Major Molecular Differences Between Mammalian Sexes are involved in Drug Metabolism and Renal Function

John L. Rinn[1*], Joel S. Rozowsky[1*], Ian J. Laurenzi[1], Petur H. Petersen[2], Kaiyong Zou, Weimin Zhong[2], Mark Gerstein[1], Michael Snyder[1,2]

1 Department of Molecular Biophysics and Biochemistry,
2 Department of Molecular, Cellular and Developmental Biology,
Yale University, New Haven, Connecticut 06520-8005, USA

Summary

Many obvious anatomical differences exist between males and females. These sexual dimorphisms are further manifested on a molecular level by different hormonal and chemical environments, many of which are well-studied. However, a comprehensive investigation of the gene expression differences between males and females has not been performed. In this study we utilize DNA microarray technology to survey the expression of 13,977 mouse genes in male and female hypothalamus, kidney, liver and reproductive tissues. We observed extensive differential gene expression not only in the reproductive tissues, but also in the kidney and liver. The majority of differentially expressed genes are involved in drug and steroid metabolism, osmotic regulation or as yet unresolved cellular roles. In contrast to the kidney and liver we observed very few molecular differences between the male and female hypothalamus in both mice and humans. We conclude that there are persistent differences in gene expression between adult males and females in the kidney and liver. Furthermore, these molecular differences have important implications for the physiological differences between males and females.

Introduction

Most mammals exhibit obvious phenotypic differences between the male and female sexes, and many of the hormonal, chemical and anatomical differences between males and females have been well investigated. We expect that the hormonal and chemical differences between males and females should ultimately result in differential gene expression, which in turn should control mammalian behavior and physiology. The extent to which genes are differentially expressed in male and female adult tissues has not been investigated previously on a systematic level.

To date only a relatively modest number of differences in gene expression between sexes have been documented, largely through the analysis of individual genes. For example, several cytochrome p450 genes have been found to exhibit sex-specific expression in the liver (Anderson, 2002; Tullis et al., 2003). However the extent this occurs in other tissues has not been fully investigated in a comprehensive manner. Moreover, it remains to be determined what other types of genes are differentially expressed by sex. Such information is expected to help elucidate basic physiological difference between sexes and may provide important insights into the documented sex-specific propensities to diseases such as anemia, hypertension and renal dysfunction (Belperio and Rhew, 2004; Depner, 2003; Schwartz, 2003; Thompson and Khalil, 2003).

To gain a better understanding of the molecular differences between mammalian sexes, we have utilized DNA microarrays to identify differences in the adult male and female

transcriptomes. The hypothalamus, kidney, liver and reproductive tissues were analyzed because of their importance in control of behavior (hypothalamus), physiology (kidney and liver) and reproduction (gonads). We found many significant differences in gene expression in the kidney, liver and reproductive tissues. A majority of the genes differentially expressed in the kidney and liver are involved in drug and steroid metabolism or osmotic regulation, raising important implications for the differential abilities of males and females to respond to drugs or acquire hypertension. In contrast to the liver and kidney very few gene expression differences we observed in the adult male and female hypothalamus.

Results

Surveying the Molecular Differences Between Sexes

We have investigated patterns of gene expression in different tissues and sexes of isogenic Swiss Webster mice using DNA microarrays. Five adult mouse tissues liver, kidney, hypothalamus, ovary and testis, were studied. For each somatic tissue, triple selected poly-A mRNA was prepared from 6 independent pools (biological replicates), 3 male and 3 female. Likewise, 3 pools were prepared from the ovary and 3 pools from the testis. Each pool of RNA was derived from 10 individuals. Furthermore, for each biological replicate two cRNAs (technical replicates) were prepared and independently hybridized to Affymetrix MOE430A chips, designed to monitor the expression of 13,997 unique genes and expressed sequence tags (ESTs). We also compared expression patterns between the human and mouse hypothalamus. In total we probed 71 microarrays resulting in over 1,500,000 gene expression data points.

To determine the consistency of our mouse data we calculated correlation coefficients (r) of both biological and technical replicates. The average correlation values (r) of biological and technical replicates ranged from 0.96 to 0.99 across the five tissues. Thus, our data are highly reproducible, presumably because the variation between biological samples was minimized due to the number of individuals represented in each RNA pool. This data set is publicly available to the research community via NCBI GEO (Accession numbers: GSE1147, GSE1148).

To asses the biological accuracy of our data set we hierarchically clustered all expression values from the 48 mouse hybridizations (Figure 1). As expected, samples within a tissue were more similar than samples from different tissues and thus clustered accordingly. To further validate the biological relevance of the data set we identified genes that were highly expressed in one and only one tissue (Figure 1B, Experimental Procedures). Presumably these genes demonstrating tissue-specific expression are important for the basic physiology of that tissue. We identified a total of 1,631 genes with tissue specific expression (Figure 1B: 152 for kidney, 342 for hypothalamus, 257 for liver, 104 for ovary and 776 for testis. As expected, genes with tissue-specific expression have functions typical of the tissue in which they were expressed (hypothalamus: neurogenesis and synaptic transmission genes, kidney: vitamin cofactor and sodium ion transport genes, liver: xenobiotic metabolism and lipid transport genes, testis: spermatogenesis genes and ovary: steroid metabolism genes. Thus, our microarray data is not only precise, but also biologically accurate.

Sex-specific Gene Expression in the Somatic Tissues

We first examined genes with sex-specific expression patterns in somatic tissues. We searched for two types of differential expression patterns in the hypothalamus, liver and kidney: I) expression that occurs in both males and females but is significant higher in one sex relative to the other, II) expression detected only in one sex and not the other.

To identify type I signatures we used Analysis of Variance (ANOVA) to determine genes that are differentially expressed with 99.9% confidence ($P < .001$). This analysis was only applied to probes considered to have significant hybridization signals using the MAS5.0 algorithm. We further required a minimum of a 3-fold differential expression between sexes. A probe was considered to have a type II pattern if the genes had significant hybridization in 5 of 6 samples and insignificant hybridization in all or all but one probe for the other sex. A total of 20 and 19 unique genes and ESTs demonstrated Type I and Type II sex-specific expression, respectively (Figures 2, 4 and 5). Five genes exhibited sex-specific expressed in all tissues and 27, 6 and 1 genes were differentially expressed uniquely in the kidney, liver and hypothalamus, respectively.

To independently verify the sex-specific expression 14 genes (33% of total) that exhibited differential expression in the kidney or liver were randomly selected and subjected to absolute-quantification-real-time PCR using the relevant RNAs. Each of the 14 exhibited strong sex-specific expression identical or similar to that observed for in the microarray data (Figures 2b, 4b, 5b). The actual quantity differences from these

experiments are listed in supplemental Table 1. Thus, we conclude that our microarray data are highly accurate and representative of normal adult female and male physiology.

The five sex-specific genes expressed in all three tissues (liver, kidney and hypothalamus) are Xist, DBY, SMCY, Eif3ay and Uty (Figures 4, 5). The latter 4 genes are encoded on the Y chromosome and are believed to be involved in tissue sexual dimorphisms (van Abeelen et al., 1989). Xist is involved in X inactivation and is typically only expressed in females. Each of these genes is documented to have sex-specific expression patterns and thus serve as internal controls for our approach.

Molecular Differences Between the Female and Male Kidney

A total of 27 genes were found to have sex-specific expression in the kidney (Figure 2). These genes primarily belong to three categories: 1) Drug and steroid metabolism 2) Osmotic regulation and 3) Uncharacterized.

Nine or one-third of the genes with sex-specific expression in the kidney are involved in drug and steroid metabolism. The dominant class encodes 5 cytochrome p450 family members, Cyp7b1, Cyp2d9, Cyp4a12, Cyp2e1, and Cyp2j13. Cytochrome p450s are a family of proteins important for drug metabolism. Interestingly, each of the sex-specific cytochrome p450s were preferentially expressed in the male kidney; in fact the largely uncharacterized Cyp2j13 to be exclusively expressed in the male kidney. In addition to genes involved in drug and steroid metabolism we observed sex-specific expression in three transferases (Ugt2b5, Ugt8 and Kat2) and the corticosteroid-binding globulin.

Together these results indicate an abundance of genes differentially expressed by sex in the kidney and many cytochrome p450s are preferentially expressed in the male kidney.

Six genes exhibiting sex-specific expression are likely to be involved in osmotic regulation of the kidney. The prolactin receptor precursor (Prlr) is primarily known for its in mammary tissue development (Bole-Feysot et al., 1998). However, Prlr also has a clear role in osmotic regulation in lower vertebrates (Brown et al., 1986; Ogawa et al., 1973) and thus may also be involved in mammalian osmo-regulation (Bole-Feysot et al., 1998). We observed a five-fold enrichment of Prlr expression in the female kidney, indicating this gene may have a sex-specific role in osmo-regulation. We also found that the Rat hypertension homologue (SaH) is expressed 10-fold more in the male kidney compared to the female kidney. Sah has been previously been linked to hypertension susceptibility in rats (Iwai et al., 1992). Three uncharacterized organic anion transporters, Slc21a1, Slc7a12 and mOATL-6, were all found to have male (Slc21a1 and Slc7a12) and female (mOATL-6) specific expression respectively. Slc21a1 is encoded on the X chromosome, Slc7a12 is similar to Slc21a1 and mOATL-6 is a predicted organic anion solute carrier. Cyp4a12, a drug and steroid metabolism gene, was preferentially expressed in both the male kidney and liver. Cyp4a12 is known to metabolize arachidonic acid into hydroxyeicosatetraenoic acids (HETEs), which affects smooth muscle tone and ultimately blood pressure. The expression pattern of Cyp4a12 suggests production of HETEs may be regulated in a sex-specific fashion.

A sixth gene implicated in osmotic regulation is the corticosteroid-binding globulin (Cbg). Cbg is a glucocorticoid binding protein that functions as the major transporter of glucocorticoids such as cortisol and progestins into the bloodstream of most vertebrates (Scrocchi et al., 1993). Cbg is known to be developmentally important and a familial null deletion is associated with hypertension and fatigue (Seralini, 1996; Tropy et al., 2001). However, Cbg has not been shown previously to have sex-specific expression; we found that it is expressed in the female but not the male kidney.

The remaining six genes, demonstrating sex-specific expression in the kidney are largely uncharacterized. Gc, Xat, MGC18894, NM_144930, Timd2, 0610033EO6Rik have all been identified by ESTs and cDNA analysis, but the genes have no documented function. A majority of them are highly expressed in at least one sex; therefore, we expect them to serve an important biological role in either males or females.

Cbg mRNA Localizes to the Cortico Medullary Junction in a Female Specific Fashion

In order to further investigate the female-specific role of Cbg in the kidney we determined the localization of Cbg mRNA *in vivo* using RNA in situ hybridization (RISH). Kidney sections were prepared from both males and females and hybridized with a probe complementary to Cbg mRNA. This analysis revealed that Cbg is specifically expressed in the corticomedullary junctions and only in females (Figure 3). This female-specific corticomedullay expression pattern was observed in three separate animal

preparations. Staining was not observed when a noncomplementary probe was used (data not shown). These results indicate that Cbg expression is cell-specific and also suggest that Cbg is expressed exclusively in the female collecting ducts, an important site for controlling osmotic pressure in the kidney.

Molecular Differences Between the Female and Male Liver

A majority of the 6 genes with sex-specific expression in the liver are involved in drug and steroid metabolism (Figure 4). Three are cytochrome p450s: Cyp4a12, Cyp2b13 and Cyp3a16. Cyp4a12 also exhibited male-specific expression in the kidney. Cyp2b13 is induced by phenobarbitol and has been documented as a female-specific testosterone dehydrogenase (Lakso et al., 1991). Cyp3a16 was found to be expressed only in the female liver, but not in the male. In addition to the cytochrome p450s we also found Hsd3b5, to exhibit male-specific expression in the liver. Wong et al. observed a similar male specific expression pattern for this gene in C57BL/6 mice (Wong and Gill, 2002). Thus, similar to the kidney, a large fraction of the genes differentially expressed between sexes in the liver are involved in drug and steroid metabolism.

Sex-specific differences in the Mouse Hypothalamus

The hypothalamus exhibited very little sex-specific transcription (Figure 5). Other than the five sex-specific genes expressed in all tissues, only one additional gene, TSIX, was found to be sex-specific in the hypothalamus. As its name implies this gene is inversely transcribed to Xist and it is expressed only in the female. The low number of differentially expressed genes in the hypothalamus is surprising considering the

hypothalamus has well documented sexual dimorphisms, such as the Sexually Dimorphic Region of the Pre Optic Area of the hypothalamus (Goldstein et al., 2001).

Sex-specific differences in the Human Hypothalamus

Because the number of sex-specific genes identified in the mouse hypothalamus was low, we set out to determine if this was the case in humans, which we postulated might have different male and female gene expression. To this end we obtained postmortum hypothalami from 7 males and 5 females, prepared RNAs from each of sample and hybridized cRNA to human U133 microarrays representing 13,624 genes and ESTs. Technical replicates were performed for each RNA sample, except for one female sample for which only one hybridization was carried out. Type I and Type II signatures from the 23 hybridizations were identified as described. Similar to the mouse, very few genes (10) were found to be differentially expressed in the human male and female hypothalamus using our stringent criterion (Figure 5). Seven genes on the Y chromosome all had male-specific expression. Xist and Tsix were the only two genes with female-specific expression in the human hypothalamus. In summary, the only genes demonstrating sex-specific expression in the hypothalamus are either encoded on the Y chromosome or involved in X chromosome inactivation. However, these genes do serve as a positive control for the microarray approach to survey molecular differences between sexes.

Molecular differences in the Gonads

We also investigated sex-specific differences in gene expression among the reproductive tissues. Although some transcriptional differences between the gonads have been well

studied, a comprehensive analysis of differential expression between the male and female gonads has not been reported (reviewed: (Merchant-Larios and Moreno-Mendoza, 2001)).

We first determined which genes were differentially expressed between the ovary and testis using the ANOVA stringency (a P value of .001) and minimum a 3-fold expression level difference between sexes described above. At this threshold over 4,000 genes are differentially expressed between ovary and testis. Increasing stringency to a P value of less than 1 e-6 reveals 882 differentially expressed genes. 534 and 358 had increased expression in the testis and ovary, respectively (Supplemental Tables 2 and 3 respectively).

We searched the Gene Ontology Database (GODB) for gene functions that were enriched in each of the reproductive tissues using GOMiner (Zeeberg et al., 2003). Briefly, GoMiner determines which functional terms or groups are enriched or depleted in the differentially expressed genes. It also gives a statistical measure (P value) of confidence that a given functional category is over or under represented in a differentially expressed gene list. We focused on GO terms with a P value less than 1e-4, the highest confidence level provided by the GOMiner program (Supplementary Table 4).

GOMiner analysis revealed three interesting functional classifications enriched in the 534 genes demonstrating testis-specific expression. First, pol II transcription machinery genes were enriched in the testis-specific list, consistent with the transcriptional demand of

active gametogenesis in the testis. Second, we observed a paucity of gene expression for immune response genes. Of the 677 immune response genes only two were differentially expressed. Chance predicts twenty-five based on the number of these genes in the genome. Finally, 42 signal transduction genes showed a significantly higher level of expression in the testis. However, only 1 of the 332 genes with G-protein coupled receptor (GPCR) activity was differentially expressed in the testis. These results indicate that extensive testis-specific signaling, but not through the induction of GPCRs.

The 358 genes expressed significantly higher in the ovary are enriched ($P < 1e-4$) for one specific functional category: monoxygenases. Monoxygenases are genes that reduce oxygen and transfer the reactive oxygen species to other molecules; many of these are members of the cytochrome p450 family (reviewed: (Nelson et al., 1996)). Thirty-four or 10% of the genes with ovary-specific expression encode monooxygenases; 8 of these are cytochrome p450s that are only expressed in the ovary and not in the testis. Thus, many genes expressed preferentially in the ovary are monooxygenases and specifically cytochrome p450s.

Lastly, we searched for genes that have high, but similar expression levels in both the ovary and testis compared to all the other tissues (i.e. gonad-specific expression). Using a one-way ANOVA (Experimental Procedures) we found only 16 genes that demonstrate increased expression in both the male and female gonads compared to the somatic tissues (Supplemental Table 5). Of particular interest is Foxj1 a forkhead transcription factor

expressed in endothelial tissues (Huang et al 2003). We speculate that Foxj1 transcribes gonad-specific genes.

Discussion

In this manuscript we have investigated the molecular differences between females and males by surveying the expression of 13,977 genes across different tissues and sexes. Our findings demonstrate that there are significant gene expression differences between sexes not only in the reproductive tissue but also in the adult kidney and liver. Genes preferentially expressed by sex are mainly involved in drug and steroid metabolism or osmotic regulation in the kidney. In contrast, there are very few molecular differences between the female and male hypothalamus of both mice and men.

We presume that if a gene is differentially expressed in one sex over an average of 3 independent dissections, spread across several months, of 10 mice each, these transcriptional features are highly stable and thus important to the cellular physiology of the respective sex in which they are preferentially expressed. Indeed a total of 39 genes were greater than 3-fold differentially expressed by sex in the kidney, liver and hypothalamus. The most extensive sex-specific expression, 27 genes, was observed in the kidney; however, six other genes are differentially expressed in the liver and five genes are differentially expressed by sex in all tissues consistent with other studies and the nature of their function. Together these results indicate that there is extensive sex-specific expression in the somatic tissues and predominantly in the kidney.

Over one-third, 12, of the 33 genes differentially expressed by sex in the kidney and liver are involved in drug and steroid metabolism. A majority of these, 7, encode cytochrome p450s. These genes are well known for their active role in the liver where they

metabolize drugs and other xenometabolites. Several examples of sex-specific expression of cytochrome p450s have been documented in the liver in response to pollutants or external stimuli, many of which were confirmed in this study as differentially expressed in healthy animals (Anderson, 2002). In addition, we identified sex-specific expression of several new cytochrome p450s, including two uncharacterized members, Cyp2j13 and Cyp3a16. Importantly, our results further demonstrate that there is extensive sex-specific expression of cytochrome p450s in the kidney, a phenomenon typically thought to occur in the liver. In total our study has revealed 15 (17%) of the annotated cytochrome p450s vary in expression between sexes, not only in the liver, but also in the ovary, testis and kidney. Since cytochrome p450s, especially those in the liver and kidney, have crucial roles in drug metabolism, these results have important implications in how males and females differentially metabolize and respond to drug treatments.

A large fraction of sex-specific expression is derived from genes that potentially regulate the osmotic state of the kidney. Cbg is of particular interest since it has a well documented role in osmotic regulation and was expressed in the female kidney, but not the male kidney. Cbg binds corticosteroids such as cortisol and aldosterone, which are know to play an important role sodium and potassium homeostasis. We found that Cbg is expressed specifically in the collecting ducts in the corticomedullary junction of the female kidney (Figure 3). These cells play an important role in sodium reabsorbtion, which affects the osmotic state of the kidney and ultimately impacts blood pressure. Specifically, in the principal cells of the collecting ducts aldosterone binds the

aldosterone receptor and allows more sodium reabsorption. Interestingly, another corticosteroid, cortisol, can also bind the aldosterone receptor. Thus, it is possible that Cbg lowers the intracellular concentration of cortisol in the collecting ducts to dampen cross-reactivity with the aldosterone receptor. Consistent with a role for Cbg in salt metabolism, a recent study of a familial null deletion of Cbg in humans was shown to be associated with hypertension (Torpy et al., 2001). Our observation that Cbg is expressed specifically in these cells and in a female-specific fashion suggest it could have an important role in osmotic regulation of the kidney and thereby affecting hypertension in females.

In addition to Cbg we observed sex-specific expression in five other genes that may differentially affect the osmotic state of the kidney. Prlr has been shown to be a key regulator of osmotic balance in lower vertebrates and potentially could been involved in osmoregulation in the mammalian kidney. Moreover, Prlr could regulate the osmotic environment in a sex-specific manner as its expression patterns suggests. Another gene strongly implicated in osmoregulation is Sah, which is expressed in mice that are susceptible to hypertension (Iwai et al., 1992). We also observed sex-specific expression in three solute carriers that could alter the osmotic environment by transporting organic ions. Together, the differential expression of these genes may explain molecular differences in the osmotic states between male and female kidney cells. Furthermore, these molecular differences may help explain recent reports of sex-specific propensities to hypertension (Port et al., 2000).

Six genes with yet unresolved cellular functions also demonstrate sex-specific expression. Considering more than half of the genes exhibiting sex-specific expression involved in drug and osmoregulation it is likely that many of these genes will also function in these processes. Since these genes are highly and differentially expressed it is likely that they have important cellular roles and also contribute to sex-specific physiologies in their respective tissues.

Both the human and mouse hypothalamus have very few sex-specific genes. We found 6 and 10 genes to be sex-specific in the mouse and human hypothalamus, respectively. In addition, all of these genes are either on the Y chromosome or involved in X chromosome inactivation. Recently Galfalvly et al. published a paper describing a strategy similar to ours, but in a different region (Brodman areas 9 and 47) of the human brain (Galfalvy et al., 2003). The sex-specific genes identified in these two studies are in strong agreement. Seven of the genes we determined to be sex-specific were also present in their list of nine genes. Thus, two independent studies found very few genes to be differentially expressed by sex in the human adult brain. In conclusion, the sexual differences in the adult hypothalumus are mediated by a small number of genes, is transient, or occurs predominantly at the protein level.

The male and female reproductive tissues exhibited the largest amount of sex-specific expression. This is largely expected, since these tissues are sexually dimorphic. However, several interesting patterns in the types of genes that are differentially expresses were apparent. For example, the testis is immuno-suppressed on a transcriptional level and the

ovary was enriched for genes that are involved in drug and steroid metabolism. In contrast to the large transcriptional differences between the gonads, we found 16 genes with high and similar expression in both gonads compared to all other somatic tissues. Of particular note is the forkhead transcription factor Foxj1. Although, this gene is primarily expressed in ciliated cells (Huang et al. 2003), we observed Foxj1 enriched expression in the brain as well as the reproductive tissues. Brain and reproductive tissues also have similarities in molecular signaling mechanisms (i.e. growth factors and hormones) (Marchetti et al., 1996). Thus, these shared signaling molecules may elicit similar changes in gene expression and potentially are regulated by Foxj1.

In this study we were limited to using whole tissues from each sex, we were also limited to one developmental time pt. Currently we are investigating the molecular differences between sexes at multiple developmental stages (e.g. puberty) as well as the molecular differences in sub-populations of cells within a tissue.

In summary, this study presents a comprehensive analysis of tissue and sex-specific expression patterns in the five healthy mammalian tissues. Our results show that there is extensive sex-specific expression in not only the reproductive tissues, but also the somatic tissues. Moreover these transcriptional differences are observed well after puberty indicating they are important to the routine physiological differences between mature males and females.

Experimental Procedures

RNA Sample Preparation and Microarray Hybridization

For mouse liver, kidney, ovary and testis triple selected poly A+ mRNA was purchased from Ambion (Ambion, TX). For human and mouse hypothalamus total RNA was purchased from the same vendor. All RNA samples had a ratio of 28S ribosomal RNA/16S ribosomal RNA greater than 1 indicating that significant degradation had not occurred. Each RNA sample was further treated with DNase using the "Poly-A pure kit" from Ambion. Affymetrix arrays were hybridized according to Affymetrix protocols. Details can be found at: http://keck.med.yale.edu/affymetrix/protocols/

Analysis of microarray data

Data was extracted from the Affymetrix Mouse MOE430A microarrays using MAS 5.0. As suggested by Affymetrix, spots with associated p-values greater than 0.05 are considered to have insignificant hybridization and are flagged as lacking signal, those with p-values less than .05 were flagged present as exhibiting significant expression. In order to test the reproducibility of this data linear regression was performed between biological and technical replicates for a given tissue and gender. Linear regression was performed between pairs of slides using Affymetrix IDs common to both microarrays. Each individual slide was normalized to its respective median intensity. The linear regression was again performed between slides on the normalized data yielding near identical regression coefficients (.96-.99) to the unnormalized data. The normalized data is used in all the following analyses. We also choose to treat both biological and technical replicates as independent data measurements since technical and biological replicates had

similar consistency. All the microarray data can be found at NCBI GEO (Accession numbers: GSE1147, GSE1148).

Identification of Tissue-specific Genes

Using the normalized microarray data we compared for each Affymetrix ID, the data of a specific tissue pair wise against each other tissue using a one-way ANOVA (t-test). Only features flagged present were used in this analysis and we required at least two data points for each tissue in order to make the comparison. We selected the Affymetrix IDs corresponding to genes expressed higher in one tissue relative to all others by requiring that the p-value for each t-test (where possible) be less than 0.001 and that there be at least a 3 fold increase in the expression intensity of the over-expressed gene in the specific tissue. This was done for all pair wise tissue combinations. If these criterion were met in all comparisons we considered that gene to demonstrate tissue-specific expression.

Identification of sex-specific genes

Type I sex-specific expression – ANOVA (t-test): Using the normalized data for a specific tissue we performed a t-test between male and female expression levels for each Affymetrix ID. We are only able to do this t-test if there are at least two spots flagged present for both male and female microarrays (absent spots are excluded). We then selected only the Affymetrix IDs corresponding to genes with p-values less than 0.001. For all tissues we also required a expression ratio of at least 3 for one gender compared to the other. We also compared the expression data between testis and ovary in a similar

fashion but in order to get a manageable list, a threshold the p-value at 1e-6 was established in addition to a minimum expression ratio of 3.

Type II sex-specific expression: Since for each tissue (Hypothalamus, Kidney and Liver) was analyzed by 6 microarrays for each gender, we required a given Affymetix Ids to be flagged present in 5 out of 6 samples in one sex and the same ID is flagged absent in all or all but one samples of the other sex. We also require that the average intensity of present spots be at least a factor of 3 compared to the average intensity of absent.

Analysis of Gonad-specific expression

Using the normalized microarray data we performed a t-tests between the expression data for testis and ovary and require that the p-value be greater than 0.1 and the ratio of mean intensities be between 0.5 and 2. We then combine the testis and ovary data and use a t-test to compare the combined data pairwise against each other tissue. We require the p-value for each of the comparisons to be less than 0.01 and that the average intensity of the combined testis and ovary data be a factor of at least 3 greater than that for each other tissue. We only perform this comparison for Affymetrix IDs where there are at least two good data point for each of the sets being compared.

RNA In Situ Hybridization

Three female and three male Swiss-Webster mice kidney sections were prepared to determine the localization of Cbg mRNA. Probes were *in vitro* transcribed from Invitrogen clone: 5100884. Both sense and anti-sense probes were transcribed and

hybridized to tissue preparations. All animal and probe preparation were performed as described by: Petersen et al, 2002

Real Time Quantitative PCR

Each of the samples hybridized to the human HGU133A GeneChip® (Affymetrix, CA) was reverse transcribed using Applied Biosytems (ABI, CA) "high capacity cDNA archive kit." These cDNAs were subsequently mixed with ABIs' "SYBR Green® PCR Master Mix," for Real-Time PCR. These primers were further used to amplify two concentrations, 15ng and 30ng, of cDNAs prepared from each of the experimental RNA pools, each in duplicate. A specific quantity was calculated for each concentration from the average of each replicate. These specific quantities were normalized within a pool by dividing each amplification by the specifc quantity of Gapdh to give a relative quantity. Finally, the relative quantities of males and females were compared. A similar comparison was performed for the mouse samples using a relative Crtical threshold (Ct) comparison. A average Ct was calculated for the triplicate reactions and normalized to Gapdh ($\Delta Ct = X - Gapdh$ {X = Ct for a given sex and gene target}). The normalized Cts were then compared between sexes ($\Delta\Delta Ct = M-F$). Finally a fold change (M/F) was calculated from the $\Delta\Delta Ct$ ($M/F = 2^{\Delta\Delta Ct}$). More details can be found at http://www.appliedbiosystems.com/

Acknowledgments

We would like to thank Shrikant Mane and Sheila Westman at the Keck Facility Affymetrix Center for assistance with microarray hybridizations. Sherman Weissman,

Michael Kashgarian and Rick Lifton provided helpful discussions. This research was supported by research and training grants from the N.I.H.

References

Anderson, G. D. (2002). Sex differences in drug metabolism: cytochrome P-450 and uridine diphosphate glucuronosyltransferase. J Gend Specif Med *5*, 25-33.

Belperio, P. S., and Rhew, D. C. (2004). Prevalence and outcomes of anemia in individuals with human immunodeficiency virus: a systematic review of the literature. Am J Med *116 Suppl 7A*, 27S-43S.

Bole-Feysot, C., Goffin, V., Edery, M., Binart, N., and Kelly, P. A. (1998). Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. Endocr Rev *19*, 225-268.

Brown, S. C., Horgan, E. A., Savage, L. M., and Brown, P. S. (1986). Changes in body water and plasma constituents during bullfrog development: effects of temperature and hormones. J Exp Zool *237*, 25-33.

Depner, T. A. (2003). Prescribing hemodialysis: the role of gender. Adv Ren Replace Ther *10*, 71-77.

Galfalvy, H. C., Erraji-Benchekroun, L., Smyrniotopoulos, P., Pavlidis, P., Ellis, S. P., Mann, J. J., Sibille, E., and Arango, V. (2003). Sex genes for genomic analysis in human brain: internal controls for comparison of probe level data extraction. BMC Bioinformatics *4*, 37.

Goldstein, J. M., Seidman, L. J., Horton, N. J., Makris, N., Kennedy, D. N., Caviness, V. S., Jr., Faraone, S. V., and Tsuang, M. T. (2001). Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. Cereb Cortex *11*, 490-497.

Iwai, N., Kurtz, T. W., and Inagami, T. (1992). Further evidence of the SA gene as a candidate gene contributing to the hypertension in spontaneously hypertensive rat. Biochem Biophys Res Commun *188*, 64-69.

Lakso, M., Masaki, R., Noshiro, M., and Negishi, M. (1991). Structures and characterization of sex-specific mouse cytochrome P-450 genes as members within a large family. Duplication boundary and evolution. Eur J Biochem *195*, 477-486.

Marchetti, B., Gallo, F., Farinella, Z., Romeo, C., and Morale, M. C. (1996). Luteinizing hormone-releasing hormone (LHRH) receptors in the neuroendocrine-immune network. Biochemical bases and implications for reproductive physiopathology. Ann N Y Acad Sci *784*, 209-236.

Merchant-Larios, H., and Moreno-Mendoza, N. (2001). Onset of sex differentiation: dialog between genes and cells. Arch Med Res *32*, 553-558.

Nelson, D. R., Koymans, L., Kamataki, T., Stegeman, J. J., Feyereisen, R., Waxman, D. J., Waterman, M. R., Gotoh, O., Coon, M. J., Estabrook, R. W*., et al.* (1996). P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics *6*, 1-42.

Ogawa, M., Yagasaki, M., and Yamazaki, F. (1973). The effect of prolactin on water influx in isolated gills of the goldfish, Carassius auratus L. Comp Biochem Physiol A *44*, 1177-1183.

Petersen, P. H., Zou, K., Hwang, J. K., Jan, Y. N., and Zhong, W. (2002). Progenitor cell maintenance requires numb and numblike during mouse neurogenesis. Nature *419*, 929-934.

Port, S., Demer, L., Jennrich, R., Walter, D., and Garfinkel, A. (2000). Systolic blood pressure and mortality. Lancet *355*, 175-180.

Schwartz, J. B. (2003). Gender-specific implications for cardiovascular medication use in the elderly optimizing therapy for older women. Cardiol Rev *11*, 275-298.

Scrocchi, L. A., Hearn, S. A., Han, V. K., and Hammond, G. L. (1993). Corticosteroid-binding globulin biosynthesis in the mouse liver and kidney during postnatal development. Endocrinology *132*, 910-916.

Seralini, G. E. (1996). Regulation factors of corticosteroid-binding globulin: lesson from ontogenesis. Horm Res *45*, 192-196.

Thompson, J., and Khalil, R. A. (2003). Gender differences in the regulation of vascular tone. Clin Exp Pharmacol Physiol *30*, 1-15.

Torpy, D. J., Bachmann, A. W., Grice, J. E., Fitzgerald, S. P., Phillips, P. J., Whitworth, J. A., and Jackson, R. V. (2001). Familial corticosteroid-binding globulin deficiency due to a novel null mutation: association with fatigue and relative hypotension. J Clin Endocrinol Metab *86*, 3692-3700.

Tullis, K. M., Krebs, C. J., Leung, J. Y., and Robins, D. M. (2003). The regulator of sex-limitation gene, rsl, enforces male-specific liver gene expression by negative regulation. Endocrinology *144*, 1854-1860.

van Abeelen, J. H., Janssens, C. J., Crusio, W. E., and Lemmens, W. A. (1989). Y-chromosomal effects on discrimination learning and hippocampal asymmetry in mice. Behav Genet *19*, 543-549.

Wong, J. S., and Gill, S. S. (2002). Gene expression changes induced in mouse liver by di(2-ethylhexyl) phthalate. Toxicol Appl Pharmacol *185*, 180-196.

Zeeberg, B. R., Feng, W., Wang, G., Wang, M. D., Fojo, A. T., Sunshine, M., Narasimhan, S., Kane, D. W., Reinhold, W. C., Lababidi, S.*, et al.* (2003). GoMiner: a resource for biological interpretation of genomic and proteomic data. Genome Biol *4*, R28.

Figure Legends

Figure 1: Hierarchical clustering of microarray data Each gene is represented by 48 individual hybridization measurements or probe intensities across the five tissues. The probe intensities were normalized within a chip to the median intensity of all probes on that array (intra-chip normalization). The median value of the 48 normalized probe intensities (i.e. the global median) thus serves as a reference to compare the expression of each gene across experiments (inter-chip comparisons). Probe intensities above and below the global median are denoted by shades of red or blue, respectively; those at the global median are colored yellow. B) As expected, expression profiles from a given tissue clustered tightly together, whereas expression profiles across different tissues exhibited wider variation. Genes demonstrating tissue specific expression were representative of the normal physiology of the tissue in which they are expressed, thus confirming the biological accuracy of our data.

Figure 2: Extensive sex-specific expression in the kidney. A) Twenty-seven genes were differentially expressed by sex in the kidney. One-quarter, 7, of these genes are drug and steroid metabolism genes. The rest are mainly comprised of osmo-regulation genes or genes with yet unresolved cellular roles. Shades of blue represent the degree to which expression in the male is lower than the female median expression level (female specific expression). Shades of red indicate the degree to which expression in the male is above the female median expression level (male specific expression). The brightness of the color represents the amount of expression. P values of differential expression between sexes are also listed. B) We independently verified the expression of 7 genes in the

kidney using quantitative-real-time PCR. Each reaction was performed in triplicate. ΔRT represents the amount of pooled cDNA was amplified in the reaction per cycle. 100% of the randomly selected genes demonstrated the same sex-specific expression pattern observed in the microarray data.

Figure 3: Cbg mRNA localizes to the proximal tubules in a female-specific fashion. Cbg mRNA expression was analyzed *in vivo* using RNA *In Situ* Hybridization (RISH) using sense and antisense probes to Cbg to female (left) and male (right) kidney tissue preparations. In corroboration with the microarray data, only the female kidney preparation showed Cbg mRNA staining. Also the staining suggests that Cbg is specifically expressed in the corticomedullary junction, a site important for osmotic regulation. Together these results suggests that Cbg may be playing a sex-specific role osmotic regulation.

Figure 4: Sex-specific expression in the liver. A) Six genes were differentially expressed by sex in the liver. A majority of which are drug and steroid metabolism genes. Shades of blue represent the degree to which expression in the male is below the female median expression level (female specific expression). Shades of red indicate the degree to which expression in the male is above the female median expression level (male specific expression). The brightness of the color represents the amount of expression. P values of differential expression between sexes are also listed. B) We independently verified 3 genes demonstrating sex-specific expression in the liver quantitative-real-time PCR. Each reaction was performed in triplicate. ΔRT represents the amount of cDNA amplified in

the reaction per cycle. 100% of the randomly selected genes demonstrated the same sex-specific expression pattern observed in the microarray data.

Figure 5: Sex-specific expression in the mouse and human hypothalamus. A) Few genes exhibited sex-specific expression in the hypothalamus of mice and men. All genes demonstrating sex-specific expression in the mouse and human hypothalamus, were either encoded on the Y chromosome or involved in X chromosome inactivation. Shades of blue represent the degree to which expression in the male is below the female median expression level (female specific expression). Shades of red indicate the degree to which expression in the male is above the female median expression level (male specific expression). .* TSIX met our stringent criterion in the hypothalamus and ovaries but was slightly under our criterion in the liver and kidney. B) We independently verified 4 genes demonstrating sex-specific expression in human hypothalamus using quantitative-real-time PCR. Each reaction was performed in triplicate. ΔRT represents the amount of cDNA amplified in the reaction per cycle. 100% of the randomly selected genes demonstrated the same sex-specific expression pattern observed in the microarray data.