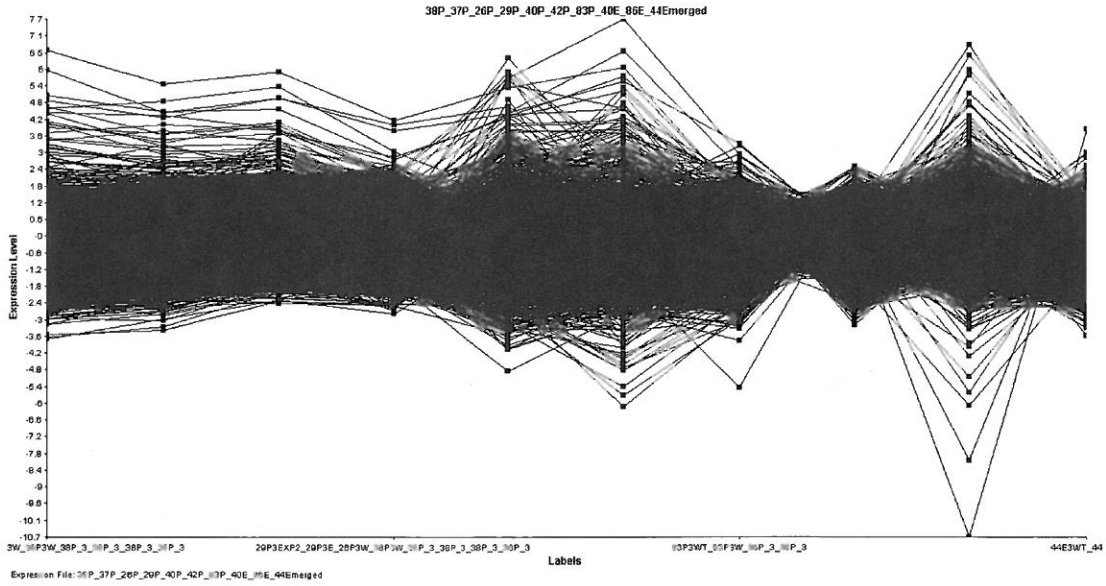


Figure 5B. Gene Expression Across the 7 chips of PtCP and the 3 chips of ethanol (in order from left to right: 738, 737, 726, 729, A40, A42, 983, 740, 486, and A44).

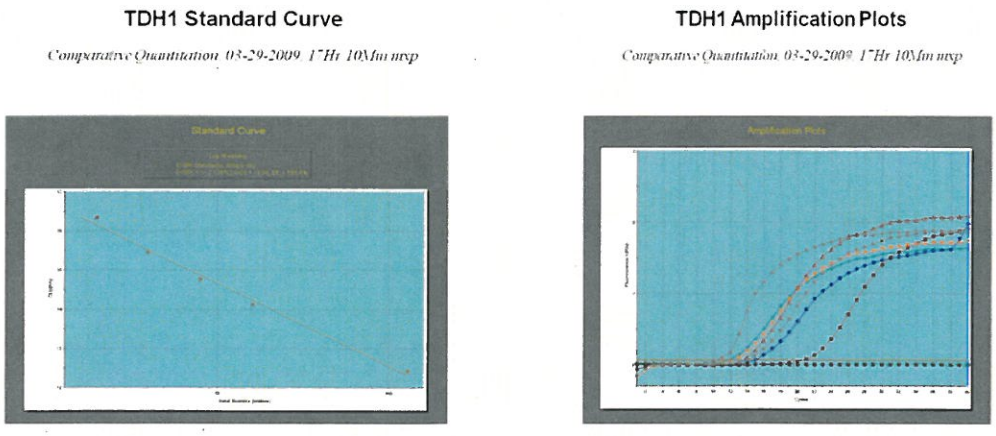


Figure 6. TDH1 RT-PCR Standard Curve (left) and TDH1 Standard Curve Amplification Plots (right)

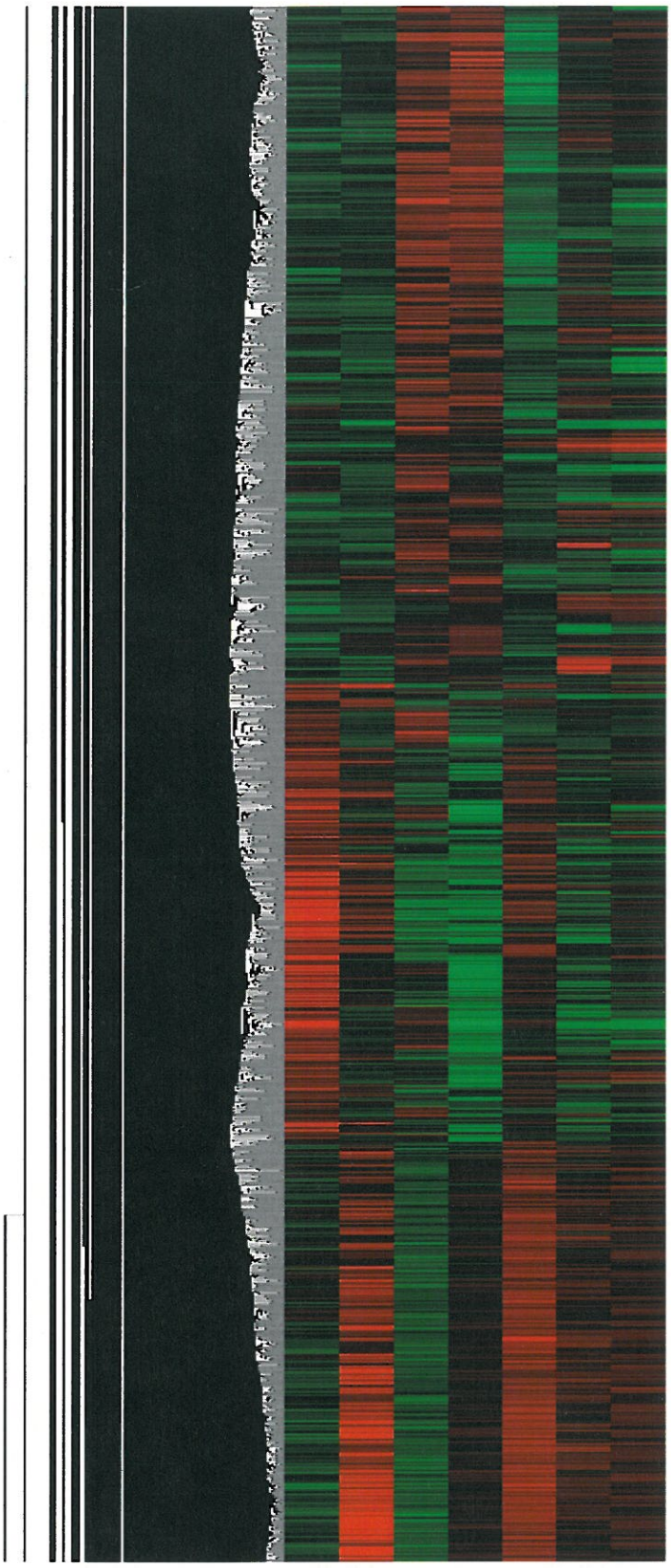


Figure 7. Dendrogram of creosote data in Java TreeView.

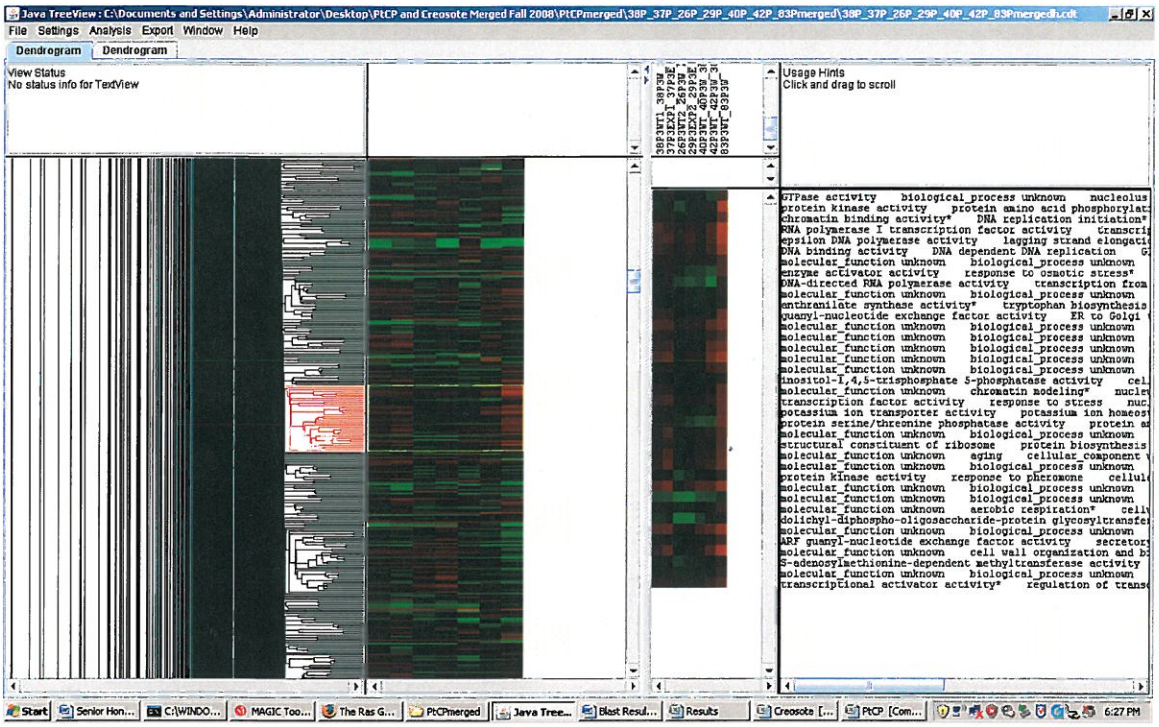


Figure 8. Dendrogram Analysis of Gene Expression Correlation. Printed screen of PtCP data shows the highlighted branches of the dendrogram most closely correlated to WAR1, a gene associated with transcription factor activity and response to stress in the nucleus.

APPENDIX 2

Table 1. Summary of non dye-swap and dye-swap chip numbers in the DNA microarray chips.

Treatment	Non-dye-swap	Dye-swap
Creosote	A41, 535, 536, 984, 985	538, 539
Methylene Chloride	A45, 537	540
PtCP	A40, A42, 726, 738, 983	729, 737
Ethanol	A44, 740	486

Table 2. Summary of total induction and repression for each treatment.

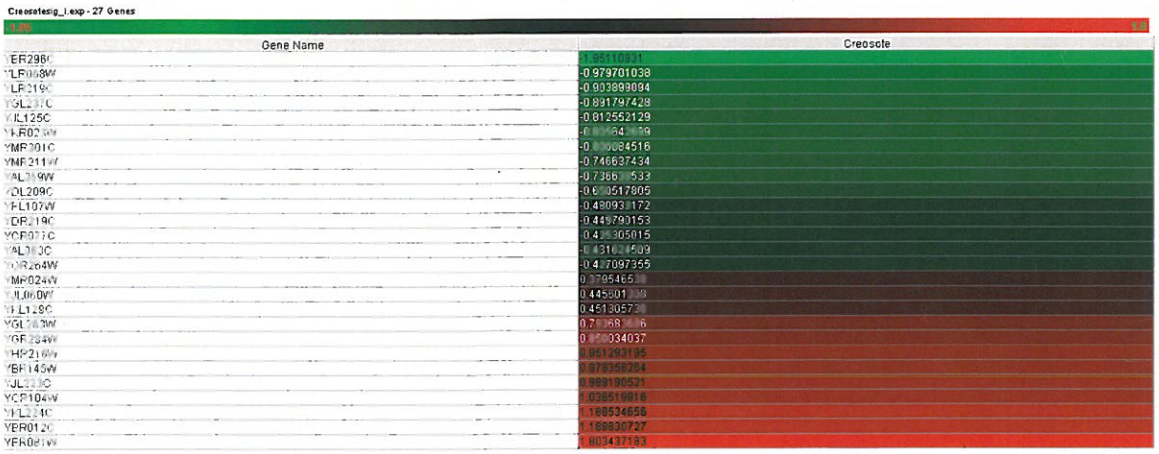
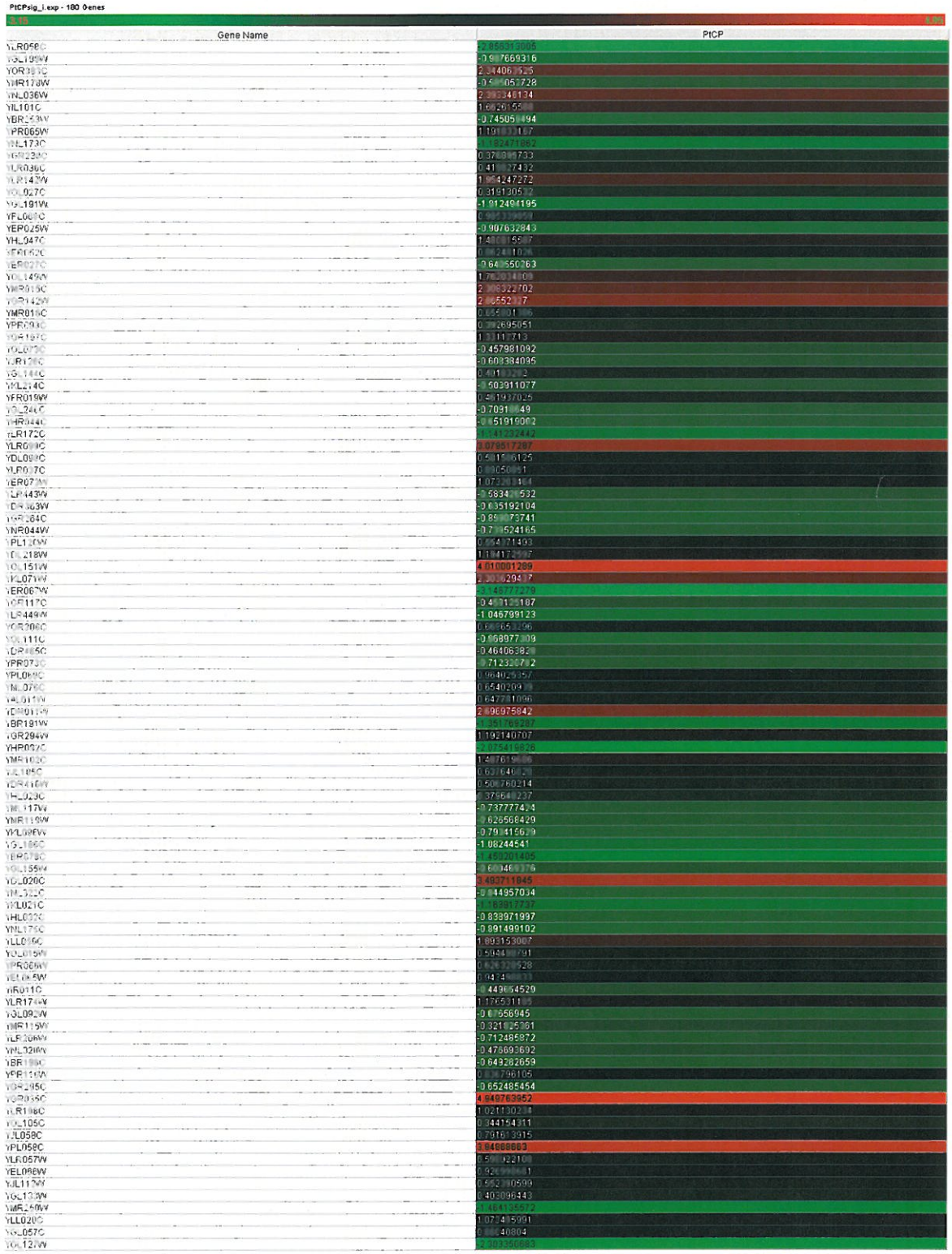


Table 3A. MAGIC Tool Table of creosote genes with significant changes in gene expression.

Treatment	Induction	Repression	Total # of Genes
Creosote	3,167	3,167	6,334
Methylene Chloride	3,226	3,108	6,334
PtCP	2,988	3,113	6,101
Ethanol	2,975	3,126	6,101

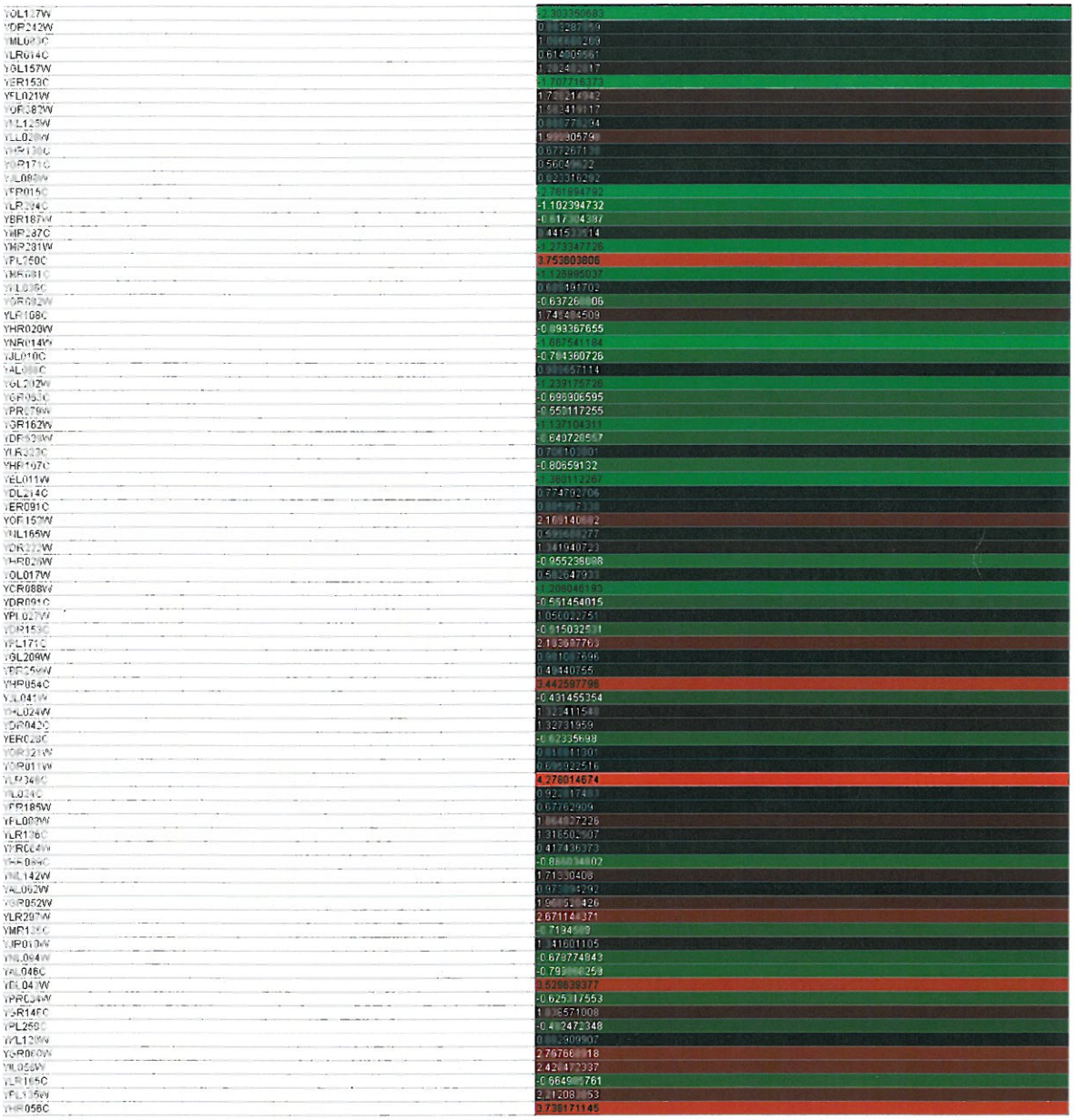
Table 3B. MAGIC Tool Table of PtCP genes with significant changes in gene expression. (Continued on next page.)



PtCPsig_l.exp - 180 Genes

Gene Name	PtCP
YLR058C	2.858313005
YGL195W	-0.9[illegible]7669316
YOR387C	2.34406[illegible]525
YMR173W	-0.5[illegible]057728
YNL036W	2.[illegible]346134
YIL101C	1.68261[illegible]
YBR[illegible]W	-0.74505[illegible]494
YPR065W	1.19[illegible]1167
YNL173C	1.[illegible]2471862
YGR[illegible]C	0.37[illegible]733
YLR036C	0.41[illegible]27432
YLR142W	1.[illegible]4247272
YOL027C	0.31913051[illegible]
YGL191W	-1.912494195
YFL0[illegible]C	0.[illegible]
YEP025W	-0.907632843
YHL047C	1.48[illegible]15[illegible]7
YER[illegible]2C	0.[illegible]1[illegible]
YER027C	0.64[illegible]650263
YOL149W	1.78[illegible]4409
YMR015C	2.[illegible]322702
YPR142W	2.[illegible]552[illegible]7
YMR015C	0.65[illegible]01[illegible]
YPR093C	0.[illegible]2695051
YDR187C	1.3[illegible]11713
YOL073C	-0.457981092
YJR126C	-0.[illegible]384095
YGL144C	0.41[illegible]2
YML214C	-0.[illegible]03911077
YER019W	0.4[illegible]1937025
YOL244C	-0.7091[illegible]49
YHR044C	-0.[illegible]51919002
YLR172C	1.141232442
YLR[illegible]C	3.079517297
YDL0[illegible]C	0.5[illegible]6125
YLR017C	0.[illegible]050[illegible]1
YER073W	1.07[illegible]164
YLR443W	-0.[illegible]532
YDR363W	-0.[illegible]5192104
YGR284C	-0.8[illegible]73741
YNR044W	-0.7[illegible]524165
YPL120W	0.[illegible]71493
YDL218W	1.1[illegible]
YOL151W	4.010001289
YKL071W	2.[illegible]29437
YER067W	3.148777279
YOR117C	-0.4[illegible]125187
YLR449W	-1.046799123
YOR206C	0.[illegible]3296
YDL111C	-0.[illegible]977[illegible]09
YDR485C	-0.46406382[illegible]
YPR073C	0.[illegible]712330[illegible]2
YPL0[illegible]C	1.[illegible]025357
YNL076C	0.[illegible]5402097[illegible]
YKL011W	0.[illegible]4[illegible]1096
YDR011W	2.[illegible]96975842
YBR181W	0.[illegible]1[illegible]6280
YGR294W	1.192140707
YHP032C	-2.[illegible]5419826
YMR102C	1.[illegible]7619[illegible]6
YJL185C	0.[illegible]6[illegible]21
YDR410W	0.50[illegible]760214
YHL023C	0.[illegible]37964[illegible]237
YML117W	-0.7377774[illegible]4
YMR119W	0.[illegible]626568429
YKL099W	-0.79[illegible]415629
YGL186C	-1.08244541
YBR073C	1.[illegible]4[illegible]1405
YGL155W	0.[illegible]60[illegible]6[illegible]76
YDL020C	3.[illegible]7[illegible]1945
YML122C	-0.[illegible]44957034
YKL021C	0.[illegible][illegible]7[illegible]
YHL032C	-0.839971997
YNL176C	-0.891499102
YLL056C	1.893153087
YOL015W	0.[illegible]5[illegible]791
YPR086W	0.[illegible]2[illegible]528
YEL065W	0.[illegible]4[illegible]
YIR011C	0.[illegible]449054520
YLR174W	1.[illegible]765[illegible]
YGL092W	-0.[illegible]656945
YMR119W	0.321[illegible]26361
YLR208W	-0.712485872
YHL028W	-0.[illegible]4[illegible]3692
YBR[illegible]	-0.649282659
YPR[illegible]W	0.[illegible]796105
YDR[illegible]C	-0.652485454
YDR035C	4.949763952
YLR[illegible]C	1.021130[illegible]
YOL105C	0.344154311
YJL058C	0.791613915
YPL058C	3.94[illegible]3
YLR057W	0.[illegible]22[illegible]
YEL066W	0.[illegible]
YJL117W	0.[illegible]0500
YGL133W	0.[illegible]96443
YMR[illegible]W	1.[illegible]35572
YLL020C	1.[illegible]5[illegible]1
YGL057C	0.[illegible]40804
YOL127W	2.[illegible]

Table 3B. MAGIC Tool Table of PtCP genes with significant changes in gene expression.



4A. Creosote gene expression in the nucleus.

Creosotesig_i.exp - 27 Genes - Group Of 2 Genes

Gene Name	Creosote	Comments	Alias	Chromosome	Location	Biological Process	Molecular Function	Cellular Component
YGL237C	-0.891797428		HAP2	7	53528	transcription*	transcriptional activator	nucleus*
YJL125C	-0.812552129		GCD14	10	186378	tRNA methylation	tRNA methyltransferas...	nucleus

Creosotesig_i.exp - 27 Genes - Group Of 1 Genes

Gene Name	Creosote	Comments	Alias	Chromosome	Location	Biological Process	Molecular Function	Cellular Component
YCR077C	-0.435305015		PAT1	3	252621	chromosome segregation*	molecular_function unknown	cytosolic small ribosomal subunit...

4B. PtCP gene expression in the nucleus and cytosol.

PtCPsig_i.exp - 180 Genes - Group Of 13 Genes

Gene Name	PtCP	Comments	Alias	Chromosome	Location	Biological Process	Molecular Function	Cellular Component
YDR153C	-0.515032531		ENT5	4	767964	Golgi to endosome tra...	clathrin binding activity	cytoplasm*
YER073W	1.073203464		ALD5	5	304027	electron transport	aldehyde dehydrogena...	mitochondrion
YGR142W	2.86552327		BTN2	7	772456	intracellular protein tra	molecular_function un	cytosol
YHR092C	2.075419826		HXT4	8	288813	hexose transport	glucose transporter act...	plasma membrane
YKL120W	0.833909907		OAC1	11	216986	sulfate transport*	oxaloacetate carrier act...	mitochondrial inner m...
YLL028W	1.099305798		TPO1	12	84803	spermine transport	spermine transporter a...	plasma membrane*
YOR011W	0.695927516		AUS1	15	349678	sterol transport	ATP-binding cassette (...	membrane
YOR306C	0.669653296		MCH5	15	891426	transport	transporter activity*	membrane
YOR382W	1.582419117		FIT2	15	1059525	siderochrome transport	molecular_function un	cell wall (sensu Fungi)
YOR383C	2.344063525		FIT3	15	1061049	siderochrome transport	molecular_function un	cell wall (sensu Fungi)
YPL058C	3.84868663		PDR12	16	450372	transport*	xenobiotic-transporting	plasma membrane
YPL259C	-0.482472348		APM1	16	52671	vesicle-mediated trans...	clathrin binding activity	AP-1 adaptor complex
YPR079W	-0.550117255		MRL1	16	698865	vacuolar transport	receptor activity	cytoplasm*

PtCPsig_i.exp - 180 Genes - Group Of 5 Genes

Gene Name	PtCP	Comments	Alias	Chromosome	Location	Biological Process	Molecular Function	Cellular Component
YBR191W	-1.351769287		RPL21A	2	606227	protein biosynthesis	structural constituent o...	cytosolic large riboso...
YGR142W	2.86552327		BTN2	7	772456	intracellular protein tra...	molecular_function un...	cytosol
YJL088W	0.823316292		ARG3	10	268495	arginine biosynthesis*	ornithine carbamoyltra...	cytosol
YLR174W	1.176531165		IDP2	12	504593	glutamate biosynthesis*	isocitrate dehydrogena...	cytosol
YOL127W	-2.303360083		RPL25	15	80347	protein biosynthesis*	structural constituent o...	cytosolic large riboso...

PtCPsig_i.exp - 180 Genes - Group Of 17 Genes

Gene Name	PtCP	Comments	Alias	Chromosome	Location	Biological Process	Molecular Function	Cellular Component
YAL062W	0.971804292		GDH3	1	31568	glutamate biosynthesis	glutamate dehydrogen...	nucleus*
YDR363W	-0.635192104		ESC2	4	1199171	chromatin silencing at ...	molecular_function un...	nucleus
YER027C	-0.640550263		GAL83	5	210231	protein amino acid pho...	SNF1/AMP-activated ...	nucleus
YFL021W	1.728214942		GAT1	6	95964	transcription initiation f...	specific RNA polymera...	nucleus
YGL133W	0.403096443		ITC1	7	257710	chromatin modeling*	molecular_function un...	nucleus
YGL209W	0.381087606		MIG2	7	95860	regulation of transcripti...	specific RNA polymera...	nucleus
YGL246C	-0.70918649		RAI1	7	38780	RNA catabolism*	enzyme regulator activity	nucleus
YGR063C	-0.696906595		SPT4	7	617825	"regulation of transcript...	Pol II transcription elon...	nucleus*
YIL101C	1.662615588		XBP1	9	177247	response to stress	transcription factor acti...	nucleus
YIR011C	-0.449654529		STS1	9	378243	ubiquitin-dependent pr...	molecular_function un...	nucleus
YKL214C	-0.503911077		YRA2	11	31694	poly(A)+ mRNA-nucleu...	RNA binding activity	nucleus
YLR014C	0.614005561		PPR1	12	174981	"regulation of transcript...	transcription factor acti...	nucleus
YLR449W	-1.046799123		FPR4	12	1030828	biological_process un...	peptidyl-prolyl cis-trans...	nucleus
YWL076C	0.654020919		WAR1	13	115347	response to stress	transcription factor acti...	nucleus
YOL017W	0.582647933		ESC8	15	292529	chromatin silencing	molecular_function un...	nucleus
YOL151W	4.010001289		GRE2	15	43692	response to stress	oxidoreductase activity*	nucleus*
YPR034W	-0.625317553		ARP7	16	639520	chromatin modeling	general RNA polymera...	nucleus*

Table 5. RT-PCR results of the 15 genes chosen for validation.

RT-PCR Results			
Treatment	**Gene Name**	**Log_2 MAGIC Tool Gene Expression Ratio**	**Log_2 RT-PCR Gene Expression Ratio**
Creosote	TDH1 (Control)	0.223	-1.089 ± 0.548 ***
	PAU12	0.85	0.851 ± 1.15
	BNA3	0.446	1.760 ± 0.763 *
	PAT1	-0.435	-0.645 ± 0.622
	HAP2	-0.892	-0.793 ± 0.426 **
	MYO3	0.451	0.756 ± 1.32
PtCP	TDH1 (Control)	-0.883	-5.935 ± 1.97 ***
	PAU12	1.192	3.816 ± 2.15
	BIT61	0.792	1.438 ± 4.13
	WSC3	0.344	2.466 ± 3.68
	MDG1	-1.182	-0.968 ± 2.12
	ITC1	0.403	2.066 ± 2.11
	RSC30	3.738	10.960 ± 1.22 *
	GAT1	1.728	4.339 ± 0.478 ***
	RPG1	-1.45	1.235 ± 11.7
	RPN4	3.494	8.428 ± 0.274
	DBF2	-0.637	0.000 ± 0.000
* ($\alpha<0.10$) ** ($\alpha<0.05$) *** ($\alpha<0.01$)			

Table 6A. BLASTn and BLASTp results from NCBI for creosote. An "X" denotes where the BLAST analysis yielded no results.

Blastn and Blastp Analysis Creosote						
Gene Name	Nucleotide Sequence	E-Value	Max Identity	Conserved Protein Domain	E-Value	Max Identity
YLR319C	Homo sapiens poly(A) binding protein, cytoplasmic 1 (PABPC1)	0	95%	AIP3 superfamily: Actin interacting protein	4E-147	100%
YBR145W	X	X	X	envelope glycoprotein	0	100%
YDR219C	Homo sapiens fibroblast growth factor receptor 2 (bacteria-expressed kinase, keratinocyte growth factor receptor, craniofacial dysostosis 1, Crouzon syndrome, Pfeiffer syndrome, Jackson-Weiss syndrome) (FGFR2)	6.00E-159	99%	X	X	X
YJL060W	Homo sapiens myeloid/lymphoid or mixed-lineage leukemia 3 (MLL3)	8.00E-169	95%	X	X	X
YMR211W	Homo sapiens cutA divalent cation tolerance homolog (E. coli) (CUTA)	3.00E-168	97%	X	X	X
YCR077C	Homo sapiens matrix metallopeptidase 11 (stromelysin 3) (MMP11)	1.00E-161	91%	X	X	X
YGL237C	Homo sapiens hemoglobin, gamma G (HBG2)	8.00E-164	94%	X	X	X
YGL263W	Homo sapiens LSM14A, SCD6 homolog A (S. cerevisiae) (LSM14A)	0	96%	X	X	X
YOR264W	X	X	X	zf-MYND superfamily: MYND-type zinc finger protein MUB1	0	100%
YER081W	Homo sapiens growth differentiation factor 15 (GDF15)	0	96%	X	X	X
YBR296C	Homo sapiens muskelin 1, intracellular mediator containing kelch motifs (MKLN1)	3.00E-38	100%	X	X	X
YDL209C	Homo sapiens adducin 3 (gamma) (ADD3)	6.00E-154	96%	X	X	X
YMR301C	Homo sapiens ubiquitin-conjugating enzyme E2H (UBC8 homolog, yeast) (UBE2H)	0	97%	X	X	X
YPL107W	Homo sapiens phospholipid scramblase 1 (PLSCR1)	5.00E-145	97%	X	X	X
YCR104W	Homo sapiens solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 6 (SLC25A6), nuclear gene encoding mitochondrial protein	5.00E-68	90%	X	X	X
YGR294W	Homo sapiens CD14 molecule (CD14)	2.00E-179	95%	X	X	X
YJL125C	Homo sapiens bromodomain adjacent to zinc finger domain, 2B (BAZ2B)	7.00E-164	97%	X	X	X

Table 6B. BLASTn and BLASTp results from NCBI for PtCP. (Continued on next page.) An "X" denotes where the BLAST analysis yielded no results.

Blastn and Blastp Analysis PtCP						
Gene Name	Nucleotide Sequence	E-Value	Max Identity	Conserved Protein Domain	E-Value	Max Identity
YHR044C	Homo sapiens 1-acylglycerol-3-phosphate O-acyltransferase 1 (lysophosphatidic acid acyltransferase, alpha) (AGPAT1)	0	95%	X	X	X
YHR054C	Homo sapiens hemoglobin, beta (HBB)	2.00E-173	97%	X	X	X
YHR056C	Homo sapiens interferon induced transmembrane protein 3 (1-8U) (IFITM3)	5.00E-130	96%	X	X	X
YHR130C	Homo sapiens tumor necrosis factor, alpha-induced protein 2 (TNFAIP2)	1.00E-124	98%	X	X	X
YIL056W	Homo sapiens coactosin-like 1 (Dictyostelium) (COTL1)	0	95%	X	X	X
YIL101C	X	X	X	AdoMet_Mtases (NADB_Rossmann superfamily): rRNA methyltransferase	5.00E-139	100%
YIR011C	Homo sapiens microtubule-associated protein 1 light chain 3 beta (MAP1LC3B)	2.00E-127	98%	X	X	X
YJL058C	Homo sapiens proteoglycan 2, bone marrow (natural killer cell activator, eosinophil granule major basic protein) (PRG2)	0	97%	X	X	X
YJL088W	Homo sapiens fibronectin type III domain containing 3B (FNDC3B)	0	94%	X	X	X
YJL113W	Homo sapiens biliverdin reductase A (BLVRA)	1.00E-151	97%	X	X	X
YJL185C	Homo sapiens polyhomeotic homolog 1 (Drosophila) (PHC1)	3.00E-153	95%	X	X	X
YJR010W	Homo sapiens zinc and ring finger 2 (ZNRF2)	2.00E-163	96%	X	X	X
YJR126C	Homo sapiens acetylcholinesterase (Yt blood group) (ACHE)	0	100%	X	X	X
YKL096W	X	X	X	Transaldolase_FSA, TIM_phosphate_binding superfamily: putative translaldolase	6.00E-98	X
YKL120W	X	X	X	PMSR superfamily: peptide methionine sulfoxide reductase	3.00E-151	100%
YLL028W	Homo sapiens ring finger protein 13 (RNF13)	6.00E-180	96%	X	X	X
YLR014C	Homo sapiens peroxiredoxin 2 (PRDX2), nuclear gene encoding mitochondrial protein, transcript variant 1	6.00E-171	94%	X	X	X
YLR036C	Homo sapiens vestigial like 1 (Drosophila) (VGLL1)	7.00E-155	91%	X	X	X
YLR037C	Homo sapiens family with sequence similarity 113, member B (FAM113B)	5.00E-17	100%	X	X	X
YLR057W	Homo sapiens DnaJ (Hsp40) homolog, subfamily C, member 13 (DNAJC13)	0	97%	X	X	X
YLR058C	Homo sapiens endothelial cell adhesion molecule (ESAM)	5.00E-161	92%	X	X	X
YLR099C	Homo sapiens signal recognition particle 14kDa (homologous Alu RNA binding protein) (SRP14)	6.00E-150	94%	X	X	X
YLR108C	Homo sapiens phospholipid scramblase 1 (PLSCR1)	5.00E-129	97%	X	X	X
YLR142W	Homo sapiens signal sequence receptor, alpha (translocon-associated protein alpha) (SSR1)	0	96%	X	X	X
YLR165C	Homo sapiens phosphatidylinositol transfer protein, alpha (PITPNA)	6.00E-159	97%	X	X	X
YLR172C	Homo sapiens keratin 19 (KRT19)	6.00E-154	97%	X	X	X
YLR174W	Homo sapiens regulator of chromosome condensation 2 (RCC2)	4.00E-177	93%	X	X	X
YLR198C	Homo sapiens septin 2 (SEPT2)	1.00E-161	93%	X	X	X
YLR206W	Homo sapiens myosin IC (MYO1C)	2.00E-175	90%	X	X	X
YLR294C	Homo sapiens KIAA1219 (KIAA1219)	0	96%	X	X	X
YLR297W	Homo sapiens ring finger protein 13 (RNF13)	6.00E-180	96%	X	X	X
YLR323C	Homo sapiens topoisomerase (DNA) II binding protein 1 (TOPBP1)	0	94%	X	X	X
YLR346C	Homo sapiens GULP, engulfment adaptor PTB domain containing 1 (GULP1)	2.00E-120	99%	X	X	X
YLR443W	Homo sapiens ribosomal protein S12 (RPS12)	1.00E-151	93%	X	X	X
YLR449W	Homo sapiens hemoglobin, beta (HBB),	9.00E-178	96%	X	X	X
YML076C	Homo sapiens RAS and EF-hand domain containing (RASEF)	0	96%	X	X	X
YML117W	Homo sapiens septin 7 (SEPT7)	5.00E-98	96%	X	X	X
YMR015C	Homo sapiens growth differentiation factor 15 (GDF15)	0	95%	X	X	X
YMR016C	Homo sapiens pregnancy specific beta-1-glycoprotein 4 (PSG4)	0	95%	X	X	X
YMR102C	Homo sapiens family with sequence similarity 3, member A (FAM3A)	0	98%	X	X	X
YMR119W	Homo sapiens zinc finger protein 770 (ZNF770)	0	95%	X	X	X
YMR250W	Homo sapiens WD repeat and FYVE domain containing 1 (WDFY1)	3.00E-53	100%	X	X	X
YMR281W	Homo sapiens dual specificity phosphatase 1 (DUSP1)	1.00E-62	97%	X	X	X
YNL036W	Homo sapiens keratin 23 (histone deacetylase inducible) (KRT23)	1.00E-162	90%	X	X	X
YNL094W	Homo sapiens coiled-coil and C2 domain containing 1B (CC2D1B)	2.00E-164	94%	X	X	X
YNL165W	Homo sapiens retinoblastoma binding protein 6 (RBBP6)	2.00E-121	95%	X	X	X
YNL173C	Homo sapiens bromodomain and WD repeat domain containing 1 (BRWD1)	0	94%	X	X	X

Table 6B. BLASTn and BLASTp results from NCBI for PtCP. (Continued on next page.) An "X" denotes where the BLAST analysis yielded no results.

Blastn and Blastp Analysis PtCP						
Gene Name	Nucleotide Sequence	E-Value	Max Identity	Conserved Protein Domain	E-Value	Max Identity
YBR191W	Homo sapiens 18S ribosomal RNA	0	85%	X	X	X
YBR198C	Homo sapiens 18S ribosomal RNA	0	85%	X	X	X
YBR253W	Homo sapiens peroxiredoxin 1 (PRDX1)	0	96%	X	X	X
YBR259W	Homo sapiens tumor protein p53 binding protein 1 (TP53BP1)	0	94%	X	X	X
YCR088W	Homo sapiens serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary) (SERPING1)	4.00E-24	97%	X	X	X
YDL020C	Homo sapiens myristoylated alanine-rich protein kinase C substrate (MARCKS)	8.00E-148	96%	X	X	X
YDL214C	Homo sapiens methyltransferase like 7A (METTL7A)	0	97%	X	X	X
YDL218W	Homo sapiens ubiquitin-conjugating enzyme E2, J2 (UBC6 homolog, yeast) (UBE2J2)	0	95%	X	X	X
YDR042C	Homo sapiens UTP3, small subunit (SSU) processome component, homolog (S. cerevisiae) (UTP3)	0	97%	X	X	X
YDR091C	Homo sapiens proteasome (prosome, macropain) 26S subunit, non-ATPase, 8 (PSMD8)	0	97%	X	X	X
YDR153C	Homo sapiens paraoxonase 2 (PON2)	4.00E-130	100%	X	X	X
YDR222W	Homo sapiens placenta-specific 4 (PLAC4)	0	98%	X	X	X
YDR363W	Homo sapiens mitochondrial ribosomal protein L13 (MRPL13), nuclear gene encoding mitochondrial protein	0	94%	X	X	X
YDR465C	X	X	X	Ferritin_like superfamily: metalloregulation DNA-binding stress protein	7.00E-85	X
YEL065W	Homo sapiens dynactin 4 (p62) (DCTN4)	0	96%	X	X	X
YEL066W	Homo sapiens tissue factor pathway inhibitor 2 (TFPI2)	0	95%	X	X	X
YER025W	Homo sapiens ATPase, Na+/K+ transporting, beta 3 polypeptide (ATP1B3)	7.00E-164	97%	X	X	X
YER027C	Homo sapiens PALM2-AKAP2 (PALM2-AKAP2)	0	97%	X	X	X
YER029C	Homo sapiens cadherin 11, type 2, OB-cadherin (osteoblast) (CDH11)	3.00E-179	90%	X	X	X
YER052C	Homo sapiens tubulin, alpha 1a (TUBA1A)	4.00E-166	95%	X	X	X
YER067W	Homo sapiens jumonji domain containing 6 (JMJD6)	9.00E-144	99%	X	X	X
YER091C	Homo sapiens solute carrier family 38, member 1 (SLC38A1)	0	94%	X	X	X
YER153C	Homo sapiens placental growth factor (PGF)	9.00E-168	97%	X	X	X
YFR015C	Homo sapiens paternally expressed 10 (PEG10)	3.00E-168	94%	X	X	X
YFR019W	Homo sapiens nudix (nucleoside diphosphate linked moiety X)-type motif 21 (NUDT21)	2.00E-165	91%	X	X	X
YGL057C	Homo sapiens heterogeneous nuclear ribonucleoprotein C (C1/C2) (HNRNPC)	4.00E-166	97%	X	X	X
YGL092W	Homo sapiens SH3 domain and tetratricopeptide repeats 2 (SH3TC2)	0	95%	X	X	X
YGL133W	Homo sapiens LSM4 homolog, U6 small nuclear RNA associated (S. cerevisiae) (LSM4)	2.00E-92	98%	X	X	X
YGL144C	Homo sapiens G protein-coupled receptor 126 (GPR126)	9.00E-147	98%	X	X	X
YGL155W	Homo sapiens protein O-linked mannose beta1,2-N-acetylglucosaminyltransferase (POMGNT1)	0	96%	X	X	X
YGL157W	Homo sapiens phosphatase and actin regulator 2 (PHACTR2)	6.00E-170	95%	X	X	X
YGL198W	Homo sapiens ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C2 (ATP6V1C2)	2.00E-164	100%	X	X	X
YGL202W	Homo sapiens slowmo homolog 2 (Drosophila) (SLMO2)	0	95%	X	X	X
YGL246C	Homo sapiens erythroblast membrane-associated protein (Scianna blood group) (ERMAP)	4.00E-23	95%	X	X	X
YGR052W	Homo sapiens adaptor-related protein complex 3, delta 1 subunit (AP3D1)	0	97%	X	X	X
YGR063C	Homo sapiens mediator complex subunit 23 (MED23)	0	98%	X	X	X
YGR146C	Homo sapiens cAMP responsive element binding protein 3-like 2 (CREB3L2)	1.00E-141	92%	X	X	X
YGR171C	Homo sapiens vacuolar protein sorting 24 homolog (S. cerevisiae) (VPS24)	2.00E-163	99%	X	X	X
YGR197C	Homo sapiens arsenate resistance protein 2 (ARS2)	3.00E-162	97%	X	X	X
YGR238C	Homo sapiens growth differentiation factor 15 (GDF15)	3.00E-177	96%	X	X	X
YGR264C	Homo sapiens tripeptidyl peptidase I (TPP1)	0	96%	X	X	X
YGR294W	Homo sapiens CD14 molecule (CD14)	2.00E-179	95%	X	X	X
YHL024W	Homo sapiens exportin, tRNA (nuclear export receptor for tRNAs) (XPOT)	0	93%	X	X	X
YHL029C	Homo sapiens flotillin 1 (FLOT1)	0	96%	X	X	X
YHL047C	Homo sapiens growth differentiation factor 15 (GDF15)	5.00E-145	97%	X	X	X
YHR020W	Homo sapiens prolyl endopeptidase-like (PREPL)	9.00E-91	95%	X	X	X
YHR026W	Homo sapiens growth differentiation factor 15 (GDF15)	5.00E-145	97%	X	X	X

Table 6B. BLASTn and BLASTp results from NCBI for PtCP. An "X" denotes where the BLAST analysis yielded no results.

Blastn and Blastp Analysis PtCP						
Gene Name	Nucleotide Sequence	E-Value	Max Identity	Conserved Protein Domain	E-Value	Max Identity
YNR014W	Homo sapiens PDZ domain containing 2 (PDZD2)	0	98%	X	X	X
YOL015W	Homo sapiens ATPase, Cu++ transporting, alpha polypeptide (Menkes syndrome) (ATP7A)	2.00E-69	98%	X	X	X
YOL017W	Homo sapiens ATPase, Cu++ transporting, alpha polypeptide (Menkes syndrome) (ATP7A)	4.00E-92	100%	X	X	X
YOL073C	Homo sapiens family with sequence similarity 3, member A (FAM3A)	0	100%	X	X	X
YOL105C	X	X	X	Herpes_UL95 superfamily [Human herpesvirus 6A]	0	100%
YOL111C	X	X	X	Herpes_U59 superfamily: tegument protein UL88 [Human herpesvirus 6A]	0	100%
YOL127W	X	X	X	EFRF1 [Human herpesvirus 6]	1.00E-55	100%
YOL149W	X	X	X	Herpes_UL31 superfamily: nuclear egress lamina protein	5.00E-154	100%
YOL151W	X	X	X	Herpes_UL33: XILF1 [Human herpesvirus 6]	1.00E-55	100%
YOR011W	X	X	X	CA superfamily: LI-cadherin	0	100%
YOR153W	X	X	X	Sema superfamily, PSI superfamily: Sema domain, immunoglobulin domain (Ig)	0	100%
YOR306C	X	X	X	PAP1 superfamily: Basic leucine zipper (bZIP) transcription factor	0	100%
YOR321W	X	X	X	CAP_N superfamily, CARP superfamily, CAP_C multidomain: CAP (cyclase-associated protein)	0	100%
YOR382W	X	X	X	Putative protein of unknown function	0	100%
YPL027W	Homo sapiens hydroxyprostaglandin dehydrogenase 15-(NAD) (HPGD)	5.00E-176	92%	X	X	X
YPL058C	Homo sapiens amino-terminal enhancer of split (AES)	8.00E-174	91%	X	X	X
YPL088W	Homo sapiens calumenin (CALU)	4.00E-171	96%	X	X	X
YPL120W	Homo sapiens reticulocalbin 2, EF-hand calcium binding domain (RCN2)	1.00E-142	91%	X	X	X
YPL135W	Homo sapiens keratin 7 (KRT7)	7.00E-149	94%	X	X	X
YPL171C	Homo sapiens tumor necrosis factor, alpha-induced protein 8 (TNFAIP8)	2.00E-163	99%	X	X	X
YPL250C	Homo sapiens WD repeat domain 68 (WDR68)	0	96%	X	X	X
YPL259C	Homo sapiens ubiquitin-like 3 (UBL3)	8.00E-159	91%	X	X	X
YPR034W	Homo sapiens UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 3 (B4GALT3)	8.00E-169	92%	X	X	X
YPR065W	Homo sapiens hydroxyprostaglandin dehydrogenase 15-(NAD) (HPGD)	2.00E-179	93%	X	X	X
YPR093C	Homo sapiens retinoblastoma binding protein 6 (RBBP6), transcript variant 3	0	95%	X	X	X
YPR116W	Homo sapiens placenta-specific 1 (PLAC1)	1.00E-146	97%	X	X	X