THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 50, Issue of December 14, pp. 47195–47201, 2001
Printed in U.S.A.

Gonadotropin-releasing Hormone Receptor-coupled Gene Network Organization*[S]

Received for publication, September 10, 2001, and in revised form, September 28, 2001
Published, JBC Papers in Press, October 1, 2001, DOI 10.1074/jbc.M108716200

Elisa Wurmbach‡§, Tony Yuen‡§, Barbara J. Ebersole‡, and Stuart C. Sealfon‡¶‖

From the ‡Departments of Neurology and the ¶Fishberg Center for Neurobiology, Mount Sinai School of Medicine, New York, New York 10029

An early gene cDNA microarray was developed to study genes that are regulated immediately following gonadotropin-releasing hormone (GnRH) receptor activation. 956 selected candidate genes were printed in triplicate, a *t* statistic-based regulation algorithm was used for data analysis, and the response to GnRH in a time course from 1 to 6 h was determined. Measurements were highly reproducible within arrays, between arrays, and between experiments. Accuracy and algorithm reliability were established by real-time polymerase chain reaction assays of 60 genes. Gene changes ranging from 1.3- to 31-fold on the microarray were confirmed by real-time polymerase chain reaction. Many of the genes were found to be highly regulated. The regulated genes identified were all elevated at 1 h of treatment and returned nearly or completely to baseline levels of expression by 3 h of treatment. This broad, robust, and transient transcriptional response to constant GnRH exposure includes modulators of signal transduction (*e.g.* Rgs2 and IκB), cytoskeletal proteins (*e.g.* γ-actin), and transcription factors (*e.g.* c-Fos, Egr1, and LRG21). The interplay of the activators, repressors, and feedback inhibitors identified embodies a combinatorial code to direct the activity of specific downstream secondary genes.

The mechanisms underlying the specificity of the transcriptional response to the activation of cell surface receptors are not well understood. The pituitary gonadotropin-releasing hormone receptor (GnRHR),[1] which mediates the biosynthesis of the gonadotropin luteinizing hormone and follicle-stimulating hormone, provides a salient example of the exquisite requirements for signaling specificity between the membrane and the genome. The pattern of downstream gene responses depends on the frequency of receptor stimulation. Specific patterns of GnRHR stimulation lead to the generation of distinct transcriptional programs. For example, prolonged GnRH stimulation favors induction of the common α-gonadotropin subunit. In contrast, a specific physiologically relevant frequency range of receptor stimulation on the order of 1 pulse/h preferentially induces the luteinizing hormone beta subunit (LHβ) gene (1, 2). Although downstream signal transduction mediators (such as JNK and ERK) and a number of transcription factors (including Egr1, SF1, and NAB1) have been implicated in modulation of the LHβ promoter (3–6), the available data do not explain why the induction of LHβ requires specific patterns of GnRHR activation.

The activation of specific secondary genes depends on the pattern of induction of primary genes, which can encode proteins involved in signal feedback and in modulating the transcription of downstream targets. To better understand this process, we have developed a microarray to identify the early gene program induced by GnRHR activation.

Microarray techniques have emerged as important approaches for the simultaneous analysis of multiple gene transcripts. These methods have proven valuable in refining cancer classification (7, 8) and for providing qualitative assessment of the global gene programs that accompany cell division, development, and the responses to specific stimuli (9, 10). However, data obtained using both commercial and custom global microarrays have been limited by the expense of the assays and by problems in quality control (11).

To overcome the potential limitations of global microarrays, we have integrated microarray technology with a massively parallel candidate gene approach that we refer to as focused microarray analysis (FMA). This approach has been used to develop an early gene microarray for the study of GnRHR responses. 956 cDNAs were carefully selected for inclusion on this microarray, including many early response genes identified in various experimental systems. The size of this array facilitates high quality array production, validation, and data generation. Several novel aspects of the approaches described could be generally useful. We have established array quality benchmarks and empirical data analysis algorithms that are fully supported by extensive independent gene measurement. We also report the capacity for FMA to quantify the level of gene regulation.

We have utilized FMA to characterize the time course from 1 to 6 h of gene responses occurring following activation of the GnRHR in the mouse LβT2 gonadotrope cell line (12). This study reveals a highly structured response induced in the genetic signaling network.

EXPERIMENTAL PROCEDURES

Cell Culture and Sample Preparation—LβT2 cells were obtained from Pamela Mellon (University of California, San Diego) and were maintained at 37 °C in 5% CO_2 in humidified air in Dulbecco's modified Eagle's medium (Mediatech) containing 10% fetal bovine serum (Gemini). For experiments, 40–50 × 10^6 cells were seeded in 15-cm dishes. The medium was replaced 24 h later with Dulbecco's modified Eagle's

* This work was supported by Grant RO1 DK46943 from the National Institutes of Health and a Howard Hughes Medical Institute award. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains a table of real-time PCR primer pairs and an illustration of test hybridization on the array.

§ These authors contributed equally to this study.

‖ To whom correspondence should be sent: Neurology Box 1137, Mount Sinai School of Medicine, New York, NY 10029; Tel.: 212-241-7075; Fax: 212-289-4107; E-mail: Stuart.Sealfon@mssm.edu.

[1] The abbreviations used are: GnRHR, gonadotropin-releasing hormone receptor; GnRH, gonadotropin releasing hormone; PCR, polymerase chain reaction; LHβ, luteinizing hormone beta subunit; FMA, focused microarray analysis.

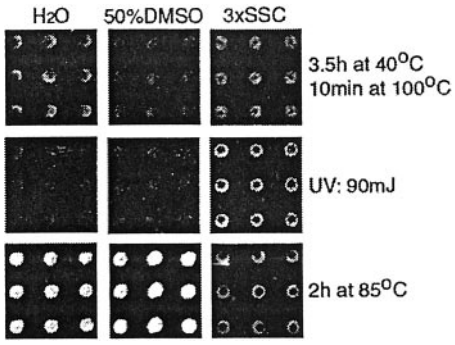

FIG. 1. **Comparison of spotting solutions and fixing protocols.** Test arrays were hybridized with a Cy5-labeled oligonucleotide against one of the PCR primers used for cDNA amplification to optimize the printing solution and attachment protocol. PCR products were suspended in the indicated solutions and were fixed to the slide using either baking (two protocols) or UV cross-linking. 3× SSC solution caused ring artifacts (*right column*). Either two-stage heating or UV cross-linking caused unreliable attachment (*top two rows*). The best signal and morphology were obtained with 50% Me_2SO (*DMSO*) solution and 2 h of baking at 85 °C (*center bottom*).

medium containing 25 mM HEPES (Mediatech) and glutamine. 18 h later, the cells were treated with 100 nM GnRH or vehicle and were returned to the CO_2 incubator for 1, 3, or 6 h. The incubation was stopped by aspirating the incubation medium and adding 10 ml of lysis buffer (4 M guanidinium thiocyanate, 25 mM sodium citrate, pH 7.0, 0.5% *N*-laurylsarcosine, and 0.1 M 2-mercaptoethanol). Total RNA was isolated according to the method of Chomczynski and Sacchi (13). Approximately 400 μg of total RNA was obtained from each plate.

Microarray Development and Quality Control—Plasmids were purified using the Qiaprep 96 Turbo Miniprep kit. Following insert amplification by PCR, products were confirmed by agarose gel electrophoresis and purified with Qiaquick 96 kit (Qiagen). The product was dried, dissolved in either 18 μl of H_2O, 50% Me_2SO, or 3× SSC, and spotted (3 hits/feature) with a GMS 417 Arrayer (Affymetrix) on CMT-GAPS-coated glass slides (Corning). DNA was fixed either by incubating the slide for 3.5 h at 40 °C followed by 10 min at 100 °C or for 2 h at 85 °C or by UV cross-linking with 90 mJ (Stratalinker, Stratagene).

The suspension solutions and fixing protocols were compared in a 3×3 design because feature morphology may be influenced by surface tension effects arising from surface and spotting solution chemistry and by attachment efficiency. We evaluated feature morphology and attachment by hybridizing a small test array with a Cy5-labeled oligonucleotide complementary to one of the amplification primers (5′-CGT TTT ACA ACG TCG TGA CTG GG-3′). The optimum morphology and intensity were reproducibly observed, independently of insert sequence or size, with the PCR product dissolved in 50% Me_2SO and fixed for 2 h at 85 °C to CMT-GAPS-coated slides, which was used for subsequent array production (Fig. 1).

The 956 clones were selected from a NIA 15K library (14) or purchased from Research Genetics. The gene selection was based on literature searches, input from collaborating laboratories, and unpublished data. A large number of genes known to be induced at early time points in other experimental systems were included as well as 67 putative "housekeeping" genes. The quality of the libraries utilized and the reliability of the clone picking and isolation were evaluated by sequencing 119 clones on the array, including randomly selected clones on each 96-well plate and all genes that were subsequently identified as regulated. 92% of the clones picked from the NIA library and 82% of the clones purchased from Research Genetics had been correctly identified. Based on the distribution of clone sources, we estimate that clone identification of the unsequenced clones on the array was 91% accurate. 98.7% of the clones generated sufficient PCR product to be detectable by gel electrophoresis.

In order to distinguish signals arising from surface artifacts and facilitate analysis, the selected genes were spotted in triplicate. Arrays were stored light-protected at room temperature until use. A test hybridization using the Cy5-labeled oligomer confirmed the presence of all spotted genes at similar concentrations (see supplementary data, Fig. 7).

RNA Labeling and Hybridization—20 μg of total RNA from each sample was labeled with either Cy3 or Cy5 using the Atlas indirect labeling kit (CLONTECH) as indicated by the manufacturer. Following array prehybridization in 6× SSC, 0.5% SDS, and 1% bovine serum albumin at 42 °C for 45 min, the probe was denatured and hybridized in 24 μl of 50% formamid, 6× SSC, 0.5% SDS, and 5× Denhardt's with 2.4 μg of salmon sperm DNA and 10 μg of poly(dA) at 42 °C (room temperature for test oligomer hybridization) for 16 h. Following 10-min washes in 0.1× SSC, 0.1% SDS, and twice in 0.1× SSC, the slide was scanned using the GMS 418 Scanner (Affymetrix).

Real-time PCR—We used a previously described protocol (15). Briefly, 5 μg of total RNA was converted into cDNA, and 1:400 (~250 pg) was utilized for 40-cycle three-step PCR in an ABI Prism 7700 in 20 mM Tris, pH 8.4, 50 mM KCl, 3 mM $MgCl_2$, 200 μM dNTPs, 0.5× SYBR green (Molecular Probes), 200 nM each primer, and 0.25 units of Platinum *Taq* (Life Technologies). Amplicon size and reaction specificity were confirmed by agarose gel electrophoresis. The number of target copies in each sample was interpolated from its detection threshold (C_T) value using a plasmid or purified PCR product standard curve included on each plate. The sequence of the 60 primer sets utilized can be found in the supplementary material. Each transcript in each sample was assayed five times, and the median C_T values were used for analysis.

Data Analysis—Scanned microarray data were exported as TIFF files to Genepix (Axon Instruments), and spot registration was manually optimized as suggested by the developer. The median background-subtracted feature intensity was utilized for further analysis. Overall differences in the signal intensity of the two wavelengths measured on each slide (λ=532 nm and λ=635 nm) were corrected using the locally linear robust scatter plot smoother implemented in the loess function in S Plus Professional (Insightful Corporation). Predictors were generated using a symmetric distribution, span = 0.75 (16).

The data quality from each array was estimated by determining the median coefficient of variation within each array for each gene, which indicates the overall variation of triplicate measurements on a given array,

$$cv = \frac{100}{\bar{R}}\sqrt{\frac{\sum_{i=1}^{n}(R_i - \bar{R})^2}{n-1}} \qquad \text{(Eq. 1)}$$

where $\bar{R}$ is the geometric mean ratio for that gene on each slide, R_i is each ratio, and n is the number of measurements.

The sources of variation in these experiments were explored using analysis of variance for each gene measured (17). Variance within and between experiments were determined.

Variance within,

$$s_R^2 = \sum_{t=1}^{k}\sum_{i=1}^{n_t}\frac{(y_{ti} - \bar{y}_t)^2}{N-k} \qquad \text{(Eq. 2)}$$

where N is the total number of measurements of each gene, k is the number of arrays, n_t is the number of measurements of each gene on each array, y_{ti} is the individual ratio, and $\bar{y}_t$ is the average ratio for that gene on each array.

Variance between,

$$s_T^2 = \frac{\sum_{t=1}^{k} n_t(\bar{y}_t - \bar{y})^2}{k-1} \qquad \text{(Eq. 3)}$$

where k is the number of arrays, n_t is the number of measurements, and $\bar{y}$ is the average ratio for that gene from all experiments.

In order to select genes for further study, t values for the log transform ratios (r) were determined for data from each slide.

$$t = \frac{\bar{r}\sqrt{n}}{s}, \quad \text{where} \quad s = \sqrt{\frac{\sum_{i=1}^{n}(r_i - \bar{r})^2}{n-1}} \qquad \text{(Eq. 4)}$$

Genes were selected according to the following algorithm: (i) ± fold-change of >1.3, (ii) $|t|$ >3, (iii) signal intensity >1% of the median signal intensity value in at least one channel, and (iv) criteria i–iii observed in at least two of three experiments.

RESULTS

Microarray Variation Analysis—The early transcriptional response to GnRH was studied in the GnRHR-expressing LβT2

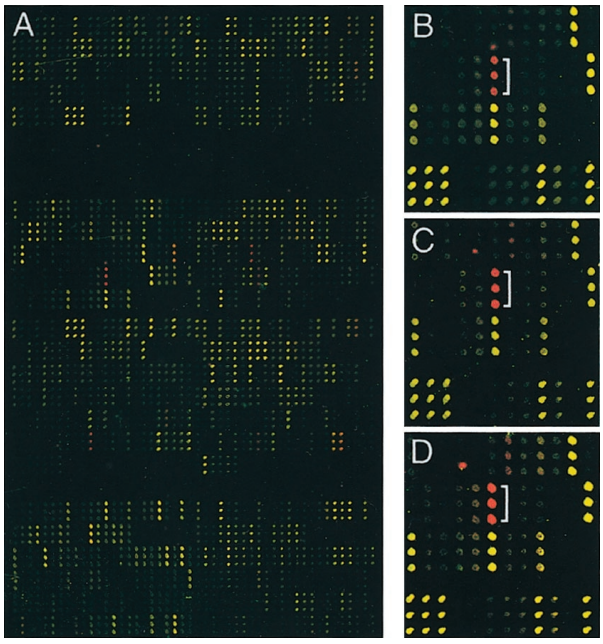

FIG. 2. **cDNA array hybridization from GnRH- (*red*) and vehicle-treated (*green*) LβT2 cells.** *A*, entire array from one experiment. Note the low background and consistency of triplicate measurements, which are present in adjacent rows. The scatter plots obtained from this and two other independent experiments are shown in Fig. 3. *B–D*, comparison of a segment of arrays taken from three independent experiments. The *brackets* indicate the triplicate features representing up-regulated *Egr1* transcript in each experiment.

gonadotrope cell line (12). Three separate experiments comparing the effects of a 1-h exposure to GnRH or vehicle were performed. Hybridization was uniform across all arrays, and background signal was low (Fig. 2). Data reproducibility and labeling bias were assessed by repeating the labeling and hybridization for one pair of samples two additional times, once with the control (Cy3) and treated (Cy5) labels reversed.

The reliability of the data generated from these five microarray experiments was substantiated by the reproducibility of measurements within arrays, between arrays, and between experiments. The pattern of hybridization observed was qualitatively similar for all experiments (Fig. 2). Because the triplicate measurements within each array were nearly identical for each of the five hybridizations (median cv, 8–13%), regulated genes could reliably be distinguished from surface artifacts (Fig. 2*D*). The reproducibility of the three measurements of each gene on each slide was utilized in our data analysis algorithm (see below). When the identical two RNA samples (control and treated) were labeled repeatedly, a high linear correlation was obtained for the ratios of the regulated genes (r=0.985). Similar analysis of a dye swap experiment also showed a very high correlation, r = 0.975. These data indicate that the relative expression obtained from these comparisons are reproducible and do not show sequence bias, as has been reported with direct labeling protocols (18).

To determine the relative sources of variability in the microarray experiments, an analysis of variance was performed on the data from these five hybridizations. There are several sources of variation in these experiments: (i) the variation between the triplicate measurements for each gene on each array, (ii) the variation between the measurement of each gene obtained by repeated labeling and hybridization, and (iii) the variation between the levels of each mRNA that occur in independent, replicated experiments. The median variance for the triplicate measurements within each array, which represents the variation of the three repeated measurements, was s_R^2 = 0.0198 for all arrays, indicating a very low level of variability. The two potential sources of variability between replicated experiments can be distinguished by comparing the analysis of replicate hybridizations of the same samples (n=3) and of repeated, independent experiments (n=3). The median sample variances between the ratios obtained with repeated labeling/hybridization of the same samples and between RNA samples obtained from independent experiments were s_T^2 = 0.073 and s_T^2 = 0.071, respectively. These data reveal similar levels of variation between the measurements obtained from repeated assays of the same samples and of samples from independent experiments. Thus the level of variation between the regulation of mRNAs in replicate experiments was so low that it does not add appreciably to the modest measurement variation resulting from the hybridization procedure.

Scatter graphs of data from the three independent stimulation experiments, normalized using the robust locally linear loess function, are shown in Fig. 3. We preferred this normalization to an overall linear correction because it compensates for variation of the correction factor with signal intensity and is largely unaffected by outliers, which include the regulated transcripts. Most transcripts are not regulated, and the normalized data are tightly grouped along y = x. The triplets corresponding to several regulated genes are evident and show similar regulation and scatter graph location in all three independent experiments. These strikingly similar results obtained in independent experiments reflect a high level of reliability and reproducibility in all aspects of these studies, including array production, cell culture and treatment, RNA extraction, labeling, hybridization, and data acquisition.

Microarray Identification of Genes Regulated by GnRH—In order to identify candidate genes whose regulation is less marked than the visually obvious triplets indicated in Fig. 3, an empirical selection algorithm was implemented. The presence of triplicate features for each cDNA allowed calculation of a t statistic for each gene on each array (see "Experimental Procedures"). The criteria utilized to select the regulated gene candidates were based on fold-change ($>\pm1.3$), the t statistic ($|t|>3$), and signal intensity (>1% of median signal intensity in at least one channel) within each experiment (see "Experimental Procedures"). Genes that met these criteria in at least two independent experiments were selected for further study. This selection strategy identified 31 candidate regulated genes, 28 of which increased and 3 of which decreased (see below). Four of these genes (*Rgs2*, *TSC22*, *PRL1*, and *Nrf2*) were represented by two different clones and spotted in different locations on the arrays. In all four cases, regulation was detected by our criteria, and the degree of change was similar (*Rgs2*, 4.0 ± 1.3, 2.9 ± 1.2; *TSC22*, 3.4 ± 0.4, 3.3 ± 0.6; *PRL1*, 2.0 ± 0.4, 1.5 ± 0.4; *Nrf2*, 1.3 ± 0.2, 1.3 ± 0.1). These data indicate that the observed changes were independent of position on the array.

Confirmation of Gene Regulation by Real-time PCR—The measurement accuracy of the microarray and the biological variability of the transcriptional program identified were evaluated through a validation study comprising real-time PCR assays of 60 transcripts, including 26 genes that met the threshold for regulation in the microarray assay. To generate accurate and precise reference measurements of these genes, five measurements were made of each gene in each RNA sample, and the measurements were calibrated with a standard curve included on each PCR plate (Fig. 4). Nine experimental pairs of vehicle and GnRH-treated cultures were tested, including the microarray assayed samples.

100% (17 of 17) of the gene changes showing >1.6-fold change on the microarray and 66% (6 of 9) of gene changes showing between 1.3- and 1.6-fold changes were confirmed by

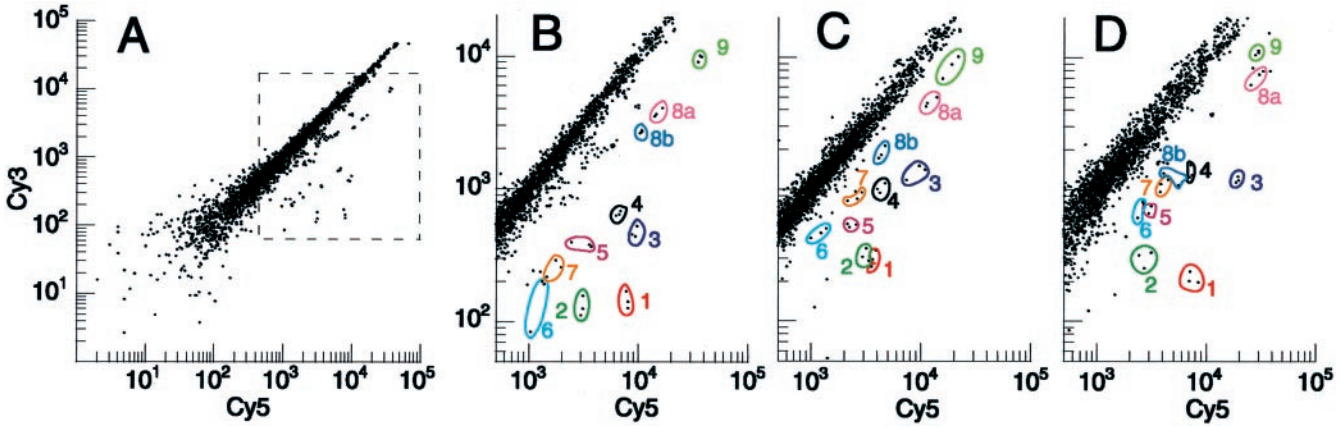

FIG. 3. **Scatter plots of independent replicate experiments.** Three experiments were performed, each of which generated samples from cells exposed to vehicle or GnRH for 1 h. Data shown were normalized using a robust locally linear loess algorithm (see "Experimental Procedures"). *A*, all data (background-subtracted signal intensity) from one experiment. *B*, *boxed area* enlarged from *panel A*. *C* and *D*, enlarged scatter plot of the corresponding region from the other two independent activation experiments. Triplicate spots representing individual up-regulated genes are identified and *encircled*. Note the consistent pattern of regulation observed in the three array experiments. *Encircled* genes: *1*) *LRG21*, *2*) *c-fos*, *3*) *Egr1*, *4*) *nur77*, *5*) *Pip92*, *6*) *Rgs2*, *7*) *c-jun*, *8a* and *8b*) two different clones for *TSC22*, and *9*) *γ-actin*.

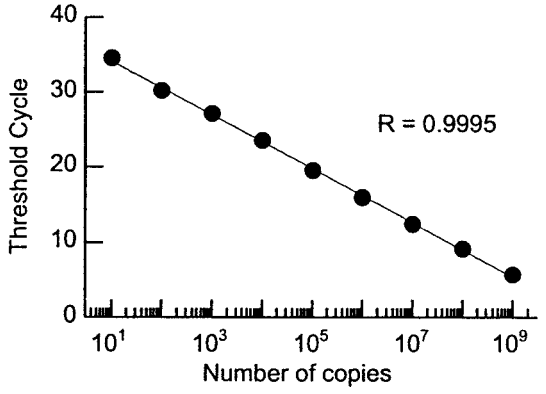

FIG. 4. **Standard curve for real-time PCR.** The standard curve for Egr1 mRNA is shown. Note the broad linear range of the assay.

TABLE I
Fold-changes ± S.E. from the early gene microarray compared to real-time PCR

Values meeting the criteria for up- or down-regulation are indicated by red or green, respectively. The identification of all clones listed was confirmed by sequencing.

Gene Name	Accession Number	Microarray (n=3)	SYBR PCR (n=9)	Microarray (n=1)	
		1h	1h	3h	6h
LRG 21	U19118	31.2 ± 16.6	96.8 ± 15.7	1.3	0.8
Egr1	NM_007913	12.9 ± 3.8	390.6 ± 83.9	1.2	1.2
c-fos	J00370	12.5 ± 4.0	52.0 ± 5.8	1.2	0.6
Nr4a1/nur77	AI322974	5.9 ± 1.5	49.4 ± 10.1	1.0	1.0
Ier2/Pip92	W14782	5.2 ± 1.2	17.1 ± 2.4	0.9	1.0
Rgs2	NM_009061	4.0 ± 1.3	8.7 ± 0.4	1.8	1.2
c-jun	NM_010591	3.8 ± 0.6	6.6 ± 0.9	1.1	1.0
TSC22	NM_009366	3.4 ± 0.4	3.3 ± 0.2	1.3	1.0
γ-actin	L21996	2.7 ± 0.3	3.9 ± 0.4	2.2	1.0
Klf-like EST	BE368139	2.3 ± 0.3	2.5 ± 0.2	0.7	0.6
Period1	AB030818	2.1 ± 0.1	4.9 ± 1.1	1.0	1.0
β-actin	NM_007393	2.0 ± 0.5	1.6 ± 0.1	1.6	1.3
PRL1	NM_011200	2.0 ± 0.4	2.1 ± 0.1	1.7	1.2
IκB	NM_010907	1.7 ± 0.2	2.6 ± 0.2	0.9	0.8
Klf4	NM_010637	1.7 ± 0.3	2.9 ± 0.3	1.1	1.3
Gem	AA177829	1.7 ± 0.2	4.3 ± 0.3	1.2	1.1
gly96	X67644	1.7 ± 0.6	3.8 ± 0.5	1.5	0.8
junD	W12943	1.6 ± 0.2	1.2 ± 0.2	1.0	1.3
Egr2	AA727313	1.5 ± 0.2	294.7 ± 35.6	1.0	1.0
transgelin	AF149291	1.4 ± 0.1	1.8 ± 0.2	1.3	2.4
NMDMC	NM_008638	1.4 ± 0.3	1.6 ± 0.3	1.1	1.0
Stat3B	U30709	1.4 ± 0.2	1.2 ± 0.1	1.0	0.8
MKP1/3CH134	W34966	1.4 ± 0.1	3.4 ± 0.4	0.9	1.2
Nrf2	U20532	1.3 ± 0.2	1.5 ± 0.1	1.2	0.9
HSP30	NM_019979	1.3 ± 0.2	1.9 ± 0.2	1.7	1.3
STY-Kinase	M38381	1.29 ± 0.04	1.7 ± 0.2	0.7	0.7
glucose transport protein	M22998	1.1 ± 0.4	2.2 ± 0.4	1.3	1.3
SCL	AJ297131	1 ± 0.2	2.4 ± 0.2	1.2	1.0
Gata2	NM_008090	0.7 ± 0.1	1 ± 0.1	1.1	1.0

PCR. All of the confirmed genes were up-regulated (Table I). Three genes, which showed a low level of regulation on the microarray, had less than 1.3-fold regulation (below the threshold for regulation) with the real-time PCR, but the actual fold-change measurements were actually relatively close (Table I). Three of 32 genes that appeared to be unregulated on the microarray were found to be up-regulated by real-time PCR (Table I). Thus the algorithm utilized to identify regulated genes was able to correctly identify regulated transcripts showing changes as low as 1.3-fold regulation on the microarray.

This data set of 5400 PCR assays provides a reference standard that allows assessment of the accuracy of the microarray and reliable quantification of the degree of regulation of these transcripts following GnRH exposure. The accuracy of the fold-change determinations obtained from the microarray was poor for transcripts showing high degrees of regulation (Table I), but the fold-change was meaningful for transcripts showing less extreme levels of induction. The power function correlation for the fold-change measurements by microarray and real-time PCR for genes showing <20-fold regulation by PCR ($n=24$) was $r = 0.87$ (Table I and Fig. 5). These data indicate that the microarray measurements can be calibrated to provide a quantitative estimate of the degree of gene regulation.

Organization of the Early Induced Gene Network—Our data reveal that GnRHR activation modulates the expression of a large number of genes. The regulated genes identified include transcription factors (*e.g. Klf4*, *Egr1*, and *Egr2*), cell signaling modulators (*e.g. Rgs2* and *IκB*), channel regulators (*gem*), and proteins contributing to cytoskeletal dynamics (γ-actin and transgelin). More than half of the induced genes are transcription factors, including both activators and repressors, with the major structural motifs encoded being leucine zipper factors (*c-jun*, *Nrf2*, *LRG21*, and *TSC22*) and zinc finger proteins (*Egr1*, *Egr2*, *Klf4*, *Klf-like EST*, and *Nr4a1*) (Table I and Fig. 6). Many immediate early genes are known to be only transiently induced by various stimuli. Analysis of samples obtained from cells exposed to GnRH for 3 and 6 h reveals that nearly all of the induced genes return to baseline levels of expression by 3 h (Table I). A commonality of many induced transcripts is that the proteins encoded, after synthesis, would contribute to subsequent down-regulation of receptor-activated signaling. This

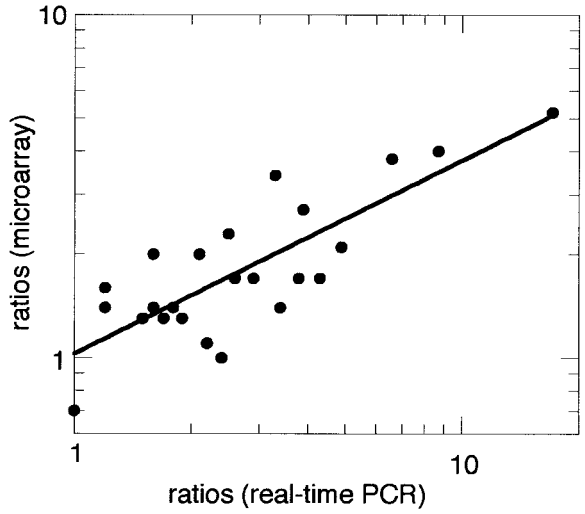

FIG. 5. **Correlation of fold-change measurements obtained by FMA and real-time PCR.** The fold-changes obtained with both techniques for genes showing less than 20 fold-changes by real-time PCR are plotted. The fold-change measurements were highly correlated. The *line* indicates a best-fit power function (r=0.87).

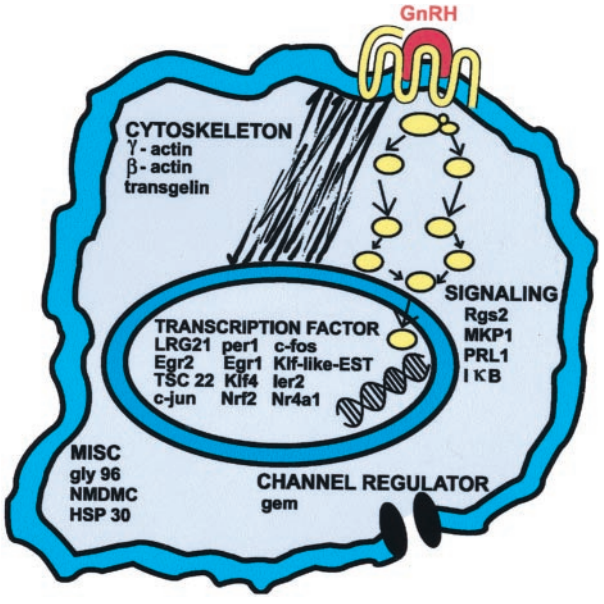

FIG. 6. **Schematic of transcripts regulated by GnRH in LβT2 cells.** All real-time PCR-confirmed microarray-identified genes regulated by GnRH (a 1-h exposure) in LβT2 cells are shown.

category includes Fos, Rgs2, IκB, MKP1, PRL1, and gem (see "Discussion").

DISCUSSION

We have identified a broad, robust, and transient transcriptional response to the activation of the GnRH receptor. The proteins encoded by the induced genes represent five categories: transcription factors, cytoskeletal proteins, signaling mediators, channel regulators, and proteins of miscellaneous or unknown function (Fig. 6). Several of the genes that we have newly discovered to be regulated by GnRH have obvious implications for the function of the GnRH signaling system. For example, Klf4 is a zinc-finger protein in the Sp/XKLF family (19). Sp1 sites in the LHβ promoter have been implicated in the response to pulsatile GnRH (20). Therefore, our data raise the possibility that Klf4 may bind to the LHβ promoter and contribute to its GnRH-mediated activation.

Notably, many induced transcripts can be recognized as encoding proteins that would down-regulate or suppress GnRHR-activated signaling. Fos is known to transrepress the induction of many genes (21). Rgs2 is a GTPase-activating protein that down-regulates the activity of $G\alpha_q$ (22), the major G-protein utilized by the GnRH receptor (23). The small G-protein gem (kir) has recently been reported to down-regulate L-type calcium channel function (24). The induction of gem by GnRH could contribute to the modulation of GnRH-induced gonadotropin release and MAP kinase activation, which both depend on L-channel activation (25). IκB inhibits NFκB signaling. MKP1 and PRL1 are both phosphatases that could attenuate GnRHR signaling mediated by kinase activation.

The pattern of gene induction observed reveals that the first wave of genes induced by GnRH represents a seamless transition between the cellular signaling network and the downstream gene responses representing the ultimate targets of this signaling. The components of this gene network we have identified that (after synthesis) can suppress more proximal GnRH signaling intermediates are likely to play a significant role in modulating the responses to receptor activation. For example, our data reveal the rapid induction of Rgs2 by GnRH, which would likely attenuate GnRH receptor signaling. We speculate that a sinusoidal induction of Rgs2 by pulsatile receptor stimulation could contribute to the well known frequency dependence of downstream secondary genes.

The activation or suppression of specific secondary gene targets is dictated by a combinatorial code. Induction of a specific secondary gene requires the presence of a particular combination of induced and constitutive factors and a relative absence of repressors for that promoter. Because the genetic network activates transcripts that encode proteins for activators and repressors as well as feedback inhibitors to signaling, it is ideally structured to generate different patterns of co-expressed activating and repressing factors in response to different stimuli. This formulation is consonant with the known control of secondary targets, such as the LHβ gene. The LHβ promoter binds transcription factors relatively weakly and requires the presence of multiple distinct activating proteins for efficient transcription (3–6, 20, 26, 27). Furthermore, this promoter is preferentially activated at specific frequencies of GnRHR stimulation. The composition of the proximal gene network that we have found to be activated by GnRH has activating and suppressing components that may contribute to frequency-dependent gene responses.

The extensive validation studies of the targeted microarray approach we have pursued indicates a high level of sensitivity and accuracy in identifying regulated transcripts. The scale of genome-wide microarrays causes several problems. One is the difficulty of quality control for both academic and commercial suppliers. Another is the increase in statistical uncertainty due to multiple hypothesis testing. For an experiment utilizing a fixed number of arrays, the statistical power of correctly assigning a gene as regulated or unregulated decreases as the size of the array increases. However, the high expense of global arrays constrains the number of arrays that should be analyzed to provide statistically acceptable sensitivity and specificity. Therefore, as an alternative, we have explored the development and use of FMA in which the specific genes are repeatedly printed on the array. Although FMA provides less broad transcriptome coverage than a global array, this limitation can be partially overcome by the careful selection of which clones are represented (see "Experimental Procedures"). FMA facilitates the generation of high quality experiments and is sensitive to small regulatory changes that can be confirmed by independent measurements.

Although microarrays have become widely used, generally accepted standards for data reproducibility and reliability have

not been established. The controllable sources of error in microarray experiments include the biological procedure, RNA isolation, labeling method, microarray fabrication, hybridization technique, and analysis algorithm. In developing the early gene array, we relied on several aspects of our experimental design and analysis to confirm data quality. Printing and attachment conditions were optimized. All genes were spotted in triplicate, which facilitates the identification of regulated genes and the exclusion of artifacts and allows statistics to be calculated for each array. cDNA identification was more than 90% accurate, and 100% of amplified clones were detectable by test hybridization on the array (see supplementary data). These data indicate a high level of microarray fabrication accuracy (see Refs. 11 and 28). The data obtained with indirect probe labeling were nearly identical in a dye swap experiment. As direct labeling has been reported to show sequence-specific biases (18), our data suggest that indirect labeling may be preferable to direct labeling protocols. Our hybridizations were uniform (Fig. 2), and we found a low median coefficient of variation for all triplicate assays within each slide. Regulated genes for which two clones were spotted on the array showed nearly identical changes, indicating that the changes observed were independent of array position. Our measurements were reproducible within slides, between slides, and between experiments. Using a cell line system, the biological variability between samples is low and does not appreciably increase the variance observed between separate arrays. Studies performed with more complex sources, such as mixed cultures or tissues, lead to scatter plots with a broader distribution around the non-regulated line in comparison with Fig. 3 and might therefore increase the degree of variability.[2]

The scatter plots of the array data showed a dense linear alignment of those transcripts that were not significantly regulated. The changes in expression observed were demonstrated on separate arrays representing independent experiments (Fig. 3). The early gene transcripts known to be regulated by GnRH in gonadotropes, including *Egr1* (3), *c-fos* (29, 30), and *c-jun* (29), were correctly identified in our microarray data.

The ultimate assessment of the experimental procedures and microarray data was provided by independent determination of the levels of 60 transcripts by real-time PCR. These assays confirmed the regulation of nearly all transcripts that met our criteria for significant regulation on the array (Table I). Three genes that did not meet our criteria on the microarray were found to be regulated when tested by real-time PCR. On the other hand, three genes that were found to be slightly regulated on the microarray did not reach the same 1.3-fold change criterion when assayed by real-time PCR (Table I). To summarize, the microarray analysis was able to successfully identify several truly but only slightly regulated genes down to 1.3-fold regulation (*e.g. Nrf2*, *MKP1*, and *transgelin*) at the cost of very few false positives and correctly identified all genes with a fold-change greater than 1.6.

The sensitivity of the microarray to regulation of a particular transcript may be influenced by nonspecific hybridization of the target sequence, by target location within the gene, and by the level of the transcript. These factors may have contributed to the gene least accurately assayed by the microarray, *Egr2*. The number of *Egr2* transcripts determined by real-time PCR in control cells was among the lowest of any genes showing regulation. The *Egr2* transcript levels were ~150 and 45,000 copies/12.5 ng of total RNA in samples from control and GnRH-treated cells, respectively. In comparison, the *Egr1* levels were 3000 and 1,500,000 copies/12.5 ng of total RNA. As would be expected from these PCR results, the intensities of the features corresponding to *Egr2* on the microarrays were extremely low (data not shown). Thus the *Egr2* measurements are inaccurate because they are so close to the limits of detection of the array. Despite this measurement inaccuracy, the microarray analysis algorithm was nonetheless able to correctly identify *Egr2* as a regulated transcript (Table I). Although the degree of regulation of *Egr1* is underestimated on the microarray, this transcript is identified as being highly regulated. The microarray is a solid-substrate hybridization, and the accuracy of fold-change measurement decreases as the level of regulation increases. This relationship is predictable for regulatory changes within a large dynamic range and can serve as the basis for developing a calibration function.[3]

With the exception of genes that are very highly regulated (higher than 20-fold by real-time PCR), the fold-change measurements obtained by the microarray are closely correlated with the real-time PCR results (r=0.87). Within this dynamic range, FMA provides a reasonable estimate of both the identification and magnitude of gene induction. As public data bases are developed to archive the results of microarray experiments, the refinement of these techniques to reduce measurement error and to facilitate assessment of the reliability of the measurement of each RNA transcript is important.

It has recently been proposed that signal transduction pathways may form complex networks that manifest emergent properties whose overall patterns of activity are relatively independent of the behavior of specific components (31). Receptor activation leads to changes in multiple, interconnected signal transduction pathways. An attempt to understand the basis for signaling specificity would therefore benefit from a parallel assessment of signaling responses. Although measuring multiple signaling pathways in response to receptor activation can be highly informative (*e.g.* see Refs. 32 and 33), many signaling assays are cumbersome and not easily quantifiable. We propose that quantification of the cell's transcriptional activity may provide an indirect reflection of the receptor-mediated changes in the cell's various signal transduction pathways. This use of parallel gene measurements to monitor cell signaling can be understood by analogy with gene reporter constructs, which are widely used to monitor the activation of G-protein-coupled receptors (*e.g.* see Ref. 34). The responses of these gene reporters depend on the identity of the promoter used. Similarly, the cell's own genes can be considered to be intrinsic gene reporters of cellular signaling with the various gene promoters showing differential sensitivity to activity in different cell signaling pathways. Quantifying transcriptional activity provides an approach to monitor the coordinated changes in the activity of multiple signal transduction pathways and may reveal differences in signaling not apparent with conventional signaling assays.

Acknowledgments—We thank Sanjay Tyagi for help in developing the real-time PCR assays, Pamela Mellon for providing the LβT2 cells, and Bob Pfeffer, Celia Sealfon, David Krakauer, and Andreas Jenny for critical reading of the manuscript. We thank Marianna Max, Charles Mobbs, Jim Wetmur, Carter Bancroft, Katie Clelland, Michael Bradbury, and Fabien Campagne for helping to select the clones for the array, Irina Ivanova for technical assistance, and Jaime Pei for graphic arts.

REFERENCES

1. Dalkin, A. C., Haisenleder, D. J., Ortolano, G. A., Ellis, T. R., and Marshall, J. C. (1989) *Endocrinology* **125,** 917–924
2. Weiss, J., Jameson, J. L., Burrin, J. M., and Crowley, W. F. (1990) *Mol. Endocrinol.* **4,** 557–564
3. Tremblay, J. J., and Drouin, J. (1999) *Mol. Cell. Biol.* **19,** 2567–2576

[2] E. Wurmbach, R. L. Pfeffer, and S. C. Sealfon, unpublished results.

[3] T. Yuen, E. Wurmbach, B. J. Ebersole, and S. C. Sealfon, manuscript in preparation.

4. Dorn, C., Ou, Q., Svaren, J., Crawford, P. A., and Sadovsky, Y. (1999) *J. Biol. Chem.* **274,** 13870–13876
5. Kaiser, U. B., Halvorson, L. M., and Chen, M. T. (2000) *Mol. Endocrinol.* **14,** 1235–1245
6. Sevetson, B. R., Svaren, J., and Milbrandt, J. (2000) *J. Biol. Chem.* **275,** 9749–9757
7. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. (1999) *Science* **286,** 531–537
8. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., Levy, R., Wilson, W., Grek, M. R., Byrd, J. C., Botstein, D., Brown, P. O., and Staudt, L. M. (2000) *Nature* **403,** 503–511
9. Lashkari, D. A., DeRisi, J. L., McCusker, J. H., Namath, A. F., Gentile, C., Hwang, S. Y., Brown, P. O., and Davis, R. W. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 13057–13062
10. Iyer, V. R., Eisen, M. B., Ross, D. T., Schuler, G., Moore, T., Lee, J. C. F., Trent, J. M., Staudt, L. M., Hudson, J., Jr., Boguski, M. S., Lashkari, D., Shalon, D., Botstein, D., and Brown, P. O. (1999) *Science* **283,** 83–87
11. Knight, J. (2001) *Nature* **410,** 860–861
12. Turgeon, J. L., Kimura, Y., Waring, D. W., and Mellon, P. L. (1996) *Mol. Endocrinol.* **10,** 439–450
13. Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159
14. Tanaka, T. S., Jaradat, S. A., Lim, M. K., Kargul, G. J., Wang, X., Grahovac, M. J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H., Wood, W. H., III, Becker, K. G., and Ko, M. S. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9127–9132
15. Yuen, T., Ebersole, B. J., Zhang, W., and Sealfon, S. C. (2001) *Methods Enzymol.* **345,** 556–569
16. Venables, W. N., and Ripley, B. D. (1999) *Modern Applied Statistics with S-PLUS,* 3rd Ed., Statistics and computing, Springer-Verlag, New York
17. Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978) *Statistics for Experimenters: an Introduction to Design, Data Analysis, and Model Building*, pp. 165–207, John Wiley & Sons, Inc., New York
18. Taniguchi, M., Miura, K., Iwao, H., and Yamanaka, S. (2001) *Genomics* **71,** 34–39
19. Philipsen, S., and Suske, G. (1999) *Nucleic Acids Res.* **27,** 2991–3000
20. Weck, J., Anderson, A. C., Jenkins, S., Fallest, P. C., and Shupnik, M. A. (2000) *Mol. Endocrinol.* **14,** 472–485
21. Ofir, R., Dwarki, V. J., Rashid, D., and Verma, I. M. (1990) *Nature* **348,** 80–82
22. De Vries, L., and Gist Farquhar, M. (1999) *Trends Cell Biol.* **9,** 138–144
23. Grosse, R., Schmid, A., Schoneberg, T., Herrlich, A., Muhn, P., Schultz, G., and Gudermann, T. (2000) *J. Biol. Chem.* **275,** 9193–9200
24. Beguin, P., Nagashima, K., Gonoi, T., Shibasaki, T., Takahashi, K., Kashima, Y., Ozaki, N., Geering, K., Iwanaga, T., and Seino, S. (2001) *Nature* **411,** 701–706
25. Mulvaney, J. M., Zhang, T., Fewtrell, C., and Roberson, M. S. (1999) *J. Biol. Chem.* **274,** 29796–29804
26. Yokoi, T., Ohmichi, M., Tasaka, K., Kimura, A., Kanda, Y., Hayakawa, J., Tahara, M., Hisamoto, K., Kurachi, H., and Murata, Y. (2000) *J. Biol. Chem.* **275,** 21639–21647
27. Quirk, C. C., Lozada, K. L., Keri, R. A., and Nilson, J. H. (2001) *Mol. Endocrinol.* **15,** 734–746
28. Taylor, E., Cogdell, D., Coombes, K., Hu, L., Ramdas, L., Tabor, A., Hamilton, S., and Zhang, W. (2001) *BioTechniques* **31,** 62–65
29. Cheng, K. W., Ngan, E. S., Kang, S. K., Chow, B. K., and Leung, P. C. (2000) *Endocrinology* **141,** 3611–3622
30. Abbas, M. M., and Evans, J. J. (2000) *Neuroendocrinology* **71,** 292–300
31. Bhalla, U. S., and Iyengar, R. (1999) *Science* **283,** 381–387
32. Mitchell, R., McCulloch, D., Lutz, E., Johnson, M., MacKenzie, C., Fennell, M., Fink, G., Zhou, W., and Sealfon, S. C. (1998) *Nature* **392,** 411–414
33. Berg, K. A., Maayani, S., Goldfarb, J., Scaramellini, C., Leff, P., and Clarke, W. P. (1998) *Mol. Pharmacol.* **54,** 94–104
34. Klein, C., Gurvich, N., Sena-Esteves, M., Bressman, S., Brin, M. F., Ebersole, B. J., Fink, S., Forsgren, L., Friedman, J., Grimes, D., Holmgren, G., Kyllerman, M., Lang, A. E., de Leon, D., Leung, J., Prioleau, C., Raymond, D., Sanner, G., Saunders-Pullman, R., Vieregge, P., Wahlstrom, J., Breakefield, X. O., Kramer, P. L., Ozelius, L. J., and Sealfon, S. C. (2000) *Ann. Neurol.* **47,** 369–373