Pure Appl. Chem., Vol. 75, Nos. 11–12, pp. 2415–2418, 2003.

Workshop 2.2
Temporal responses to estrogen in the uterus*

Kirsten C. Fertuck, on behalf of Tim R. Zacharewski

Biochemistry and Molecular Biology, National Food Safety and Toxicology Center, and Institute for Environmental Toxicology, Michigan State University, East Lansing, MI 48824-1319, USA

Abstract: The estrogen receptor acts primarily as an estrogen-inducible transcription factor in target organs, however, until recently the identification of individual responsive transcripts has been cumbersome. In the present study, oligonucleotide GeneChip microarrays were used to describe and analyze the levels of approximately 6500 transcripts in the uteri of immature, ovariectomized mice at various times following oral exposure to 100 µg/kg of ethynyl estradiol. The most reproducible responses were identified and subjected to K-means clustering, and functional annotations were obtained for the transcripts within each cluster. Connections were made between the observed transcriptional responses and the known physiological responses to estrogen exposure, focusing here on effects on uterine cell cycle progression leading to estrogen-stimulated uterine growth.

INTRODUCTION

Efforts to identify compounds that act as estrogenic endocrine disruptors in vivo often use the rodent uterus as a model, because the rapid and robust stimulation of uterine tissue proliferation is a very characteristic response to estrogen exposure [1]. A dominant role of the estrogen receptor is as an estrogen-inducible transcription factor, and the modulation of transcript levels of a number of genes both prior to and during this proliferative response have been described. Some of these have then been used as indicators of the estrogenic action of exogenous compounds, with lactoferrin and complement component C3 being particularly well characterized estrogen-inducible markers [2]. However, more detailed examinations have indicated that estrogens may vary in the specific transcript profiles that they induce, which may be reflective of differences in their mechanism of action [3,4]. More study is therefore needed in order to identify the responses that are truly indicative of an estrogenic response, with the hope of eventually identifying a reliable subset that signals adverse responses to estrogenic exposure.

While the rat is a very commonly used species in uterine studies, mouse was used as the animal model of choice in the present study for several reasons. First, results from these studies can complement and be compared to results of other murine uterotrophic and microarray studies within the laboratory. As well, maintenance and dosing is relatively cost-effective, public databases of sequence and functional information are quite well annotated, and results can be compared with those of possible future studies using mice with targeted disruptions in genes of interest. However, a major drawback is the small tissue size and consequently low RNA yield from the mouse uteri compared with larger species.

*Report from a SCOPE/IUPAC project: Implication of Endocrine Active Substances for Human and Wildlife (J. Miyamoto and J. Burger, editors). Other reports are published in this issue, *Pure Appl. Chem.* **75**, 1617–2615 (2003).

APPROACH

In the present study, duplicate immature, ovariectomized C57BL/6 mice were treated with 100 μg/kg 17α-ethynyl estradiol (EE), an orally active pharmaceutical estrogen, or with vehicle for 2, 8, 12, or 24 h. Additional groups were administered three daily doses of EE or vehicle before sacrifice, which induced the approximately 5-fold increase in uterine weight characteristic of the widely used uterotrophic assay for estrogenic effect [1,5]. RNA was extracted from individual uteri, and samples were separately hybridized to Affymetrix Mu11KSubA GeneChip arrays (Santa Clara, CA) using the methods provided by the manufacturer. The Affymetrix MicroArray Suite 5.0 software was used to obtain a Signal intensity value for each probe set, which is the set of sequences that report the relative abundance of one transcript.

The five-step approach that was used to analyze the resulting microarray data is described below. Using this approach, the most reproducible, treatment-induced responses were identified, the temporal response patterns were characterized, and the associated gene functions were identified and discussed.

Screening for active responses

In a method that will be published separately, a modification of the empirical Bayes approach [6] was employed to analyze all data at all time points, yielding one summary statistic per probe set indicative of the degree and reproducibility of the response over time relative to that of the time-matched controls. A cut-off value of 0.99 was arbitrarily assigned, and all probe sets with associated p1z values greater than the cut-off value, which consisted of approximately 15 % of all measured responses, were retained for further analysis.

Screening active responses for treatment effect

Probe set responses that were identified as being active in the first screening step were then further screened using an analysis of variance. In this method, responses that were significant at the arbitrarily chosen $p < 0.01$ level in either Treatment or in the Treatment*Time interaction were considered to have a significant treatment effect, with the approximately 400 responses representing 45 % of those retained by the first screening step. The remaining responses with solely a significant Time effect were rejected since this is suggestive of either vehicle or circadian responses.

Clustering

For the 400 responses that passed the second stage of the screening procedure, the averaged duplicate Signal responses at each time point, normalized by the corresponding time-matched control values, were clustered by the K-means method in GeneSpring (Redwood City, CA). Though the 3×24 h treatment group was the only group to receive multiple doses, this group was included with all others for clustering purposes since the relationship of probe set responses before and during estrogen-induced proliferation is of interest.

In K-means clustering the number of clusters must be specified by the user. By examining results of between 2 and 15 clusters, it was judged that seven K-means clusters appeared to most fully describe the 400 responses. These seven clusters could be generally described as follows: upregulated only at 2 h; upregulated only at 8 h; upregulated at 8 and 3×24 h; upregulated at 12 h; upregulated only at 24 h; upregulated at 24 and 3×24 h; and upregulated at 3×24 h only.

Annotation

Following the assignment of the 400 responses to the seven clusters, the identities of the responses were determined and functions were ascribed where possible, using public repositories of gene information and customized Perl scripts that return hyperlinked Excel output files. First, the GenBank accession ID associated with each probe set was used to query the National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov>) UniGene database. In approximately 60 % of cases a gene name, abbreviation, and UniGene ID were returned, while in most other cases the accession ID corresponded to an unnamed expressed sequence tag (EST). In 23 cases, the accession ID was not contained within the UniGene database, but a BLAST search of the probe set sequence in the GenBank databases allowed identification of a putative gene name in 18 cases.

Additional information for probe sets with associated gene names was then obtained from the NCBI LocusLink database, including mRNA and protein RefSeq IDs, chromosomal location, NCBI PubMed references, and Gene Ontology names and IDs. In addition, the LocusLink ID for the human homolog was obtained where available, which then provided links to human databases such as the NCBI Online Mendelian Inheritance in Man (OMIM) database and the Weizmann Institute GeneCards database (<http://bioinfo.weizmann.ac.il/cards>).

These scripts were also used to annotate hand-curated published information for uterine estrogen-responsive genes, and UniGene ID was then used as the identifier with which to allow comparison of the present experimental results with published results.

A future interest is to expand the capabilities of dbZach, the in-house database used for microarray data analysis and interpretation, to support GeneChip data. This database currently supports only cDNA microarray data, but its features are also applicable to other array types. dbZach stores LocusLink, UniGene, and Gene Ontology information, as well as toxicology data of interest, and links it automatically to stored microarray data and sample annotation information to facilitate analysis.

Interpretation

For simplicity, we focus here on effects of estrogen on the transcription of genes related to cell cycle progression. A well studied effect of estrogen exposure in the uterus is the stimulation of synchronized cell division which results in greatly increased uterine tissue mass within a relatively short period of time. This cell cycle progression, in which shortly after estrogen exposure quiescent cells are recruited from G_0 to G_1, followed by the initiation of DNA synthesis in S phase at approximately 15 h and the progression to G_2/M beginning at approximately 24 h [7,8], requires the coordinated regulation of expression of specific genes. Some of these specific responses were confirmed from previous studies, while others have not been previously reported in this system but based on their known function can be reconciled with the temporal pattern of gene expression observed.

For example GADD45α, which was highly induced at 2–8 h, is known to be critical for the DNA damage and repair checkpoint prior to S phase and to be induced only during G_1. Cyclin D2, which is also required for the G_1/S transition, was also upregulated during this time, while a number of other transcripts for proteins involved in the start of DNA synthesis in S phase were found in clusters that were upregulated at approximately 8–12 h. Other genes required for S phase, as well as some required for M phase progression, were upregulated later (24 h). No genes associated with cell cycle progression were identified in the cluster that was upregulated at 3×24 h, the time at which maximal growth has been attained.

In some cases, the response obtained did not confirm results reported in the literature, or two different probe sets interrogating different regions of the same transcript yielded highly different results. In these cases, the specific sequences being probed, and the characteristics and performance of the probes composing the probe set were examined in more detail. The Affymetrix database NetAffx (<http://www.affymetrix.com>) was used to obtain specific probe sequence information, and to exam-

ine the number and overlap of the probes within the set, while dChip (<http://biosun1.harvard.edu/complab/dchip>) was used to examine the performance of the different probes within the set. These sources of information, in addition to the suffix flag codes provided with the probe set IDs, was helpful in explaining unexpected or discordant results.

CONCLUSIONS

In the present study, a statistically rigorous method of reducing temporal microarray responses to estrogen in the mouse uterus to the most active and reproducible responses was presented. Automated scripts were used to obtain data from local copies of public database information, both for mouse genes and for the human homologs where available. These scripts were also used to annotate literature search information for estrogen-induced responses in the uterus, and all of the assembled information was used in the subsequent physiological interpretation. This interpretation, which is ongoing, aims to link the observed temporal transcriptional responses to the tissue-level changes that lead to the rapid, coordinated, and extensive uterine proliferation following estrogen exposure. It is anticipated that the large-scale characterization of transcriptional responses to estrogen in the uterus will be valuable in better understanding the physiological effects of estrogen exposure, which in turn can be used to evaluate the presence and toxicological significance of estrogenic effects of exogenous compounds.

ACKNOWLEDGMENTS

Assistance with scripting by Ms. Shraddha Pai and Ms. Raeka Aiyar is gratefully acknowledged. Financial support from ES 011271, as well from fellowships from Michigan State University and the Society of Toxicology, have also been gratefully received.

REFERENCES

1. E. Padilla-Banks, W. N. Jefferson, R. R. Newbold. *Environ. Health Perspect*. **109**, 821–826 (2001).
2. W. N. Jefferson, E. Padilla-Banks, R. R. Newbold. *Reprod. Toxicol*. **14**, 103–110 (2000).
3. D. A. Zajchowski, K. Kauser, D. Zhu, L. Webster, S. Aberle, F. A. White, H. L. Liu, R. Humm, J. MacRobbie, P. Ponte, C. Hegele-Hartung, R. Knauthe, K. H. Fritzemeier, R. Vergona, G. M. Rubanyi. *J. Biol. Chem*. **275**, 15885–15894 (2000).
4. H. K. Hamadeh, P. R. Bushel, S. Jayadev, K. Martin, O. DiSorbo, S. Sieber, L. Bennett, R. Tennant, R. Stoll, J. C. Barrett, K. Blanchard, R. S. Paules, C. A. Afshari. *Toxicol. Sci*. **67**, 219–231 (2002).
5. H. D. Lauson, C. G. Heller, J. B. Golden, E. L. Sevringhaus. *Endocrinology* **24**, 35–44 (1939).
6. B. Efron, R. Tibshirani, J. D. Storey, V. Tusher. *J. Am. Stat. Assoc*. **96**, 1151–1160 (2001).
7. S. V. Cheng, B. S. MacDonald, B. F. Clark, J. W. Pollard. *Exp. Cell Res*. **160**, 459–470 (1985).
8. J. Z. Guo and J. Gorski. *Mol. Endocrinol*. **2**, 693–700 (1988).