BIOLOGY OF REPRODUCTION **73,** 404–413 (2005)
Published online before print 4 May 2005.
DOI 10.1095/biolreprod.105.039719

The Mouse Epididymal Transcriptome: Transcriptional Profiling of Segmental Gene Expression in the Epididymis[1]

Daniel S. Johnston,[2,3,4] Scott A. Jelinsky,[3,5] Hyun J. Bang,[6] Paul DiCandeloro,[4] Ewa Wilson,[5] Gregory S. Kopf,[4] and Terry T. Turner[6,7]

Contraception,[4] Woman's Health & Bone, Wyeth Research, Collegeville, Pennsylvania 19426
Biological Technologies,[5] Molecular Profiling and Biomarker Discovery, Wyeth Research, Cambridge, Massachusetts 02140
Departments of Urology[6] and Cell Biology,[7] University of Virginia Health Science System, Charlottesville, Virginia 22908

ABSTRACT

Maturation of spermatozoa, including the acquisition of motility and the ability to undergo capacitation, occurs during transit through the dynamic environment of the epididymis. The microenvironments created along the length of the epididymal tubule are essential to the molecular modifications of spermatozoa that result in fertile gametes. The secretory and resorptive processes of the epithelial cells that line this tubule generate these microenvironments. In the current study, 10 morphologically distinct segments of the mouse epididymis were identified by microdissection. We hypothesized that the changing environments of the epididymal lumen are established by differential gene expression among these segments. RNA isolated from each of the 10 segments was analyzed by microarray analysis. More than 17 000 genes are expressed in the mouse epididymis, compared with about 12 000 genes identified from whole epididymal samples. Screening a panel of normal mouse tissues identified both epididymal-selective and epididymal-specific transcripts. In addition, this study identified 2168 genes that are up-regulated or down-regulated by greater than 4-fold between at least two different segments. The expression patterns of these genes identify distinct patterns of segmental regulation. Using principal component analysis, we determined that the 10 segments form 6 different transcriptional units. These analyses elucidate the changes in gene expression along the length of the epididymis for 17 000 expressed transcripts and provide a powerful resource for the research community in future studies of the biological factors that mediate epididymal sperm maturation.

epididymis, gene regulation, male reproductive tract, sperm maturation, sperm motility and transport

INTRODUCTION

The mammalian epididymis is a segmented organ comprised of a single, highly coiled tubule conventionally divided into caput, corpus, and cauda regions (Fig. 1). As mammalian spermatozoa pass from the testis into the epididymis they possess a specialized and distinct morphology, but are infertile. Through a number of resorptive, secretory, and contractile processes, each region of the epididymis contributes to the concentration, maturation, transport, or storage of sperm [1, 2]. The hallmarks of the maturational process are the acquisition of motility and the competence to undergo capacitation, leading to the ability to fertilize an egg [3, 4].

Spermatozoa are transcriptionally and translationally inactive. Therefore, extratesticular maturation of spermatozoa is not under the control of the germinal genome; rather, it is mediated by factors within the epididymal fluid. The composition of the luminal environment of the epididymis changes continually due to both secretion and removal of specific ions, small organic molecules, and proteins from the luminal compartment of the organ. The regulation of the processes that establish the discrete epididymal microenvironments remain largely unknown and are the object of considerable investigation [5].

Most studies to date have used relatively large epididymal regions (caput, corpus, cauda) as boundaries for analysis of intraepididymal gene and protein expression patterns [6, 7]. In describing the localization of epididymal gene or protein expression, these regions have also been further divided (e.g., proximal, mid, and distal caput). The epididymal regions can also be definitively subdivided into discrete, intraregional segments or lobules of coiled tubule bounded by connective tissue septae [8, 9]. Over the past decade it has become evident that many protein localization and gene expression patterns within these regions are restricted to one or more segments [10]. This implies that these segments are distinct regulatory subunits of the epididymis [10, 11] and may play an important role in directing the tightly regulated composition of the epididymal tubule fluid.

The reported diversity, complexity, and potential importance of segment-specific gene expression to overall epididymal function has led us to investigate global gene expression patterns in each intraregional segment of the mouse epididymis using the Affymetrix GeneChip Mouse Expression Set 430. Transcriptional profiling of the epididymis has been previously reported for both the rat [6, 12] and mouse [13, 14]. However, the current work aims to increase the completeness of the mouse epididymal transcriptome with the use of a genome-wide array and to provide greater sensitivity by investigating each intraregional segment.

In the present study we have determined by physical microdissection the number of segments that comprise the

[1]Supported by National Institutes of Health grant DK45179 to T.T.T.
[2]Correspondence: Daniel S. Johnston, Wyeth Research, 500 Arcola Rd. N3169, Collegeville, PA 19426. FAX: 484 865 9367;
e-mail: johnstd2@wyeth.com
[3]These authors contributed equally to this work.

Received: 8 February 2005.
First decision: 28 February 2005.
Accepted: 25 April 2005.
© 2005 by the Society for the Study of Reproduction, Inc.
ISSN: 0006-3363. http://www.biolreprod.org

entire epididymis of C57BL/6 mice and have evaluated the expression of more than 39 000 mouse transcripts in each segment. Our analysis also defines the intrasegmental differences in gene expression and establishes distinct transcriptional profiles among the segments. This characterization of the mouse epididymal transcriptome provides a high-resolution, physiologically relevant index of genes expressed along the full length of the mouse epididymis, and will be a powerful resource for future studies defining the components and maturation events that occur within segment-specific epididymal microenvironments.

MATERIALS AND METHODS

Animals and Tissue Collection

Protocols for the use of animals in this experiment were approved by the University of Virginia Animal Care and Use Committee and were in accordance with the National Institutes of Health's standards established by the *Guidelines for the Care and Use of Experimental Animals.* Adult male C57BL/6 mice were obtained from University of Virginia vivarium sources and maintained on a 12L:12D light cycle with food and water ad libitum. The mice were anesthetized by i.p. injection of sodium pentobarbital (0.1 mg/g body weight) and subjected to a midventral laparotomy. The testes and epididymides were exposed with traction and the epididymides removed after ligation of the spermatic and vasal arteries. Epididymides were immediately placed in ice-cold saline in a Petri dish and carefully defatted. This and all subsequent manipulations were performed with the use of a dissecting microscope while keeping the dissection medium on ice at all times. Each defatted epididymis was transferred to fresh saline and the epididymal tunica albuginea was removed using sharp microdissection. Removal of the tunica and blunt microdissection along the plane of the segment-dividing connective tissue septea resulted in the separation of all detectable epididymal segments. The precise location of the segments and the track of the connective tissue septea borders varied somewhat from animal to animal, but the general appearance of the segments and their connective tissue septea were relatively consistent. The appearance of individual segments in situ in a detunicated epididymis has been illustrated previously [15]. As each segment was isolated it was immediately placed in RNALater (>10× tissue volume; Ambion, Austin, TX) in a 1.5-ml Eppendorf tube on ice. All 10 segments of each epididymis were dissected cold and placed in separate tubes of RNALater within 30 min of epididymal extirpation. This procedure was repeated until the tube for each epididymal segment contained at least five representatives of that segment. These pooled segments comprised one sample for subsequent RNA extraction. When all of the segments comprising one sample were collected, the samples were immediately stored at −80°C. This procedure was repeated until each epididymal segment was represented by 5–8 samples for RNA extraction.

RNA Preparation

The samples were thawed at room temperature and the RNALater was removed. The samples were washed twice with 1 ml of ice-cold TRIzol. Tissue was homogenized using a glass mortar and pestle (total capacity 200 μl, Kimble/Kontes, Vineland, NJ) in 150 μl of TRIzol (Gibco BRL, Gaithersburg, MD) and RNA was extracted according to the manufacturer's instructions. RNA was further purified with an RNAeasy column (Qiagen, Valencia, CA). Tissue samples from segment 7 and from whole epididymis of larger volumes were homogenized with a PowerGen 700 automatic homogenizer (Fisher Scientific, Hampton, NH) in 600 μl of TRIzol and total RNA was isolated according to the manufacturer's instructions. RNA from these samples was further purified with and RNAeasy column. For all samples, RNA quantity was determined by absorbance at 260 nm (NanoDrop, Wilmington, DE) and quality was determined using an Agilent Bioanalyzer (Palo Alto, CA).

Microarray Processing

Five micrograms of total RNA was used to generate biotin-labeled cRNA using an oligo T7 primer in a reverse transcription (RT) reaction followed by an in vitro transcription reaction with biotin-labeled UTP and CTP. Ten micrograms of cRNA was fragmented and hybridized to MOE430A and MOE430B arrays (Affymetrix, Santa Clara, CA). Hybridized arrays were stained according to manufacturer's protocols on a Fluidics Station 450 and scanned on an Affymetrix scanner 3000. All array images were visually inspected for defects and quality. Arrays with excessive background, low signal intensity, or major defects within the array were eliminated from further analysis. This, plus the procedural loss of two samples of segment 7 resulted in the number of replicates shown in Table 1. Signal values were determined using GeneChip Operating System 1.0 (GCOS, Affymetrix). For each array, all probe sets were normalized to a mean signal intensity value of 100. The default GCOS statistical values were used for all analyses. Signal values and absolute detection calls were imported into Genesis 2.0 (GeneLogic, Gaithersburg, MD) for analysis.

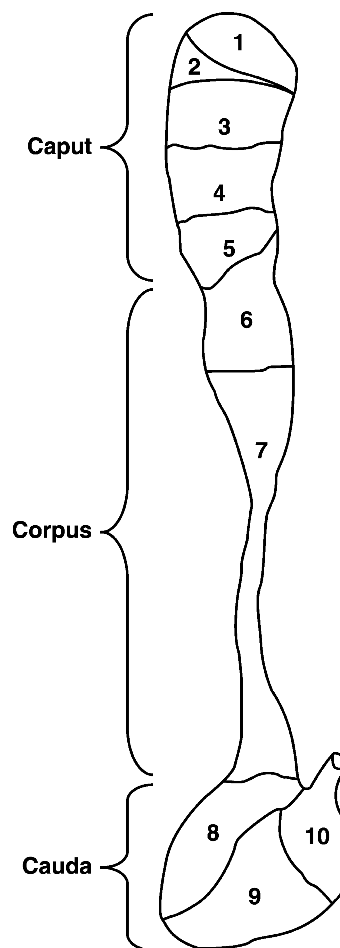

FIG. 1. Segment structure of the mouse epididymis. A typical pattern of epididymal segmentation is shown, and these regions were microdissected for the studies performed in this report.

A gene was considered detectable if the mean expression in any segment was greater than 50 signal units and the percentage of samples with a Present (P) call as determined by GCOS default settings was greater than or equal to 66%. For example, in segments 7 and 8, a gene was considered detected if it received a P call in two out of the three arrays analyzed. Normalized signal values were transformed to the log base10 and an analysis of variance (ANOVA) was performed. A gene was considered to be segmentally regulated if the difference between two segments met the following criteria: 1) the gene had to be detected in at least 66% of the samples of at least one of the segments, 2) the fold change between at least two segments was at least 4.0, and 3) the *P* value based on an ANOVA test had to be ≤0.01. These conditions were met by 2168 qualifiers, which were used for further analysis. For some computations, the expression values for each qualifier were normalized to a mean of zero and a standard deviation of 1 (z-score normalization). This allowed a direct comparison of patterns within the data without respect to absolute expression levels.

Principal Component Analysis

Principal component analysis (PCA) was performed on the transformed data and visualized using Spotfire 7.2 (Somerville, MA). PCA is a statis-

TABLE 1. Number of qualifiers detected in the 10 mouse epididymal segments.

Segment	Qualifiers expressed[a]	Additional qualifier detected[b]	Increased sensitivity[c] (%)	Genes significantly different[d]	Genes selective to epididymis[e]	Genes unique to epididymis[f]	No. of Replicates
1	11959	1654	16	713	224	32	7
2	12018	1532	15	724	238	31	8
3	11746	1417	14	648	232	33	5
4	11829	1544	15	652	221	36	5
5	12080	1712	17	690	223	40	4
6	12360	1914	18	677	218	39	4
7	12268	1638	15	721	210	34	3
8	13241	2600	24	750	201	33	3
9	11419	1509	15	776	188	29	5
10	12176	1962	19	864	181	26	5
Whole	11863	0	0	842	257	37	4
Total unique	17096	5233	44	2186	307	75	53

[a] Expression greater than 50 signal units and 66% Present call.
[b] Qualifiers detected in segments but called absent in whole epididymis.
[c] Present increase in sensitivity per segment.
[d] Unique qualifiers with differentially expression (4.0-fold) in specified segment versus any other segment.
[e] Qualifiers expressed in specified segment and 3.0-fold greater than expression in any of 22 normal tissues.
[f] Genes selective to epididymis and called present 0% or expressed less than 50 signal units in all the 22 normal tissues.

tical method for reducing multidimensionality data sets by transforming the data to a new set of variables that illustrate the differences in the data [16]. These reduced variables allow visualization of multidimensional data sets and aids in the functional and biological interpretation of the overall data variation.

Identification of Defensin Genes

Defensin gene family members were identified through a Unigene gene annotation database search [17]. Twenty-eight qualifiers were annotated as defensin or defensin-related cryptin molecules. Thirteen of these qualifiers representing 12 unique genes were expressed in the epididymis as described above and were analyzed further.

Tissue Specificity

Genes expressed in any segment were screened for tissue selectivity by comparing to a Wyeth Research database of MOE430 genechip profiling data from 22 normal tissues. The data were acquired from a source generated by GeneLogic and Wyeth Research (Cambridge, MA) using standardized procedures and internal controls to minimize variation. All tissues were from C57BL/6 mice. Each tissue comprised between 3 and 12 replicas. All transcriptional profiling data were normalized to a mean signal intensity value of 100 in GCOS (Affymetrix). Genes were considered epididymis-selective if the following conditions were met: 1) the qualifier had to have a 66% P call in at least one epididymal segment, 2) the expression of the qualifiers in the epididymis had to be ≥50 signal units, and 3) the mean expression of the qualifiers in any segment had to be 3-fold higher than the mean expression in any of 22 other tissues. Genes were considered to be epididymis-specific if the following conditions were met: 1) the qualifier had to be epididymal-selective, 2) the qualifier was never detected (0% P call) in any of the 22 other tissues, or 3) the qualifier had be expressed below 50 signal units in all 22 tissues.

First-Strand cDNA Synthesis

Complementary DNA samples for real-time polymerase chain reaction (PCR) analyses were synthesized by random priming in a final volume of 20 μl using the Superscript III First Strand Synthesis System for RT-PCR (Invitrogen, Gaithersburg, MD) according to the manufacturer's instructions. The cDNA samples corresponding to each of the epididymal segment samples as well as whole epididymis were diluted to a final volume of 200 μl to produce a 10× cDNA stock.

Primers for Real-Time RT-PCR Analyses

Real-time RT-PCR analysis was carried out using FAM-labeled flourogenic LUX primers from Invitrogen [18]. Primers were chosen from published sequences for *Clu* (Clusterin), *Cst8* (Cystatin 8, cysteine-related epididymal-specific [CRES]), *Crisp1* (cysteine-rich secretory protein 1, protein DE, acidic epididymal glycoprotein), *Gpx5* (glutathione peroxidase 5), *Adam28* (a disintegrin and metalloproteinase domain 28), *Emb* (embigin), *Myom2* (myomesin 2), *Smpx* (small muscle protein, x-linked), *Srd5a2* (steroid 5 alpha-reductase 2), and *Ros1* (ros1 proto-oncogene). All primers were designed using the Web-based version of LUX Designer (available at http://www.Invitrogen.com/lux). The primers chosen for each transcript and the GenBank accession number for each sequence are as follows:

Clu: forward primer, 5′-gaacatgGGGATGAGGTGTTGAGCATGtTC-3′; reverse primer, 5′-CGCCAGGAGCTGAACGACT-3′ (NM_013492).
Cst8: forward primer, 5′-TGCCAATGTGAAGCAGTGTG-3′; reverse primer, 5′-cacagacTTGGTATTCCATTCGGTCTGtG-3′ (NM_009978).
Crisp1: forward primer, 5′-caactgAGGAGAACCGTGTGCCAGtTG-3′; reverse primer, 5′-GGTGCATAGCCCATCTTCACAG-3′ (NM_009638).
Gpx5: forward primer, 5′-gaccagAAAGAGCAATGGAGGCTGGtC-3′; reverse primer, 5′-TCAGGAACTGACAAGGCATTTG-3′ (NM_010343).
Adam28: forward primer, 5′-cacgatCAGTCTTCCACTTCTCCATCGtG-3′; reverse primer, 5′-CACTTTGGCGCTGGATCACTA-3′ (NM_176991).
Emb: forward primer, 5′-cacgttAGACAAAGATGCCCTCAACGtG-3′; reverse primer, 5′-TGTCCACCCTGTACCCACAGA-3′ (NM_010330).
Myom2: forward primer, 5′-CAGCGAGCACTTCTTGGTGAA-3′; reverse primer, 5′-gactcgTTGATGCTGTACTTGCCCGAGtC-3′ (NM_008664).
Smpx: forward primer, 5′-AAACCTCCTGAATGCCCAAGA-3′; reverse primer, 5′-cacgaaCACTTCCCAGGTTGTAAATTCGtG-3′(NM_025357).
Srd5a2: forward primer, 5′-AAAGCCACTGCCTTCTGCAT-3′; reverse primer, 5′-ctacctCCGCGCAATAAACCAGGtAG-3′(NM_053188).
Ros1: forward primer, 5′-CCGGTGGAGTCATCAACAAA-3′; reverse primer, 5′-gaactcaCTGGCATCGTGTCATCTGAGtTC−3′(NM_011282).

Fluorophore-labeled primers were synthesized by Invitrogen and nonlabeled primers were synthesized by Wyeth Research (Collegeville, PA).

Real-Time PCR Analyses

Quantitative real-time PCR reactions (qRT-PCR) were conducted using Platinum Quantitative PCR SuperMix-UDG (Invitrogen). The reagent was diluted to the appropriate working concentration in a final volume of 25 μl and added to a 15-μl mixture containing 200 nM of each primer, 1 μl of ROX reference dye, and 10 μl of the cDNA stock solution to a final reaction volume of 50 μl.

PCR reactions were run on an automated fluorometer (ABI Prism 7700 Sequence Detection System; Applied Biosystems, Foster City, CA) in a 96-well format. PCR conditions for all reactions was as follows: 1 cycle at 48°C for 30 min, 95°C for 10 min followed by 40 cycles at 95°C for 15 sec, and 60°C for 1 min. Relative expression was determined by using the C_T method [19] using Sequence Detector software, version.1.6.3 (Applied Biosystems). The cDNA samples of each segment and of whole epididymis were evaluated in triplicate with the primer pair for each gene. Results were normalized to 18S ribosomal RNA expression and expressed as a ratio of expression of each gene in each segment compared to the expression of that gene in the whole epididymis.

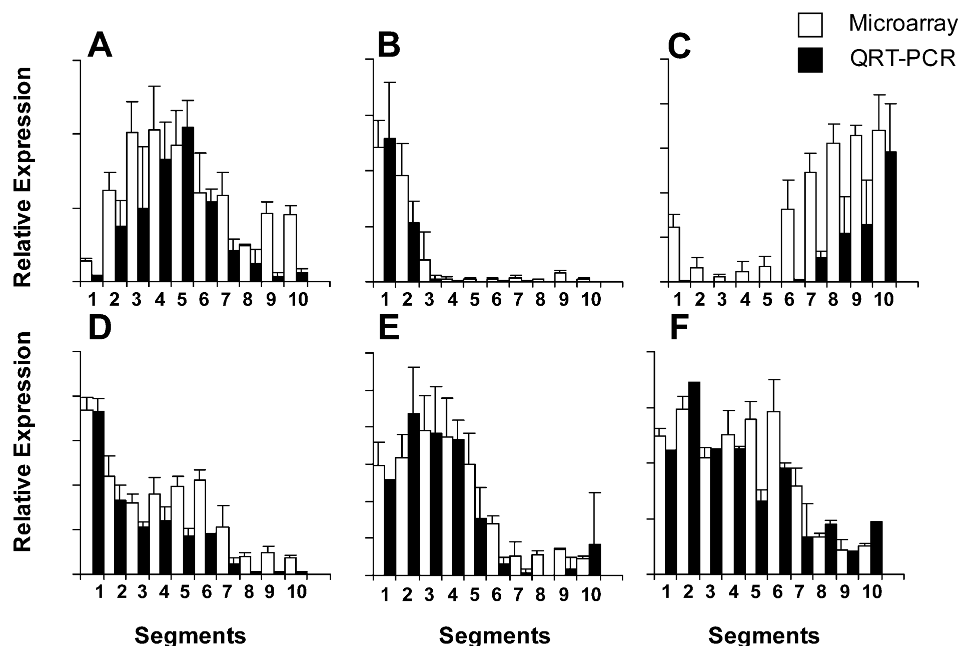

FIG. 2. Confirmation of microarray results by quantitative RT-PCR using selected genes known to be expressed in the epididymis. Messenger RNA expression levels of (**A**) *Clu,* (**B**) *Cst8,* (**C**) *Crisp1,* (**D**) *Ros1,* (**E**) *Gpx5,* and (**F**) *Srd5a2* in the 10 segments of the mouse epididymis were investigated and compared using both microarray results and qRT-PCR data. Values represent the mean relative expression ± SD; n ≥ 3.

RESULTS

Segmentation of the Mouse Epididymis

To determine the transcriptome of individual epididymal segments two experimental points needed to be established to permit unequivocal interpretation. First, the number of segments that comprise the epididymis of the species under study had to be determined. Second, the segments used for the analysis had to be accurately identified, isolated, and preserved to prevent contamination, loss, or degradation of RNA.

Extending our histology-based epididymal segment determination method used previously [11], the current study employed microdissection to determine the number of physical segments. Approximately 20 mouse epididymides were carefully dissected with the aid of a dissecting microscope. Although all operators (D.S.J., H.J.B., T.T.T.) noticed slight variation between epididymides with respect to the relative size and shape of specific segments, 10 distinct segments were unequivocally and consistently identified (Fig. 1).

Relevant to the importance of quickly identifying, isolating, and preserving the samples, the isolation procedure was rehearsed until each epididymis could be removed from an anesthetized animal, detunicated in ice-cold saline, and all segments isolated and placed in RNALater in less than 30 min. Once the operators were practiced, all 10 samples could routinely be collected in less than 25 min.

Microarray Analysis of the Mouse Epididymal Transcriptome

Segmental analysis of the entire mouse epididymis identified 17 096 distinct qualifiers (Table 1) corresponding to approximately 43% of qualifiers on the array set. In contrast, only 11 863 transcripts were detected when the whole epididymis was profiled. Therefore, investigating the epididymal transcriptome segment by segment yielded a 46% increase in sensitivity for detecting genes expressed in the epididymis. The additional 5233 qualifiers detected by segmental analysis were low-abundance genes with a median expression of 80 signal units in the whole epididymal sample. Segmental analysis increased the signal value for these genes an average of 1.7-fold in at least one segment. Within each individual segment we observed between a 14% (Table 1, segment 3) and 24% (Table 1, segment 8) increase in the number of transcripts identified compared to the entire epididymis and an average increase of 16.8%.

Many genes expressed in the epididymis are well-characterized as to both their level of expression and region-dependent expression pattern (i.e., caput, corpus, cauda). We have selected a number of these genes for qRT-PCR analysis to authenticate the results from the microarray. Clusterin (*Clu* [20]), cystatin 8 (*Cst8* [21]), cysteine-rich secretory protein (*Crisp1* [22]), glutathione peroxidase 5 (*Gpx5* [23]), ros1 proto-oncogene (*Ros1* [13]), and steroid 5α-reductase 2 (*Srd5a2* [24]) all showed consistency between microarray and qRT-PCR analysis (Fig. 2). The segment-dependent expression patterns for these genes also agree with the protein or transcriptional patterns, so far as available, reported in the literature for mouse epididymal *Clu* [20], *Cst8* [21], *Crisp1* [22], *Gpx5* [23], and *Ros-1* [13]. Of interest, *Srd5a2* expression in the mouse epididymis appears to follow the general pattern reported for the rat epididymis [24]. These comparisons validate the accuracy of both the microarray results and the microdissection methods and lend credence to each genes segmental expression pattern.

A second set of genes was selected for qRT-PCR analysis as examples of genes not previously reported to be expressed in the mouse epididymis but that exhibit tightly regulated expression patterns in different regions of this organ. The genes selected are examples showing distinct expression patterns moving from extreme proximal caput (Fig. 3A; segment 1; *Adam28*) to corpus (Fig 3B; segment 7; *Myom2*) to cauda (Fig. 3C; segments 8, 9, and 10; *Smpx*). Also illustrated is a biphasic pattern typified by Embigin (*Emb*; Fig. 3D), with expression observed in the caput and cauda regions. Quantitative RT-PCR analysis of these genes correlated closely with the results from the microarray analysis (Fig. 3) and further validates this microarray analysis

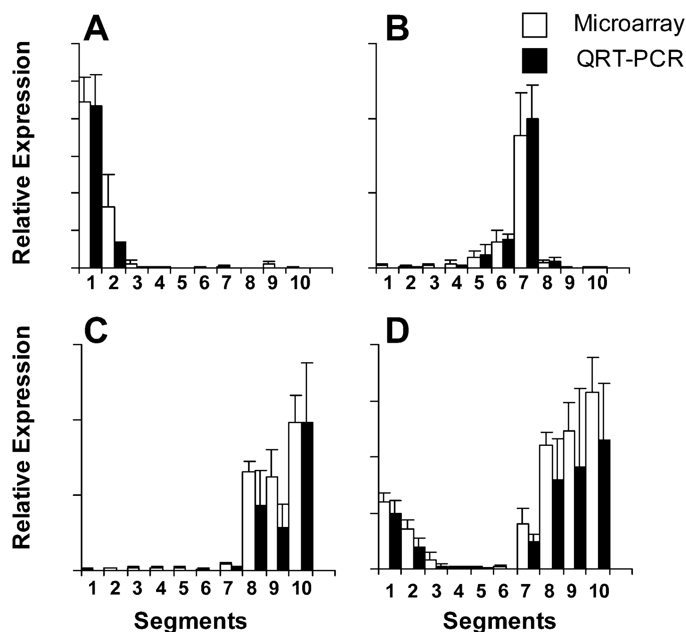

FIG. 3. Confirmation of microarray results using genes not previously known to be expressed in the epididymis, but that show distinct expression patterns. Messenger RNA levels for (**A**) *Adam28,* (**B**) *Myom2,* (**C**) *Smpx,* and (**D**) *Emb* exhibit distinct expression patterns across the 10 segments of the mouse epididymis as determined by both microarray analysis and qRT-PCR. Values represent the mean relative expression ± SD; n ≥ 3.

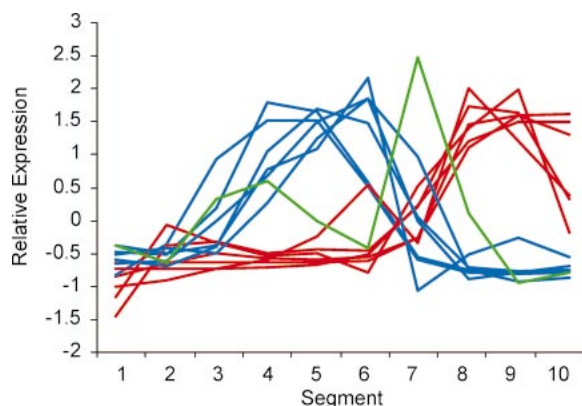

FIG. 4. Segmental expression of defensin genes in the epididymis. Gene ontology annotation was used to select antimicrobial genes as described in *Materials and Methods.* The relative expression pattern of 12 defensin genes along the length of the epididymis occurs in three distinct groups as indicted by color. The genes comprising each group are as follows (the segment showing maximal expression of the gene and the mean signal value are indicated in parentheses): Blue: *Defb12* (4; 2788), *Defb15* (5; 5274), *Defb19* (5; 138), *Defb13* (6; 225), *Defcr-rs1* (6; 284), *Defcr-rs10* (6; 383); Green *Defcr15* (7; 109); and Red: *Defb10* (8; 671), *Defb1* (two tilings: 8; 697 and 9; 624), *Defb11* (9; 3971), *Defb9* (9; 1514), *Defb2* (10; 5930).

of segmented epididymal gene expression. The data in both Figures 2 and 3 illustrate the sharp differences that commonly occur between adjacent segment gene expression profiles. These results underscore the consistency of the dissection technique and validate the need for analyses of epididymal gene expression at a higher resolution than caput, corpus, and cauda regions.

An extremely large number of genes display differences in relative expression between segments. Specifically, 6126, 2186, 148, and 58 qualifiers were differentially regulated 2-, 4-, 50-, and 100-fold or more, respectively. More than 60 genes exhibit at least a 100-fold difference in relative expression between any two segments. Among the highest in this category are defensin beta 9 (1832-fold), adam28 (1790-fold), and defensin beta 11 (981-fold). Furthermore, our results suggest that segment-dependent regulation of gene expression is a prominent theme in the epididymis; approximately 30% of detected genes (6126/17 096) display a significant difference in expression level between two segments (at least 2.0-fold; $P < 0.01$). In defining segmentally regulated genes for detailed analysis, however, we have chosen to include only qualifiers with signal values that change by more than 4.0-fold ($P < 0.01$) between any two segments. Fitting this criterion are 2186 qualifiers, representing 12.8% of the qualifiers expressed in the epididymis.

In addition to examining the pattern of expression of individual genes of interest, the data can be used to elucidate the pattern of expression of entire gene families. For example, the relative expression pattern of 12 defensin genes along the length of the epididymis occurs in three distinct patterns (Fig. 4). The first group of defensin genes has a high relative expression in segments 4, 5, and 6, a second group consisting of only one gene is expressed in segment 7, and the final group has high relative expression in segments 8, 9, and 10. The identities of each of the defensin genes shown in Figure 4 are provided in supplemental Figure 1 (available online at http://www.biolreprod.org). Although the relative expression of all of the defensin genes investigated in the analysis are very low in segments 1 and 2, some members of the defensin family of genes shown in Figure 4 are significantly expressed in each segment of the epididymis.

Evaluation of Tissue Specificity of Gene Expression

Comparison of the gene expression data from this experiment to an internal database of microarray results allowed for the identification of genes that exhibit a higher level of expression in the epididymis when compared to 22 other tissues. In the current analysis these genes were categorized as either epididymis-selective or epididymis-specific on the basis of criteria defined in *Materials and Methods.* Three-hundred seven genes were determined to be epididymis-selective (Table 1), including genes that have been well-studied in the field such as *Clu* (clusterin; Fig. 5a) and *Prom1* (prominin 1; Fig. 5b). In addition, 75 genes were epididymis-specific, including a number of genes historically studied in the epididymis (e.g., *Cst11* (cystatin 11; Fig. 5c), *Lcn5* (lipocalin 5, epididymal retinoic-acid binding protein; Fig. 5d), and *Gpx5* (glutathione peroxidase; Fig. 5d). In addition, this analysis also identified epididymal specificity for known genes not previously shown to be expressed in the epididymis (e.g., *Defb11*; defensin beta 11, Fig. 5e), and showed epididymal specificity for uncharacterized genes not previously known to be expressed in the epididymis (expressed sequence tag AW455861; Fig. 5f). At the level of sensitivity of these analyses a number of defensins (2, 9, 11, 13, and 15) were, interestingly, epididymal-specific. A list of 49 qualifiers determined to be epididymal-specific is shown in supplemental data (available online at http://www.biolreprod.org).

Cluster Analysis of Segmental Gene Expression

To obtain an overview of segmentally regulated gene expression patterns in the mouse epididymis, the 2186

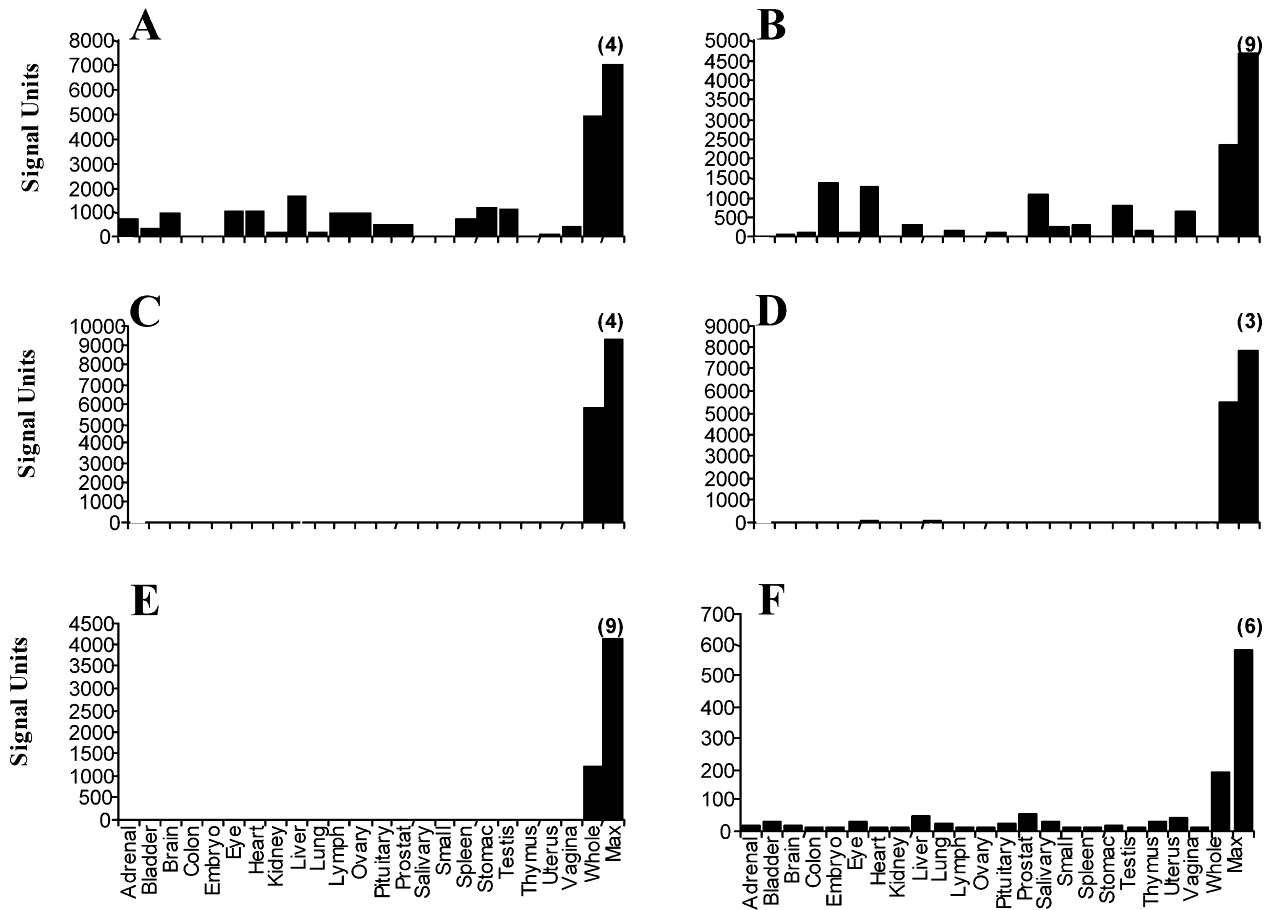

FIG. 5. Tissue expression profiles of epididymal expressed genes. Criteria for selection as either epididymis-selective or epididymal-specific are defined in *Materials and Methods.* Shown are the expression profiles of two epididymal-selective genes, *Clu* (**A**) and *Prom1* (**B**); and four epididymal specific genes, *Lcn5* (**C**), *Gpx5* (**D**), *Def11* (**E**), and expressed sequence tag AW455861 (**F**). Whole refers to entire epididymis and Max refers to the signal value obtained from the epididymal segment (segment number in parentheses) with the highest expression of that particular gene.

genes defined as segmentally regulated (Table 1) were clustered into 10 groups on the basis of epididymis segment with the highest expression. The graphical output of this cluster analysis (Treeview [25]) reveals numerous distinct expression patterns along the epididymis (Fig. 6). Many of the groups show that relatively high levels of expression can persist over multiple segments, but many others show high expression restricted to one segment. The clearest examples of the latter are genes with the highest expression in segments 1 and 7 (Fig. 6). Further insight into segmental gene expression is given by the averaged z-score normalized expression values for each of the 10 data groups (Fig. 7). The 2168 transcripts with a 4-fold difference in expression across the epididymis were grouped into 10 categories based on the segment with the highest expression. These analyses clearly demonstrate that the epididymis regulates gene expression in a spatially coordinated and segment-dependent manner.

Principal Component Analysis

Principal component analysis determined whether distinct groups of samples could be identified within the 49 tissue samples prepared from the 10 different segments. The PCA analysis reduced the data from the individual samples sets onto a two-dimensional space spanned by the two leading principal components, represented by the two axes on the graph (Fig. 8). Four whole epididymal samples were also included in the analysis. Projection of the samples onto the first principal component captured 41% of the variability and projection onto the second principal component explains 20.4% of the variability.

Individual segments clustered closely with one another, indicating little variation between replica segments (Fig. 8). This two-dimensional PCA analysis grouped the 10 epididymal segments into 6 distinct groups or transcriptional units; segments 1, 2, 3, and 7 each demonstrated a distinct expression pattern and formed their own unique groups. Further homologies between segments 4–6 and 8–10 comprised two additional groups. The whole epididymal samples formed a seventh group that is separate from all other groups.

DISCUSSION

In the present study, the segment-dependent expression patterns of approximately 39 000 genes were profiled in the mouse epididymis. This experimental approach allows for expression profiling of nearly the entire genome with increased sensitivity and higher spatial resolution than allowed by the conventional dissection of the epididymis into the caput, corpus, and cauda regions. From this approach we identified 2168 genes that are differentially expressed along the length of the epididymis and categorize the 10 anatomically distinct segments into 6 transcriptionally related entities.

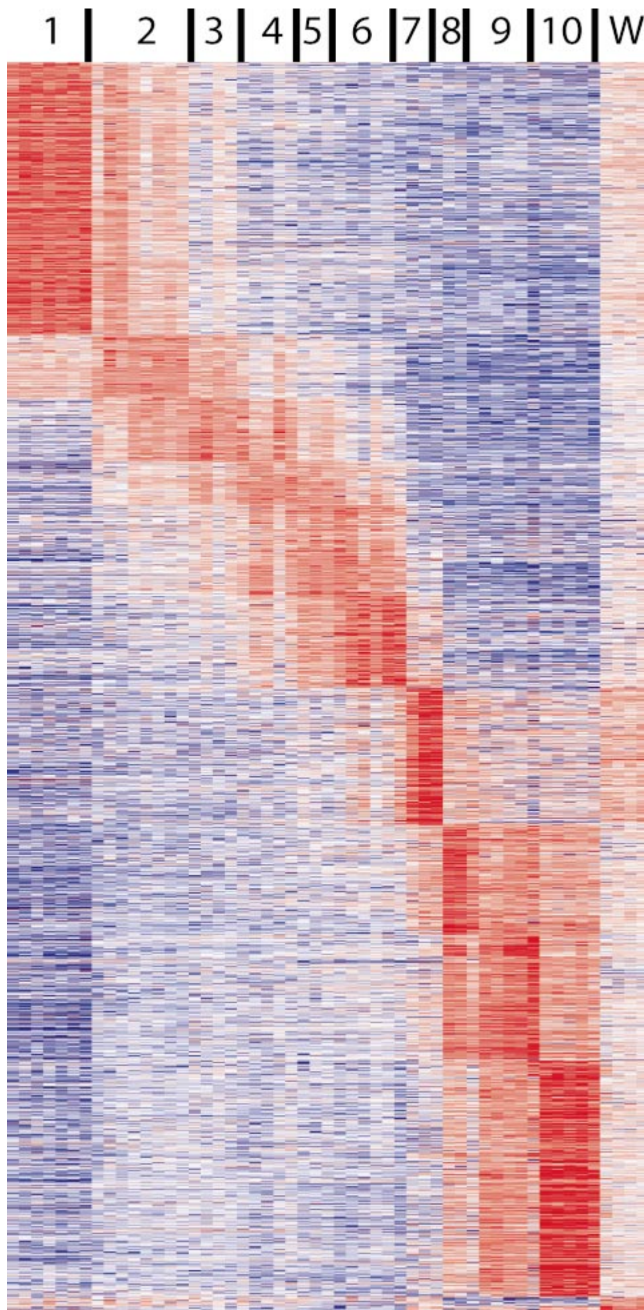

FIG. 6. Expression of differentially expressed transcripts within the epididymis. The 2168 qualifiers that showed a 4-fold difference in expression pattern between any two segments were ordered based on the segment with the highest expression. Data were visualized in Treeview [25]. For each gene, relatively high expression is shown in red, relatively low expression is shown in blue, and median expression is shown in white. The numbers correspond to the epididymal segment number; W refers to whole epididymis samples.

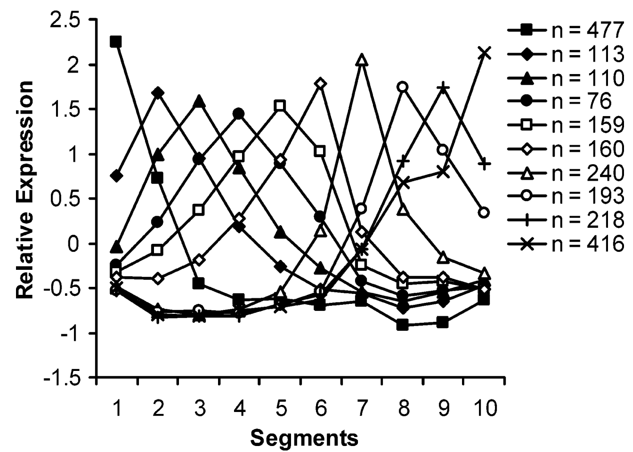

FIG. 7. General gene expression patterns across the 10 segments of the epididymis. The 2168 transcripts with a 4-fold difference in expression across the epididymis were grouped into 10 categories based on the segment with the highest expression. Shown are average expression patterns for genes grouped by their highest expression in each segment. Values are z-scored, normalized average intensity scores. N, number of genes represented in the expression pattern.

The first step in the accurate study of epididymal segment-dependent gene expression was the unequivocal determination of the number segments in the epididymis of the C57BL/6 mouse. Previous determinations have been made by inspecting hematoxylin-eosin-stained longitudinal sections of mouse epididymides and counting apparent connective tissue septae and semisepta [11, 26]. This method determined that there were either eight [26] or nine [11] distinct connective tissue septae but it required a subjective determination between what might be a thinner than usual connective tissue septae or a more prominent than usual semiseptum. In the current study we used a microdissection procedure based on the surface view of the detunicated epididymis and the natural cleavage planes present in the tissue. For example, distinguishing between segments 9 and 10 (where the connective tissue septae are not consistently obvious histologically [27]) is facilitated in dissection by a cleavage plane consistently between coils of tubule of markedly different diameters. The microdissection approach accurately defined the physically evident epididymal segments and was crucial to the consistency and reliability of the gene expression profiling results.

Another potential concern in isolating segment 1 is the possible contamination by a distal remnant of efferent ducts that had not been precisely dissected. At the magnifications

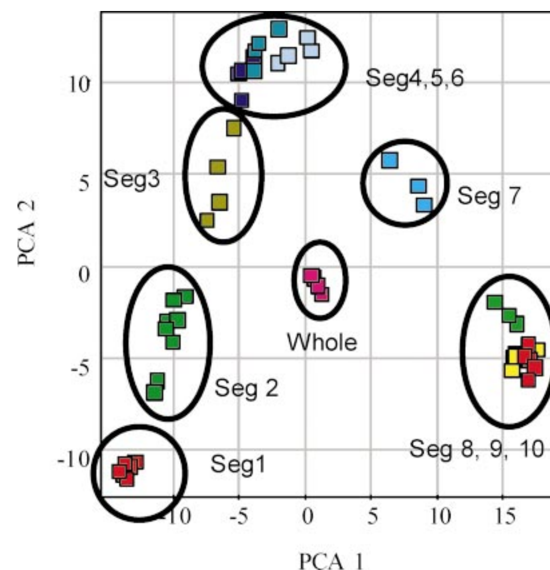

FIG. 8. Principal component analysis of gene expression profiles from epididymal segments. Log base10-transformed data from each individual sample were combined and PCA analyses were conducted on the 2168 segment-regulated qualifiers. Samples are color-coded and labeled by segment. The first two principal components are shown. Six distinct clusters are identified.

used during microdissection, the efferent ducts are distinct and their junction with the epididymis is separable by careful decapsulation and dissection along cleavage planes. We expect any contamination of segment 1 by efferent duct to be insignificant.

Several studies have investigated gene expression in the epididymis; however, they have been limited in both the number of transcripts analyzed and confined to conventional segmentation approaches. Studies have ranged from 1) commercially purchased microarrays of 1200 cDNAs to profile the caput, corpus, and caudal regions [6] to 2) the set of 15 247 mouse cDNAs derived from preimplantation and peri-implantation embryos to examine gene expression in initial segments and the remainder of the epididymis (midcaput through cauda) [14]. The largest study to date used the Affymetrix MGU74AV2 GeneChip oligonucleotide microarray platform to examine the epididymis of wild-type and castrated C57BL/6 mice segmented into caput, corpus, and cauda regions [7]. However, data presented in that study focused on androgen-regulated genes and reported little of the overall expression patterns in the wild-type epididymis. The current study aims to provide global expression profiling in each of the 10 segments of the mouse epididymis using the Affymetrix MOE430 GeneChip set and to communicate these data in a searchable Web-based format.

The microarray data generated in this study were validated both by correlation with our own qRT-PCR and with published expression patterns of well-studied epididymal transcripts. The relative levels of gene expression determined by both qRT-PCR and microarray analysis are very similar for all 10 genes examined (Figs. 2 and 3). Furthermore, the microarray data were consistent with published reports of either gene or protein expression in mice for *Cst8* [21] *Ros1* [13], *Gpx5* [23], *Clu* [20], and *Crisp1* [22]. For example, *Cst8* mRNA has been shown to be restricted to the proximal region (corresponding to segment 1) of the mouse epididymis by Northern blot and in situ analyses [21]. The results of our microarray and qRT-PCR analyses demonstrate a strong *Cst8* signal in segment 1 with a reduced signal in segment 2, and undetectable in some samples of segment 3 (see Fig. 2b). These results and similar comparisons to other gene protein localizations available in the literature demonstrate both the precision of the dissection method and the accuracy of the microarray analysis. These validations were fundamental to the additional analyses presented in the paper and support the accuracy of their results.

The data were used to identify genes that were expressed either predominantly or uniquely in the epididymis. More than 300 genes were shown to be expressed more than 3-fold higher in the epididymis than in any of the 22 other tissues examined, and 75 genes were shown to be expressed uniquely in the epididymis (Table 1 and Fig. 5). These qualifiers were found to be to selective/specific with regard to the sensitivity of the microarray technology, and the panel of tissues used. Evaluation of these transcripts using additional tissues, more-sensitive assays, or both could alter their designation as epididymal selective/specific. This analysis provides an important dataset of genes and gene products to be evaluated further to understand the role they may play in epididymal function.

An obvious use of the microarray data generated in this study is to identify genes that are differentially expressed in a segment-by-segment fashion. The epididymis, an organ with a highly specialized role in sperm maturation, displays an incredibly complex pattern of gene regulation. Our analysis of differentially regulated genes in the epididymis demonstrates that more than 2186 genes, approximately 13% of the genes expressed in the epididymis, differ in their level of expression in different segments by at least 4-fold. The large number of regulated genes within a grossly dissected single organ is, to our knowledge, unprecedented. By comparison, only 1186 genes are differentially regulated (>4.0-fold, $P > 0.01$) between the kidney and the liver (unpublished data), two distinct organs with very different functions. Similarly, only 2478 and 1075 genes are differentially regulated (>4.0-fold, $P > 0.01$) between colon and heart and between eye and bladder, respectively (unpublished data). This suggests that sperm maturation, transport, and storage in the epididymis are highly complex events. Interestingly, segments 1 and 10 (corresponding to the proximal and distal ends of the organ) express the greatest numbers of differentially regulated genes (926 more then 4-fold, 2465 by more then 2-fold). These segments are morphologically distinct from their adjacent segments (i.e., tubule diameter, epithelial height, etc.) and also are the developmental boundaries of the organ. These observations support the hypothesis that these segments play specialized roles in the posttesticular maturation, storage of spermatozoa, or both.

An interesting result within the cluster analysis data was the clear demonstration of coordinated and regulated gene expression along the length of the epididymis (Figs. 4 and 6–8). While androgens and other factors synthesized in the testis have been shown to mediate transcription in the epididymis [7, 28], the observed expression patterns resembling entrained waves for defensin genes (Fig. 4) and differentially regulated genes in general (Figs. 6–8) are more complex than can be accounted for by testosterone or lumicrine factors alone.

The intricate patterns of expression may be facilitated by the presence of distinct intrasegmental microenvironments within each segment. Previously, we have used the microinjection of specific-sized dyes and radiolabeled compounds to show that connective tissue septae significantly restrict the movement of molecules with molecular masses as low as approximately 1 kDa, implying that the segment-specific interstitial microenvironments are maintained in the intact organ [15]. The present study corroborates the multifactorial coordination of gene expression in the epididymis.

In addition to allowing higher spatial resolution of segment-dependent gene expression, our experimental design also provides increased sensitivity in detecting low-abundance mRNA species. Summation of the qualifiers identified from the 10 individual segments enabled the detection of 5233 more qualifiers than when the starting mRNA is derived from the whole epididymis. This 44% increase in detection is likely due to a decreased dilution of low-abundance mRNA species. These may be genes that are either expressed in low abundance throughout the epididymis, or genes with expression that is limited to one or more individual segments. Surprisingly, approximately 40% of the 2186 genes exhibiting segment-dependent expression were not detected in the whole epididymis mRNA sample.

Principal component analysis of the 48 segment samples and 4 whole epididymal samples showed that individual segments clustered closely with one another, indicating little variation between replica segments. The small variation that was present, most noticeably in segments 2 and 3, is not believed to be due to inexact segmental division or to

technical variation in the genechip analysis, but rather to true biological variations within the sample. The PCA categorized the 10 segments of the epididymis into 6 distinct units corresponding to segments 1, 2, 3, 4–6, 7, and 8–10 (Fig. 8). This finding provides unique insight into the potential biological significance of distinct regions of the epididymis. Investigation into the regulatory regions of the genes specifically expressed in each region may elucidate the common mechanisms of gene regulation in each unit. Identification of which genes are uniquely or highly expressed in each unit may provide powerful insight into the role of each unit in modulating sperm maturation, transport, or storage. This analysis may be extended further to identify genes whose protein products are expected to be secreted and therefore may ultimately end up contributing to the luminal environment and affect sperm maturation.

The data generated from this study allow the rapid elucidation of the regulation pattern of not only single genes, but families of genes as well. In this report, we have demonstrated the pattern of expression for several members of the defensin family of genes along the epididymis. This family of genes shows a coordinated pattern of expression with specific defensin genes being highly expressed in a specific segment in a proximal-to-distal fashion (Fig. 4). Similar analyses can be performed on any categorized families of genes, including HOX genes, ion channels, and transcription factors. Clearly, regulated expression of enzymes with counteracting activities (i.e., proteases/protease inhibitors and glycosyltranferases/glycosidases) is crucial to establishing the timing and balance of these events that are critical to the process of epididymal maturation.

We have generated a robust gene expression dataset by taking advantage of the naturally occurring segmentation of the mouse epididymis and generated the most comprehensive epididymal transcriptome to date. In particular, we have used these data to demonstrate and investigate the enormous amount of highly regulated gene expression that takes place along the length of the epididymis. We have also identified genes expressed either uniquely or predominately in the epididymis (Fig. 5). The products of these epididymal-specific/selective genes may prove to be targets for contraception or may provide insight to the unique responsibilities of the epididymis in supporting male fertility.

Approximately 90% of the data from this investigation will be made publicly available in a searchable format via the Mammalian Reproductive Genetics database (http://mrg.genetics.washington.edu) when this report goes to press. The remaining 10% will be added within 18 mo of the publication date. The data can be queried for numerous purposes, including examination of segmented expression profiles of individual genes expressed within the organ or to determine the expression patterns of entire families of genes to assist in determining their role in sperm maturation, transport, or storage in the epididymis. This resource can assist interested investigators to develop new hypotheses about the cell and molecular biology of the epididymis, which will lead to new insights into epididymal function.

ACKNOWLEDGMENT

The authors thank Dr. Michael A. Nolan for critically reading the manuscript.

REFERENCES

1. Robaire B. Efferent ducts, epididymis, and vas deferens: structure, functions and their regulation. In: Knobil E, Neil J (eds.), Physiology of Reproduction, vol. 1. New York: Raven Press; 1988:999–1080.
2. Cornwall GA, Lareyre J, Matusik R, Hinton BT, Orgebin-Crist MC. Gene expression and epididymal function. In: Robaire B, Hinton BT (eds.), The Epididymis: From Molecules to Clinical Practice. New York: Kluwer Academic Press/Plenum Press Publishers; 2002:169–199.
3. Turner TT. On the epididymis and its role in the development of the fertile ejaculate. J Androl 1995; 16:292–298.
4. Visconti PE, Bailey JL, Moore GD, Pan D, Olds-Clarke P, Kopf GS. Capacitation of mouse spermatozoa. I. Correlation between the capacitation state and protein tyrosine phosphorylation. Development 1995; 121:1129–1137.
5. Turner TT, Hinton BT. The Third International Conference on the Epididymis. Charlottesville, VA: The Van Doren Company; 2003.
6. Jervis KM, Robaire B. Dynamic changes in gene expression along the rat epididymis. Biol Reprod 2001; 65:696–703.
7. Chauvin TR, Griswold MD. Androgen-regulated genes in the murine epididymis. Biol Reprod 2004; 71:560–569.
8. Setchell B, Brooks D. Anatomy, vasculature, innervation, and fluids of the male reproductive tract. In: Knobil E, Neil J (eds.), The Physiology of Reproduction, vol. 1. New York: Raven Press; 1988:753–836.
9. Abou-Haila A, Fain-Maurel MA. Regional differences of the proximal part of mouse epididymis: morphological and histochemical characterization. Anat Rec 1984; 209:197–208.
10. Kirchhoff C. Gene expression in the epididymis. Int Rev Cytol 2002; 188:133–202.
11. Bomgardner D, Hinton BT, Turner TT. 5′ hox genes and meis 1, a hox-DNA binding cofactor, are expressed in the adult mouse epididymis. Biol Reprod 2003; 68:644–650.
12. Ezer N, Robaire B. Gene expression is differentially regulated in the epididymis after orchidectomy. Endocrinology 2003; 144:975–988.
13. Cooper TG, Wagenfeld A, Cornwall GA, Hsia N, Chu ST, Orgebin-Crist MC, Drevet J, Vernet P, Avram C, Nieschlag E, Yeung CH. Gene and protein expression in the epididymis of infertile c-ros receptor tyrosine kinase-deficient mice. Biol Reprod 2003; 69:1750–1762.
14. Hsia N, Cornwall GA. DNA microarray analysis of region-specific gene expression in the mouse epididymis. Biol Reprod 2004; 70:448–457.
15. Turner TT, Bomgardner D, Jacobs JP, Nguyen QA. Association of segmentation of the epididymal interstitium with segmented tubule function in rats and mice. Reproduction 2003; 125:871–878.
16. Alter O, Brown PO, Botstein D. Singular value decomposition for genome-wide expression data processing and modeling. Proc Natl Acad Sci U S A 2000; 97:10101–10106.
17. Pontius J, Wagner L, Schuler G. Unigene: a unified view of the transcriptome. In: The NCBI Handbook. Bethesda, MD: National Center for Biotechnology Information; 2003.
18. Lowe B, Avila HA, Bloom FR, Gleeson M, Kusser W. Quantitation of gene expression in neural precursors by reverse-transcription polymerase chain reaction using self-quenched, fluorogenic primers. Anal Biochem 2003; 315:95–105.
19. Bustin SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J Mol Endocrinol 2000; 25:169–193.
20. Ahuja HS, Tenniswood M, Zakeri ZF. Differential expression of clusterin in the testis and epididymis of postnatal and germ cell deficient mice. J Androl 1996; 17:491–501.
21. Cornwall GA, Orgebin-Crist MC, Hann SR. The CRES gene: a unique testis-regulated gene related to the cystatin family is highly restricted in its expression to the proximal region of the mouse epididymis. Mol Endocrinol 1992; 6:1653–1664.
22. Eberspaecher U, Roosterman D, Kratzschmar J, Haendler B, Habenicht UF, Becker A, Quensel C, Petri T, Schleuning WD, Donner P. Mouse androgen-dependent epididymal glycoprotein CRISP-1 (DE/AEG): isolation, biochemical characterization, and expression in recombinant form. Mol Reprod Dev 1995; 42:157–172.
23. Drevet J. Regulation of gene expression in the epididymis. In: Robaire B, Chemes H, Morales C (eds.), Andrology in the 21st Century. Englewood, NJ: Medimond Publishing Co.; 2001.
24. Viger RS, Robaire B. The mRNAs for the steroid 5 alpha-reductase isozymes, types 1 and 2, are differentially regulated in the rat epididymis. J Androl 1996; 17:27–34.
25. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95:14863–14868.
26. Bomgardner D, Hinton BT, Turner TT. Hox transcription factors may

play a role in regulating segmental function of the adult epididymis. J Androl 2001; 22:527–531.
27. Turner TT, Bomgardner D, Jacobs JP. Sonic hedgehog pathway genes are expressed and transcribed in the adult mouse epididymis. J Androl 2004; 25:514–522.
28. Hinton BT, Kirby JL, Rodriguez CM, Lye RJ, Traon BV, Yang L. Signal transduction pathways to gene expression. In: Hinton BT, Turner TT (eds.), The Third International Conference on the Epididymis. Charlottesville, VA: The Van Doren Company; 2003:103–113.