LCN6, a novel human epididymal lipocalin
Katherine G Hamil†1,3, Qiang Liu†1,3,7, P Sivashanmugam2,3,8, M Anbalagan1,3,9, Suresh Yenugu1,3, Rama Soundararajan1,3,10, Gail Grossman2,3, AJ Rao4, Charles E Birse5, Stephen M Ruben5,11, Richard T Richardson2,3, Yong-Lian Zhang6, Michael G O’Rand2,3, Peter Petrusz2,3, Frank S French1,3 and Susan H Hall*1,3

Address: 1Department of Pediatrics, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599, USA, 2Department of Cell and Developmental Biology, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599, USA, 3Laboratories for Reproductive Biology, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599, USA, 4Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India, 5Human Genome Sciences, Inc, Rockville, Maryland 20850, USA, 6State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China, 7Present address: State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Science, Chinese Academy of Sciences, Shanghai 200031, China, 8Present address: Department of Urology, Duke University, Durham, North Carolina 27708, USA, 9Present address: Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India, 10Present address: Department of Medicine, University of California, San Francisco 94143, USA and 11Present address: Celera Genomics, Rockville, Maryland 20850, USA

Email: Katherine G Hamil – khamil@med.unc.edu; Qiang Liu – qliu@sibs.ac.cn; P Sivashanmugam – sivas001@mc.duke.edu; M Anbalagan – anbu@biochem.isc.ernet.in; Suresh Yenugu – suresh1@med.unc.edu; Rama Soundararajan – ramas_18in@yahoo.com; Gail Grossman – gghgrossm@med.unc.edu; AJ Rao – ajrao@biochem.isc.ernet.in; Charles E Birse – charlie_birse@bpsci.com; Stephen M Ruben – steve.ruben@celera.com; Richard T Richardson – rtrich@med.unc.edu; Yong-Lian Zhang – yonglz@sunm.shcnc.ac.cn; Michael G O’Rand – morand@unc.edu; Peter Petrusz – petrusz@med.unc.edu; Frank S French – fsfrench@med.unc.edu; Susan H Hall* – shh@med.unc.edu

* Corresponding author †Equal contributors

Published: 14 November 2003
Received: 21 October 2003
Accepted: 14 November 2003

Reproductive Biology and Endocrinology 2003, 1:112

Abstract

Background: The lipocalin (LCN) family of structurally conserved hydrophobic ligand binding proteins is represented in all major taxonomic groups from prokaryotes to primates. The importance of lipocalins in reproduction and the similarity to known epididymal lipocalins prompted us to characterize the novel human epididymal LCN6.

Methods and Results: LCN6 cDNA was identified by database analysis in a comprehensive human library sequencing program. Macaca mulatta (rhesus monkey) cDNA was obtained from an epididymis cDNA library and is 93% homologous to the human. The gene is located on chromosome 9q34 adjacent LCN8 and LCN5. LCN6 amino acid sequence is most closely related to LCN5, but the LCN6 beta-barrel structure is best modeled on mouse major urinary protein 1, a pheromone binding protein. Northern blot analysis of RNAs isolated from 25 human tissues revealed predominant expression of a 1.0 kb mRNA in the epididymis. No other transcript was detected except for weak expression of a larger hybridizing mRNA in urinary bladder. Northern hybridization analysis of LCN6 mRNA expression in sham-operated, castrated and testosterone replaced rhesus monkeys suggests mRNA levels are little affected 6 days after castration. Immunohistochemical staining revealed that LCN6 protein is abundant in the caput epithelium and lumen. Immunofluorescent staining of human spermatozoa shows LCN6 located on the head and
tail of spermatozoa with the highest concentration of LCN6 on the post-acrosomal region of the head, where it appeared aggregated into large patches.

**Conclusions:** LCN6 is a novel lipocalin closely related to Lcn5 and Lcn8 and these three genes are likely products of gene duplication events that predate rodent-primate divergence. Predominant expression in the epididymis and location on sperm surface are consistent with a role for LCN6 in male fertility.

**Background**
Sperm maturation in epididymis requires successive sperm surface modifications that promote the development of capacities for directional swimming and egg fertilization [1]. Although numerous secreted epididymal proteins that interact with spermatozoa have been identified and characterized, our understanding of sperm maturation is far from complete [2,3]. Possible lipocalin involvement in sperm maturation was highlighted by recent reports describing the spermatozoan-association of Lcn5 also known as mouse epididymal protein 10 (Mep10) and epididymal retinoic acid binding protein (Erabp) [4], bull prostaglandin D synthase (PTGDS) [5], rat androgen-regulated secretory protein B [6], lizard lipocalin sperm binding protein [7] and mouse Lcn2/24p3 [8]. Other lipocalins are implicated in reproductive processes including mouse Lcn8 also known as epididy-

\[ \beta \]-barrel is closed at one end by amino acid side chains and open at the opposite end permitting specific entry and binding of retinoids, steroids, terpenoids, fatty acids and a range of aromatic and aliphatic compounds [18]. Ligands are often tightly bound with equilibrium dissociation constants in the 1–100 nanomolar range [20–22] and some in the 1–100 micromolar range [23,24]. The rodent epididymal lipocalins, Lcn5 and Lcn8 bind retinoic acid, a vital regulatory factor in the male reproductive tract [25]. Lcn5 and 8 may transport retinoic acid from the proximal regions of the epididymis to either spermatozoa or epithelium at more distal locations [26]. Rodent lipocalins are abbreviated Lcn and human lipocalins are abbreviated in uppercase, LCN.

To characterize the novel lipocalin LCN6, we analyzed its sequence, expression in epididymis and localization on spermatozoa. We show that LCN6 expression is highly epididymis-specific, but appears to be regulated little if at all by androgen but may be regulated by testis factors in corpus. We demonstrate LCN6 location on spermatozoa, consistent with a role in fertility.

**Materials and Methods**

**DNA Sequencing and Analysis**
DNA was sequenced at the UNC-CH Automated DNA Sequencing Facility using an ABI PRISM Model 377 DNA Sequencer (PE Applied Biosystems {ABI, Foster City, CA}) and the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq(R)DNA Polymerase FS. Primers were synthesized on an automated Applied Biosystems DNA synthesizer Model 394 using standard cyanoethyl phosphoramidite chemistry. Both strands of the original clone were sequenced. Data were initially received into Chromas, version 1.61 (Technelysium Pty. Ltd, Helensvale, Queensland, Australia) and further analyzed using the programs contained in the Wisconsin Package Version 10.1, Genetics Computer Group (GCG), (Madison, Wisconsin). Chromosomal locations were visualized using the Entrez genome Homo sapiens Map Viewer accessible through http://www.ncbi.nlm.nih.gov. Amino acid sequence analyses also utilized the Proteomics Tools at http://www.expasy.ch/ which is the ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (Geneva, Switzerland).

**RNA isolation and analysis**
Total RNA was isolated by the method of Chirgwin et al. [27]. Poly A+ RNA was prepared using the Poly(A) Quik mRNA isolation kit (Stratagene, La Jolla, CA) following the manufacturer’s recommendations or by standard oligo d(T) affinity chromatography. RNAs were denatured...
using glyoxal and Northern hybridizations were performed as previously described [28].

**Isolation of human and monkey cDNAs by RTPCR**

Total RNA (2.5 μg) isolated from *Macaca mulatta* epididymis was reverse transcribed using SuperScript™ reverse transcriptase (GIBCO/BRL, Grand Island, NY) according to the manufacturer's recommendations. The monkey LCN6 cDNA was amplified by PCR using the forward primer F1: CAGTGTAGACCTGGGAGGATG and the reverse primer 13R: GCGGTACCTACTGTGACAGGAAGC at denaturing, reannealing and extension temperatures of 95 C, 45 C and 72 C respectively. Gel purified 187 bp cDNA products were ligated into pGEM-T-easy (Promega) and subjected to automated sequencing using the M13 F and R primers. Three RTPCR clones were sequenced and the consensus reported. In addition, a human caput cDNA library [29] was screened, clones were isolated confirming the sequence of the original clone from HGS. The macaque sequence was confirmed by analyzing independent cDNAs amplified using a rhesus epididymal cDNA library as template [30].

The full length monkey sequence was obtained by 3’ rapid amplification of cDNA ends (RACE). Monkey epididymal total RNA (2 μg) was reverse transcribed using 0.5 μg Oligo dT-adaptor primer: GACTCGAGTCACTCGA(T17) [31] and SuperScript™ reverse transcriptase. The 750 bp cDNAs was amplified by PCR using the adaptor primer: GACTCGAGTCACTCGA as the reverse primer and the forward primer mf2: ACTAGTGAT\_\_\_\_TCAACITCDA. Three clones were sequenced and the consensus reported. The macaque cDNA sequence was confirmed by sequencing independent cDNAs amplified using a rhesus epididymal cDNA library as template [30].

**Tissue Sources**

For analysis of androgen regulation, male rhesus monkeys (*Macaca mulatta*) of similar age, weight and testicular size underwent subcapsular orchietomy [32] or sham operation. One orchietomized monkey was immediately injected im with testosterone enanthate 30 mg/kg body weight (400 mg total), the other with vehicle. This was a single injection of this long lasting testosterone. Epididymis were removed 6 days later and frozen in liquid nitrogen. Serum samples for testosterone radioimmunoassay were taken just prior to surgery on day 0 and again on day 6. All animals used in these studies were maintained in accordance with the NIH Guide for the Care and Use of Laboratory Animals. The protocol follows accepted veterinary medical practice and was approved by the University of North Carolina Animal Care and Use Committee and is in compliance with the Helsinki Declaration. The animals were given analgesics and were monitored closely following surgery.

Other tissues from rhesus monkeys 10–12 years of age with proven breeding history were provided by Dr. Stephen Pearson, Covance Research Products Inc., Alice, Texas and Dr. Catherine VandeVoort, California Regional Primate Center, Davis California. Human testes and epididymides from prostate cancer patients ranging in age from 58 to 83 were made available by Dr. James L. Mohler, Department of Urology/Surgery, University of North Carolina at Chapel Hill. Other human tissues were obtained from the Tissue Procurement Core Facility of the Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill. Human tissues are accompanied by identifying information and cannot be traced to the donor. All human tissues were obtained after informed consent and an institutional human research committee approved the investigation.

**Recombinant human LCN6 protein used to raise antibody in rabbit**

The cDNA corresponding to the mature human LCN6 protein (Val21 to stop) was amplified by PCR from a human caput/corpus cDNA library [27] constructed in ρZAPII (Stratagene, LaJolla, CA) using forward primer 91F: GCGGATCCGTTGTTGGGGAAGACTG and reverse primer 91R: GCGGTACCTTCTGACGCTGGGCCCCTCTCAC. The gel purified amplification product was digested with Bam HI and Kpn I and ligated into pQE30 (Qiagen, Valencia, CA). The His-tagged protein was expressed and purified according to the manufacturer’s recommendations as previously reported [33]. Recombinant mature human LCN6 (2 mg) was sent to Bethyl Laboratories, Montgomery, Texas to raise rabbit antibodies.

**Immunohistochemical staining**

Tissues for immunohistochemistry were fixed in Bouin’s solution (75 ml saturated picric acid, 5 ml glacial acetic acid, 25 ml 37% formaldehyde) promptly after excision and embedded in paraffin [34]. As a pre-treatment, sections were heated in a microwave oven in 0.01 M citrate pH 6.0 [35]. Antiseras raised in rabbits against recombinant human His-tagged mature LCN6 were diluted 1:800 to 1:1000. Control stainings were performed using antisera preabsorbed with recombinant protein. The double peroxidase-antiperoxidase method [36] was used to demonstrate immunoreactive LCN6 using dianminobenzidine as chromogen resulting in a dark brown reaction product. Photographs were taken with a SPOT Cooled Color digital imaging system (Diagnostic Instruments, Inc, Sterling Heights, MI) attached to a Nikon Eclipse E600 (Southern Micro Instruments, Marietta, Georgia). Photographs were prepared using SPOT image processing software. Images were arranged using PhotoShop (Adobe Systems Inc, San Jose, CA).
Cryopreserved spermatozoa were washed in phosphate buffered saline (PBS) (150 mM NaCl, 10 mM sodium phosphate, pH 7.5) and fixed in 2% paraformaldehyde for 15 minutes. Spermatozoa were washed three times in PBS containing 50 mM glycine and were smeared on glass slides and stored at -20 °C. On the day of the staining, spermatozoa were rehydrated in PBS for 15 minutes followed by blocking in 4% normal goat serum in PBS for 15 minutes. Spermatozoa were incubated with affinity purified specific antibody or the same antibody preincubated overnight with an affinity resin to remove specific antibodies and separated using Handee Mini-Spin columns (Pierce, Rockford, IL). These antisera were diluted 1:5 in 1% normal goat serum in PBS / 0.1% sodium azide. After washing four times in PBS, spermatozoa were incubated using 1:200 fluorescein-conjugated goat anti-rabbit IgG (Vector Laboratories, Burlingame, CA) for 30 minutes. Spermatozoa were washed four times in PBS and mounted using ProLong anti-fade kit (Molecular Probes, Eugene OR). Spermatozoan images were taken using a Zeiss Axioskop microscope with a Zeiss AxioiCam digital camera.

**Molecular modeling**

Fold-recognition services based on sequence-derived properties provided by 3D-PSSM [37], GenTHREADER [38], Fugue profile library search [39], and the Bioinbgu server [40] were used to predict the structure of hLCN6. Representative structures from the lipocalin family as defined by the structural classification of proteins database (SCOP) were evaluated as templates [41]. Of these structures, bovine lipocalin allergen (B17.pdb) [42], pig odorant binding protein (HQP.pdb) [43], and mouse major urinary protein 1 (MUP.pdb) [12] in Protein Data Bank [44] were structurally closest to LCN6. The root mean square deviations when the templates were superimposed ranged from 0.88 to 1.10 indicating strong structural similarity in the protein core. A model of LCN6 was built based on MUP.pdb using the Modeler module of the Insight II molecular modeling system from Accelrys Inc., (San Diego, CA and http://www.accelrys.com). The self-compatibility score indicating compatibility of the predicted side chain environments with their natural preferences was calculated using the Profiles 3-D module of Insight II. The overall score was 50.5, similar to the typical score of 64.7 for a native protein of this size and well above 29.1, a low score that would indicate an incorrect structure. The figure was created using SPOCK [45] in the Structural Bioinformatics Core Facility, University of North Carolina at Chapel Hill under the direction of Dr. Brenda Temple.

**Results**

To investigate novel proteins involved in sperm maturation, the expressed sequence tag (EST) database of Human Genome Sciences Inc, Rockville, MD was searched for epididymis-specific cDNA clones. From more than 130 clones obtained, a cDNA encoding a novel lipocalin, LCN6 was selected for analysis in part because of its close relationship to two well studied rodent epididymal lipocalins, Lcn5 and Lcn8. The human LCN6 gene corresponds to the 5′ half of Unigene cluster Hs.98132, LOC158062 on chromosome 9q34 next to the human orthologs of Lcn5 and Lcn8, in a region rich in lipocalin genes (Fig. 1). The Locus158062 and Unigene cluster information are not shown in Fig. 1, but are available at the National Center for Biotechnology Information (NCBI) website http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?The human LCN6 sequence is based on more than 10 clones we isolated during library screening.

The relative positions of LCN6 and representative related genes are indicated in Fig. 1 in a 9 megabase section of chromosome 9q34 located one megabase from the telomere. The LCN6 gene spans 4.5 kilobases and contains 7 exons (Figs. 1 & 2) which is a highly conserved gene structure in the lipocalin family [46]. LCN6 introns generally do not interrupt the coding sequence of beta strands but instead interrupt coding for connecting loops, another conserved feature of lipocalin gene structure [46,47] (Figs. 2, 3 and 4). Based on the human genome Build 34, Version 1 the gene adjacent to LCN6, 5 kb toward the telomere at LOC138307, Unigene Hs.32991 is similar to the mouse Lcn8 gene [9]. An additional 2.0 kb farther is transcription unit Hs.413902, similar to the rodent Lcn5 gene [26]. Approximately 180 kb toward the telomere from LCN6, is the gene encoding the complement C8 gamma subunit (C8G), and the prostaglandin D2 synthase (PTGDS) gene [48] is located another 30 kb beyond C8G. One megabase (Mb) closer to the centromere than LCN6 are the genes for PAEP [49], odorant binding protein 2A (OBP2A) [50] and LCN1 (tear lipocalin). Another 7 megabases beyond LCN1 closer to the centromere is the LCN2 [51] gene, also known as neutrophil gelatinase-associated lipocalin (NGAL) or in mouse, 24p3. All of these lipocalin genes are expressed in the male reproductive tract [8,10,12,23,50,52]. The mouse orthologue of each of these genes is located on mouse chromosome 2.

The open reading frame of human LCN6 encodes a protein of 163 amino acids with a predicted cleavage site releasing a 20 amino acid N-terminal signal peptide [53] (Figs. 2 and 3) and a mature protein with a predicted molecular weight of 16.0 kDa. The three element lipocalin signature motif, GXWY, TDYXXY and R is conserved in rhesus monkey, but R120 is replaced by L120 in human. A ProSite search [54] revealed a consensus cAMP/cGMP-dependent protein kinase phosphorylation site at human and rhesus Ser73, three casein kinase II phosphorylation sites at Ser64, Thr101 and Ser118. No glycosylation sites were predicted. The rhesus LCN6 is 93% identical to the
human and includes a 17 amino acid C-terminal extension containing the second cysteine found in many lipocalins, but lacking in the human LCN6 due to the early stop codon position. This stop codon is present in the human genome database (Fig. 2) and was further verified by sequencing several independent RTPCR products derived from different human donors (data not shown). Human LCN6 protein is 40% similar to rat Lcn5 protein [6,26], 34% to 36% similar to mouse Lcn5 and to human PTGDS [55] and 30% to 32% similar to human LCN2/NGAL [51] and mouse Mup1 [14].

Thus, the similarity of the LCN6 amino acid sequence to other lipocalins is low, nevertheless strong conservation of the lipocalin 3-dimensional structure is predicted by computer analyses. Based on the similarity of the predicted human LCN6 structure to that of mouse Mup1 previously determined by X-ray diffraction [14], a model of the human LCN6 was calculated (Fig. 4). The predicted basket-like β-barrel structure of LCN6 closely matches that of Mup1, however the relatively short C-terminus of human LCN6 lacks the region that in Mup1 contains the cysteine that forms a disulfide bond with a cysteine in the β-strand B. The conserved lipocalin signature amino acids are located close to each other at the closed end of the β-barrels in LCN6 and Mup1. These signature amino acids were proposed to form a receptor-interacting domain [18].

To determine whether LCN6 expression is predominantly in the epididymis as reported for Lcn5 and Lcn8, total RNAs from epididymis and 25 other human tissues and 4 cell lines were analyzed by Northern blot hybridization (Fig. 5). A strongly hybridizing 1.0 kb species was detected only in RNA from epididymis. In addition, an approximately 2.0 kb weakly hybridizing RNA in urinary bladder was near the limits of detection. This larger mRNA in urinary bladder may represent an alternative transcript containing a partial intron or C-terminal extension similar to LCN6 ESTs BI828733 from brain medulla or BG256703 from a testis embryonal carcinoma cell line. These two ESTs are variants that do not correctly encode LCN6.
Human chromosomal sequence aligned with LCN6 mRNA and amino acid sequences

The predicted TATA box and polyA addition sites are underlined. Exons are in upper case letters, introns in lower case. The ^ indicates the predicted transcription start site. The amino acids are indicated in single letter and three letter abbreviations. The signal peptide amino acids are brown, bold and italicized. Numbers in parentheses indicate amino acids of the mature protein. The lipocalin signature amino acids are red and bold. The gene sequence was extracted from GenBank accession number AL355987. The human cDNA sequence is available at accession number AF303084 and at EST BU502206.

Figure 2
Human chromosomal sequence aligned with LCN6 mRNA and amino acid sequences. The predicted TATA box and polyA addition sites are underlined. Exons are in upper case letters, introns in lower case. The ^ indicates the predicted transcription start site. The amino acids are indicated in single letter and three letter abbreviations. The signal peptide amino acids are brown, bold and italicized. Numbers in parentheses indicate amino acids of the mature protein. The lipocalin signature amino acids are red and bold. The gene sequence was extracted from GenBank accession number AL355987. The human cDNA sequence is available at accession number AF303084 and at EST BU502206.
Figure 3  
Amino acid alignment of mature LCN6 and related lipocalin. Lipocalin signature amino acids are in red and cysteines in orange. The predicted casein kinase II phosphorylation sites at S64, T101 and S118 are indicated by dotted underline. Predicted protein kinase A phosphorylation site at Ser73 is magenta and double underlined. Predicted protein kinase C phosphorylation sites at S53 and T118 are magenta and single underlined (solid underline for S53 and dotted for S118). Amino acids known to form α-helices in mouse Mup1 and rat Lcn5 and predicted in LCN6 are green. β-strands are blue and underlined and labeled A to I corresponding to the β-strands in Fig. 4. The rhesus LCN6 (rhesLCN6) sequence is available at GenBank accession number AF303085. Rat Lcn5 sequence is from accession number X59832 and structural information from 1EPA.pdb and (61). Mouse Lcn8 accession number is NM_033145. Mouse major urinary protein 1 (Mup1) is from M16355. Intron locations in human LCN6 are indicated by the filled triangles (▼). Numbering refers to LCN6 amino acid positions in the mature protein.

| Amino Acid | mature LCN6 | rhesLCN6 | humLCN6 | rhesLcn5 | mouseLcn8 | mouseMup1 |
|------------|-------------|-----------|----------|----------|-----------|-----------|
| C          | VTLTPENNLR  | TLLSQQHLGG| CDQSVMLIK| VEK, ENLIA | LTFSGNVTV | IHVL, ENSLV |
| D          | E           | D         | E        | E        | D         | E         |
| E          |             |           |          |          |           |           |
| F          |             |           |          |          |           |           |
| G          |             |           |          |          |           |           |

Currently more than 30 ESTs derived from LOC158062 indicate expression in many organs. However many of these fail to encode LCN6 either because they are splicing variants or they originate from the 3′ non-LCN6 half of this locus and thus do not indicate LCN6 expression in those tissues.

To determine if LCN6 is regulated by testosterone as reported for Lcn5 in the mouse [56] or testis factors as reported for mouse Lcn8, RNA was obtained from caput, corpus and cauda epididymis of rhesus monkeys that were sham operated, castrated 6 days and castrated but given a single injection of 400 mg testosterone enanthate immediately following testis removal (Fig. 6). The concentration of LCN6 mRNA in caput total RNA samples appeared little affected following 6 days castration and testosterone replacement. Low level mRNA expression in the corpus of the sham operated animal was abolished by castration and not maintained in the testosterone replaced animal suggesting regulation by testis factors other than androgen. This maintenance of similar LCN6 mRNA concentrations under the three treatment conditions contrasts with the androgen regulation of human epididymal (HE) protein 2 [29], epididymis specific clone (ESC) 42 [30] and cystatin (CST) 11 [57] we previously reported in this same experiment. Morning serum testosterone levels prior to testis removal in sham-operated, castrated and castrated-androgen-replaced were 3.0, 3.8 and 1.4 ng/ml respectively. Six days after surgery, morning serum testosterone levels ranged from 5 ng/ml in the morning to 18 ng/ml during the nocturnal surge [58].

LCN6 protein is most abundant in late stage efferent ducts and caput epithelium in human (Fig. 7). LCN6 is also present in the lumen in association with spermatozoa, consistent with secretion from epithelial cells as suggested
by the presence of a predicted hydrophobic signal peptide. The protein was also detected at lower levels in corpus and cauda (data not shown). Labeling of human ejaculate spermatozoa with affinity-purified anti-LCN6 demonstrated that LCN6 was present on all spermatozoa (Fig. 8). LCN6 was localized on the head and tail of spermatozoa with the highest concentration of LCN6 seen on the postacrosomal region of the head, where it appeared aggregated into large patches. Along the tail smaller more discrete focal points of LCN6 were observed. Interestingly, the large patches in the postacrosomal region of the head appear in the region of the spermatozoon thought to fuse with the oocyte plasma membrane. The LCN6 protein on spermatozoa is not likely to originate in the testis since the mRNA was not detected by Northern blot analysis in testis RNA (Fig. 5).

**Discussion**

LCN6 is the first human member to be described of a gene cluster that includes LCN5 and LCN8 in a lipocalin-rich region of chromosome 9q34. The similar gene clusters on mouse chromosome 2A3 and on rat 3p13 indicate that these genes diverged from a common ancestor prior to the separation of rodent and primate evolutionary lines. The mouse and rat Lcn5 and mouse Lcn8 are the only genes in these conserved clusters previously described. They are exclusively expressed in the epididymis as judged by Northern blot analysis. Our Northern results showing epididymis-specific expression of fully processed LCN6 mRNA might appear to be contradicted by the ESTs derived from LOC158062 that were isolated from libraries from many other organs. This locus represents a fusion of two GenomeScan-predicted models, LCN6 and an adjacent lipocalin-like gene based on BG256703, a bridging EST derived from a testis embryonal carcinoma cell line. However, our results indicate that the LCN6 gene can function independently of the 3' half of the locus, at least in the epididymis. Separate orthologs of LCN6 and of the adjacent gene are predicted in mouse and rat and we isolated the rat Lcn6 as an independently expressed sequence.
Most of the ESTs listed from this combined locus are derived from the 3' lipocalin-like adjacent gene. The LCN6-specific GenBank ESTs are all from epididymis, "whole testis" which includes epididymis, or different cell lines, thus are consistent with our Northern blot tissue distribution analysis. In addition, GenBank contains LCN6 splicing variant ESTs from brain medulla and prostate. These splicing variants may not be translated, a mechanism previously reported that restricts GnRH expression to the hypothalamus [59].

Expression of LCN6 in distal efferent ducts and caput epididymis and localization of the protein on the surface of spermatozoa are consistent with a role in spermatozoa maturation. That role could be to carry ligands from the proximal epididymis to receptors on distal epithelial cells, a mechanism suggested for Lcn5 transport of retinoic acid [26]. A similar model was proposed to explain the regulation of proenkephalin gene expression in the rat caput by an unidentified spermatozoa-associated factor [60]. Moreover, the ligand might be delivered to receptors in the female tract. Delivery of ligand could result in changes in gene expression in the recipient cells.

The ligand(s) of LCN6 have not been identified, but may be similar to retinoic acid which is a known ligand bound by Lcn5. The Lcn5 and LCN6 proteins have diverged substantially in linear sequence, yet the amino acids known and predicted to line the binding pocket and entrance of LCN6 and rat Lcn5 [61] are more closely related than the 40% similarity of the whole proteins. Of the 23 amino acids that form the ligand binding cavity in rat Lcn5 [61], 9 or 39% are identical in LCN6 and six others are similar for a total of 65% similarity in the ligand binding cavity. Furthermore, 3 of the 5 aromatic amino acids that are deepest in the binding cavity, forming a semicircle surrounding the β-ionone ring of the retinoic acid in the holo-Lcn5, are identical in LCN6. The other two amino acids not shown).
acids, Phe6 and Phe11 in Lcn5 are replaced by leucine in LCN6. At the entrance to the binding cavity in Lcn5 are charged amino acids Glu17, Glu63, Arg80, Lys85 and Lys115. These are replaced in LCN6 by Val/Ile, Ser, Ser, Leu and Glu respectively. The effects of these amino acid differences remain to be evaluated by ligand binding studies and X-ray diffraction but presumably the range of ligand structures that can be accommodated in the cavities and their orientations with respect to the amino acids lining the cavity are probably different in LCN6 and Lcn5.

Ligand binding properties of human LCN6 may be affected by the lack of the cysteine near the C-terminus that is strongly conserved in lipocalins. Where present, this cysteine forms a disulfide bridge with a cysteine located in β-strand D and anchors the C-terminus to the β-barrel (Fig 4B). In studies on LCN1, the corresponding intact disulfide bond diminishes the affinity for retinol and restricts the displacement of native lipids by retinol probably by conferring rigidity on the barrel structure and restricting movement of the aromatic side chains involved in ligand binding [62]. Disulfide reduction in LCN1 was accompanied by alteration in ligand-induced conformational changes. The absence of this disulfide bridge in human LCN6 may similarly allow a broader range of ligands to fit into the human pocket than can fit into the monkey form.

The molecular basis of LCN6 interaction with spermatozoa is unknown, but its presence on spermatozoa in discrete patches raises the possibility of interaction with specific receptors. Each known spermatozoan surface lipocalin is distributed according to a particular pattern suggesting a specific molecular interaction. The location of LCN6 generally on post-acrosomal human spermatozoa contrasts with 24p3 on mouse spermatozoa which is predominantly on the anterior acrosomal region [8]. On bull spermatozoa, PTGDS is concentrated on the apical ridge of the acrosomal cap [5]. On surfaces of cell types other than spermatozoa, several lipocalins interact with cell surface receptors [63]. Cellular responses were demonstrated for 24p3/Lcn2 protein which induced apoptosis by a receptor-mediated pathway in murine FL5.12 12 B lymphocytic cell line [64] and Lcn-1 was internalized after binding a specific receptor on human NT2 embryonal carcinoma cells [65]. Whether internal responses are generated in spermatozoa after lipocalin binding has not been reported.

Regional differences in expression site and differences in the influence of the testis on expression of Lcn5, Lcn8 and LCN6 may be factors driving the evolutionary maintenance in rodent and human of this gene cluster that is derived by gene duplication [9]. The LCN6 expression in distal efferent ducts, initial segment and proximal caput is similar to that of rat Lcn5 [66] and mouse Lcn8 [67] but differs from the mouse Lcn5 which is expressed in distal caput and corpus and accumulates in caudal fluid [4,26]. The more distal expression of mouse Lcn5 positions it to affect sperm maturation downstream of where Lcn8 becomes available to spermatozoa in the lumen. Unlike the rodent Lcn5, monkey LCN6 mRNA levels appear to be little regulated by androgen and unlike mouse Lcn8, monkey LCN6 regulation by other testis factors, appears minimal 6 days after testis removal. This is a surprising result for an epididymal protein associated with spermatozoa since spermatogenesis and sperm maturation in the epididymis both depend on androgen. The lack of androgen regulation of LCN6 expression may reflect functions apart from sperm maturation such as maintenance of healthy epididymal epithelium under changing physiological conditions. Since a number of lipocalins have demonstrated roles in host defense [68] and the C-terminus of the monkey LCN6 protein is cationic, we tested recombinant human and monkey LCN6 for antibacterial activity and found they had little or no effect on E. coli [33] suggesting that killing bacteria is not a likely role for LCN6.
Conclusions

LCN6 is a novel lipocalin closely related to Lcn5 and Lcn8. The LCN5,6 and 8 genes are likely products of ancient gene duplication events that predate rodent-primate divergence. Predominant expression in the epididymis and location on sperm surface are consistent with a role for LCN6 in male fertility.

Author's contributions

KH and QL obtained the full length human and monkey cDNAs. KH screened libraries and subcloned the human cDNA in expression vectors and prepared the human recombinant protein, found the gene in the database and constructed Figures 2 and 3. KH and QL wrote portions of the paper. QL performed the initial sequence analyses, designing primers and performing 5’ and 3’ RACE. He obtained the full length rat cDNA. PS performed sperm immunofluorescent staining. MA prepared the monkey recombinant LCN6 protein and with SY tested the human and monkey recombinant proteins for antibacterial activity. RS and GG performed the immunohistochemical staining of the human, monkey and rat epididymides. GG

Figure 7

LCN6 protein in human efferent ducts and caput. Panels A and B human efferent ducts from an 83 yr old donor. A, LCN6 protein appears brown, distinct from the toluidine blue counterstain. B, A control staining for which antibody was preincubated with 1 mg/ml recombinant human LCN6 protein. Panels C and D human caput. C, LCN6 appears brown. D, Antibody preincubated with 100 µg/ml human LCN6. Photographs were taken using a 20× objective.
also contributed to the immunostaining of sperm. AJR provided paraffin-embedded monkey tissues for the immunostaining and advised MA and RS including assistance in interpretation of their results. CEB and SMR provided consultation, analyzed the HGS databases for epididymis-specific clones, obtained their sequences, provided sequence and supporting clonal information as well as the actual cDNA clones, one of which was the original LCN6 clone that started this project. RR raised the initial antisera in mice, assisted with the sperm staining and interpretation and provided consultation. Y-LZ served as advisor to LQ and provided consultation, data interpretation and participated in writing the paper. MGO’R assisted with the sperm staining and interpretation and manuscript writing. PP assisted with the epididymis immunohistochemistry, interpretation and preparation of figures. FSF and SHH wrote the majority of the manuscript. SHH coordinated the work and supervised preparation of the figures. All authors read, commented upon and approved the final manuscript.

Acknowledgments
Support for this project [CIG-96-06-A] was provided by the CICCR Program of the Contraceptive Research and Development Program, Eastern Virginia Medical School. The views expressed by the authors do not necessarily reflect the views of CONRAD or CICCR. This work was also supported by grants from The Andrew W. Mellon Foundation and by National Institutes of Health grants R37-HD04466, through cooperative agreement U54-HD35041 as part of the Specialized Cooperative Centers Program in Reproduction Research, and by the Fogarty International Center Training and Research in Population and Health grant D43TW/HD00627. We thank Zang De-Ying for excellent technical assistance. Special thanks to Betty F. Horton and Richard L. Pippin for preparation of the figures.

References
1. Cooper TG: Epididymal proteins and sperm maturation. In Spermatogenesis-Fertilization-Contraception Edited by: Nieschlag E, Habe- nicht UF. Berlin/New York:Springer-Verlag; 1992:285-318.
2. Cuasnicu PS, Cohen DJ, Ellerman DA, Buso D, DaRos VG, Morgenfeld MM: Changes in specific sperm proteins during epididy- mal maturation. In The Epididymis From Molecules to Clinical Practice Edited by: Robaire B, Hinton BT. New York: Kluwer Academic/Plenum; 2002:389-403.
3. Dacheux JL, Gatti JL, Dacheux F: Contribution of epididymal secretory proteins for spermatozoa maturation. Microsc Res Tech 2003, 61:7-17.
4. Rankin TL, Tsuruta KJ, Holland MK, Griswold MD, Orgebin-Crist MC: Isolation, immunolocalization and sperm-association of three proteins of 18, 25 and 29 kilodaltons secreted by the mouse epididymis. Biol Reprod 1992, 46:747-766.
5. Gerena RL, Irikura D, Eguchi N, Urade Y, Killian GJ: Immunocyto- chemical localization of lipocalin-type prostaglandin D synthase in the bull testis and epididymis and on ejaculated sperm. Biol Reprod 2000, 62:547-556.
6. Brooks DE, Means AR, Wright EJ, Singh SP, Tiver KK: Molecular cloning of the cDNA for two major androgen-dependent secretory proteins of 18.5 kilodaltons synthesized by the rat epididymis, J Biol Chem 1986, 261:4956-4961.
7. Morel L, Dufaure JP, Depeiges A: The lipocalin sperm coating liz- ard epididymal secretory protein family: mRNAs structural analysis and sequential expression during the annual cycle of the lizard, Lacerta vivipara, J Mol Endocrinol 2000, 24:127-133.
8. Chu ST, Lee YC, Nein KM, Chen YH: Expression, immunolocalization and sperm-association of a protein derived from 24p3 gene in mouse epididymis, Mol Reprod Dev 2000, 57:26-36.
9. Lareyre J-J, Winfrey VP, Kasper S, Ong DE, Matusik RJ, Olson GE, Orgebin-Crist M-C: Gene duplication gives rise to a new 17-kilodalton lipocalin that shows epididymal region-specific localization of LCN6 on human spermatozoa Panel A, Immunofluorescent staining of spermatozoa with anti-LCN6. Panel B, Control immunofluorescent staining of spermatozoa with anti-LCN6 preincubated with LCN6 protein attached to affinity resin to remove specific antibodies. Original images were taken using a 63× objective.
expression and testicular factor(s) regulation. Endocrinology 2001, 142:1296-1308.

10. Halttunen M, Kamarainen M, Koistinen H, Koistinen R, Oehninger S, Patankar MS, Seppala M, Clark GF. Gender-specific glycosylation of human glycodelin affects its contraceptive activity. J Biol Chem 1996, 271:32159-32167.

11. Liu Q, Hamil KG, Sivashanmugam P, Grossman G, Soundararajan R, Rao AJ, Richardson RT, Zhang Y-L, O’Rand MG, Petrusz P, French FS, Hall SH: HE2 and HE2, new members of an epididymis-specific family of androgen-regulated proteins in the human. Endocrinology 2000, 141:1205-1212.

12. Hamil KG, Sivashanmugam P, Richardson RT, Grossman G, Ruben SM, French FS, Hall SH: Antibacterial properties of the sperm-binding proteins and peptides of human epididymis 2 (HE2) family; salt sensitivity, structural dependence and their interaction with outer and cytoplasmic membranes of Escherichia coli. Biochem J 2003, 372:473-483.

13. Ong DE, Newcomer ME, Lareyre JJ, Orgebin-Crist MC: Utilisation des micro-organismes pour aviver ou restaurer les antigènes avant la réaction immunohistochimique. [Use of microorganisms for enhancing or restoring antigens before immunohistochemical staining]. [French]. Ann Pathol 1993, 13:188-189.

14. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Epitope-mapping of the lipocalin protein family. Mol Biol Evol 2000, 17:117-124.

15. Redl B, Oehninger S, Patankar MS, Seppala M, Clark GF: The structural properties of porcine odorant-binding protein. J Mol Biol 2000, 299:499-520.

16. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Characterization of puberty onset in female mice by male urinary proteins. J Physiol 1998, 506:517-522.

17. Devireddy LR, Teodoro JG, Richard FA, Green MR: Induction of apoptosis by a secreted lipocalin that is transcriptionally regulated by IL-3 deprivation. Science 2001, 293:829-834.

18. Sivashanmugam P, Grossman G, O’Rand MG, Petrusz P, French FS, Hall SH: HE2, a novel protein containing a trefoil-like motif in monkey and human. Endocrinology 2001, 142:4529-4539.

19. Frohman MA, Dash MK, Martin GR: Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 1988, 85:8986-9002.

20. Zhang X-Z, Donovan MP, Williams BT, Mohler JL: Comparison of subcapsular and total orchietomy for treatment of metastatic prostate cancer. J Urol 1996, 147:402-404.

21. Sanchez D, Ganfornina MD, Gutierrez G, Mar On A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2003, 20:775-785.

22. Tung M, Polois P, Vincent F, Spinelli S, Campanacci V, Grolli S, Romani R, Camillau C: Mammalian odorant binding proteins. Biochim Biophys Acta 2000, 1482:229-340.

23. Liu Q, Hamil KG, Sivashanmugam P, Grossman G, Soundararajan R, Rao AJ, Richardson RT, Zhang Y-L, O’Rand MG, Petrusz P, French FS, Hall SH: Primates: a novel protein containing a trefoil-like motif in monkey and human. Endocrinology 2001, 142:1245-1253.

24. Mucignat-Caretta C, Caretta A, Cavaggioni A: Acceleration of puberty onset in female mice by male urinary proteins. J Physiol 1995, 486:517-522.

25. Livera G, Rouiller-Fabre V, Pairault C, Levacher C, Habert R: Anti-bacterial properties of the sperm-binding proteins and peptides of human epididymis 2 (HE2) family; salt sensitivity, structural dependence and their interaction with outer and cytoplasmic membranes of Escherichia coli. Biochem J 2003, 372:473-483.

26. Boscik I, Groom CR, Flower DR, Wright CE, Phillips SEV, Cavaggioni A, Findlay JBC, North ACT: Chromosomal location, exon/intron organization and testicular factor(s) regulation. J Mol Biol 2001, 310:243-257.

27. Ong DE, Newcomer ME, Lareyre JJ, Orgebin-Crist MC: Hybrid fold recognition method for genomic sequences. J Mol Biol 1999, 287:797-815.

28. Seki S, Blundell TL, Miizuichi K: FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J Mol Biol 2001, 310:243-257.

29. Fischer D: Hybrid fold recognition: combining sequence derived properties with evolutionary information. Pacif Symp Biocomp 2000,1:119-130.

30. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

31. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2003, 20:775-785.

32. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

33. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

34. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

35. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

36. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

37. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

38. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

39. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

40. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

41. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

42. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

43. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

44. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

45. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

46. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

47. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

48. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

49. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.

50. Sanchez D, Ganfornina MD, Gutierrez G, MarOn A: Exon-intron structure and evolution of the lipocalin gene family. Mol Biol Evol 2000, 17:117-124.
54. Hofmann K, Bucher P, Falquet L, Bairoch A: The PROSITE database, its status in 1999. Nucl Acids Res 1999, 27:215-219.
55. Tokugawa Y, Kunishige I, Kubota Y, Shimoya K, Noburaga T, Kumra T, Saij F, Murata Y, Fujin N, Oda H, Uraya Y, Hayashi O: Lipocalin-type prostaglandin D synthase in human male reproductive organs and seminal plasma. Biol Reprod 1998, 58:600-607.
56. Lareyre JJ, Matzel MG, Kasper S, Newcomer ME, Ong DE, Matusik RJ, Orgebin-Crist MC: Structure and putative function of a murine epididymal retinoic acid-binding protein (mE-RABP). J Reprod Fertil 1998, Suppl 53:59-65.
57. Hamil KG, Liu Q, Sivaswamugam P, Yenugu S, Soundararajan R, Grossman G, Richardson RT, Zhang YL, O’Rand MG, Petrusz P, French FS, Hall SH: Cystatin I: a new member of the cystatin type 2 family. Endocrinology 2002, 143:2787-2796.
58. Chambers KC, Resko JA, Phoenix CH: Correlation of diurnal changes in hormones with sexual behavior and age in male Rhesus macaques. Neurobiol Aging 1982, 3:37-42.
59. Son GH, Jung H, Seong JY, Choe Y, Geum D, Kim K: Excision of the first intron from the gonadotropin-releasing hormone (GnRH) transcript serves as a key regulatory step for GnRH biosynthesis. J Biol Chem 2003, 278:18037-18044.
60. Garrett JE, Garrett SH, Douglass J: A spermatozoa-associated factor regulates proenkephalin gene expression in the rat epididymis. Mol Endocrinol 1990, 4:108-118.
61. Newcomer ME: Structure of the epididymal retinoic acid binding protein at the 2.1 Å resolution. Structure 1993, 1:7-18.
62. Glasgow BJ, Abduragimov AR, Yusifov TN, Gaysyrov OK, Horwitz J, Farghly WL, Flauk KF: A conserved disulfide motif in human tear lipocalins influences ligand binding. Biochem 1998, 37:2215-2225.
63. Flower DR: Beyond the superfamily: the lipocalin receptors. Biochim Biophys Acta 2000, 1482:327-336.
64. Devireddy LR, Teodoro JG, Richard FA, Green MR: Induction of apoptosis by a secreted lipocalin that is transcriptionally regulated by IL-3 deprivation. Science 2001, 293:829-833.
65. Wójnar P, Lechner M, Redl B: Antisense down-regulation of lipocalin-interacting membrane receptor expression inhibits cellular internalization of lipocalin-1 in human NT2 cells. J Biol Chem 2003, 278:16209-16215.
66. Garreth SH, Garrett JE, Douglass J: In situ histochemical analysis of region-specific gene expression in the adult rat epididymis. Mol Reprod Dev 1991, 30:1-17.
67. Fouchecourt S, Lareyre JJ, Chaurand P, DaGue BB, Suzuki K, Ong DE, Olson GE, Matusik RJ, Orgebin-Crist MC: Identification, immunolocalization, regulation, and postnatal development of the lipocalin EP17 (epididymal protein of 17 kilodaltons) in the mouse and rat epididymis. Endocrinology 2003, 144:887-900.
68. Logdberg L, Wester L: Immunocalins: a lipocalin subfamily that modulates immune and inflammatory responses. Biochim Biophys Acta 2000, 1482:284-297.