J Am Soc Nephrol 15: 2344–2357, 2004

Identifying the Molecular Phenotype of Renal Progenitor Cells

GRANT A. CHALLEN, GEMMA MARTINEZ, MELISSA J. DAVIS, DARRIN F. TAYLOR, MARK CROWE, ROHAN D. TEASDALE, SEAN M. GRIMMOND, and MELISSA H. LITTLE
Institute for Molecular Bioscience, The University Of Queensland, Brisbane, Queensland, Australia

Abstract. Although many of the molecular interactions in kidney development are now well understood, the molecules involved in the specification of the metanephric mesenchyme from surrounding intermediate mesoderm and, hence, the formation of the renal progenitor population are poorly characterized. In this study, cDNA microarrays were used to identify genes enriched in the murine embryonic day 10.5 (E10.5) uninduced metanephric mesenchyme, the renal progenitor population, in comparison with more rostral derivatives of the intermediate mesoderm. Microarray data were analyzed using R statistical software to determine accurately genes differentially expressed between these populations. Microarray outliers were biologically verified, and the spatial expression pattern of these genes at E10.5 and subsequent stages of early kidney development was determined by RNA *in situ* hybridization. This approach identified 21 genes preferentially expressed by the E10.5 metanephric mesenchyme, including Ewing sarcoma homolog, 14-3-3 θ, retinoic acid receptor-α, stearoyl-CoA desaturase 2, CD24, and cadherin-11, that may be important in formation of renal progenitor cells. Cell surface proteins such as CD24 and cadherin-11 that were strongly and specifically expressed in the uninduced metanephric mesenchyme and mark the renal progenitor population may prove useful in the purification of renal progenitor cells by FACS. These findings may assist in the isolation and characterization of potential renal stem cells for use in cellular therapies for kidney disease.

There is increasing evidence from a number of organ systems that cells with at least multipotentiality (1,2) and possibly pluripotentiality (3) exist in adult organs that previously were thought not to contain such populations. Although cell division is infrequent in the adult kidney, this organ possesses the capacity for regeneration as evidenced by the cellular proliferation observed during recovery from conditions such as acute tubular necrosis (4). Although cell lineage relationships during renal repair and regeneration are poorly defined, it is possible that these processes recapitulate some aspects of embryonic renal development. The permanent kidney (metanephros) arises *via* reciprocal interactions between two tissues, the ureteric bud (UB) and the metanephric mesenchyme (MM) (5). Each of these tissues is initially derived from the intermediate mesoderm (IM), although the UB develops as a caudal outgrowth from the nephric duct (ND), whereas the MM develops from the nephrogenic cord, the same block of tissue that gives rise to the pronephroi and the mesonephroi (5). After invasion by the UB, cells of the MM are induced to differentiate into specific renal lineages. The central dogma of kidney development suggests that the UB forms the ureter and collecting duct system of the mature kidney, whereas the MM gives rise to the remaining portions of the nephrons, from Bowman's capsule to distal tubule (6). Although many of the interactions between the UB and MM are now well characterized, the processes by which the MM initially differentiates from surrounding IM and becomes committed to a renal fate remain poorly understood.

The uninduced MM is composed of a few thousand morphologically similar mesenchymal cells. Tissue recombination experiments have shown the MM is the only embryonic cell population that can be induced to undergo nephrogenesis (5). The uninduced MM has been proposed to be the renal progenitor population because it has the ability to differentiate into many more differentiated cell types than UB precursors (7). Several reports have generated evidence that support this notion. Lineage-tracing studies have suggested that uninduced MM cells not only have the potential to develop into all of the epithelial regions of the nephron but also can be incorporated into UB-derived collecting duct epithelia (8). Further evidence for the developmental potential of this tissue was recognized when embryonic porcine metanephroi that were transplanted into immunodeficient mice developed nonrenal derivatives such as cartilage and bone in addition to mature glomeruli and tubules (9). Several other studies of cell lines isolated from embryonic kidneys have indicated that the MM displays multipotentiality (7,10,11), although *in vivo* experiments have yet to complement these findings. What is not presently known is whether the multipotentiality of the uninduced MM results from a homogeneous progenitor cell population that is able to form all renal derivatives or if the MM already contains a mixture of progenitor cells that are committed to forming the different lineages found in the mature kidney (Figure 1A) (12). In either case, we refer to the cells that compose the committed

Received February 10, 2004. Accepted June 4, 2004.
Correspondence to Dr. Melissa H. Little, Institute For Molecular Bioscience, Queensland Bioscience Precinct, 306 Carmody Road, The University of Queensland, St. Lucia, Brisbane, QLD, 4072. Phone: +61-7-3346-2054; Fax: +61-7-3346-2101; E-mail: M.Little@imb.uq.edu.au
1046-6673/1509-2344
Journal of the American Society of Nephrology

DOI: 10.1097/01.ASN.0000136779.17837.8F

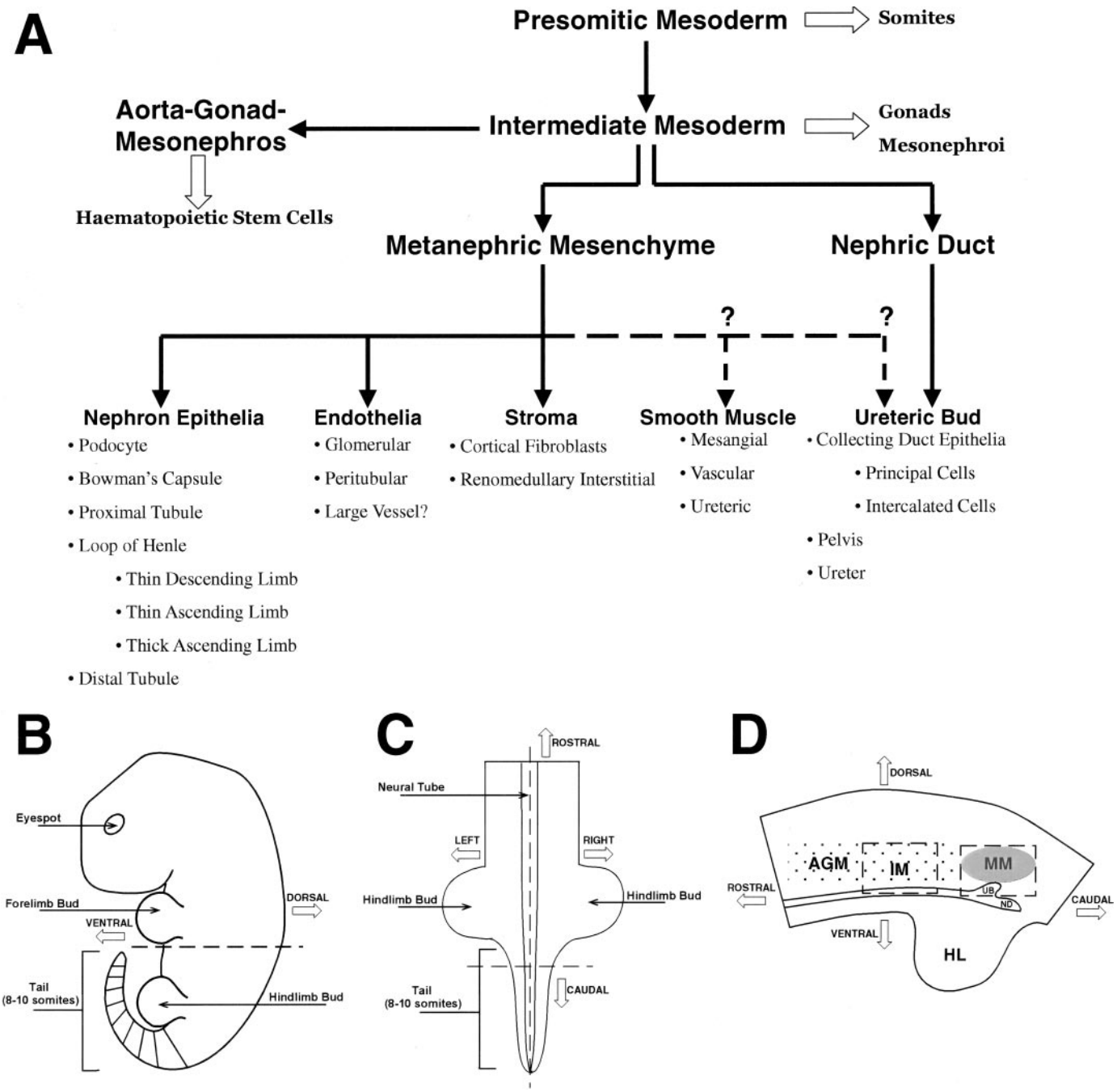

Figure 1. (A) Origins of cell lineages in kidney development. Solid lines indicate known differentiation pathways, and dashed lines suggest possible pathways based on experimental evidence. The uninduced metanephric mesenchyme (MM) is proposed to be the renal progenitor population because it has the ability to give rise to many more differentiated cell types than the ureteric bud (UB). (B through D) Isolation of tissue for microarray analysis. Dashed lines indicate planes of dissection. Embryonic day 10.5 (E10.5) embryos were cut transversely below the forelimbs (B), and the hind portion then was cut longitudinally down the neural tube (C) with the tail removed. This produced two tissue pieces, each with the nephric duct (ND) and MM visible (D). The region of uninduced MM was located directly dorsal to the emerging UB. The dashed lines indicate the regions of tissue that were collected for microarray analysis. To ensure that the entire MM was isolated, the region indicated by the more caudal box was collected, leading to possible contamination from UB tissue. A region of intermediate mesoderm-derived tissue indicated by the rostral box, comprising ND, nephrogenic cord, and aorta-gonad mesonephros region, was taken from the same embryo.

but uninduced MM (embryonic day 10.5 [E10.5]) as the renal progenitor population.

Microarrays have been used to generate temporal profiles of gene expression over the course of metanephric development (13,14) and to analyze the expression profiles of discrete kidney subcompartments (15) and renal cell lines (16). Although these data represent important steps in the creation of a catalogue of gene expression during kidney development and

the establishment of a baseline for the future examination of mutant mice, to date, no study has applied microarray technology to the renal progenitor population. A greater understanding of the processes that leads to commitment of progenitors to renal differentiation and the identification of specific cell surface markers of renal progenitors is essential in the search for renal stem cells. The two main goals of this study were (*1*) to identify the earliest markers of commitment to renal differentiation and (*2*) to determine the cell surface molecule phenotype of renal progenitor cells. By comparing the gene expression profile of uninduced MM to more rostral IM-derived tissue, we have defined some of the earliest genes expressed that distinguish the uninduced MM from surrounding nonmetanephrogenic tissue. These genes represent markers that signify commitment of the renal progenitors to a metanephric fate and, as such, may prove useful in the search for renal stem cells in the mature postnatal kidney or in the induction of nonrenal stem cells to adopt a renal fate.

Materials and Methods

Tissue Collection and RNA Isolation

Naturally mated outbred female CD1 mice were culled by cervical dislocation (Animal Ethics Committee approval number IMB/479/02/NIH). MM and rostral IM tissue was dissected from E10.5 embryos and snap-frozen on dry ice. Embryos were defined as E10.5 by the presence of 8 to 10 tail somites. E10.5 embryos were dissected transversely below the forelimbs and longitudinally along the neural tube to expose the ND, UB, and MM (Figure 1). Uninduced MM was identified spatially as the area immediately dorsal to the caudal terminus of the ND, from which the primitive UB had just begun to emerge. The emerging UB and uninduced MM are located medial to the caudal third of the hindlimb bud at E10.5, approximately halfway between the dorsal and ventral sides of the embryo and adjacent but medial to the mesonephric duct. IM tissue rostral to the MM, comprising the mesonephros and genital ridge, was dissected from the same embryos that were used for MM collection. Dissections were performed in cold PBS, and pooled tissue was stored at −80°C. Tissue was collected from 65 embryos, representing eight litters, and pooled for RNA isolation. Total RNA was prepared using Trizoll (Gibco BRL) extraction and cleaned with RNeasy mini kits (Qiagen) including on-column DNase digestion (Qiagen). RNA concentration and quality were determined by both spectrophotometry and on bioanalyzer RNA micro-fluidic chips (Agilent) and visualized by agarose gel electrophoresis.

Target Amplification and Microarrays

Total RNA was linearly amplified using the messageAMP aRNA kit (Ambion). Briefly, 1000 ng of total RNA was reverse transcribed into cDNA using a T7 promoter-dT primer and amplified through an *in vitro* transcription reaction (12 h) using T7 RNA polymerase to produce antisense RNA (aRNA). Amplification reactions yielded a minimum of 20 μg of aRNA. The quality and integrity of the aRNA were assessed by spectrophotometry and agarose gel electrophoresis before labeling. Five micrograms of each aRNA sample was reverse transcribed using random hexamers (Promega) into cDNA incorporating either Cy5- or Cy3-labeled dUTP (Amersham) and hybridized to microarray chips for 16 h at 45°C. Microarray analysis was performed in triplicate, including duplicate hybridizations using aRNA from independent amplification reactions (technical replicate for amplification) and dye swap replicates using the same pool of aRNA (technical replicate for labeling).

Arrays were produced by the SRC Microarray Facility, University of Queensland (ARC Centre for Functional and Applied Genomics). Experiments were performed using chips arrayed with the National Institutes of Health–National Institute of Aging (NIA) 15K mouse cDNA clone set (http://lgsun.grc.nia.nih.gov/cDNA/15k.html) in addition to a selection of custom clones submitted from the Institute for Molecular Bioscience (University of Queensland, Brisbane) and contained 15,989 elements in total (http://microarray.imb.uq.edu.au/). Every element was spotted in duplicate on each chip. The NIA 15K mouse clone set contained 15,247 expressed sequence tags (EST) derived from pre- and peri-implantation embryo, E12.5 female gonad/mesonephros, and newborn ovary cDNA libraries (17), thus making it an ideal gene set for this experiment. A total of 14,428 of these EST sequences map to Unigene/TIGR/Ensembl and represent 11,834 distinct transcripts after removal of duplicates.

Array Analysis and Bioinformatics

Hybridized slides were scanned with a GMS 418 array scanner (Genetic MicroSystems), and images were analyzed with Imagene 5.5 (Biodiscovery). Microarray data were uploaded into BioArray Software Environment (BASE) 1.2.10 (18) and analyzed with R statistical software using the LIMMA package (http://bioinf.wehi.edu.au/limma/) with scripts developed by Ola Spjuth of the Linnaeus Centre for Bioinformatics (http://www.lcb.uu.se/baseplugins.php). Mean foreground signals were taken for each spot and normalized within each array using print-tip lowess without background correction to give a mean value of zero for the log ratios of the two channels within each print block. The final normalized values were used for B-statistics calculations. B-statistics analysis included an allowance for the correlation between adjacent duplicate spots printed on the same array. Differential expression was defined using a robust statistical method rather than simple fold change. All genes were ranked using the B-statistic method, whereby both fold change and variance of signals in replicates are used to determine the likelihood that genes are truly differentially expressed. Genes with a B score >0 have a $>50\%$ probability of being truly differentially expressed. This analysis was executed using the Bio-conductor package that has been implemented as a plug-in tool in BASE. Normalized data were also exported from BASE into Genespring 5.1 for visualization and generation of lists of differentially expressed genes. Full documentation of cDNA array fabrication, gene content, experimental procedures, and all results is available for download in accordance with MIAME guidelines at http://kidney.scgap.org/base/index.phtml.

Representative sequences for differentially expressed EST were extracted from the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/). Using BLAST (19), each NIA EST sequence was mapped to an identical full-length RIKEN representative transcript/protein sequence (RTPS 6.3) (20). BLAST hits were scored on the basis of a combination of three features: e-value, identity, and coverage—calculated from the alignment. The RIKEN RTPS 6.3 set was annotated with a prediction of secretory status and membrane organization (21), and proteins were classified as belonging to one of six classes on the basis of the presence or absence of endoplasmic reticulum signal peptides and helical transmembrane domains (Table 1). Signal peptides were predicted by a best of three independent prediction methods, whereas transmembrane domains were annotated by five prediction methods, with a minimum of three positive predictions required for the region of a domain to be predicted. Potential conflicts between predictions of

Table 1. Membrane organization classification of proteins[a]

Class[b]	Description	SP[c]	TMD[d]
A	Soluble, intracellular protein	No	No
B	Soluble, secreted protein	Yes	No
C	Type I membrane protein	Yes	Yes
D	Type II membrane protein	No	Yes
E	Multispan membrane protein	Yes or No	>1
F	Unclassified: conflicting predictions	?	?
0	No prediction	?	?

[a] Proteins are classified into classes A through E on the basis of the presence or absence of signal peptides and transmembrane domains. Classification into class F occurs when the membrane organization cannot be resolved as a result of conflicts in the bioinformatic prediction pipeline. Proteins are designated class 0 in cases in which the RIKEN RPS transcript represents only a partial open reading frame. Full-length sequences are required for membrane organization prediction.
[b] Membrane organization class.
[c] Presence of a signal peptide.
[d] Presence of a transmembrane domain.

signal peptide and a transmembrane domain at the amino terminal of the sequence were resolved by the application of an N-terminal filtering program (22).

Metanephric Culture and Tissue Preparation

For RNA *in situ* hybridization, embryos were collected from outbred CD1 mice as above at days 10.5 and 12.5 of gestation. E10.5 embryos were cut transversely below the forelimbs and longitudinally down the midline to expose the ND, UB, and MM as for initial tissue collection. At E12.5, complete urogenital (UG) tracts were collected. For explant culture, metanephroi were isolated from E12.5 embryos and grown as explants for 2 d at 5% CO_2 at 37°C on 3.0-μm polycarbonate transwell filters (Costar) in MEM supplemented with 10% FCS and 20 mM glutamine. Tissue for whole-mount *in situ* hybridization was fixed overnight in 4% PFA/PBS at 4°C.

RNA In Situ Hybridization

Expression patterns were analyzed by RNA *in situ* hybridization using digoxigenin-labeled sense and antisense riboprobes. Probes were synthesized as described previously (23) using pSPORT1 constructs that contained the NIA EST of interest (SRC Microarray facility). Probes were not fragmented by hydrolysis and were purified using Sephadex columns (Roche) after digestion of the vector with DNaseI (Promega) for 15 min at 37°C. Whole-mount *in situ* hybridizations were performed as described previously (24) with minor modifications. All probes were hybridized at 65°C, and post-antibody washes were reduced to 30 min. Tissue was mounted in Mount-Quick aqueous (Daido Sangyo), and photographs were taken using an Olympus AX70 compound microscope with Kodak Elite Ektachrome 160T color reversal film.

Results

The microarray expression profile of murine E10.5 uninduced MM, the renal progenitor time point, was examined by comparison with adjacent rostral IM, comprising the portion of the nephrogenic cord that contained the mesonephros and presumptive genital ridge. In this way, markers identifying the MM from UB and surrounding IM at the point of MM commitment were determined. A major aim was to find cell surface markers ubiquitously expressed by this mass of progenitors that may be useful in identifying other cells throughout development of the metanephros and in the mature kidney with a cell surface profile similar to that of the renal progenitor population. Cells in the adult kidney that show a molecular phenotype similar to the renal progenitor time point may represent a residual multipotent progenitor population.

Gene Expression Profile Analysis

Until recently, it has been difficult to perform meaningful expression profile studies in embryologic systems because of the limited amount of nucleic acid available from small amounts of tissue. This problem was circumvented in this study by using mRNA linear amplification. aRNA produced by *in vitro* transcription has been shown to have a correlation coefficient >0.95 to total RNA and aRNA produced using different amounts of template total RNA (25,26). To minimize nonlinearity and ensure consistency, all samples were amplified under exactly the same conditions at the same time.

The array experiment was performed with amplification replicates and included a dye reversal experiment to account for any dye bias. The three arrays were normalized to each other to give a comparable range of log ratios (Figure 2). Using R statistical software, genes were ranked in decreasing order of duplicate-correlated B scores, with the highest B scores representing differentially expressed genes showing the highest statistical confidence (a B score >0 represents a >50% probability that the gene is differentially expressed). The B score was calculated for each spot on the array using normalized signal and background intensities regardless of signal strength compared with other spots on the array. By doing this, we were confidently able to identify differentially expressed genes that had low absolute signal strength (but significant compared with background noise) that may represent low-abundance transcripts. The 40 genes with the highest B-score values are shown in Table 2, and the entire list of ranked genes from this analysis is available at http://kidney.scgap.org/extra_data/Rankedgenes.txt.

Normalized data were imported into Genespring 5.1 for visualization and further analysis. The magnitude (M) difference of a gene between samples was determined during B-statistic calculation, and the fold change for each gene was determined in Genespring. Elements on the microarrays for which the target signal intensity fell below 400 relative fluorescence units (RFU) were considered not reliable for fold-change calculations. Ratio data from duplicate spots on each array were determined and averaged across experimental replicates, and a mean ratio of 1.80-fold was used as the cutoff for differential expression. Outliers of genes differentially expressed between samples were defined using the combination of a B score >0 and a fold change >1.80-fold. Genes that showed increased differential expression in the MM are listed in Table 3, whereas genes that were more highly expressed in IM are listed in Table 4. The target intensity signals of these outliers ranged from 455 to 42452 RFU, suggesting that this

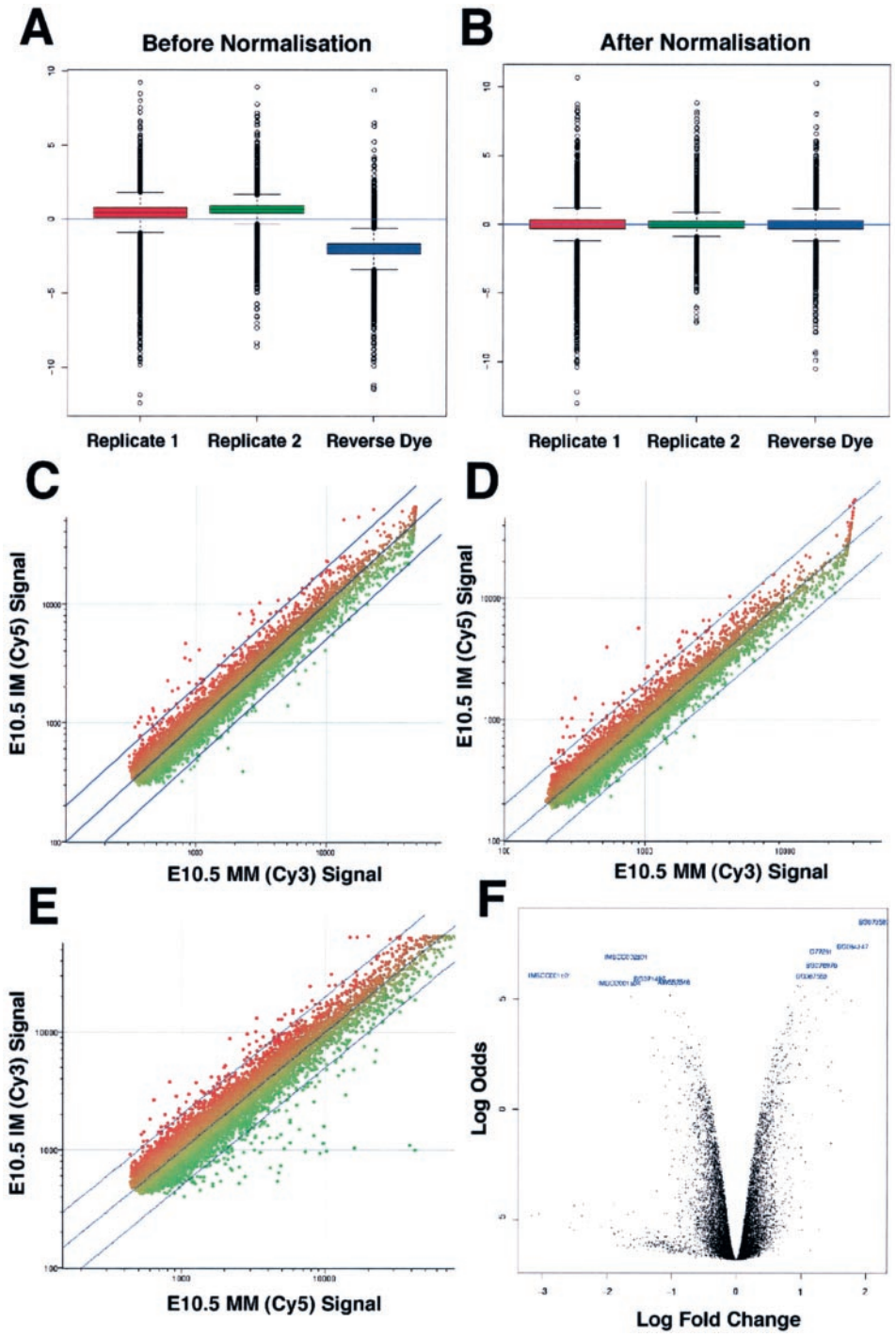

Figure 2. Analysis of microarray data. (A and B) The three array comparisons were normalized to each other to give a comparable range of log ratios. Box plot representations of the individual hybridizations before (A) and after (B) print-tip lowess normalization. (C through E) Genespring scatterplots of each hybridization. For amplification replicates (C and D), E10.5 MM antisense RNA (aRNA) was labeled with Cy3, and E10.5 intermediate mesoderm (IM) aRNA was labeled with Cy5. (E) The sample labeling was reversed in a dye-swap experiment to account for any dye bias. The outer lines represent 1.80-fold differences in expression between samples. Spots located outside the lines were identified as outliers differentially expressed between samples. (F) Genes were ranked in decreasing order of the duplicate-correlated B value, with the highest B values indicating those with the most significant changes in expression level.

approach was identifying both low- and high-expressing genes as outliers.

A total of 24 (21 nonredundant) elements with a B score >0 were found to have an average increase in differential expression >1.80-fold greater in the uninduced MM compared with the surrounding IM (Table 3). The sequence of each differen-

Table 2. List of the 40 highest ranked elements (representing 38 nonredundant genes) showing differential expression between intermediate mesoderm and metanephric mesenchyme at E10.5 according to statistical significance (B-score ranking)[a]

Probe ID[b]	Description	M[c]	*t*[d]	*P* Value[e]	B Score[f]	Average Fold Change[g]
BG073597	Troponin T1, skeletal, slow (Tnnt1)	2.12	27.4	0.00043	8.51	0.23
BG084347	Embryonal stem cell specific gene 1 (esg1)	1.79	20.6	0.00139	7.39	0.29
C77281	Catenin src	1.30	19.6	0.00139	7.17	0.40
IMBCC002d01	GATA binding protein 3 (Gata3)	−1.71	−18.6	0.00151	6.92	3.70
BG076976	Lectin, galactose binding, soluble 1 (Galectin-1)	1.31	17.2	0.00204	6.53	0.40
IMBCC001e01	ISL1 transcription factor, LIM/homeodomain (Islet1)	−2.88	−15.8	0.00240	6.07	4.48
BG067553	Unknown EST	1.16	15.7	0.00240	6.05	0.46
BG071497	Ewing sarcoma homolog (Ewsh)	−1.34	−15.3	0.00240	5.92	2.53
AW552546	Aldehyde dehydrogenase family 1, subfamily A2	−0.97	−14.9	0.00240	5.77	1.75
IMBCC001a04	Heat shock protein, 60 kD (Hspd1)	−1.82	−14.8	0.00240	5.73	3.76
BG065122	Ectodermal-neural cortex 1	1.02	14.5	0.00240	5.61	0.51
BG067391	H3053H09-3 NIA Mouse 15K cDNA Clone 3′ mRNA sequence	0.94	14.5	0.00240	5.61	0.55
BG077268	t-complex protein 1 (Tcp1)	−0.94	−14.4	0.00240	5.58	1.85
BG065742	H3034F12-3 NIA Mouse 15K cDNA Clone 3′ mRNA sequence	0.98	14.3	0.00240	5.56	0.49
AW550681	Brain protein 16	1.40	14.3	0.00240	5.54	0.39
BG073712	Hypothetical retroviral GAG p10	1.54	13.8	0.00263	5.34	0.37
BG067309	ESTs	0.89	13.8	0.00263	5.33	0.53
BG064703	RIKEN cDNA 1700012O15 gene	1.26	13.7	0.00263	5.30	0.44
AW553287	Periostin, osteoblast specific factor (OSF-2)	1.48	13.5	0.00263	5.22	0.37
BG085206	Serine protease inhibitor, Kunitz type 2 (Spint-2)	−1.03	−13.4	0.00263	5.17	1.97
IMBCC001c22	H19 fetal liver mRNA (H19)	−1.03	−13.3	0.00263	5.14	2.00
BG067247	H3052B12-3 NIA Mouse 15K cDNA Clone 3′ mRNA sequence	1.05	13.3	0.00263	5.11	0.49
BG082455	Claudin-6	−1.63	−13.2	0.00263	5.10	2.95
BG063173	Actinin, α-1	1.38	12.9	0.00273	4.95	0.40
BG075073	Thymosin, β4, X chromosome	0.83	12.9	0.00273	4.95	0.63
BG063671	Hypothetical G-protein β WD-40	0.99	12.8	0.00273	4.93	0.49
BG071895	Embryonal stem cell specific gene 1 (esg1)	1.77	12.8	0.00273	4.92	0.29
BG067061	Expressed sequence C81543	0.82	12.5	0.00313	4.77	0.54
BG076108	RIKEN cDNA 2010107E04 gene	0.79	12.2	0.00350	4.64	0.61
BG080332	H3052D04-5 NIA Mouse 15K cDNA Clone 5′ mRNA sequence	1.11	12.0	0.00384	4.53	0.44
BG072301	Hypothetical protein	−1.10	−11.8	0.00390	4.44	1.99
BG064604	RIKEN cDNA 2310061A22 gene	1.65	11.8	0.00390	4.40	0.35
C87546	Serine/threonine kinase 11	0.86	11.7	0.00390	4.39	0.52
BG073524	Serine protease inhibitor, Kunitz type 2 (Spint-2)	−0.96	−11.7	0.00390	4.39	1.87
BG075311	Transketolase	−0.81	−11.7	0.00390	4.37	1.73
BG064116	Axotrophin	0.84	11.5	0.00390	4.28	0.56
BG064362	Fibroblast growth factor inducible 14	0.99	11.5	0.00390	4.27	0.50
BG088677	Paternally expressed 3 (peg3)	0.82	11.5	0.00390	4.26	0.58
BG076069	CD24 antigen (CD24, heat stable antigen)	−0.91	−11.5	0.00390	4.25	1.90
BG080423	RIKEN cDNA B130024B19 gene	0.86	11.4	0.00390	4.23	0.57

[a] Genes with positive magnitude (M) values were more highly expressed in intermediate mesoderm compared with metanephric mesenchyme (MM). Genes with negative M values were more highly expressed in E10.5 MM. The average fold change of a gene was derived from the ratio of expression in the E10.5 metanephric mesenchyme compared with the intermediate mesoderm (Genespring).

[b] National Institute on Aging probe identification.

[c] Magnitude measurement: mean $\log_2$ channel 1 (Cy5 − E10.5 IM)/channel 2 (Cy3 − E10.5 MM).

[d] Penalized *t* test to determine differential expression.

[e] Probability that the observed differential expression is due to chance.

[f] Log odds score that the gene is differentially expressed.

[g] Average fold change (E10.5 MM/IM) of gene between two samples (Genespring).

Table 3. Genes differentially expressed >1.80-fold (B scores >0) in E10.5 MM compared with surrounding intermediate mesoderm from microarray analysis[a]

ProbeID	Average Fold Change[b]	B Score	Description	RTS[c]	RPS[d]	Membrane Organization[e]	RefSeq	Unigene
IMBcc001e01	4.48	6.07	ISL1 transcription factor, LIM/homeodomain (Islet1)	TB3026	PB3026	A	NM_021459	Mm.42242
IMBcc001a04	3.76	5.73	Heat shock protein, 60 kD (Hspd1)	TA2816	PA2816	F	NM_010477	Mm.299398
IMBcc002d01	3.70	6.92	GATA binding protein 3 (Gata3)	TB2292	PA2292	A	NM_008091	Mm.289671
BG082455	2.95	5.10	Caudin-6	TA7109	PA7109	E	NM_018777	Mm.86421
BG071497	2.53	5.92	Ewing sarcoma homolog (Ewsh)	TB1980	PA1980	A	NM_007968	Mm.142822
IMBcc001d13	2.43	1.71	Transformation-related protein 53 (p53)	TB5764	PA5764	A	NM_011640	Mm.222
BG085134	2.43	3.24	CD24 antigen (CD24, heat-stable antigen)	TB1208	PA1208	C	NM_009846	Mm.29742
BG072306	2.43	2.68	Unknown EST	TF34306	no_prot	—	—	—
IMBcc001c05	2.39	1.05	Cadherin-11	TB1259	PA1259	C	NM_009866	Mm.1571
BG084914	2.37	2.36	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide (14-3-3 θ)	TB6022	PA6022	A	NM_011739	Mm.291458
BG073496	2.22	2.45	Retinol dehydrogenase 10 (all-trans) (Rdh10)	TB19324	PC19324	0	NM_133832	Mm.274376
IMBcc001d15	2.14	1.34	Retinoic acid receptor α (RAR-α)	TB4770	PA4770	A	NM_009024	Mm.103336
IMBcc002p03	2.05	0.16	Bone morphogenetic protein 7 (BMP-7); osteogenic protein 1	TB1012	PA1012	B	NM_007557	Mm.595
IMBcc001c22	2.00	5.14	H19 fetal liver mRNA (H19)	TB2578	PB2578	F	X58196	Mm.14802
BG072301	1.99	4.44	Hypothetical protein	TF27449	PC27449	E	—	Mm.173932
BG085206	1.97	5.17	Serine protease inhibitor, Kunitz type 2 (Spint-2)	TB5324	PA5324	C	NM_011464	Mm.295230
BG066641	1.94	3.24	Stearoyl-CoA desaturase 2 (Scd2)	TB4988	PA4988	E	NM_009128	Mm.193096
BG078410	1.92	2.33	Enolase 1, α non-neuron (Eno1)	TA1895	PC1895	A	NM_023119	Mm.70666
BG076069	1.90	4.25	CD24 antigen (CD24, heat-stable antigen)	TB1208	PA1208	C	NM_009846	Mm.29742
BG063464	1.89	2.06	RIKEN cDNA 1600029D21 gene	TB16827	PC16827	B	NM_029639	Mm.29959
BG073524	1.87	4.39	Serine protease inhibitor, Kunitz type 2 (Spint-2)	TB5324	PA5324	C	NM_011464	Mm.295230
IMBcc002c07	1.86	0.30	H19 fetal liver mRNA (H19)	TB2578	PB2578	F	X58196	Mm.14802
BG077268	1.85	5.58	t-complex protein 1 (Tcp1)	TB5536	PA5536	A	NM_013686	Mm.32019
BG071792	1.83	2.40	RIKEN cDNA 1110034A24 gene	TF19664	PC19664	F	NM_027269	Mm.107180

[a] Average fold change indicates the relative increase in expression of a gene in the MM compared with the intermediate mesoderm. The membrane organization prediction was adopted for each gene. "no_prot" indicates where no annotated protein open reading frame is associated with a particular RTS transcript. Membrane organization class 0 indicates cases in which only a partial open reading frame is present in the RPS. No prediction could be made in either of these cases.

[b] Average fold change in differential expression of triplicate hybridizations (Genespring).

[c] Representative transcript from RIKEN RTS 6.3 set to which NIA sequence mapped.

[d] Representative protein sequence from RIKEN RPS 6.3 set that RTS encoded.

[e] Membrane organization class prediction.

Table 4. Genes differentially expressed >1.80-fold (B scores >0) in intermediate mesoderm compared with E10.5 MM from NIA microarray analysis[a]

ProbeID	Average Fold Change[b]	B Score	Description	RTS	RPS	Membrane Organization[c]	RefSeq	UniGene
BG073597	4.31	8.51	Troponin T1, skeletal, slow (Tnnt1)	TB5715	PB5715	A	NM_172894	Mm.258670
BG071895	3.50	4.92	Embryonal stem cell specific gene 1 (esg1)	TB1952	PB1952	A	NM_025274	Mm.139314
BG084347	3.41	7.39	Embryonal stem cell specific gene 1 (esg1)	TB1952	PB1952	A	NM_025274	Mm.139314
BG064604	2.84	4.40	RIKEN cDNA 2310061A22 gene	TB8850	PB8850	E	NM_025872	Mm.3676
BG072148	2.70	1.92	RIKEN cDNA 2700055K07 gene	TB9696	PB9696	A	NM_026481	Mm.29358
AW553287	2.69	5.23	Periostin, osteoblast-specific factor (OSF-2)	TB6789	PB6789	B	NM_015784	Mm.236067
BG073712	2.68	5.34	Hypothetical retroviral GAG p10	TD60120	PC60120	A	AI464860	Mm.259177
C87237	2.68	0.39	Zinc finger protein 294	TF32582	PC32582	0	XM_128374	Mm.249005
BG067621	2.61	2.14	Similar to hypothetical protein 2 (rRNA external transcribed spacer)	TB4860	no_prot	—	—	—
BG081752	2.57	1.08	H3068F10-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3068F10 5′, mRNA sequence	TB1670	PA1670	A	—	—
AW550681	2.54	5.54	Brain protein 16	TB7882	PB7882	A	NM_021555	Mm.133202
BG063004	2.52	2.90	Lectin, galactose binding, soluble 1 (Galectin-1)	TB3297	PA3297	A	NM_008495	Mm.43831
BG063173	2.52	4.95	Actinin, α1	TB360	PC360	0	NM_011501	Mm.253564
BG076976	2.49	6.53	Lectin, galactose binding, soluble 1 (Galectin-1)	TB3297	PA3297	A	NM_008495	Mm.43831
C77281	2.48	7.17	Catenin src	TB1145	PA1145	A	NM_007615	Mm.35738
BG071484	2.37	3.60	RIKEN cDNA 6720475J19 gene	TB9838	PB9838	A	AU043178	Mm.291581
BG064703	2.29	5.30	RIKEN cDNA C330018L13 gene	TF16014	no_prot	—	NM_080562	Mm.27090
BG076248	2.25	2.12	Ribosomal protein L41	TB9680	PA9680	A	NM_018860	Mm.13859
BG080332	2.25	4.53	H3052D04-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3052D04 5′, mRNA sequence	no_match	—	—	—	—
BG067593	2.21	4.14	H3056B01-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3056B01 3′, mRNA sequence	no_match	—	—	—	—
BG076798	2.19	0.86	45S pre rRNA gene	TB1670	PA1670	A	—	—
BG075084	2.19	3.23	Crystallin, α B	TB1507	PA1507	A	NM_009964	Mm.178
BG067553	2.19	6.05	Unknown EST	TF30718	no_prot	—	—	—
BG067442	2.09	3.94	H3054E03-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3054E03 3′, mRNA sequence	no_match	—	—	—	—
BG067527	2.09	4.06	H3055D06-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3055D06 3′, mRNA sequence	no_match	—	—	—	—
BG085154	2.09	3.51	Tissue inhibitor of metalloproteinase 3 (timp3)	TB5658	PA5658	B	NM_011595	Mm.4871
BG065742	2.06	5.56	H3034F12-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3034F12 3′, mRNA sequence	no_match	—	—	—	—
BG067247	2.06	5.11	H3052B12-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3052B12 3′, mRNA sequence	no_match	—	—	—	—
BG087417	2.05	4.07	H3139C10-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3139C10 5′, mRNA sequence	no_match	—	—	—	—
BG073218	2.05	2.69	Placenta specific homeobox 1 (psx1)	TB4571	PB4571	A	NM_008955	Mm.10105
BG063671	2.04	4.93	RIKEN cDNA 2410118I19 gene	TD14432	PC14432	A	AW538657	Mm.270376
BG088871	2.02	1.43	GATA binding protein 4 (Gata4)	TB2293	PA2293	A	NM_008092	Mm.161558
BG086908	2.02	4.23	H3132A10-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3132A10 5′, mRNA sequence	no_match	—	—	—	—
BG086394	2.02	3.95	Ribosomal protein L41	TB9680	PA9680	A	NM_018860	Mm.290786
BG064362	1.99	4.28	Fibroblast growth factor inducible 14	TB2137	PA2137	A	NM_008015	Mm.289662
BG065122	1.95	5.61	Ectodermal-neural cortex 1	TF39142	no_prot	—	NM_007930	Mm.241073
BG077141	1.94	2.02	Mannosidase, β A, lysosomal	TB11401	PB11401	B	NM_027288	Mm.280536
BG064691	1.94	3.55	Protein phosphatase 1B, magnesium dependent, β isoform	TB4465	PA4465	A	NM_011151	Mm.249695
BG065276	1.92	2.07	Podocalyxin-like (podxl)	no_match	—	—	NM_013723	Mm.89918
C87546	1.91	4.39	Serine/threonine kinase 11	TB5401	PB5401	A	NM_011492	Mm.135645
BG066813	1.88	2.86	Unknown EST	TF26922	no_prot	—	—	—
BG067648	1.88	2.79	H3056F09-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3056F09 3′, mRNA sequence	no_match	—	—	—	—
BG067309	1.87	5.31	EST	TF10751	no_prot	—	BG067309	Mm.173004
BG067061	1.85	4.77	Expressed sequence C81543	no_match	—	—	BG080135	Mm.172902
BG067391	1.83	5.61	H3053H09-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3053H09 3′, mRNA sequence	no_match	—	—	—	—

[a] Average fold change indicates the relative increase in expression of a gene in the intermediate mesoderm compared with the MM. The membrane organization prediction was adopted for each gene (see Table 3 for details). "no_match" indicates that the NIA sequence could not be mapped to a corresponding RIKEN full-length RTS transcript. No prediction could be made in this case.

[b] Average fold change in differential expression of triplicate hybridizations (Genespring).

[c] Membrane organization class prediction on the basis of RPS.

tially expressed EST was mapped to a full-length protein sequence from the RIKEN RTPS 6.3 set, and the membrane organization predictions were adopted for each gene of interest. Of particular interest were candidate genes from class C (type I membrane proteins), class D (type II membrane proteins), and class E (multispan membrane proteins) for their potential utility in antibody-based fluorescence-activated cell sorting (FACS) for the purification of renal progenitors. Six genes fell into this category.

Molecules that Identify the Renal Progenitor Time Point

The microarray comparison was done to identify molecules that spatially define the uninduced MM from surrounding tissue at E10.5. This experiment did not make any assumptions about what happened to the temporal expression and distribution of these molecules during subsequent important stages of early metanephric development. These questions were approached using *in situ* hybridization analysis. Initially, the expression patterns of these genes in E10.5 embryo whole mounts was analyzed to observe expression at the renal progenitor time point and verify the microarray data. RNA *in situ* analysis was also performed on E12.5 urogenital tracts to observe how specific expression of the gene was in the metanephroi compared with the organs developing around it. Analysis of metanephric explants was also used to identify what individual renal cell types expressed the genes of interest.

The temporospatial expression patterns of genes in Table 3 were determined and appear in Figure 3. Whole-mount RNA *in situ* analysis at E10.5 and E12.5 indicates that few of these are restricted to the MM, and although differential expression in the MM was verified, this was over a background of broad or ubiquitous expression. Such genes included heat-shock protein 60 kD, p53, retinol dehydrogenase 10, bone morphogenetic protein 7, H19, and enolase 1. In later renal development, as assessed in metanephric explant culture, specific expression patterns were observed for most of these genes. Islet1 was shown to be highly expressed in the urogenital sinus, suggesting that the differential expression between MM and IM was due to contamination of tissue caudal to the MM. Gata3 and RIKEN cDNA 1600029D21 gene were expressed in the ND and emerging UB at E10.5, and expression was also observed in UB derivatives in metanephric explants.

Four of the differentially expressed genes were clearly verified as being expressed within the MM itself at a higher level than surrounding tissue. These were Ewing sarcoma homolog (Ewsh), tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein 14-3-3 θ, retinoic acid receptor-α (RAR-α), and stearoyl-CoA desaturase 2 (Scd2). In explant cultures, 14-3-3 θ and Scd2 showed similar expression patterns, being broadly expressed by epithelial segments, RAR-α was expressed in the renal interstitium, particularly the interstitium at the core of the explant, and Ewsh was expressed in cells that were undergoing mesenchymal-epithelial transition around UB tips.

Cell Surface Markers of Renal Progenitor Cells

Six transmembrane proteins showed a significant increase in expression in the E10.5 MM compared with the surrounding IM (Table 3). Scd2 is known to be an intracellular transmembrane protein localized to the endoplasmic reticulum and was therefore not considered in the category of potential cell surface marker of renal progenitors. The *in situ* expression patterns of these genes were examined across the same tissue types (Figure 4) as for the nontransmembrane molecules in Figure 3.

Two transmembrane genes claudin-6 and spint-2 were identified as outliers because of strong and very specific expression in the ND and emerging UB, with the UB expression resulting in the differential between the two populations being compared. The *in situ* expression of these two genes was very specific to the ND and primordial UB in the E10.5 embryo and to ND/UB derivatives in the E12.5 UG tract (spint-2 also showed strong expression in the gonads at E12.5) and metanephric explants. Expression was also observed in the mesonephroi at E12.5.

Although present in other areas of the embryo, expression of CD24 was strikingly specific to the uninduced MM population in the nephrogenic cord of the E10.5 embryo. CD24 was specifically expressed in the kidneys of the E12.5 UG tract, although weak expression was also observed in the ND at this time point, made difficult to see as a result of the strong expression in the kidneys. In explants, CD24 was expressed in epithelial cells of both UB and MM lineages, although not in the lower limbs of the S-shaped bodies that give rise to the glomeruli. Cadherin-11 showed expression throughout the mesoderm of the embryo at E10.5 but particularly the renal progenitor population. Expression of cadherin-11 became more widespread at E12.5 and was strongly expressed throughout the renal interstitium of the explant, particularly the cells surrounding the UB tips. The NIA EST BG072301 was mapped to a full-length RIKEN RTPS that was bioinformatically predicted to encode for a protein that contains multiple transmembrane domains. No other information about this molecule is available in the literature. There was apparently ubiquitous expression of this molecule across the embryo at E10.5 and E12.5, although the levels of expression were elevated in the E10.5 MM. In explants, expression was observed in the interstitium but particularly the mesenchymal cells bordering epithelial structures.

Genes Enriched in the Mesoderm Surrounding the Metanephric Mesenchyme

Microarray analysis identified 45 elements, representing 42 nonredundant genes, with a B score >0 and expressed >1.80-fold greater in the IM than the E10.5 MM (Table 4). Genes encoding secreted molecules expressed specifically in the tissue surrounding the MM may be important in directing MM toward a renal fate. Secreted molecules (class B) such as periostin, and tissue inhibitor of metalloproteinase 3, which were more strongly expressed in the IM, may be involved in signaling to the MM to direct early renal cell fate decisions. Membrane organization could not be established for 20 of the

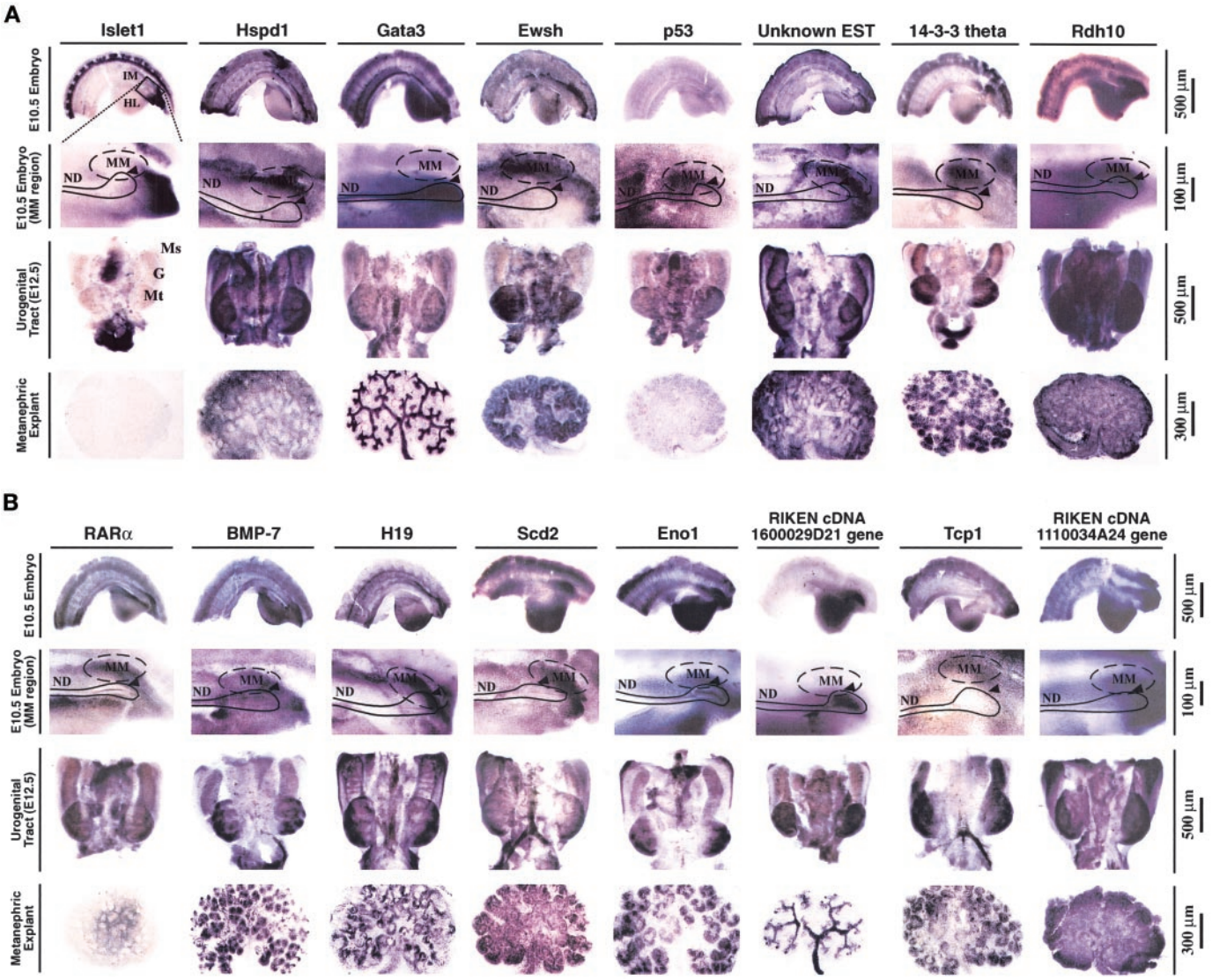

Figure 3. RNA *in situ* hybridizations of genes enriched in uninduced MM at E10.5 in the mouse. The expression pattern of each gene was surveyed across E10.5 embryos (HL, hind limb bud; arrows represent budding site of primitive ureteric bud), E12.5 urogenital tracts (G, gonad; Mt, metanephros; Ms, mesonephros), and metanephric explants. (A) Islet1 was not expressed in renal derivatives at any time point but was strongly expressed in the urogenital sinus. Hspd1 expression was widespread at E10.5 and E12.5, but stronger expression was observed in the MM at E10.5 and in the interstitium of metanephric explants. Gata3 was strongly expressed in the ND and UB at E10.5 and in UB derivatives in metanephric explants. Ewsh was strongly and specifically expressed in the MM at E10.5, the metanephroi at E12.5 and mesenchymal cells undergoing mesenchymal-epithelial transition in explants. p53 was weak and ubiquitous in all tissues analyzed. Unknown EST (BG072306) was widespread at E10.5 and E12.5 but showed enrichment in the MM at E10.5 and was expressed in the interstitium of explants. 14-3-3 θ was strongly and specifically expressed in the MM at E10.5, in the metanephroi at E12.5 (weakly in gonads and ND), and in condensing mesenchyme in explants (weakly UB). Rdh 10 was strongly expressed in the MM at E10.5 but was not specific as widespread expression was observed at this time point and E12.5 and explants. (B) RARα was observed in the MM at E10.5 and the primary interstitium of explants but was not specific at E12.5. BMP-7 was enriched in the MM at E10.5 and metanephroi at E12.5 and expressed strongly in condensing mesenchyme in explants (weakly in UB). H19 expression was specific to the MM at E10.5 but not to the metanephroi at E12.5, although expression was highly upregulated in mesenchymal cells surrounding UB tips in explants. Scd2 was strongly expressed in the MM at E10.5 but broadly expressed throughout the UG tract at E12.5 and in epithelial segments of metanephric explants. Eno1 was strongly expressed in the MM at E10.5 and weakly in the ND and UB, strong again in the metanephroi at E12.5, and weakly in the ND and strongly in condensing mesenchyme in explants. RIKEN cDNA 1600029D21 gene was expressed weakly in the ND and MM at E10.5 but was highly upregulated in the emerging UB, was strongly expressed in the metanephroi at E12.5 (weakly in ND and mesonephroi), but strongly expressed in UB derivatives in explants (weakly in mesenchymal caps). Tcp1 was not specific at E10.5 and expressed in gonads and metanephroi at E12.5 with expression maintained in condensing mesenchyme in metanephric explants. RIKEN cDNA 1110034A24 gene was strongly but not specifically expressed in the MM at E10.5 and was also broadly expressed in E12.5 UG tracts and metanephric explants.

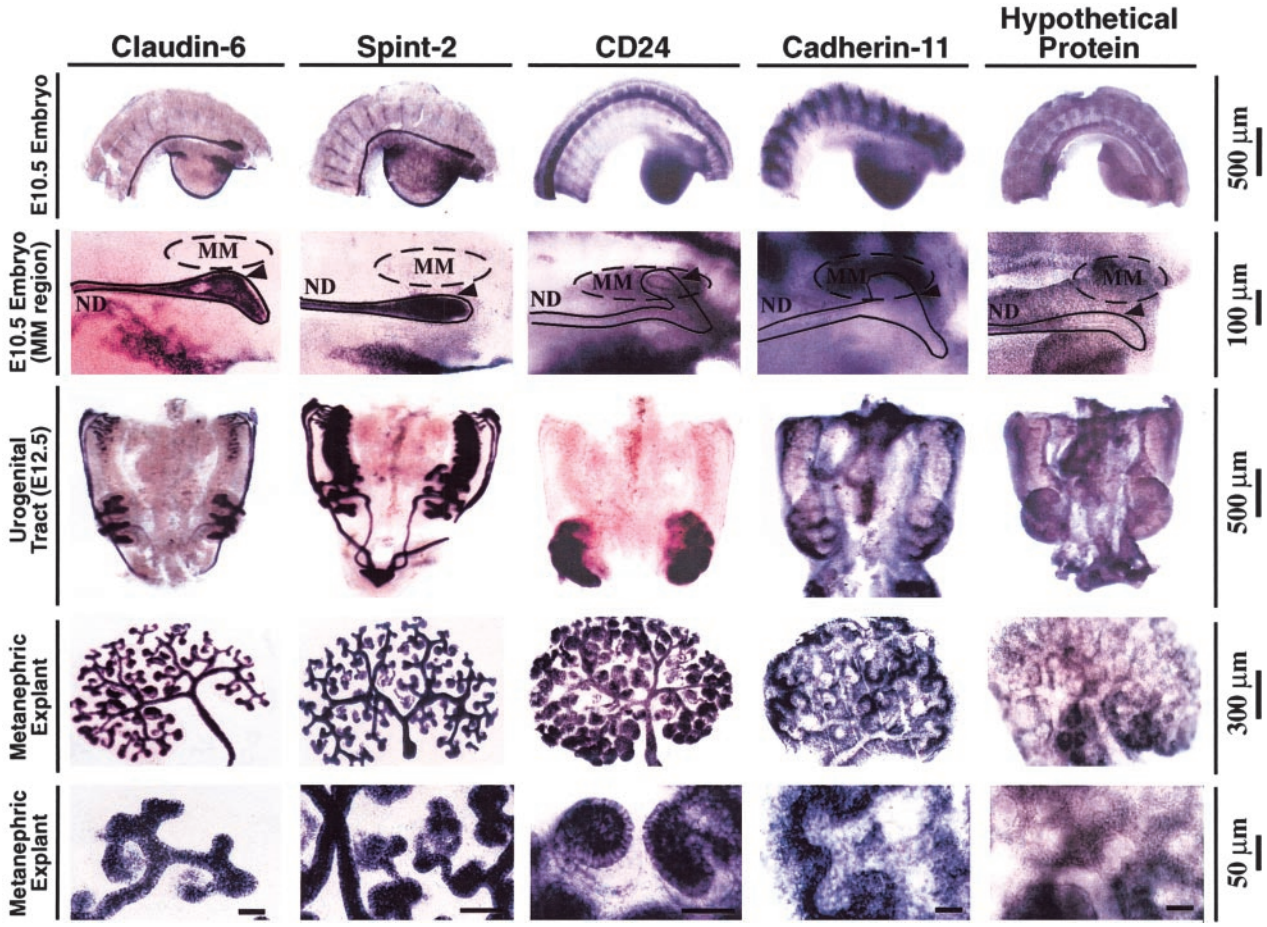

Figure 4. Whole-mount RNA *in situ* hybridizations of cell surface proteins representing renal progenitor cell markers. The expression pattern of each gene was surveyed across E10.5 embryos (arrows represent budding site of primitive UB), E12.5 urogenital tracts, and metanephric explants. Claudin-6 was highly specific to the ND and UB at E10.5, mesonephroi, ND and UB derivatives at E12.5, and UB derivatives in metanephric explants. Spint-2 was highly specific to the ND and UB at E10.5, the sex cords of the gonads, mesonephroi, ND and UB derivatives at E12.5, and UB derivatives in explants. CD24 was strikingly specific to the MM at E10.5 (in the nephrogenic cord) and the metanephroi at E12.5 with expression observed in epithelial segments of both MM and UB origin in explants. Cadherin-11 was strongly but not specifically expressed in the MM at E10.5 and in the metanephroi at E12.5 with expression observed in the interstitium of metanephric explants. Hypothetical protein (BG072301) was ubiquitous at E10.5 and E12.5, although enriched in the MM with expression observed in the interstitium of explants.

42 genes because the NIA sequence did not confidently map to a corresponding full-length protein-coding RIKEN transcript.

Discussion

The goal of this project was to find genes that identify renal progenitor cells in the uninduced MM at E10.5. Microarray analysis identified a subset of genes that were upregulated in the uninduced MM in comparison with adjacent rostral IM, whose expression was not broadly expressed at E10.5 and whose expression persisted during development but became restricted or reduced in expression over time. These markers include Ewsh, 14-3-3 θ, Scd2, RAR-α, CD24, and cadherin-11. The intracellular markers may be crucial in specification of the MM and provide useful tools for identification of nonrenal stem cells being induced to adopt a kidney fate. The cell surface markers CD24 and cadherin-11 may prove useful in purification of potential renal progenitor cell populations by FACS.

CD24 was strongly and specifically expressed in all uninduced MM cells at E10.5, and cadherin-11 was also strongly expressed by this population. Although these molecules both seem to mark the renal progenitor time point, their expression patterns diverged greatly as kidney development progressed. In metanephric explants, CD24 expression was observed in all epithelial structures of the developing kidney except for the lower limbs of the S-shaped bodies, whereas cadherin-11 was expressed by mesenchymal cells of the renal interstitium, most strongly by those surrounding the UB tips, but not in epithelial cells. The fact that CD24 marks cell types of both MM and UB derivatives suggests that it identifies renal progenitors that are committed to differentiating into epithelial segments of the nephron, whereas cadherin-11 may identify progenitor cells

that are destined to form the primary renal interstitium. There is some indirect evidence to suggest that these molecules may mark a renal progenitor cell population. CD24 is strongly expressed in Wilms' tumors (27) and renal cell carcinomas (28), and cadherin-11 expression has also been detected in renal cell carcinoma cell lines (29) and Wilms' tumors and is located on the long arm of chromosome 16, a region known to demonstrate allele loss in such tumors (30). Expression of CD24 and cadherin-11 by renal tumor cells may indicate that these cells are reverting to a more primitive or embryonic state, a condition analogous to the uninduced MM at E10.5.

Ideally, markers of renal progenitors would be restricted to MM expression. Although the expression of CD24 and cadherin-11 during embryogenesis as a whole was not examined in this study, their expression is not restricted to the MM at E10.5 and has been described in other locations during development (31,32). This is not surprising, and in all other progenitor or stem cell populations, the markers used for enrichment are not restricted to those populations. Rather, it is a combination of expression and lack of expression that is used for isolation. Neither CD24 nor cadherin-11 is recognized as a positive marker of embryonic, neural, or hematopoietic stem cells. Indeed, murine neural stem cells are regarded as $CD24^{lo}$ (3). However, CD24 has been identified as a marker of spermatogonial stem cells (33), and links have been made between the expression of cadherin-11 and osteogenic progenitors (34). The latter is interesting given the differentiation of bone and cartilage in Wilms' tumors. These molecules may represent cell surface markers of renal progenitor cells conserved between mouse and human, but their expression is not restricted to the uninduced MM. Therefore, as for other stem cell systems, it will be necessary to use these markers in combination with others to enrich for renal progenitors.

These experiments made no *a priori* assumptions about what happened to the expression and localization of these MM markers as metanephric development progressed, simply asking what defines the E10.5 MM renal progenitor time point as an entity from the tissue that surrounds it. Although the dogma has suggested that kidney development ceases at birth, the growing evidence of resident stem cell populations in a variety of adult organs suggests that such a population may also exist in the kidney. The location of such a population is unknown, as is the phenotype of such cells. Specific markers of renal progenitor cells that maintained expression in a rare adult stem cell population should decrease in abundance and become localized to specific cell types as the kidney differentiates and the progenitor pool becomes depleted. CD24 is known to remain expressed in the distal tubules (35) of the adult kidney, but this does not rule out the possibility that it is also expressed in a minor stem cell population of the adult kidney. Further analysis of the later expression patterns of these renal progenitor time-point genes may help to determine such a subcompartment. FACS analysis using combinations of the cell surface markers identified in this experiment could reveal whether minor populations of progenitor-like cells exist in the adult kidney. FACS analysis may also be applied to determine whether the renal progenitor time point represents a homogeneous population of progenitors or a collection of cell types with distinct renal lineage potential.

Several genes that are known to be involved in early metanephric development were identified as outliers in this microarray experiment. For example, Gata3 expression is crucial in early metanephric development as the MM of Gata3 null embryo fails to differentiate and renal hypoplasia occurs (36) and RAR-α double-null mutants have reduced UB growth (37). Gene knockout models have suggested crucial roles for molecules such as WT-1 (38), Lim1 (39), Eya1 (40), and Pax2 (41) in early metanephric development. Expression of these genes is thought to be some of the earliest signs of commitment of the metanephric mesenchyme to a renal fate. Many of these genes were not represented in the NIA clone set, and some genes that were present did not produce a signal strength that would allow a reliable determination of differential expression to be made compared with background noise, resulting in a B score <0. However, two genes involved in early metanephric development, the transcription factor WT-1 and the secreted morphogen wnt-4, showed higher expression in the IM compared with the MM in microarrays (fold change just below 1.80-fold threshold), findings verified by *in situ* hybridization (results not shown). This highlights that a comparison between IM and MM at E10.5 essentially represents a spatial comparison between mesonephros/presumptive genital ridge and MM. Hence, some genes that play important roles in metanephric development will initially be expressed to a higher degree in the mesonephros/genital ridge than the MM. This also shows how microarray experiments can be used only as a guide, and although they are efficient ways to survey thousands of genes, they must be used in combination with other techniques to uncover the roles of various genes in developmental processes.

In summary, for the first time, we have catalogued the expression profile of the MM at the point of commitment to a metanephric lineage. This analysis has identified Ewsh, RAR-α, 14-3-3 θ, Scd2, CD24, and cadherin-11 as molecules that are highly expressed by renal progenitors. In addition, CD24 and cadherin-11 are cell surface markers at this time point. By distinguishing some of the earliest genes expressed by the uninduced MM, this study has not only identified novel molecules involved in metanephric development but also provided tools for the reverse transcription–PCR–based identification of embryonic stem (ES) cells that adopt a renal fate. The existence of a renal stem cell has not been established, and although recent discoveries in stem cell biology suggest that they might exist, no markers for renal stem or progenitor cells have been discovered, prohibiting the identification and isolation of such cells. By defining a combination of cell surface proteins that specifically mark the renal progenitor time point, this approach will facilitate purification of cells with this phenotype from mixed populations, such as kidneys at various stages of development or differentiating ES cell cultures, using antibody-based FACS.

Acknowledgments

This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health

(DK63400) as part of the Stem Cell Genome Anatomy Project (http://www.scgap.org/). This work was performed as part of the Renal Regeneration Consortium. Grant Challen holds an Australian Postgraduate Award. Sean Grimmond and Rohan Teasdale are both recipients of NHMRC Career Development awards. Melissa Little is an NHMRC Senior Research Fellow.

We thank Chris Johns and the Australian Research Council Centre for Functional and Applied Genomics for the production of microarray chips.

References

1. Gussoni E, Soneoka Y, Strickland CD, Buzney EA, Khan MK, Flint AF, Kunkel LM, Mulligan RC: Dystrophin expression in the mdx mouse restored by stem cell transplantation. *Nature* 401: 390–394, 1999
2. Forbes S, Vig P, Poulsom R, Thomas H, Alison M: Hepatic stem cells. *J Pathol* 197: 510–518, 2002
3. Rietze RL, Valcanis H, Brooker GF, Thomas T, Voss AK, Bartlett PF: Purification of a pluripotent neural stem cell from the adult mouse brain. *Nature* 412: 736–739, 2001
4. Shankland SJ, Wolf G: Cell cycle regulatory proteins in renal disease: Role in hypertrophy, proliferation and apoptosis. *Am J Physiol* 278: F515–F529, 2000
5. Saxen L: *Organogenesis of the Kidney*, Cambridge, Cambridge University Press, 1987
6. Horster MF, Braun GS, Huber SM: Embryonic renal epithelia: Induction, nephrogenesis, and cell differentiation. *Physiol Rev* 79: 1157–1191, 1999
7. Herzlinger D, Kosek D, Mikawa T: Metanephric mesenchyme contains multipotent stem cells whose fate is restricted after induction. *Development* 114: 565–572, 1992
8. Qiao J, Cohen D, Herzlinger D: The metanephric blastema differentiates into collecting system and nephron epithelia in vitro. *Development* 121: 3207–3214, 1995
9. Dekel B, Burakove T, Arditti F, Reich-Zeliger S, Milstein O, Aviel-Ronen S, Rechavi G, Friedman N, Kaminski N, Paswell J, Reisner Y: Human and porcine early kidney precursors as a new source for transplantation. *Nat Med* 9: 53–60, 2003
10. Herzlinger D, Barasch J, Koseki C, Al-Awqati Q: Immortalisation of renal stem cells. *J Am Soc Nephrol* 2: 438, 1991
11. Oliver JA, Barasch J, Yang J, Herzlinger D, Al-Awqati Q: Metanephric mesenchyme contains embryonic renal stem cells. *Am J Physiol* 283: F799–F809, 2002
12. Al-Awqati Q, Oliver JA: Stem cells in the kidney. *Kidney Int* 61: 387–395, 2002
13. Stuart RO, Bush KT, Nigam SK: Changes in global gene expression patterns during development and maturation of the rat kidney. *Proc Natl Acad Sci U S A* 98: 5649–5654, 2001
14. Schwab K, Patterson LT, Aronow BJ, Luckas R, Liang, H-C, Potter SS: A catalogue of gene expression in the developing kidney. *Kidney Int* 64: 1588–1604, 2003
15. Stuart RO, Bush KT, Nigam SK: Changes in gene expression patterns in the ureteric bud and metanephric mesenchyme in models of kidney development. *Kidney Int* 64: 1997–2008, 2003
16. Valerius MT, Patterson LT, Witte DP, Potter SS: Microarray analysis of novel cell lines representing two stages of metanephric mesenchyme differentiation. *Mech Dev* 112: 1–14, 2002
17. Tanaka TS, Jaradat SA, Lim MK, Kargul GJ, Wang X, Grahovac MJ, Pantano Sano SY, Piao Y, Nagaraja R, Doi H, Wood WH, Becker KG, Ko MSH: Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc Natl Acad Sci U S A* 97: 9127–9132, 2000
18. Saal LH, Troein C, Vallon-Christersson J, Gruvberger S, Borg A, Peterson C: Bioarray software environment—A platform for comprehensive management of microarray data. *Genome Biol* 3:SOFTWARE0003, 2002
19. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *J Mol Biol* 215: 403–410, 1990
20. FANTOM Consortium RIKEN Genome Exploration Research Group Phase I, and IITeam: Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* 420: 563–573, 2002
21. Kanapin A, Batalov S, Davis MJ, Gough J, Grimmond S, Kawaji H, Magrane M, Matsuda H, Schonbach C, Teasdale RT, RIKEN GER Group, GSL members, Yuan Z: Mouse proteome analysis. *Gen Res* 13: 1335–1344, 2003
22. Yuan Z, Davis MJ, Zhang F, Teasdale RD: Computational differentiation of N-terminal signal peptides and transmembrane helices. *Biochem Biophys Res Commun* 312: 1278–1283, 2003
23. Holmes GP, Negus K, Burridge L, Raman S, Algar E, Yamada T, Little MH: Distinct but overlapping expression patterns of two vertebrate slit homologs implies functional roles in CNS development and organogenesis. *Mech Dev* 79: 57–72, 1998
24. Christiansen JH, Dennis CL, Wicking CA, Monkley SJ, Wilkinson DJ, Wainwright BJ: Murine Wnt-11 and Wnt-12 have temporally and spatially restricted expression patterns during embryonic development. *Mech Dev* 51: 341–350, 1995
25. Luo L, Salunga RC, Guo H, Brittner A, Joy KC, Galindo JE, Xiao H, Rogers KE, Wan JS, Jackson MR, Erlander MG: Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat Med* 5: 117–122, 1999
26. Baugh LR, Hill EL, Hunter CP: Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acids Res* 25: e29, 2001
27. Droz D, Rousseau-Merck M-F, Jaubert F, Diebold N, Nezelof C, Adafer E, Mouly H: Cell differentiation in Wilms' tumor (nephroblastoma): An immunohistochemical study. *Hum Pathol* 21: 536–544, 1990
28. Droz D, Zachar D, Charbit L, Chretien Y, Iris L: Expression of the human nephron differentiation molecules in renal cell carcinomas. *Am J Pathol* 137: 895–905, 1990
29. Shimazui T, Giroldi LA, Bringuier PP, Oosterwijk E, Schalken JA: Complex cadherin expression in renal cell carcinoma. *Cancer Res* 56: 3234–3247, 1996
30. Schulz S, Becker KF, Braungart E, Reichmuth C, Klamt B, Becker I, Atkinson M, Gessler M, Hofler H: Molecular analysis of E-cadherin and cadherin-11 in Wilms' tumours. *J Pathol* 191: 162–169, 2000
31. Magnaldo T, Barrandon Y: CD24 (heat stable antigen, nectadrin) a novel keratinocyte differentiation marker, is preferentially expressed in areas of the hair follicle containing the colony-forming cells. *J Cell Sci* 109: 3035–3045, 1996
32. Simonneau L, Kitagawa M, Suzuki S, Thiery JP: Cadherin 11 expression marks the mesenchymal phenotype: Towards new functions for cadherins? *Cell Adhes Commun* 3: 115–130, 1995
33. Kubota H, Avarbock MR, Brinster RL: Spermatogonial stem cells share some, but not all, phenotypic and functional characteristics with other stem cells. *Proc Natl Acad Sci U S A* 100: 6487–6492, 2003
34. Cheng SL, Lecanda F, Davidson MK, Warlow PM, Zhang SF, Zhang L, Suzuki S, St John T, Civitelli R: Human osteoblasts

express a repertoire of cadherins, which are critical for BMP-2 induced osteogenic differentiation. *J Bone Miner Res* 13: 633–644, 1998
35. Platt JL, Lebien TW, Michael AF: Stages of renal ontogenesis identified by monoclonal antibodies reactive with lymphohemopoietic differentiation antigens. *J Exp Med* 157: 155–172, 1983
36. Lim KC, Lakshmanan G, Crawford SE, Gu Y, Grosveld F, Engel JD: Gata3 loss leads to embryonic lethality due to noradrenaline deficiency of the sympathetic nervous system. *Nat Genet* 25: 209–212, 2000
37. Mendelsohn C, Batourina E, Fung S, Gilbert T, Dodd J: Stromal cells mediate retinoid-dependent functions essential for renal development. *Development* 126: 1139–1148, 1999
38. Kreidberg JA, Sariola H, Loring JM, Maeda M, Pelletier J, Housman D, Jaenisch R: WT-1 is required for early kidney development. *Cell* 74: 679–691, 1993
39. Fujii T, Pichel JG, Taira M, Toyama R, Dawid IB, Westphal H: Expression patterns of the murine LIM class homeobox gene Lim1 in the developing brain and excretory system. *Dev Dyn* 199: 211–230, 1994
40. Xu PX, Adams J, Peter H, Brown MC, Heaney S, Mass R: Eya1-deficient mice lack ears and kidneys and show abnormal apoptosis of organ primordia. *Nat Genet* 23: 113–117, 1999
41. Torres M, Gomez-Pardo E, Dressler GR, Gruss P: Pax-2 controls multiple steps of urogenital development. *Development* 121: 4057–4065, 1995