Abstract. The bladder exstrophy-epispadias complex (BEEC) represents a spectrum of urological abnormalities where part, or all, of the distal urinary tract fails to close during development, becoming exposed on the outer abdominal wall. While the etiology of BEEC remains unknown, strong evidence exists that genetic factors are implicated. To understand the pathways regulating embryonic bladder development and to identify high-probability BEEC candidate genes, we conducted a genome-wide expression profiling (GWEP) study using normal and exstrophic human urinary bladders and human and mouse embryologic bladder-precursor tissues. We identified 162 genes differentially expressed in both embryonic and postnatal human samples. Pathway analysis of these genes revealed 11 biological networks with top functions related to skeletal and muscular system development, cellular assembly and development, organ morphology, or connective tissue disorders. The two most down-regulated genes desmin (DES, fold-change, -74.7) and desmuslin (DMN, fold-change, -53.0) are involved in desmosome mediated cell-cell adhesion and cytoskeletal architecture. Intriguingly, the sixth most overexpressed gene was desmoplakin (DSP, fold-change, +48.8), the most abundant desmosomal protein. We found 30% of the candidate genes to be directly associated with desmosome structure/function or cytoskeletal assembly, pointing to desmosomal and/or cytoskeletal deregulation as an etiologic factor for BEEC. Further findings indicate that p63, PERP, SYNPO2 and the Wnt pathway may also contribute to BEEC etiology. This study provides the first expression profile of urogenital genes during bladder development and points to the high-probability candidate genes for BEEC.

Introduction

The bladder exstrophy-epispadias complex (BEEC) represents a rare birth defect thought to be a clinical spectrum ranging from isolated epispadias (EP), to the classic bladder exstrophy (CBE) and to its most severe form, cloacal exstrophy (CE) often referred to as the OEOIS (omphalocele, exstrophy of the bladder, imperforate anus, and spinal defects) complex (1-3). Prevalence at birth ranges from 1/30,000 for CBE to
1/200,000 for CE (4). There is a greater proportion of affected males than females, ranging from 1.8:1 to 6:1 (5,6). While the etiology of BEEC remains unknown, strong evidence exists linking its development to genetic effects. The etiological association of genetic determinants is supported by the discrepant male to female ratio (7); a 400-fold increase of the recurrence risk for offsprings of affected individuals (8); observations of rare multiplex families (9); and much higher concordance rates (62% vs. 11%) among monozygotic as compared to dizygotic twins (10). Its non-Mendelian pattern of segregation and the incomplete concordance among MZ twins implies a complex trait model of inheritance.

Typically, the patients with BEEC do not have developmental delays, and with the exception of CE, there are few anomalies outside of the urogenital system (5,6). The associated anomalies observed in BEEC are mesoderm-derived, including omphalocele, hernias, kidney agenesis, imperforate anus, and pelvic dysplasia. This suggests that the genetic insult may be limited to the mesodermal lineage. Indeed, Wood et al (11) proposed that BEEC occurred due to the overall deficiency of the somitic and lateral plate mesoderm in the infra-umbilical region. Since this mesodermal cell population forms the muscular components of the bladder and infra-umbilical body wall, it is plausible that a deficiency of this cell population would account for this phenotype.

The emerging picture of BEEC as a complex trait determined by mutations in genes regulating the proliferation and differentiation of the mesoderm and the morphogenesis of the smooth muscle of the urinary bladder motivated the expression studies described in this paper. DNA microarrays have proven to be a powerful technological platform for genome-wide expression profiling (12). This approach has been used successfully for identifying causative genes for various diseases such as craniosynostosis (13), asthma (14-16), heart disease (17-19), and hypospadias (20). To better understand the molecular profile of BEEC and to identify potential candidate genes, we conducted a genome-wide expression profiling study using human urinary bladder at two time-points: human embryonic mesenchyme (EM) surrounding the urogenital sinus at 8 weeks of gestation, and postnatal normal bladders (NB) or exstrophic bladders (EB). To further investigate whether differentially expressed bladder derived genes were expressed in the same relevant organogenesis period in both humans and the mouse, we also studied the gene expression profile of urinary bladders pooled from the same mouse litter at gestational day (GD) 13. The concept behind this study was to identify genes with at least two-fold expression difference between normal and exstrophic bladder smooth muscle that are also expressed in the embryonic period when BEEC occurs. These BEEC candidate genes would be a valid target for further molecular and association studies.

**Materials and methods**

**Tissue collection.** This study was approved by the Johns Hopkins Medicine Institutional Review Boards and was conducted in accordance with their guidelines. Prior to sample collection, informed consent was obtained from all individuals or their legal guardians. Mouse tissue samples were obtained under approved animal protection protocol. Upon collection all samples were placed in RNA later® tissue collection: RNA stabilization solution (Ambion, Austin, TX) and stored at -80°C until RNA isolation. Procedural steps such as tissue freezing and RNA extraction were performed without time delay between stages.

*Prenatal human bladder.* Bladder-precursor EM surrounding the urogenital sinus was obtained from 3 human embryos electively aborted by dilatation and curettage (D&C) at 8 weeks of gestation. The dissections were carried out under a dissecting microscope in warm buffered saline solution, with tissue extraction completed within 15 min of termination of the pregnancy. Specimens were then transported in tissue culture media on ice to the laboratory, where 3 primary cultures of the mesenchymal tissue were commenced.

*Postnatal human bladder.* Three postnatal EB and three postnatal NB (control) tissue samples were collected. EB samples were obtained during surgical repair and consisted of discarded bladder tissue from patients with CBE. Surrounding fat and connective tissue were separated from the smooth muscle. NB samples were obtained during ureter re-implantation and consisted of anonymously collected and de-identified bladder tissue from vesico-ureteral reflux patients who did not have BEEC. Since the bladder size, function and morphology were intact, we assumed these samples to be normal bladder tissue. After removing the inner urothelial and outer serosal layers, the EB and NB specimens were placed immediately in RNA later solution (Ambion). In addition, EB and NB samples were race- and gender-matched, with 2 female pairs and 1 male pair. All samples were collected from Caucasian individuals younger than 3 years of age.

*Mouse bladder.* The urinary bladder and the surrounding mesenchyme were microdissected from C57BL/6 mouse (Charles Rivers Laboratory) embryos at GD13, corresponding to human EM samples. The microdissected target tissues were immediately placed in RNA later stabilizing solution, and housed in a -20°C cold block throughout the dissection period. Only tissues collected within 25 min post-removal of embryos from the uterus were used for RNA isolation to reduce sample degradation. The target tissues from the same litter were pooled together and were transferred to a -80°C freezer for storage until RNA isolation.

*RNA isolation, purification, quantification, and quality control.* Total RNA was extracted using a standard TRIzol protocol (Invitrogen, Carlsbad, CA) and cleaned up with the RNeasy Micro kit (Qiagen, Valencia, CA). The RNA concentration and integrity were measured and assessed by the RNA 6000 Nano Assay on the 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Prior to target preparation, all samples had verified optimal rRNA 18S:28S ratios of 1:2 and clean run patterns. Post-hybridization analysis of the chip (image analysis and 3’/5’ ratios of housekeeping genes) raised no significant concern about the integrity of the starting RNA.
Microarray preparation, hybridization and data acquisition. RNA was processed using the RNA amplification protocol, as described in the Affymetrix (Santa Clara, CA) expression manual. Briefly, 5 μg of total RNA were used to synthesize first-strand cDNA using oligonucleotide probes with 24 oligo-dT plus T7 promoter as primer (Promega, Boulder, CO) and the SuperScript Choice System for cDNA Synthesis (Invitrogen). Following double-stranded cDNA synthesis, the product was purified by phenol-chloroform extraction, and biotinylated antisense cRNA was generated through in vitro transcription using the BioArray HighYield RNA Transcript Labeling Kit (Enzo Biochem, Farmingdale, NY). Biotinylated cRNA (15 μg) samples were fragmented at 94°C for 35 min (100 mM TAE, pH 8.2, 500 mM KOAc, 150 mM MgOAc). All human samples were hybridized to the Affymetrix Human Genome U133 Plus 2.0 Arrays, and the mouse GD13 sample was hybridized to the Affymetrix Mouse Genome 430 2.0 Array. For each sample, 10 μg of total fragmented cRNA was hybridized to the microarray for 16 h at 45°C with constant rotation (60 rpm). The Affymetrix Fluidics Station 450 was used to wash and stain the chips, to remove the non-hybridized target and to stain the biotinylated cRNA by incubation with a streptavidin-phycoerythrin conjugate. The staining was amplified using goat IgG as a blocking reagent, and a biotinylated anti-streptavidin goat antibody, followed by a second staining step with a streptavidin-phycoerythrin conjugate. Fluorescence was detected using an Affymetrix GeneChip Scanner 3000, and image analysis of each GeneChip was done through the Affymetrix GeneChip Operating System Software (GCOS) 1.1.1, using the standard default settings. For comparison between different chips, a global scaling was used, scaling all probe sets to a user-defined target intensity of 150.

Quantitative real-time PCR (qPCR). Reactions for qPCR were performed on an ABI Prism 7900HT Fast real-time PCR system (Applied Biosystems, Foster City, CA, USA) using SYBR-Green for detection. Due to the fact that RNA from the individuals analyzed by expression arrays was not available, the qPCR validation of our results was conducted using bladder tissues from 6 other CBE patients (2 female, 4 males) and 4 controls including 2 tissue samples from healthy males, total bladder RNA (male sample, Ambion) and pooled fetal female bladders RNA harvested between 18 and 23 gestation weeks (Stratagene, TX), respectively. The qPCR was performed for 4 selected genes of interest including 2 down-regulated (DES, DMN) and 2 up-regulated genes (AQP3, PERP). Each qPCR assay included 20 ng/ml DNA samples in triplicate. Reaction mixtures (10 μl) contained 0.2 mM of each primer and 5 μl of Power SYBR-Green PCR Master Mix (PE-Applied Biosystems) with cycling conditions as follows: initiation, 50°C for 2 min; denaturation, 95°C for 10 min, followed by 40 cycles at 95°C for 15 sec, and a combined annealing and extension step at 60°C for 60 sec. The threshold cycles (Ct) values were normalized using the Ct value of 3 reference genes (AQP3, CFTR and RNase P subunit p38) for each DNA sample. Relative quantification was done using the comparative Ct method, also known as the 2-ΔΔCt method (21). Fold-change was defined as 2-ΔΔCt, where Ct was the average Ct from 3 replicates, ΔΔCt =Ct gene of interest -Ct reference gene, and ΔΔCt = ΔCt case - ΔCt control. ΔCt control was calculated by averaging the values from the controls. The sequence of primers used to confirm gene expression alterations in microarray analysis is available upon request.

Data analysis. Raw data were normalized by the Robust Multichip Average (RMA) method. Fold-change analysis was applied to the normalized data to identify transcripts differentially expressed between EB and NB samples. Transcripts were selected based on a minimal change of 2-fold, which has been widely used as a threshold in microarray analysis (22,23). Transcripts with at least 2-fold increase or decrease of expression in EB as compared to NB were further filtered by whether they were expressed in all 3 prenatal human bladders, and they were considered to be expressed if they showed a Present call in the chip analysis from the Affymetrix GCOS in the EM samples. The BEEC candidate gene list consisted of the transcripts that were at least 2-fold overexpressed or underexpressed in common among all 3 pairs of NB/EB samples and were also prenatally expressed. BEEC candidate genes were categorized into groups based on their functional characteristics using the Onto-Tools (24), an ensemble of web-accessible, ontology-based tools for the functional design and interpretation of high-throughput gene expression experiments through the use of Gene Ontology (GO, http://www.geneontology.org). Ingenuity Pathways analysis (Ingenuity® Systems, http://www.ingenuity.com) was also applied to candidate genes to identify biological networks, based on a comprehensive database representing a large number of individual protein-protein and gene-complex interactions. The software computes a score for each network according to the fit of Focus Molecules using the Fisher's exact test. The score, derived from a P-value from the test, reflects the likelihood of Focus Molecules in a network by chance: a 1 in 100 chance is scored 2; 1 in 1,000 is scored 3, etc. Therefore, scores of ≥2 have ≥99% confidence (and ≥3 have ≥99.9%) of not being generated by random chance alone.

We observed that the genes with the highest degree of differential expression between EB and NB (DES, DMN, DSP) were related to the desmosome, motivating us to investigate whether other candidate genes are also associated with desmosomal function and/or cytokeratin regulation. Every candidate gene was reviewed in depth and assigned to one of the following 6 categories: 1) significant or known association with desmosomal structure/function; 2) lesser known or secondary association with desmosome; 3) association with any cell-cell connectivity function (other types of cell-cell junctions); 4) association with connecting cellular components (e.g., actin, ECM); 5) weak association with any of the above functions; 6) almost certainly no association with the above functions. A gene was defined as an adhesion-related gene if it was assigned to one of the first four categories. For each gene, the following sources were examined and the relevant information was collected: the NCBI Gene database (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene) that includes functional information, gene ontology and potential interacting genes; the UCSC Genome browser (http://genome.ucsc.edu/) for any pertinent localization information, and the PubMed literature search browser (http://www.ncbi.nlm.nih.gov/pubmed/) to determine any
previous work that would associate the gene of interest with cell connectivity.

In addition to gene ontology analyses, the expression of candidate genes in midgestation mouse embryos was examined by searching the Molecular Anatomy of the Mouse Embryo Project (MAMEP) database (http://mamep.molgen.mpg.de; Spörle, Werber, Neidhardt, and Herrmann, unpublished data). The MAMEP database represents the largest collection of gene expression data on mouse embryos encountering a broad collection of images of gene expression patterns in mouse embryos at GD9 to GD10, generated by whole mount in situ hybridization (25). Therefore, the expression pattern of candidate genes at these developmental stages can be taken as an indication of their functional relevance during early development of the bladder primordium and its associated tissues. To further validate candidate genes, we analyzed their chromosomal localization for physical overlap with regions of putative linkage/linkage disequilibrium identified by our recent work and breakpoints of chromosome rearrangements observed in BEEC patients (7,26,27).

Results

**BEEC candidate genes.** We analyzed the genome-wide RNA expression profile of 3 race- and gender-matched EB and NB samples. Overall, 54,675 probes representing 47,650 genes were analyzed. Based on fold-change analysis, a total of 584 transcripts, corresponding to 422 distinct genes, showed at least 2-fold difference of expression in all 3 NB/EB samples, including 301 overexpressed and 283 underexpressed transcripts (Table I). Among them, 207 transcripts corresponding to 162 genes were also consistently expressed in all 3 EM samples. These genes formed our BEEC candidate gene list.

**Quantitative real-time PCR (qPCR) analysis.** To validate the microarray analysis, we performed qPCR for the 2 most underexpressed genes in EB, *DES* and *DMN*, and the 2 overexpressed genes, *AQP3* and *PERP*, using bladder tissues from 6 other CBE patients (2 female, 4 males) and 4 controls. Fig. 1 shows the fold-change in *DES*, *AQP3* and *PERP*. The fold-change values for *DMN* for samples CBE1-CBE4 were too small to be presented graphically. Three out of 4 CBE samples tested for *DES* and 4 out of 4 samples tested for *DMN* showed underexpression as compared with the controls, with fold-change ranging from 0.06-0.65 for *DES* and 0.00004-0.0001 for *DMN*, respectively. Thus, the qPCR confirmed that *DES* and *DMN* were differentially expressed as expected by the microarray analysis. For *AQP3* and *PERP*, not all samples showed the same pattern of expression: 2 out of 4 tested CBE samples were overexpressed, with fold-change values 4.98 and 29.46 for *AQP3* and 2.84 and 3.50 for *PERP*, respectively. The fact that not all tested CBE samples confirmed the expression patterns for *DES*, *AQP3*, and *PERP* may be because the samples used in the qPCR were different from those used for the microarray analysis. Furthermore, as a complex trait, BEEC is likely to be a genetically heterogeneous condition with individuals harboring genetic defects in different, but interacting genes.

Table I. Differentially expressed transcripts consistently present in three embryonic mesenchyme samples.

<table>
<thead>
<tr>
<th>Profile NB vs. EB</th>
<th>Observed</th>
<th>EM1</th>
<th>EM2</th>
<th>EM3</th>
<th>Intersection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up-regulated</td>
<td>301</td>
<td>148</td>
<td>132</td>
<td>125</td>
<td>102</td>
</tr>
<tr>
<td>Down-regulated</td>
<td>283</td>
<td>159</td>
<td>146</td>
<td>123</td>
<td>105</td>
</tr>
<tr>
<td>Total</td>
<td>584</td>
<td>307</td>
<td>278</td>
<td>248</td>
<td>207</td>
</tr>
</tbody>
</table>

NB, normal bladder; EB, exstrophic bladder; EM, embryonic mesenchyme; intersection, number of transcripts in common, among all 3 EM samples. 584 transcripts correspond to 422 distinct genes; 207 transcripts correspond to 162 distinct genes.

![Figure 1](755-765.qxd 7/4/2011 01:12 ÌÌ  ™ÂÏ›‰·758)

Figure 1. qPCR data of selected genes including one down-regulated (*DES*) and two up-regulated (*PERP*, *AQP3*) genes in microarray analysis. The Y-axis shows the fold-change and the X-axis shows the samples: NB is the average of three controls, C1 represents one control and CBE1-CBE6 are cases. Fold-change < 1 represents down-regulation and fold change > 1 represents up-regulation.
Biological network analyses. Biological network analyses of BEEC candidate genes were performed using the Ingenuity Pathways analysis software. The 162 BEEC candidate genes fell into 11 potential networks, 9 of which included at least 10 candidate genes (Table II). Eight of the 11 networks have top functions related to skeletal and muscular system development and function, cellular assembly and development, tissue development, organ morphology, embryonic development, skeletal and muscular disorders or connective tissue disorders. The largest network consists of 24 candidate genes (including the most underexpressed genes, DES, and DMN) and has a score of 43, indicating there is almost 100% chance that the network was not randomly generated. The main functions of this network include skeletal and muscular system development and function, tissue morphology, cellular assembly and organization. In particular, 16 of these 24 candidates are located in the cytoplasm (Fig. 3), where the functions for cell expansion, growth and replication are carried out. In the EB samples, 12 (75%) of the 16 genes were underexpressed compared to the NB samples, and these genes are closely involved in muscular system development and function and in tissue morphology.

Genes related to desmosomal function and/or cytoskeletal regulation. The 2 most underexpressed genes in EB, DES (fold-change, -74.7) and DMN (fold-change, -53.0), encode muscle-specific intermediate filament (IF) proteins, which directly interact with the C-terminal domain of desmoplakin, encoded by the sixth most overexpressed gene in EB samples (DSP, fold-change, +48.8). Keratin proteins constitute a major component of the IF in all epithelia, and we observed that 3 type I cytokeratin genes were up-regulated in the EB: KRT10 (fold-change, +48.7), KRT18 (fold-change, +8.7) and KRT19 (fold-change, +28.2).

The observation that the genes with the highest degree of differential expression between EB and NB (DES, DMN, DSP) were related to the desmosome motivated us to investigate whether other candidate genes are also associated with desmosomal function and/or cytoskeletal regulation. By categorizing the 162 BEEC candidate genes, we found that 30.2% (n=49) of them are associated with desmosomal structure/function or with cytoskeletal regulation. Moreover, 34 of the 49 embryologically expressed genes (69.3%) were underexpressed in exstrophic human bladders, implying deficient cytoskeletal organization and more specifically, compromised desmosome-mediated intercellular adhesion in the exstrophic smooth muscle.

Mapping of differentially expressed genes to BEEC candidate loci. To further validate the candidate genes, we analyzed their chromosomal localization for physical overlap with regions of putative linkage/linkage disequilibrium identified by our recent work and breakpoints of chromosome rearrangements observed in BEEC patients. Seven candidate genes, SLC8A1, FLNC, CALD1, SYNPO2, RRM2, ROD1, and ASS1, showed partial or complete physical overlap with regions of putative linkage or breakpoints of chromosome rearrangements in BEEC patients (7,26). All these genes were integrated in 5 of our top 10 biological networks: CALD1 in network 1, ASS1, SLC8A1 and SYNPO2 in network 2, RRM2 in network 3, FLNC in network 6, and ASS1 and ROD1 in network 9 (Table II). CALD1 encodes caldesmon 1, a calmodulin- and actin-binding protein playing an essential role in the regulation of smooth muscle and cytoskeleton contraction. RRM2 encodes 1 of 2 non-identical subunits for ribonucleotide reductase and functions downstream of β-catenin as an inhibitor of Wnt signaling. SLC8A1 functions as a sodium/calcium exchanger and it rapidly transports Ca^{2+} during excitation-contraction coupling. SYNPO2 has an actin-binding and actin-bundling activity. ROD1 may play a role in the regulation of differentiation, and FLNC is a muscle-specific filamin playing a central role in muscle cells. It may be involved in reorganizing the actin cytoskeleton in response to signaling events and may also display structural functions at the Z-discs in muscle cells.

Correlation with mouse expression data. A total of 148 mouse orthologous transcripts of these 162 candidate genes were present on the GeneChip Mouse Genome 430 2.0 array, and approximately 80% (n=118) of these mouse orthologs
Table II. Result of Ingenuity pathways analysis based on 162 candidate genes.

<table>
<thead>
<tr>
<th>Molecules in network</th>
<th>Score</th>
<th>Focus molecules</th>
<th>Top functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Actin, ACTN1, α-Actinin, CALD1, CaMKII, CSRP1, DES, DSP, DTNA, ERK, ERMP1, F Actin, Fascin, FLNA, GEM, KLF5, LPP, McI, MYL9 (includes EG:10398), NEXN, Pak, PALLD, PDGF-AA, PDLIM1, PDLIM7, PP1B, PP1R12A, PRKACB, Rock, S100A11, SYNM, TNS1, TPM1, TPM2, Tropomyosin</td>
<td>43  24</td>
<td>Skeletal and muscular system development and function; tissue morphology; cellular assembly and organization</td>
<td></td>
</tr>
<tr>
<td>2 14-3-3, Ap1, AQP3, ASS1, BAG2, CDC4, CRABP2, CTSC, DHC2, ERK1/2, FABP5, FHL1, FSH, GROE1, HSPB6, IgG, IL1, IL12 (complex), Jnk, Mapk, MIP12, P38, MAPK, PDGF BB, PIHLD2, Pkc(s), SC4MOL, SFN, SLC16A1, SLC8A1, SOX9, SYNP2, Tgf β, TXNDC17, WNT5A, ZAK</td>
<td>39  21</td>
<td>Cell death; cell signaling; embryonic development</td>
<td></td>
</tr>
<tr>
<td>3 ACTC1, Akt, ALP, CA12, Caspase, CDC2, Cyclin A, DTL, E2f, EFNB2, Growth hormone, hCG, IgF, Histone h3, Histone h4, Insulin, KIAA0101, KRT10, KRT18, MCM4, MCM5, Mek, NUSAP1, PCNA, PP2A, Rb, RNA polymerase II, RRM2, SACS, Shc, TOP2A, TSPO, TYMS, UHRF1, Vegf</td>
<td>33  19</td>
<td>Cancer; DNA replication; recombination and repair; genetic disorder</td>
<td></td>
</tr>
<tr>
<td>4 ALDH1B1, Calpain, CD24, Collagen(s), DDR1, FAK, FERMT2, Fibrinogen, FNI, HSPA1A, IFN β, IGFBP2, INPP5A, Integrin, Integrin α2 B1, Integrin α3 B1, Integrin α4 B1, Integrin α, ITG2A, ITGA5, ITGA6, Laminin, MAP2K1/2, NPxβ (complex), PI3K, PRDX4, Rac, Ras, Ras homolog, RHOB, SLC25A4, TFRC, TGM2, TLR1, TXNRD1</td>
<td>32  18</td>
<td>Cell-to-cell signaling and interaction; cancer; cellular assembly and organization</td>
<td></td>
</tr>
<tr>
<td>5 ACTC1, AKR1B1, AMMECR1, APRT, ATP5J2, BAT1, β-estradiol, CA12, CDC4, copper, CSTB, DDX39, FAM129A, FGFI, FOLH1, GGH, GJB1, GJC1, GRNLiA, MIRLET7B (includes EG:406884), MYC, MYO9A, PSMC2, RFC3, RNF115, SARNP, SHMT1, SMT1, SLIT2, TIMM11, TMED3, TUBG1, UQCRQ, VHL</td>
<td>25  15</td>
<td>Tissue development; organ morphology, amino acid metabolism</td>
<td></td>
</tr>
<tr>
<td>6 CAMK2G, CDC42EP3, CFL2, CHD3, CSNK2A2, DKK1, EWSR1, FLNC, HDAC1, HNF4A, JUB, JUN, KRR1, MIR297-2, MLF1IP, MREG, MYO22, NUDT5, PEX13, PEX14, PHB2, Pka, PLEKHO1, PRUN2, PTBP2, SGCB, SGCD, SGCE, SGCG, SNAI2, SVIL, TMEM35, TSPAN6, UMP3, WNT11</td>
<td>25  15</td>
<td>Developmental disorder; skeletal and muscular disorders; tissue development</td>
<td></td>
</tr>
<tr>
<td>7 5-hydroxytryptamine, BMP1, CNRK, copper, CSNK2A2, ETHE1, FOS, H-d-neuronal intranuclear inclusions, HSP90AB1, HTT, ITGA2, KBTBD10, KCNE4, KLF5, LTBP2, MIR17 (includes EG:406952), MSRB3 (includes EG:253827), MYCN, NANS, PERP, RAB27B, RNF144B, RPH3A, RPL12 (includes EG:6136), RPSA, SCHIP1, SLMAP, STARD13, STRN3, TGFBI, TM4SF1, TMED9, TP53, TUBA4A, UNC13D</td>
<td>20  13</td>
<td>Nervous system development and function; neurological disease; cellular development</td>
<td></td>
</tr>
<tr>
<td>8 ABL1M1, AFF3, AIM1 (includes EG:202), ARHGDA1A, C60RF115, CDH3, CDH5, DSG2, DSP, ERBB2, ERBB2IP (includes EG:55914), ETV4, Fascin, FER (includes EG:2241), IKS1, Itgam-Itgb2, ITGAX, ITGGB2, JAM3, JUP, KIAA1217, MUC1, MYO10, NCK2, PECAM1, PKN2, RAB8B, RIB2, SH3YL1, SMARCA4, SNAI2, VIM, ZEB1, ZNF185, ZWINT (includes EG:11130)</td>
<td>18  12</td>
<td>Cell-to-cell signaling and interaction; tissue development; cellular movement</td>
<td></td>
</tr>
<tr>
<td>9 ACAN, ARFGAP3, ASS1, CD3E, CLIC4, CTSC, DEFB4 (includes EG:1673), DVL1, ESM1, GPR126, GPSN2, H1N1, HSPG2 (includes EG:3339), IFNG, MARCKSL1, MGP, MMP10, MUC1, NDFIP2, NFA57, NTF2, PAPP, PECAM1, PRL, RAB23, ROD1, SCUBE1, SCUBE3, SEMA3C, SLC29A1, SMTN, SNAI2, SOD1, TNF, TNFAIP2</td>
<td>18  12</td>
<td>Connective tissue disorders; inflammatory disease; skeletal and muscular disorders</td>
<td></td>
</tr>
</tbody>
</table>
Table II. Continued.

<table>
<thead>
<tr>
<th>Molecules in network</th>
<th>Score</th>
<th>Focus molecules</th>
<th>Top functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 ADCY, amino acids, ARAF, Ck2, CSNK2A2, DAPK1, <strong>DMPK</strong>, DUSP7, DYRK1A, geranylgeranyl pyrophosphate, GTP, HTR1B, Importin α/β, Importin β, KIF5B, MAP3K8, MC2R, NME2, <strong>NUP50</strong>, PAK3, PKD2 (includes EG:5311), PKN1, PKN2, PPIR8, PTPLA, PTPN2, RANBP2, RNF7, ROCK1, RRAD, <strong>S100A16</strong>, SSB, STIP1, TCOF1, TGM2</td>
<td>9</td>
<td>7</td>
<td>Amino acid metabolism; post-translational modification; small molecule; biochemistry</td>
</tr>
<tr>
<td>11 <strong>KCTD1</strong>, NFYB</td>
<td>2</td>
<td>1</td>
<td>Cellular development; gene expression</td>
</tr>
</tbody>
</table>

Genes in bold are among the list of 162 differentially expressed candidate genes. *The significance score of each pathway is calculated by a hypergeometric test.

Figure 3. The largest pathway from Ingenuity pathways analysis. A potential network involving 24 candidate genes (filled shapes) according to cellular process. The extent of shading is indicative of the magnitude of regulation. Red and green shadings indicate up-regulation and down-regulation, respectively. The genes shown in transparent circles indicate those from the Ingenuity pathways knowledge base that have been determined as interacting with the list of input candidate genes, on the basis of direct physical, enzymatic, and transcriptional interactions between orthologous mammalian genes from published, peer-reviewed articles. Node shapes indicate function. Diamond designates an enzyme; square, growth factor; triangle, kinase; circle, other. ➝ indicates physical interactions; | indicates inhibition. Sixteen of the 24 potential candidate genes are present in the cytoplasm.
were expressed at GD13, the time-point corresponding to the human EM. This high degree of concordance is especially significant, since it would be expected that most of the genes that regulate or participate in embryonic bladder development are expressed at the same developmentally relevant period in both humans and the mouse. Moreover, the 118 orthologous genes corresponding to these 148 transcripts were also investigated for their association with desmosomal structure/function or cytoskeletal regulation, and 28.8% (n=34) of these were associated with cell-cell adhesion functions. This percentage is essentially the same as that found in the human data set (30.2%).

**Tissue expression pattern of BEEC candidate genes in midgestation mouse embryos.** Candidate genes were also examined for their expression in midgestation mouse embryos by search of the MAMEP database. By search of the whole mount *in situ* expression data available in the MAMEP database, 5 of the 87 top ranked up-regulated genes (*PHLD2*, *WNT5A*, *CRABP2*, *SCAMOL*, *MYO10*) and 11 of the 74 top ranked down-regulated genes (*DES*, *TNS1*, *TPM1*, *CSRP1*, *FN1*, *ACTC1*, *ELNA*, *FNLC*, *LPP*, *CALD1*, *ACTN1*) were found to be expressed in the ventrocaudal mesoderm, the allantoic/umbilical region and/or the hindgut endoderm including the cloacal membrane at GD9-GD10. These tissue expression patterns suggest that these genes are involved in the formation of the intra-umbilical region and additionally confirm the relevance of our microarray results. In particular, the overexpressed gene *PHLD2* and the underexpressed genes *TPM1*, *CSRP1* and *TNS1* show a strong and defined expression in the ventrolateral area surrounding the cloacal membrane and the basement of the allantois including the bladder primordia. *CALD1* and *ITGA5* expression is more confined to the allantois, while *LPP* and *CRABP2* exhibit expression domains in the area surrounding the hindgut region and cloaca. *FN1* and *WNT5A* transcription takes place in the paraxial and lateral mesoderm in the caudal end of the embryo, the sites of induction and formation of the trunk progenitor tissues. In accordance to their functions in muscle development, *DES* and *FNLC* are expressed in the differentiating somites of the trunk. Additionally, *DES* shows an expression domain in the cloacal membrane and the ureteric cord, while *FNLC* message is detected in the ventrolateral tissues that are associated with the cloaca.

**Discussion**

BEEC represents a complex group of birth defects caused by failed development of the urogenital system during the first trimester of gestation with a strong evidence for involvement of genetic factors in its etiology. Analyses of genes suggested as candidates by rare chromosomal aberrations, co-occurring syndromes or syndromes with phenotypic overlap, seem rational, but yet have failed to identify causative mutations (5,28-33). Comparative genomic hybridization in a CBE family has only identified a known copy-number variant, most likely unrelated to the phenotype (34) and linkage studies have not identified strong candidate loci (26,27).

We aimed to identify high-probability BEEC candidate genes by expression profiling of normal and exstrophic human urinary bladders, and by focusing on differentially expressed genes present during the relevant bladder embryologic developmental period in humans and the mouse. We identified 422 genes differentially expressed in EB as compared to NB, and among them, 162 were expressed in EM samples, with most of them (73%) also found in mouse embryonic bladder samples corresponding to the human embryonic bladder development period. This finding indicates that most of the genes regulating or participating in embryonic bladder development are expressed in the same relevant period in both human and mouse. In addition, the search of the whole mount *in situ* expression data in the MAMEP database show that 16 candidate genes are expressed in the ventrocaudal mesoderm, the allantoic/umbilical region and/or the hindgut endoderm including the cloacal membrane at GD9-GD10, suggesting that these genes are involved in the formation of the infra-umbilical region, and confirming the relevance of the microarray results. Pathway analysis shows that most of the networks composed of the 162 genes have top functions related to cellular assembly, growth or development, skeletal and muscular system development, organ morphology or connective tissue morphology.

Our analysis of the 162 candidate genes revealed 2 important clues for further studies of BEEC etiology. First, about 30% of candidate genes identified were associated with desmosomal structure/function or cytoskeletal assembly, and 69% of these genes were underexpressed in exstrophic human bladders, suggesting desmosomal and/or cytoskeletal deficiency and/or deregulation. Secondly, *p63*, *PERP*, *SYNPO2* and genes of the Wnt pathway were implicated in BEEC as discussed below.

**Desmosomal and/or cytoskeletal deregulation in BEEC.**

The desmosome, a rivet-like cellular junctional structure, is primarily located in epithelial and muscle cells and is specialized for cell-cell adhesion. In addition to maintaining the structural integrity of tissues by resisting shear forces, desmosomes are critical for the formation of organs and tissues during embryogenesis (35). The molecular blueprint of the functional desmosome varies between tissues, but generally consists of 3 principal components including desmosomal cadherins (desmocollins and desmogleins), armadillo type of junctional proteins (plakophilins and plakoglobin), and plakin proteins (desmoplakin, plectin, envoplakin and periplakin) (36,37). Plakin proteins serve as an anchor of the IF. As an integral part of the cytoskeleton IF confer resistance to physical stress by distributing mechanical forces throughout a tissue and participate in the formation of supracellular scaffolding that maintains the integrity of tissues. In addition to these classical mechanical functions of the desmosome, there is evidence that mechanical forces transmitted along the length of the IF to the nucleus regulate gene expression (38).

Keratin proteins constitute a major component of the IF in all epithelia and 3 type I cytokeratin genes, *KRT10*, *KRT18* and *KRT19*, were found to be up-regulated in the EB. Even more intriguing was that the 2 most down-regulated genes in EB, *DES* and *DMN* encode muscle-specific IF proteins that directly interact with the C-terminal domain of desmoplakin, encoded by *DSP*, the sixth most overexpressed gene in EB.
samples. Evaluation of DES in the MAMEP database reveals that it is expressed in the differentiating somites of the trunk, the cloacal membrane and the urogenital cord, in accordance to its function in muscle development. These findings indicate that abnormal formation and/or anchoring of the IF to the desmoplakin component of the desmosome may be the specific cellular defect in BEEC. Our observations warrant morphological analysis of the exstrophic bladder desmosomes by immunofluorescence and electron microscopy.

**BEEC candidate genes - p63, PERP, SYNPO2 and genes of the Wnt pathway.** Recently, p63 has been considered as a CBE candidate gene based on a murine knockout model (39,40), however no mutations were found in 15 CBE patients and 5 patients with CE (Ching BJ, et al, ASHG Annual Meeting, 478, 2007). We observed that PERP, a tetraspan protein that is a direct p63 downstream target, was 22.2-fold overexpressed in all three human bladder sets and present in all EM samples. Intriguingly, PERP, specifically localizes to the desmosomes where it extends through the plasma membrane to link neighboring cells. Whole-mount in situ hybridization studies have demonstrated prominent PERP expression in the bladder (41). Furthermore, we have observed overexpression of some transcription-activating p63 isoforms in exstrophic bladders and underexpression or complete lack of other AN-p63 isoforms, supporting possible desmosomal defects in BEEC (42).

Another highly up-regulated gene was WNT5A (14-fold overexpression). During chick cloacal formation, in the mesoderm of the cloacal region, overexpression of the constitutively active form of mouse Gli2 has been reported to induce expression of WNT5A, Ptcsh1, Gli2, bmp-4, and hoxd-13, previously shown to play a role in hindgut patterning. Interestingly, p63 expression in the cloacal endoderm is also up-regulated by Gli2 (43), suggesting that both WNT5A and p63 are regulated by a common upstream pathway. WNT5A is strongly expressed in the paraxial and lateral mesoderm in the caudal end of the embryo, and is essential in the caudal extension of the body axis as well as the formation of the external genitalia (44). WNT5A has been integrated in network 2 in this study (Table II) and an independent expression profiling of exstrophic bladders has also detected many components of the Wnt pathway (22). Further support for a causal interplay of WNT5A in the expression of BEEC, arises from the physical overlap of the RRM2 gene with a region of suggestive linkage (LOD=1.45 in a dominant model with 80% penetrance) from our own previous work (26). RRM2, found to be up-regulated in the EB tissues, functions downstream of β-catenin as an inhibitor of Wnt signaling (45).

SYNPO2, encoding myopodin or synaptodomin 2 and found to be integrated in network 2 (Table II), is a dual compartment protein displaying actin-bundling activity. It redistributes between the nucleus and the cytoplasm in a differentiation-dependent and stress-induced manner (46). In bladder cancer, previous studies showed differential nuclear localization of myopodin among bladder cancer cells while sharing structural cytoplasmic expression patterns (47). During myotube differentiation, SYNPO2 binds to stress fibers in a punctuated pattern before incorporation in the Z-disc. Vertebrate Z-discs are remarkably complex structures serving several important functions, e.g. stabilizing and anchoring actin filaments during contraction and providing an attachment for other cytoskeletal proteins, such as desmin, nebulin, and titin. It appears that Z-discs of both skeletal and cardiac muscle respond dynamically to mechanical stress through transitions in their lattice structure (48). In this context, FLNC and CALDI, both down-regulated in the EB tissues, physically overlap regions of putative linkage (26,27), and both encode proteins needed for the stability of the cytoskeleton (49). The function of CALDI is also critical for the assembly of the desmosome (50). Moreover, according to the MAMEP database, FLNC and CALDI show strong expression in the differentiating somites of the allantois at GD9 and GD10, respectively. In light of the increased risk for adenocarcinomas of the exstrophic bladder (51), our observations warrant further investigations of p53, WNT5A and SYNPO2 in bladder carcinomas of BEEC patients.

**Validation of microarray data.** qPCR has been widely used to corroborate microarray expression data. Logically, we attempted to validate our observations by qPCR. However, lack of sufficient RNA from the samples used for the original hybridization experiments forced us to use other exstrophic and control specimens for the qPCR. We found a strong correlation between microarray expression and qPCR data for the two most underexpressed genes, DMN and DES, and for the most overexpressed gene, AQP3. We also documented that another overexpressed integral desmosomal component, PERP, a direct p63 downstream target is overexpressed by qPCR in 2 of the 4 tested exstrophic samples (Fig. 1). Due to the limited number and quantity of available exstrophic bladder specimens we could not perform qPCR for more genes. Despite this shortcoming, the consistency of the microarray results, among 3 sets of EB/NB samples corroborated by qPCR analysis strongly support the derived candidate gene list as a valid target for further molecular studies. We believe that our data unequivocally implicate desmosomal and/or cytoskeletal deregulation in the etiology of BEEC and the existing p63-/- mouse model is in agreement with our conclusions, as evident by the p63/PERP (i.e. desmosomal) interactions.

In summary, this study provides the first in depth expression signature of BEEC and has identified a set of candidate urogenital genes expressed during the relevant developmental window of bladder development. The resulting set of 162 genes clearly shows over-representation of genes related to desmosomal and cytoskeletal assembly and comprises a reasonable initial set for future candidate gene studies. In addition, genome-wide association studies (GWAS) and profiling of gene expression are two major technological breakthroughs that allow hypothesis-free identification of candidate genes. Thus the current gene expression study provides a powerful basis for future GWAS or candidate gene association studies on human BEEC to identify relevant genetic contributors.

**Acknowledgements**

SAB is partially supported by the Children's Miracle Network (CMN) endowed chair in pediatric genetics. This project has...
been partially supported through grant M01-RR00052 from the NCRR/NIH, and by a CMN grant CNRSB06 to SAB. LW, MD, ML and HR are members of the ‘Network for Systematic Investigation of the Molecular Causes, Clinical Implications and Psychosocial Outcome of Congenital Uro-Rectal Malformations (CURE-Net)’ supported by a research grant (01GM08107) from the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF).

References


