

Mycoplasma infection significantly alters microarray gene expression profiles

Crispin J. Miller, Heba S. Kassem, Stuart D. Pepper, Yvonne Hey, Timothy H. Ward, and Geoffrey P. Margison

Paterson Institute for Cancer Research, Manchester, UK

BioTechniques 35:812-814 (October 2003)

Mycoplasmas are the smallest and simplest self-replicating prokaryotes. Most are parasites, and their contamination of primary and continuous eukaryotic cell lines represents a significant problem in research, diagnosis, and biotechnological production (1). Unlike bacteria and fungi, they rarely produce turbid growth or cellular damage, and most are resistant to the antibiotics commonly used in long-term cell cultures (1,2). These factors combine to make them a stubborn contaminant, and recent surveys have shown that they affect up to 87% of cell lines. Mycoplasmas have also been known to infect the respiratory, gastrointestinal, and urogenital tracts of many patients, often without any apparent illness (3). Their ability to cause chromosomal rearrangements has led some to postulate them as a cause of cancer. Others have argued that they are cofactors for a variety of conditions including arthritis, Crohn's disease, and acquired immunodeficiency syndrome (AIDS) (4).

Mycoplasma genomes are small and have limited options for metabolism and replication. Genomic analysis of *Mycoplasma pneumoniae* and *Mycoplasma genitalium*, for example, show that they have no genes involved in amino acid biosynthesis and very few responsible for that of cofactors (2,5,6). Instead, they must depend on their host cells for the supply of essential amino acids and vitamins. Competition for nutrients can disrupt host cell integrity and alter function. To colonize a host cell, mycoplasma must adhere to the cell's surface, thereby interfering with its receptors, altering its transport mechanisms, and triggering specific signal cascades. Its lack of a cell wall can allow fusion between the mycoplasma and host, and when this occurs, a variety of components are delivered, in-

cluding hydrolytic enzymes, cytokines, chemokines, prostaglandins, and active oxygen and nitrogen metabolites. Mycoplasmas also show mitogenic activity. *Mycoplasma penetrans*, for instance, has components that display properties similar to leukocyte integrins (2).

Thus, mycoplasma infection has diverse and far-reaching consequences, from unsafe biological products, to unreliable experimental results (1). Irrespective of their role as a benign colonist, a cofactor, or even a cause of disease, their ubiquity has profound implications for diagnosis, classification, or characterization on the basis of patient samples or cell lines. Studies of the effects of mycoplasmas are often difficult to perform, because their presence is generally undesirable, and there is an understandable reluctance to introduce them into the laboratory intentionally.

During the course of a microarray experiment, routine screening identified mycoplasma contamination within one replicate set of experimental controls (after RNA extraction, but before hybridization to the chips). Rather than simply dispose of the samples, we took the opportunity to explore the effects of mycoplasma infection by hybridizing the contaminated samples to arrays and comparing their expression profiles to the replacement, uncontaminated, controls. There was a certain amount of serendipity in this, as we chose to retrospectively analyze the results of a compromised experiment, rather than to deliberately infect a cell line with mycoplasma.

The design of the original experiment, which was intended to test the response of MCF7 human breast adenocarcinoma cells to a pair of drug treatments, is shown in Figure 1. There were three separate control groups: (i) cells that were simply resuspended for 2 h (CSC); (ii) cells that were

treated for 2 h with dimethyl sulfoxide (DMSO) (DC); and (iii) cells that were treated for 4 h with DMSO (2DC). Mycoplasma contamination resulted in one sample from each of these groups becoming infected, and an extra set of replicates was generated to replace the infected ones. In order to investigate the effects of contamination, the three contaminated samples were compared to the nine uncontaminated replicates (three from each group).

MCF-7 cells were cultured in RPMI with 10% fetal calf serum (FCS) (both from PAA Laboratories, Yeovil, UK) and L-glutamine (Sigma, Gillingham, UK). No antibiotics or antifungal treatments were used. During routine mycoplasma screening using the Hoechst technique (7), the first batch of cells were shown to be positive. This identified the presence of mycoplasma, but not the specific strain. Retrospective analysis was not possible, because the contaminated samples were immediately destroyed. Subsequent batches of cells grown from fresh stocks were shown to be clear by repeated PCR screening both before and after the experimental procedure.

All experiments were performed using HgU133A oligonucleotide arrays (Affymetrix, High Wycombe, UK) as described at (<http://www.affymetrix.com/products/arrays/specific/hgu133.affx>). Total RNA from each sample was used to prepare biotinylated target RNA, with minor modifications from the manufacturer's recommendations at (http://www.affymetrix.com/support/technical/manual/expression_manual.affx). Briefly, 10 µg of mRNA were used to generate first-strand cDNA by using a T7-linked oligo(dT) primer. After second-strand synthesis, in vitro transcription was performed with biotinylated UTP and CTP (Enzo Diagnostics, Farmingdale, NY, USA), resulting in approximately 100-fold linear amplification of RNA. A complete description of biochemical procedures is available at (http://bioinf.picr.man.ac.uk/mbcf/downloads/GeneChip_Target_Prep_Protocol_CRUK_v_2.pdf). The target cDNA generated from each sample was processed as per the manufacturer's recommendation using a GeneChip® Instrument System (Affymetrix; http://www.affymetrix.com/support/technical/manual/expression_manual.affx). Briefly, eukaryotic "spike" controls were added

to 10 µg fragmented cRNA before an overnight hybridization. Arrays were then washed and stained with streptavidin-phycoerythrin before being scanned on a GeneChip scanner (Affymetrix). A com-

plete description of these procedures is available at (http://bioinf.picr.man.ac.uk/mbcf/downloads/GeneChip_Hyb_Wash_Scan_Protocol_v_2.pdf).

The quality and amount of starting

RNA were confirmed by agarose gel electrophoresis. After scanning, array images were assessed by eye to confirm scanner alignment and the absence of significant bubbles or scratches on the chip surface. Ratios (3'/5') for GAPDH and β-actin were confirmed to be within acceptable limits (0.78–0.94 and 0.91–1.08, respectively), and BioB spike controls were found to be present on all chips, with BioC, BioD, and CreX also present in increasing intensity. When scaled to a target intensity of 100 [using Microarray Analysis Suite™ (MAS) version 5.0 array analysis software (Affymetrix)], scaling factors for all arrays were within acceptable limits (0.53–1.23), as were background, Q values, and mean intensities. Further details of quality control metrics can be found on the supplementary Web site at (<http://bioinf.picr.man.ac.uk/mbcf/experiments/index.shtml>).

Nine control and three infected samples were assayed to determine changes in expression levels using U133A microarrays. All data analysis and normalization were performed using the MAS version 5.0 software. All arrays were scaled to a target intensity of 100 before a Mann-Whitney *U*-test, with a certainty level of $P \leq 0.02$, was used to identify genes that showed statistically signifi-

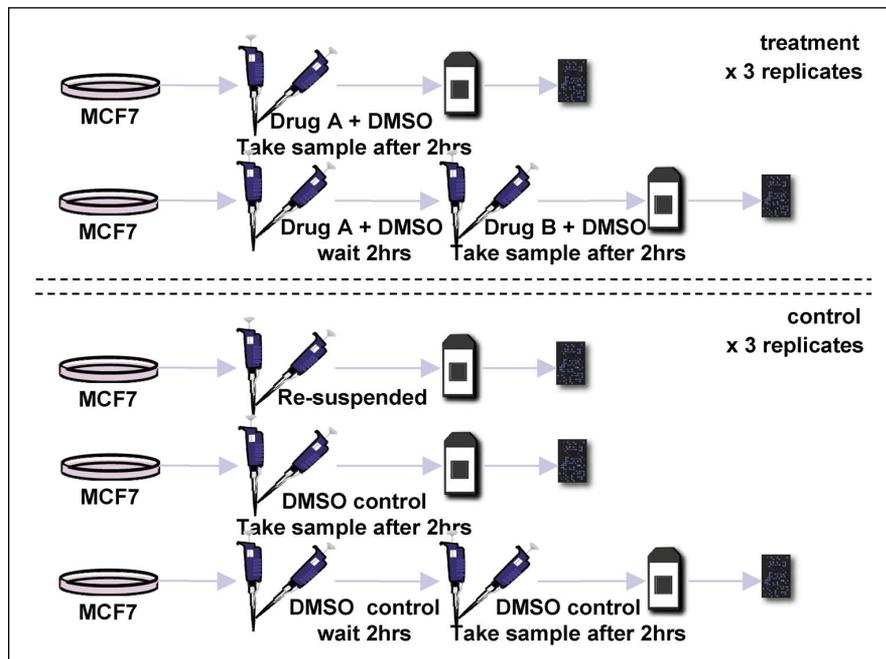


Figure 1. The original experiment that provided the mycoplasma-contaminated cell lines. The experiment was designed to test the response of MCF7 human breast adenocarcinoma cells to a pair of drug treatments. During routine screening, one set of control replicates was found to be contaminated with mycoplasma. These samples were retrospectively analyzed to investigate the effects of mycoplasma infection on gene expression.

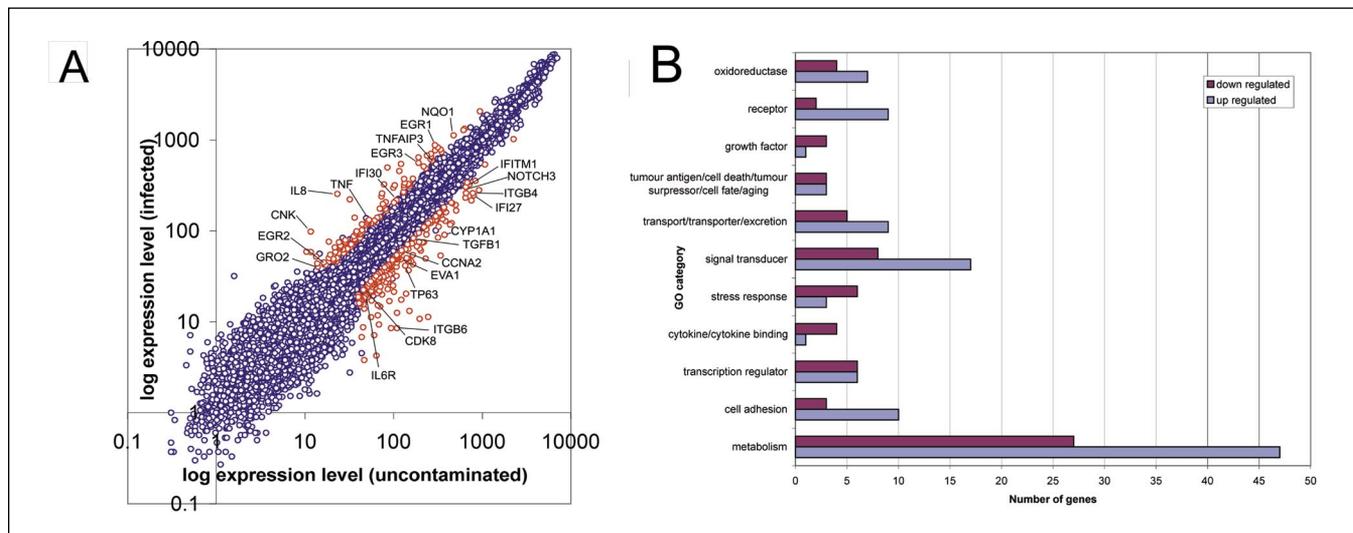


Figure 2. Analysis of significantly changing mRNA levels between mycoplasma-infected and noninfected cell cultures. Nine control and three infected samples were assayed to determine changes in expression levels using U133A microarrays. Expression levels were generated using Microarray Analysis Suite (MAS) version 5.0. (A) Mean expression levels were calculated for control versus infected replicates (blue circles). A Mann-Whitney *U*-test with a certainty level of $P \leq 0.02$ was used to identify genes that showed statistically significant variation between sample groups. Of this set, genes showing an absolute fold change ≥ 2.0 were considered significant (red circles). Such a deliberately stringent threshold was used to minimize false positives. Even so, more than 200 genes were identified as being up- or down-regulated, demonstrating that mycoplasma infection has a profound effect on gene expression patterns. (B) Functional assignments of the genes selected in panel A according to the Gene Ontology hierarchy (9). Detailed experimental methods and software can be found on the supplementary Web site (<http://bioinf.picr.man.ac.uk/mbcf/experiments/index.shtml>).

cant variation between sample groups. Of this set, genes showing an absolute fold change ≥ 2.0 in their mean expression levels between replicates were considered significant. Such a deliberately stringent threshold was used to minimize false positives. Even so, more than 200 genes were identified as being up- or down-regulated out of the approximately 22,000 probe sets. All data can be found at the supplementary Web site.

This analysis demonstrated that mycoplasmas alter the expression of hundreds of genes in cultured human cells. A variety of pathways are affected, and both up-regulation (up to 10-fold) and down-regulation (up to 12-fold) were seen. We were able to demonstrate that mycoplasma infection alters the expression not only of many cytokines (2), but also stress-response genes, transport proteins, receptors, ion channels, growth-factors, oxidases, tumor suppressors, and oncogenes (Figure 2). Comparison between the three control groups identified no significant changing genes with $P \leq 0.02$, although the different group sizes makes direct comparisons on P-score alone difficult. A less stringent P-score cut-off of ≤ 0.05 and a fold change > 2 identifies 22 changing genes between CSC and DC, 9 between CSC and 2DC, and 36 between DC and 2DC. The maximum fold change within any of these sets was 4.22. Thus, although other interactions were found between the different control groups, these changes were fewer in number and were relatively small with respect to those seen between the infected and clean samples.

It is reasonable to assume that the effects of mycoplasmas in vivo will be similarly wide-ranging to that in vitro, with potentially major consequences for the diagnosis and characterization of diseases using expression array technology. Given these substantial effects on expression patterns and the virulence of mycoplasma, sample screening immediately prior to RNA isolation should be mandatory for all array studies.

The establishment of public microarray repositories means that these issues have consequences far beyond the original experiment (8). Database annotations should include details of mycoplasma screening, detection, and treatment, so that comparisons between experiments

on the same cell line can be trusted. Meanwhile, a systematic comparison of expression profiles from various microorganism-infected, noninfected, and treated ("cured") cell lines would provide infection fingerprints that could facilitate a retrospective assessment and adjustment of data already generated.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of Cancer Research UK.

REFERENCES

1. **Rawadi, G. and O. Dussurget.** 1995. Advances in PCR-based detection of mycoplasmas contaminating cell cultures. *PCR Methods Appl.* 4:199-208.
2. **Rottem, S. and Y. Naot.** 1998. Subversion and exploitation of host cells by mycoplasmas. *Trends Microbiol.* 6:436-440.
3. **Zhang, S., D.J. Wear, and S. Lo.** 2000. Mycoplasmal infections alter gene expression in cultured human prostatic and cervical epithelial cells. *FEMS Immunol. Med. Microbiol.* 27:43-50.
4. **Baseman, J.B. and J.G. Tully.** 1997. Mycoplasmas: sophisticated, reemerging, and burdened by their notoriety. *Emerg. Infect. Dis.* 3:21-32.
5. **Himmelreich, R., H. Plagens, H. Hilbert, B. Reiner, and R. Herrmann.** 1997. Comparative analysis of the genomes of the bacteria *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. *Nucleic Acids Res.* 25:701-712.
6. **Pollack, J.D., M.V. Williams, and R.N. McElhaney.** 1997. The comparative metabolism of the mollicutes (Mycoplasmas): the utility for taxonomic classification and the relationship of putative gene annotation and phylogeny to enzymatic function in the smallest free-living cells. *Crit. Rev. Microbiol.* 23: 269-354.
7. **Chen, T.R.** 1997. In situ detection of mycoplasma contamination in cell cultures by fluorescent Hoechst 33258 stain. *Exp. Cell Res.* 104:255-262.
8. **Miller, C.J. and T.K. Attwood.** 2003. Bioinformatics goes back to the future. *Nat. Rev. Mol. Cell Biol.* 4:157-162.
9. **The Gene Ontology Consortium.** 2001. Creating the gene ontology resource: design and implementation. *Genome Res.* 11:1425-1433.

Received 29 April 2003; accepted 9 June 2003.

Address correspondence to Crispin J. Miller, Bioinformatics Group, Paterson Institute for Cancer Research, Christie Hospital Trust, Manchester, M20 4BX, UK. e-mail: cmiller@picr.man.ac.uk