

Olga Soutourina · Lydie Cheval · Alain Doucet

## Global analysis of gene expression in mammalian kidney

Received: 9 August 2004 / Accepted: 21 October 2004 / Published online: 21 December 2004  
© Springer-Verlag 2004

**Abstract** The past decade has seen the achievement of sequencing of human, rat and mouse genomes and the development of high-throughput methods for quantitative monitoring of gene expression. The aim of the beginning post-genomic era is to determine the function of all these genes, a challenge in which the community of physiologists should have a leading role. In this short review, we discuss the relevance, feasibility and impact of transcriptome analysis in renal physiology. Comparison of transcriptomes demonstrates that cell lines are poor physiological models, making it necessary to work on native kidney tissue. Palliating kidney heterogeneity therefore requires the development of methods for transcriptome analysis sensitive enough to be compatible with microdissected nephron segments. Axial comparison of transcriptomes along the human nephron unexpectedly points out that the segmentation of nephrons concerns not only genes involved in solute and water transport functions and their regulation, but also genes related to the control of cell division, differentiation and apoptosis. Comparison of transcriptome of a same nephron segment from mice under different pathophysiological conditions outlines the wide pleiotropy of kidney function regulations. Both types of comparative studies also identify yet unknown transcripts specifically expressed along the nephron or under pathophysiological conditions.

**Keywords** SAGE · Single nephron segments · Water and solute transport · Cell cycle · Apoptosis · Differentiation

### Introduction

Since the completion of the deciphering of complete mammalian genomes [19, 31, 62, 67] and the accession to the post-genomic era, one of the main objectives of biology is to determine the function of all genes. The community of physiologists should play a leading role in this challenge, inasmuch as they develop and use adapted tools for large-scale analysis. This is now possible, because the post-genomic era has not only brought pieces of information on the identification of entire genomes, but it has also seen the development of sophisticated methods to study the expression of these genomes [49, 60]. The scope of this short review is to discuss the relevance, feasibility and impact of analysis of transcriptomes in renal physiology. We will first make some general considerations about the nature of transcriptomes, their methods of analysis and the specific requirements in renal physiology and then, based on our own experience, consider some foreseeable contributions of transcriptome analysis in renal physiology and pathophysiology.

### Rationale for transcriptome research in renal physiology

Conversely to cell biology, which addresses the mechanisms of universal cell functions that allow cell division, nutrition and death, physiology is focused mainly on the cell-specific functions which emerge with cell differentiation and organogenesis during development of superior multicellular organisms. This cell specification results from expression of specific sets of genes in different cell types and tissues. Thus, establishing the catalogue of the genes which are expressed (i.e. the transcriptome) in the different tissues or cell types of multicellular organisms appears as a prerequisite to analyse the functional specificities of these cellular systems. Until recently, functional specificities of tissues were addressed through the study of individual genes and often necessitated, as a first step, the identification and cloning of these genes. Not only has this molecular identification step become useless with the

O. Soutourina · L. Cheval · A. Doucet (✉)  
Laboratoire de Physiologie et Génomique Rénales,  
CNRS-UPMC, UMR 7134, Institut des Cordeliers,  
IFR 58, 15 rue de l'Ecole de Médecine,  
75270 Paris cedex 6, France  
E-mail: alain.doucet@bhdc.jussieu.fr  
Tel.: +33-1-55427851  
Fax: +33-1-46334172

availability of complete genomes, but these protein-by-protein studies have demonstrated their limit; indeed, functional properties of individual proteins may vary according to their environment, in particular the nature of the proteins with which they associate to build supramolecular complexes. This also stresses the interest to index the whole protein equipment of individual cells.

Similarly, it is now well established that physiological regulations are pleiotropic processes that involve a large number of molecular partners. As a consequence, it is not only necessary to know the molecular equipment of specific cells, but also to get access to the variations of this equipment under different functional states.

Analysis of transcriptomes for physiological purpose may have also consequences in genome studies. Firstly, it may not only provide experimental confirmation of gene prediction but also may lead to the discovery of new genes which remain undetected by genome analysis. For example, although the complete genome sequence of *Saccharomyces cerevisiae* was available [20], analysis of its transcriptome allowed the discovery of several hundreds of transcripts which were not predicted by computerised genome annotation [61]. Secondly, it may help to discover the function of orphan genes, as physiologists are well qualified to formulate hypothesis regarding the function of a gene in view of its tissue distribution and/or of its regulation of expression and to experimentally verify their hypothesis.

#### Nature and properties of transcriptomes

Mammalian genomes are currently considered to consist of  $\approx 30,000$  genes [19, 31, 62, 67]. The exact number of genes is not known precisely because, although complete genome sequences contain all the information about genes, the exact identification of genes, i.e. the exon/intron limits and the genes boundaries, are not precisely known. These  $\approx 30,000$  genes can be transcribed in  $> 100,000$  mRNAs, as most mammalian genes consist of numerous exons which generate a great number of splice variants. Each cell likely contains 300,000 mRNA molecules which correspond to  $\approx 15,000$  different molecular species [24]. Transcripts are 1.5 kb long on average and therefore a cell transcriptome corresponds to  $\approx 5 \times 10^8$  bases. In addition, the abundance of each transcript varies from a few copies to  $> 10,000$  copies per cell [24]. Thus, methods for analysis of transcriptome must combine the potency of analysing a large number of different molecular species with a dynamic range greater than three orders of magnitude. Conversely to the genome which, except for acquired mutations, is identical in all somatic cells and stable during life span, cell transcriptomes vary not only from cell to cell, but also with time along with functional adaptations. In other words, transcriptomes provide a quantitative and dynamic view of genome expression, and therefore are related to cell function potential.

Proteomes are wider than transcriptomes, since each transcript may encode several proteins, and each protein

exists under several different functional states. From a physiological point of view, it is often argued that transcriptome analysis would not be relevant, because only proteins are the basis of cell functions and transcripts may not be translated. Unfortunately, several factors still restrict proteome analysis. Although new methods such as Isotope Coded Affinity Tagging (ICAT) [22] or protein microarrays [52] have been developed recently, proteome analysis remains today mainly based on separation and characterisation of proteins by 2D gel electrophoresis, followed by characterisation of the peptide spots through proteolysis, sequencing and/or mass spectroscopy. As yet, 2D gel electrophoresis suffers several limitations: (1) it is not an exhaustive method, as only abundant proteins can be detected (a few hundreds per sample); (2) membrane proteins are hardly detectable, a major limitation in renal physiology, since the main functions of renal cells rely on ion and water transporting proteins displaying multiple membrane spanning domains; and (3) compulsory precautionary measures are required to achieve a sufficient degree of reproducibility for comparative studies.

Besides their wider spectrum and their relative easiness of use as compared to proteome studies, transcriptome analysis presents additional advantage, as it is becoming quite clear that, besides their role as vector of information from gene to proteins, transcripts play important functions independently of their own translation. As a matter of fact, increasing amount of works demonstrate regulation of transcript expression via endogenous antisense transcripts [32]. Thus, it seems legitimate to carry out studies on transcriptomes, especially as long as proteome analysis will remain technically limited and demanding.

#### Prerequisites for transcriptome analysis in kidney

It has been recognised for decades that understanding kidney functions and dysfunctions requires circumventing the cellular heterogeneity of the organ [38]. Even when considering exclusively the glomerulo-tubular complexes of kidney, this heterogeneity is high: each nephron consists of successive tubule segments with specific functions, and some segments may consist of several intermingled cell types with distinct functions. Two strategies have been developed to palliate this heterogeneity: the physical separation of glomeruli and nephron subsegments by surgical or immunological dissection and the establishment of cellular models such as primary cell culture or immortalised cell lines. As long as physiological problems are addressed, only native tissue is relevant. Indeed, despite their advantage for addressing questions relevant to cell biology, primary cell cultures as well as established cell lines undergo important loss of function (both qualitatively and quantitatively) and some gain of function through dedifferentiation/transdifferentiation, which rule out their use for physiological purposes. For example, comparison of the transcriptomes of native mouse collecting ducts [7] and

immortalised collecting duct principal cells [45] reveals marked differences, although the same method was applied for analysis (Table 1). As a matter of fact, (1) mRNAs for specific markers of principal cells such as aquaporin 2 (the vasopressin regulated water channel) [47] or type 2 11 $\beta$ -hydroxysteroid dehydrogenase (an enzyme that contributes to the aldosterone sensitivity of the collecting duct) [1] are 200- and 20-fold less abundant in immortalised principal cells than in collecting ducts; (2) conversely, mRNAs for pendrin, a marker of collecting duct  $\beta$  intercalated cells normally absent from principal cells [46], and for urokinase, which is prefer-

entially expressed in thick ascending limbs and proximal tubules [65], are expressed at relatively high level in cultured cells; and (3) transcripts expressed at high level ( $\geq 1\%$  of all mRNAs) include 40 transcripts for ribosomal proteins in immortalised cells versus two transcripts in native collecting ducts. This index of high protein metabolism may imply that immortalised cells are not fully quiescent and fully differentiated.

This absolute necessity to analyse transcriptomes on native kidney cells generates a major constraint that results from the tiny quantities of tissue obtainable either by surgical or immunological dissection relative to the amounts required for classical transcriptome analysis methods. As a matter of fact, a 1-mm long segment of nephron contains approximately 500 cells, i.e. 0.1 ng of poly(A) mRNA, whereas classical transcriptome methods such as DNA microarrays or chips and serial analysis of gene expression (SAGE) [60] require approximately 2–10  $\mu$ g of poly(A) mRNA. However, it is now possible to palliate in part this problem by using down scaled adaptations of SAGE. Two types of adaptations have been introduced that reduce by 250- to 1000-fold the required amount: (1) addition of an early amplification step in the library generation procedure in SAR-SAGE [63], SAGE-Lite [42] and aRNA-longSAGE [25], and (2) improving the yield of mRNA extraction and cDNA synthesis in SAGE adaptation for downsized extracts (SADE) [64], miniSAGE [71] and microSAGE [9]. SADE was specifically developed for transcriptome analysis on  $\approx 100$  mm of nephron segments, an amount easily accessible through surgical microdissection, and all results presently available on renal transcriptome at the level of individual structures were obtained with this method.

**Table 1** Comparative expression of a selection of genes in mouse native collecting duct and immortalized collecting duct cells. This table compares the abundance of serial analysis of gene expression mRNA tags in two libraries constructed from microdissected mouse outer medullary collecting ducts (OMCD) and from a mouse kidney cortical collecting duct cell line (mpkCCD), after normalization to 20,000 tags in the two libraries. Three sets of transcripts were selected for comparison: (1) transcripts corresponding to functional cell markers of collecting duct principal cells, including aquaporins 2 and 3 (AQP2 and AQP3), 11 $\beta$ -hydroxysteroid dehydrogenase type 2 (11OHD2), the apical potassium secreting channel (ROMK), and two subunits of Na,K-ATPase and of the apical amiloride-sensitive sodium channel (ENaC); (2) transcripts corresponding to “housekeeping” functions equally expressed in OMCDs and mpkCCD cells; and (3) transcripts normally absent from collecting duct principal cells and/or expressed at higher level in mpkCCD cells than in OMCD. Data are from Cheval et al. [7] and Robert-Nicoud et al. [45]

Tag identification	Tag abundance	
	OMCD	mpkCCD
AQP2	328	1
AQP3	11	0
11OHD2	32	2
ROMK	12	0
Na,K-ATPase $\beta_1$ subunit	49	15
Na,K-ATPase $\gamma$ subunit	4	1
ENaC $\alpha$ subunit	8	0
ENaC $\beta$ subunit	8	2
Cyclin D1	358	257
$\beta$ actin, cytoplasmic	46	34
Cytochrome c oxidase, subunit VI a, polypeptide 1	31	24
Ribosomal protein L11	27	37
ATP synthase, F1 complex $\beta$ subunit	25	29
H3 histone, family 3A	19	25
Heat shock protein 1, $\beta$	14	21
Pendrin	0	3
Plasminogen activator, urokinase	1	20
Ribosomal protein, large, P1	18	115
Ribosomal protein S29	14	233
Ribosomal protein S16	11	63
Ribosomal protein S28	11	107
Ribosomal protein L36	4	30
Ribosomal protein S18	4	26
Ribosomal protein, large P2	4	23
Ribosomal protein L36a	3	26
Ribosomal protein L4	3	30
Ribosomal protein S26	2	20
Ribosomal protein L24	0	21

### Principle, potentialities and limits of SAGE

SAGE consists in characterising each cDNA obtained by retrotranscription of tissue mRNAs by a short (10 bp) informative nucleotide sequence, or tag. Several transcript-specific tags are concatenated into long DNA molecules which are cloned and sequenced. Computer-assisted analysis of the sequences permits to extract, classify and count the different tags, and database queries allow identification of them. The method is quantitative, as the relative abundance of the different tags in the library reflects the abundance of the cognate transcripts in the biological sample.

Without going into the details of SAGE procedure, it is worth considering some aspects which have implications for data analysis.

### Deepness of analysis

Admitting that a cell contains 300,000 mRNA molecules, sequencing of 30,000 tags allows analysis of 10% of the transcriptome: under this condition, transcripts present at 3,000–30 copies per cell are detected (300–3 tags expected) with a probability  $\geq 95\%$ , whereas

transcripts present at three copies per cell will most often not be detected (26% probability of being detected).

Using the equation of Clarke and Carbon [8], one calculates that detecting a transcript present at one copy per cell with a probability of 95% would require to sequence 900,000 tags. Besides the cost of this procedure, increasing the deepness of analysis beyond  $\approx 50,000$  tags per library is not a satisfactory method for increasing the sensitivity of SAGE. Indeed, SAGE is based on single lecture, single-strand DNA sequencing, a procedure that conveys an error rate of  $\leq 10^{-2}$  ( $\leq$  one mismatch each 100 nucleotides), meaning that up to 10% of the 10-bp tags contain a sequence error. This means that for a tag counted 100-fold in a library, there are 10 erroneous tags that differ by a single base in their sequence (30 possibilities). Considering that sequencing errors are random in nature, the probability for observing the same error at the same site in two samples of a same tag is negligible (1/30), and therefore the large majority of erroneous tags due to sequencing errors are present only once in the library. Practically, this means that one can get rid of sequencing errors just by dismissing from data analysis the tags counted only once. This would no longer hold for a tag counted 1,000-fold (i.e. the same tag but obtained at a deepness of analysis of 500,000 instead of 50,000 tags). Given these limitations, one must keep in mind that SAGE will not detect transcripts expressed at low level reliably.

### Tag identification

The specific identification of a tag to a transcript relies on (1) the nucleotide sequence of tags, as a 10-bp oligonucleotide theoretically discriminates  $> 10^6$  different molecular species, and (2) the location of the tag relative to the poly(A) tail of the transcript. Indeed (Fig. 1), due to the construction procedure, the tag corresponds to the 10-bp located downstream the 3'-most site of cleavage by a restriction enzyme (anchoring enzyme) with a 4-bp recognition site [usually *Nla*III (CATG recognition site) or, in our SADE adaptation, *Sau*3A (GATC recognition site)]. Note that cDNAs lacking recognition site for the anchoring enzyme do not generate tags and are therefore artefactually excluded from the transcriptome analysis.

Thus, tag identification through sequence matching in database may lead to different classes of annotation: (1) cDNA or EST annotations (with or without corresponding protein) for tags matching a cDNA or an EST sequence at the right position [note that this requires that authors have deposited the full length cDNA sequence containing the poly(A) in databases]; (2) unreli-

able match annotations for tags matching a cDNA or EST sequence without possible localisation with regards to the poly(A) tail; and (3) no match annotation for tags without matching sequence in GenBank. In addition, cDNA and EST annotation may not be univocal, as two different transcripts may correspond to a same tag. The efficiency of tag identification can be improved by using the longSAGE adaptation of SAGE which theoretically generates 17-bp tags and therefore permits better discrimination among transcripts [48]. Furthermore, it allows pertinent identification of no match tags through comparison with genomic sequences.

SAGE tag identification has been facilitated by the development of facilities freely accessible on the Web such as SAGEmap (<http://www.ncbi.nlm.nih.gov/SAGE/>) or SAGEgenie (<http://cgap.nci.nih.gov/SAGE/>) (note that presently SAGEgenie is restricted to the analysis of SAGE tags generated using *Nla*III as anchoring enzyme).

### Gene discovery

As already mentioned, sequencing of whole genomes is not sufficient to identify all the genes of a species and, *a fortiori* to identify all transcripts. Because SAGE is based on sequencing of cDNA fragments, it offers this possibility to identify unknown transcripts (the no match tags mentioned above). Interestingly, although the sequence information contained in a tag is reduced, it proved to be sufficient to permit PCR cloning of cognate full-length cDNA [58].

Our laboratory currently masters both large scale microdissection of nephron segments and the SADE technique, allowing us to routinely generate transcriptomes from 500–1,000 microdissected segments of nephron. In the following, we will discuss the relevance and outcomes of recent findings obtained using this approach. Two applications have been considered: (1) the axial comparison of transcriptomes along the nephron to define the molecular basis of renal functional heterogeneity and (2) the physiological regulation and pathophysiological dysregulation of gene expression in specific nephron segments.

### Segmental analysis of gene expression along the human nephron

Analysis of approximately 50,000 tags in each of eight substructures of the human nephron [glomerulus (Glo), proximal convoluted tubule (PCT) and proximal



**Fig. 1** Schematic localisation of a serial analysis of gene expression (SAGE) tag along cDNA. The **bold line** corresponds to the coding region, the **AE** boxes to the anchoring enzyme recognition sites [GATC in case of *Sau*3A utilised in SAGE adaptation for downsized extracts (SADE)]. The tag (**XX . . . XX box**) corresponds

to the ten base pairs located downstream of the 3' -most anchoring enzyme site. Although in most cases the tag is located in the 3'UTR region (as shown on the diagram), it can also be within the coding region



straight tubule (PST), medullary thick ascending limb of Henle's loop (MTAL) and cortical thick ascending limb of Henle's loop (CTAL), distal convoluted tubule (DCT) and cortical collecting duct (CCD) and outer medullary collecting duct (OMCD)] produced >92,000 different tags including >32,000 tags counted at least twice and therefore unlikely corresponding to sequence errors (see above) [5]. Data confirm the large dynamic range of gene expression with (1) the most abundant tag counted >4,000-fold and the less abundant ones only once, and (2) the 92 most abundant tags (1‰ of all molecular species) accounting for >22% of all tags. In view of such an abundant source of data, the question arises as to how to mine relevant pieces of information. Two different strategies can be used, on a non-exclusive manner, according to the questions that are addressed: global, non-selective analysis of data and gene-targeted analysis. Following are some examples of such strategies and pieces of information that can be derived.

#### *Clustering analysis of human renal SAGE libraries*

As the less *a priori*-based method for transcriptome data exploitation, one can perform general clustering analysis of SAGE data, using software freely available on the Web [4, 13]. Such an analysis of human kidney data reveals several specific clusters of genes that roughly correspond to genes differentially expressed in the different nephron portions. On that basis, it is possible to establish kinship links among renal structures. As shown by the hierarchical tree on top of Fig. 2, clustering analysis reveals the distinct positions of the Glom and of the PCT and PST versus the other tubular segments. It also shows a close relationship between the PCT and PST and of the CCD and OMCD. Unexpectedly, it indicates a closer relationship between MTAL and DCT than between MTAL and CTAL.

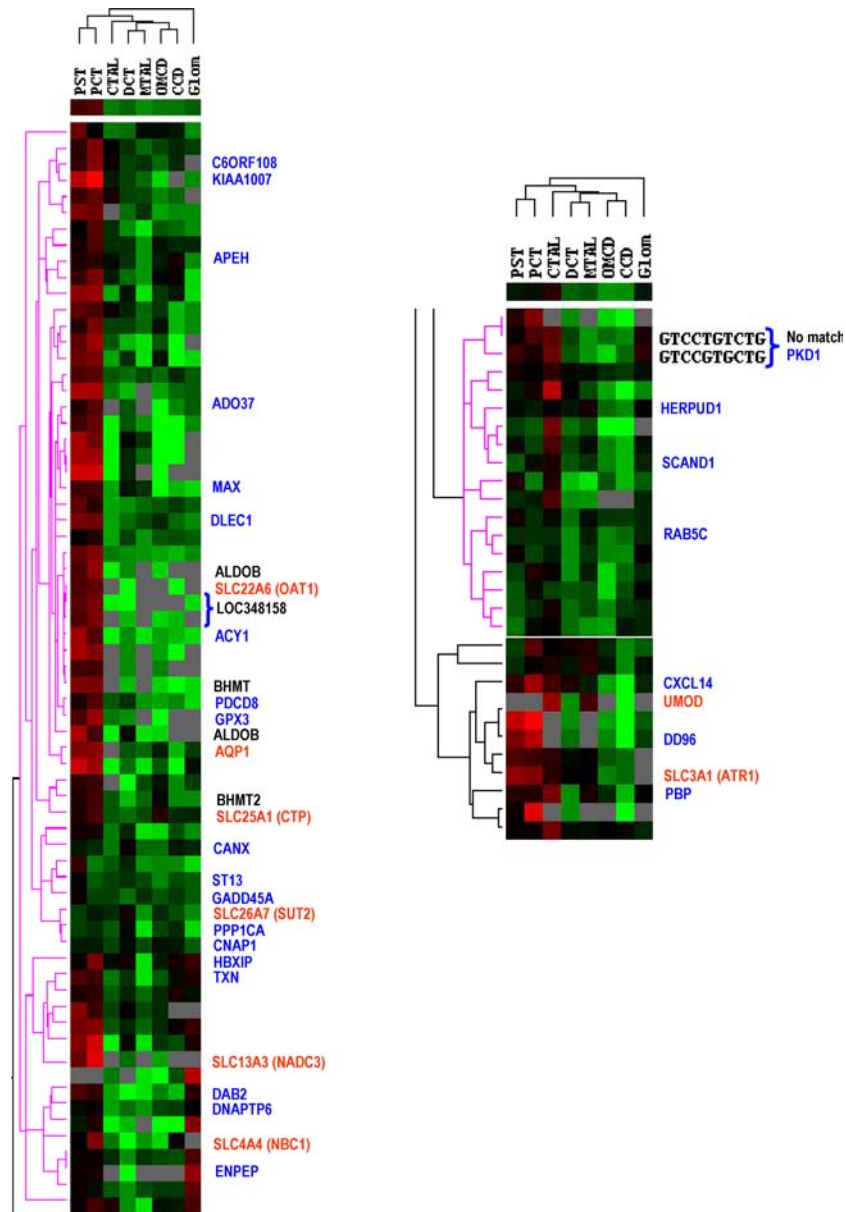
As an example of segment specific gene clusters, Fig. 2 shows the genes displaying maximum expression in the PCT and PST. Moreover, we can see a continuous transition from strictly proximal tubule-specific genes to less specific genes also expressed in Glom (at the bottom of left panel) or in the MTAL and CTAL (right panel). Identification of tags in these proximal tubule-specific gene clusters confirms the preferential expression of solute carriers [organic anion transporter (OAT1) SLC22A6] [30] and sodium-dependent dicarboxylate transporter (NADC3, SLC13A3) [66] and water channels (AQP1) [72] known to be expressed preferentially or exclusively in proximal tubules. Unexpectedly, a rather large proportion of proximal tubule-specific tags correspond to genes involved in general cell functions (shown in blue), including an established or predicted role in the regulation of cell cycle, proliferation and differentiation (C6ORF108, KIA1007, ADO37, MAX, DLEC1, ACY1), apoptosis (PDCD8), and stress responses (GPX3, GADD45A). These findings demonstrate that the axial heterogeneity of kidney tubule concerns not

only expression of genes underlying the specific physiological functions of the different nephron segments, but also genes involved in cell division and differentiation (see below). Finally, it is interesting to note the close clustering association between two tags with related sequence (gtccGTGctg vs gtccTGTctg). The first one was identified as polycystic kidney disease type I protein (PKD1) [23], whereas the second did not match to any GenBank sequence. This may suggest the existence of a frequent, but yet non-identified, polymorphism in this part of the gene, and it would be interesting to confirm the existence of this polymorphism and to evaluate whether or not it is associated with polycystic disease.

Results of clustering analysis are quite dependent on the choice of parameters and algorithms. It is noticeable, however, that such clustering provides relevant information as long as (1) it reveals specific expression of physiologically relevant genes in a given tissue, and (2) tags corresponding to a same gene (due to splice variants and/or polymorphism) such as ALDOB or LOC348158 or to functionally related genes (such as BHMT and BHMT2) are closely segregated in the same clusters.

#### *Search for expression of specific genes along the nephron*

Conversely, another way of mining transcriptome data is to localise expression along the nephron of specific genes of interest. The relevance of this approach is based on the confirmation by SAGE data of known segmental expression of genes. As a matter of fact, the two panels on top of Fig. 3 illustrate that the abundance of tags in the Glom and the different nephron segments are well correlated with the known expression of the related proteins, whether proteins are expressed along the whole nephron or are specific of a single segment or a subgroup of renal structures. For example, tags of the  $\alpha_1$ ,  $\beta_1$  and  $\gamma$  subunits of Na,K-ATPase are found along the whole nephron, but their abundance culminates in thick ascending limb and DCT, where the functional expression of Na,K-ATPase is highest [27]. Similarly, tags for specific markers such as podocalyxin (Glom) [28], aquaporin 1 (AQP1, proximal tubule) [72], furosemide-sensitive Na,K,2Cl cotransporter (NKCC2, TAL) [70], thiazide-sensitive Na,Cl cotransporter (NCCT, DCT) [36] or aquaporin 3 (AQP3, collecting duct) [12] are exclusively detected in the appropriate libraries. The predictational value of tag abundance along the nephron for determining the profile of protein expression remains valuable even for tags found at low level in the SAGE database. For example, at the time of the cloning of the non-erythroid rhesus-associated proteins Rhcg and Rhbg [34, 35], SAGE data available in our laboratory indicated a preferential abundance of the corresponding tags in the distal nephron, a finding that was later verified both at the transcript level by RT-PCR on microdissected nephron segments and at the protein level by immunolocalisation on kidney slices [14, 44].



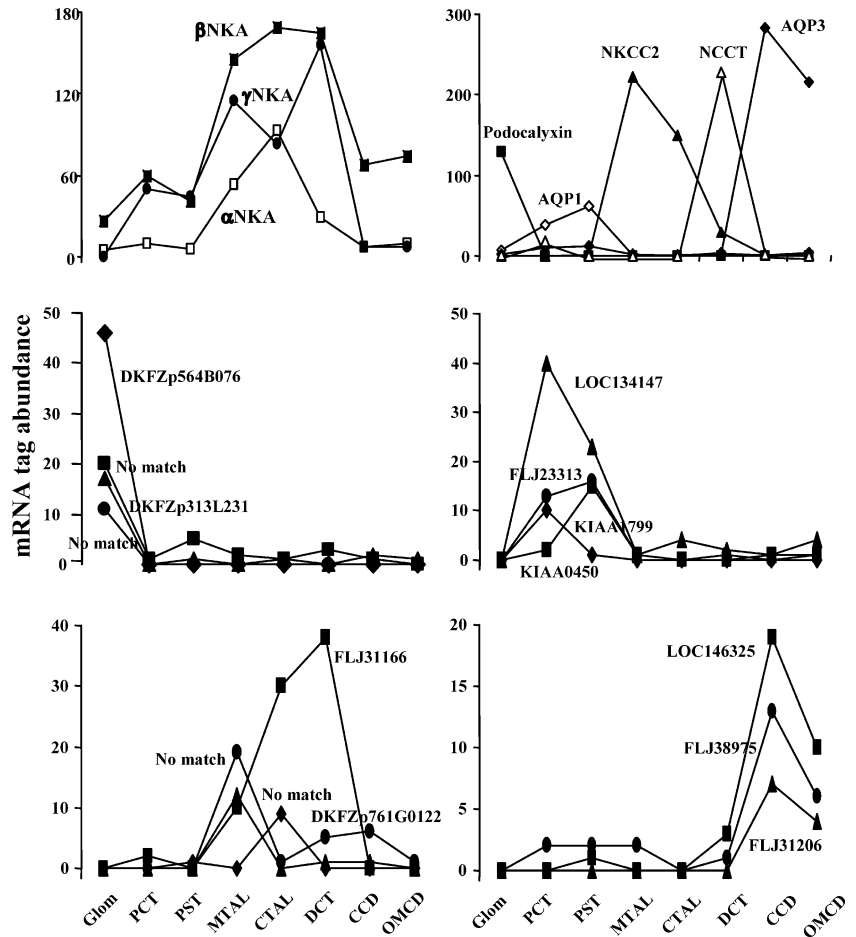
**Fig. 2** Hierarchical clustering analysis of SAGE data for eight human nephron segment libraries. Dendrogram on the *top* represents relationships among SAGE libraries corresponding to eight substructures of human nephron: *Glom* glomerulus, *PCT* proximal convoluted tubule, *PST* proximal straight tubule, *MTAL* medullary thick ascending limb of Henle's loop, *CTAL* cortical thick ascending limb of Henle's loop, *DCT* distal convoluted tubule, *CCD* cortical collecting duct and *OMCD* outer medullary collecting duct. Only the proximal tubule-specific cluster with continuous transition to the less specific genes also expressed in *Glom* (*left panel*), and in *MTAL* and *CTAL* (*right panel*) is included in the figure. Each *row* represents a gene/SAGE tag, whereas each *column* corresponds to a SAGE library/nephron portion. The expression level in each library is presented by *panel colour*, ranging from *green* for minimum expression to *red* for maximum expression, *black* corresponding to median expression level and *grey boxes* corresponding to the absence of tag count in a

given library. The *colour intensity* represents the magnitude of the deviation from the median. At the *right* of the panel shown in *red*, the names of SAGE tag identifications corresponding to solute carrier(s) or other gene(s) implicated in tissue transport and permeability (the gene symbols are from the Human Genome Organization nomenclature and usual alternative names in *brackets*), in *blue* the names of genes having more general function implicated in cell proliferation and differentiation and in *black* the tag pairs corresponding to the same or similar genes as an internal control of clustering procedure. Only genes with total tag counts of at least 40 are included in analysis. The uncentred Pearson correlation was used for distance calculation, and the average-linkage clustering was performed on logarithmically transformed median centred SAGE data, using Michael Eisen software program Cluster, version 2.12, followed by tree diagrams visualisation with TreeView, version 1.47 [13]

Because those segment-specific proteins are associated with important functions whose alterations are often associated with diseases, it will be important to cha-

racterise the function of genes of yet unknown function expressed in a segment-specific manner. Mining the human nephron segment database reveals a quite large

**Fig. 3** Patterns of distribution of SADE mRNA tags along the different portions of human nephron. Tag abundance indicates mRNA tag counts in libraries normalised to 50,000 tags. The two *top panels* show the distribution of the tags corresponding to the  $\alpha_1$ ,  $\beta_1$  and  $\gamma$  subunits of Na,K-ATPase (*left panel*  $\alpha$ ,  $\beta$  and  $\gamma$  NKA) and to markers of the different renal structures (*right panel*): podocalyxin (Glom), aquaporin 1 (*AQP1* proximal tubule), furosemide-sensitive Na,K, 2Cl cotransporter (*NKCC2* TAL), thiazide-sensitive Na,Cl cotransporter (*NCCT* DCT) and aquaporin 3 (*AQP3* collecting duct). The *lower four panels* show distribution of non-identified tags corresponding to hypothetical proteins (*DKFZ*, *LOC*, *KIAA* or *FLJ HUGO* nomenclature) or without GenBank identification (no match) that are preferentially expressed in one or a subgroup of renal structures



number of tags without functional identification, including 12 no-match tags, 25 tags identified as cDNAs without known protein and 42 tags corresponding to hypothetical proteins. The four panels at the bottom of Fig. 3 illustrate the specific distribution of some of these tags.

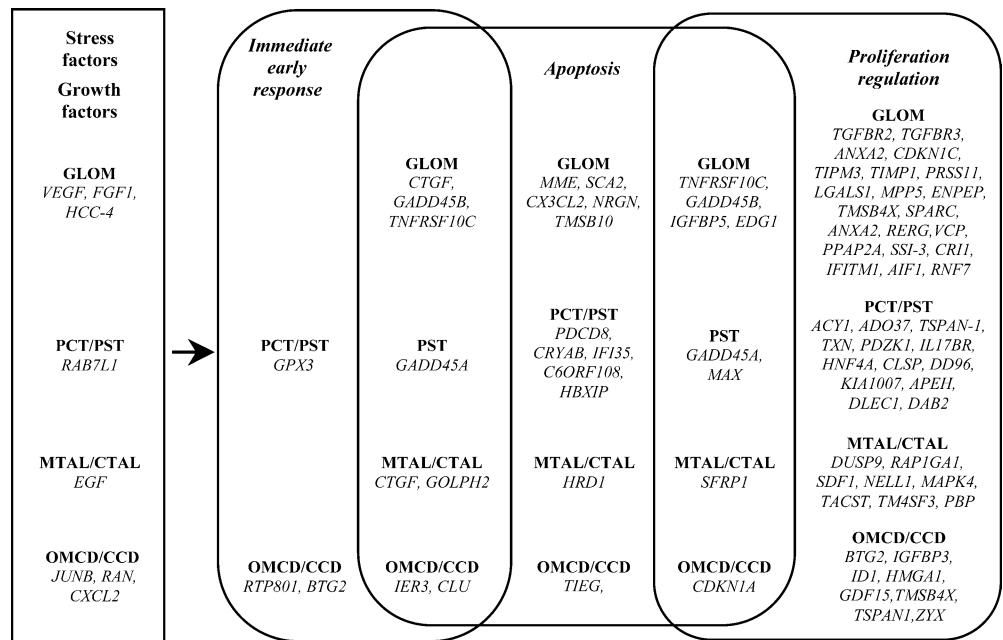
#### *Axial heterogeneity of expression of genes not involved in tubular transport properties.*

Under physiological conditions, the rate of cell division in adult kidney is low [2], since basal production of epithelial cells only compensates for the loss of the epithelial cells into the urine. This balance between mitosis and cell loss is tightly controlled to prevent both kidney size reduction and hyperplasia. In contrast, under many pathological conditions, a marked change in cell turnover rate occurs as a result of both induced cell death and regeneration processes. As a matter of fact, recent results demonstrate the role of cell cycle regulating proteins in diverse glomerular diseases, as well as apoptosis of damaged cells, dedifferentiation and proliferation of surviving tubular epithelial cells in proximal tubules during recovery from ischemia/reperfusion injury [2, 21, 37, 40]. Thus, identification of the genes which control these processes in the kidney presents an important pathophysiological relevance.

As mentioned above, besides genes involved in transport and permeability properties (i.e. solute carriers, water and ion channels, ion-transport ATPases), the analysis of differential gene expression along the human nephron has revealed a new, yet unidentified, specificity in expression pattern in the different nephron portions for the genes participating in cell differentiation and proliferation control. Figure 4 shows different classes of such genes in the different clusters of renal structures.

In adult human tissue, proliferative responses may be induced by various external and intrinsic stress signals, including DNA damage, toxic compounds, radiation or oxidative stress, induction of oncogenes and growth factors. Such responses may be observed in kidney shortly after injury, e.g. acute renal failure, and result in both cell death and regenerative cell replication [43]. At this very initial step, there is already a cell-specific response with, for example, glomeruli expressing vascular endothelial growth factor and fibroblast growth factor, whereas the thick ascending limbs express epidermal growth factor, consistent with specific functions assigned to these different growth factors [17, 37, 40]. The first step after sensing particular stress stimuli is the rapid induction of transcription of immediate early response genes in association with the entry of cells from quiescence into the cell cycle. Here again, the different nephron segments express specific genes belonging to

**Fig. 4** Distribution of genes for cell differentiation and proliferation control differentially expressed along the human nephron. In the *left box* are shown the external and intrinsic stimuli inducing proliferative cellular responses, with examples of segment specific genes in different nephron portions coding for growth factors. On the *right* the genes are regrouped according to their established or predicted role in indicated cellular processes (*vertical alignment*). The genes on the overlap of the *two oval boxes* represent the functional class sharing the both functions. The genes differentially expressed in different nephron segments, i.e. in GLOM and in PCT/PST segments, MTAL/CTAL and CCD/OMCD segments are *horizontally aligned*



this functional class, e.g. DNA-damage-inducible  $\beta$  and  $\alpha$  protein genes (GADD45B and GADD45A) in the Glom and proximal tubules, respectively [26, 33, 57]. These two early induced genes participate to the regulation of apoptosis, a feature also shared by immediate early response 3 [69] and clusterin [55], which are specifically expressed in collecting ducts. The next step in proliferative response is placed under the control of cell cycle progression, and it appears that the different renal structures also express specific regulators acting at the G1/S and G2/M phase transition checkpoints [41, 53]. For example, Glom specifically express the cyclin-dependent kinase inhibitor collecting ductKN1C, which controls G1/S transition, whereas the collecting ducts express CDKN1A (p21) [68] and BTG2, which control the same step [11, 54].

The large number of genes involved in these processes and their high level of expression may appear paradoxical for normal kidney tissue. It should be stressed, however, that the SAGE libraries were constructed from nephron segments dissected from healthy fragments of kidneys surgically removed from donors with kidney tumours, making it possible that the tissue was not totally normal. In addition, surgical constraints in humans induce periods of warm and cold hypoxia before tissue microdissection which may act as a stress signal for proliferation. In any case, the main observation remains that axial heterogeneity of nephron segments concerns not only genes involved in physiological functions, but also involved in the control of cell division, differentiation and death. Because clustering of all tags or of the tags corresponding to those genes participating in cell differentiation and proliferation control reveal the same kinship among renal structures, one might suggest the existence of a causal relationship between expression of specific genes of cell differentiation/proliferation and

establishment of a final differentiation status of the different renal cell types.

#### Regulation of gene expression

One of the objectives of physiology is to decipher the mechanisms underlying functional regulations. In relation with their key role in regulating the *milieu intérieur*, mammalian kidneys display remarkable aptitudes to adapt their transport functions in response to a wide variety of stresses, including food and water intake, drugs, hormonal environment, diseases and genetical alterations. Within the kidney, the distal tubule system plays a central role, as these nephron segments are the main targets for the regulation of water, sodium, potassium and acid/base balances, and are important contributors to calcium and magnesium balances.

When considering exclusively sodium balance, dysregulations of the distal tubule system are responsible for widespread pathologies including hypertension and oedema diseases. Indeed, all the gene products which have been shown to alter blood pressure in human, or whose deletion or overexpression modifies blood pressure in mouse or rat, are directly related to the transport of sodium in the distal nephron or to its regulation [18]. When blood pressure is not altered, abnormal sodium retention by the distal nephron may promote oedema formation [59].

Now, if the cellular mechanisms of sodium transport in the distal nephron are well known, in particular with the molecular and functional characterisation of the apical sodium entry pathways (thiazide-sensitive NaCl cotransporter and epithelial sodium channel ENaC) and of basolateral Na,K-ATPase, the mechanisms of regulation and the basis of pathological dysfunctions remain



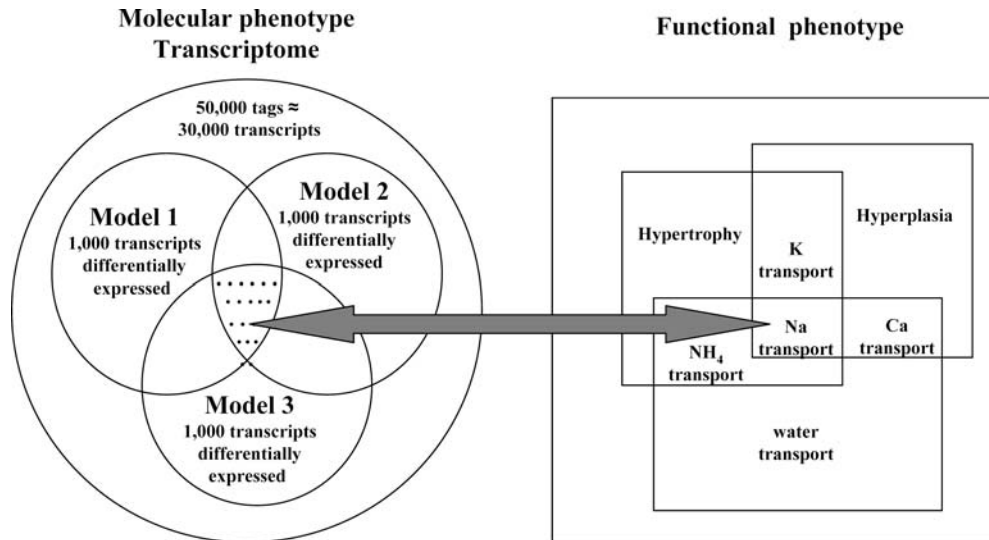
not completely understood. For example, aldosterone has been known for decades to be the main positive regulator of distal tubule sodium transport, but except for the early demonstration of its actions on sodium transporters [15, 29, 56], only a small number of aldosterone-induced genes have been identified [3, 6, 51]. The same holds for vasopressin, which controls not only water but also sodium balance. Similarly, for oedema-inducing disease such as nephrotic syndrome, the only available information regarding the dysregulation of collecting duct sodium transport concerns its independence towards aldosterone and vasopressin [10].

In conclusion, available data indicate that (1) the molecular mechanisms of the genomic effects of aldosterone and of vasopressin in the distal nephron remain mostly unknown, (2) the molecular disorders responsible for distal sodium retention in hypertensive patients are known only in the minute fraction of patients with a mendelian form of the disease, and (3) the independence of sodium retention towards aldosterone and vasopressin in oedema disease demonstrates the existence of yet unknown pathways that control distal tubule sodium transport. Thus, a promising application of renal transcriptome analysis will be to compare profiles of gene expression along the distal nephron in normal and in conditions of altered sodium transport.

As a preliminary step, one had to question the ability of SAGE to identify transcripts differentially expressed in a given nephron segment microdissected from animals under different physiological conditions. Indeed, these adaptations are likely to induce changes in gene expression profile of lesser amplitude than the differences prevailing among distinct nephron segments.

Comparison of SAGE libraries constructed from OMCD of normal mice and mice fed a potassium-depleted diet for only 3 days [7] revealed the feasibility of this strategy. As a matter of fact, sequencing of  $\approx 20,000$  tags in each library demonstrated statistically different levels of expression for  $\approx 200$  tags. Interestingly, 7% of the differentially expressed tags have no match in GenBank, further demonstrating that transcripts of physiological interest remain to be discovered. Also interesting, functionally identified tags overexpressed during potassium depletion are involved in different functional classes, including not only ion and water transporters, but also proteins involved in the control of cell cycle, apoptosis and differentiation. Kinetical correlation between the variations of expression of some of these genes and the development of cell hypertrophy and hyperplasia suggests their involvement in the control of the switch from hyperplasia to hypertrophy. In view of the above-mentioned objective to identify genes involved in the regulation of a given cell function, these results also outline the difficulty in data mining, since it seems arbitrary to decide whether a gene whose expression is modulated for example during potassium depletion is correlated with potassium transport or with any secondary cell adaptation.

One means to distinguish the genes specifically related to maintenance of the sodium balance from those related to other functions may be to compare transcriptional and functional adaptations to several "stresses" that, among various adaptative responses, all induce changes in sodium transport. As schematised on Fig. 5, analysis of transcriptomes of collecting ducts in response to adaptation to three different stresses allows characteri-



**Fig. 5** Strategy for characterising the sets of genes involved in a given cell function in distal nephron segments. The strategy is based on multiple comparisons of changes in molecular (transcriptome) and functional phenotypes of distal nephron segments in response to different adaptations (models 1–3). Adaptations are chosen to induce a common change of the function of interest (e.g. sodium transport) along with various alterations of the other functions of

distal nephron segments (including transport functions and regulation of cell cycle and differentiation). Based on previous experience, sequencing of 50,000 tags in each condition should allow the characterisation of 30,000 different tags in each library, among which 1,000 should reach statistical difference of abundance when comparing two libraries

sation of three subsets of transcripts whose expression is modified. Functional analysis of collecting ducts under the same three adaptations allows revelation of various functional and morphologic changes (with altered sodium transport as a common feature to all three adaptations). If adaptations were chosen so that sodium transport is the only functional change common to all three conditions, the group of transcripts common to the three adaptations can be defined as underlying sodium transport. This type of analysis, which aims at characterising new physiologic processes underlying regulation of sodium transport and at identifying putative targets involved in the alterations of sodium balance that induce hypertension or oedema, is currently developed in our laboratory. Besides sodium transport, it will provide profusion of data regarding other physiological processes originating in the collecting duct.

---

## Conclusion

### Methods

Coupled to microdissection, SAGE appears as a reliable method to identify gene expression profiles along the nephron in humans as well as in laboratory animals and to study changes in expression during pathophysiological processes in well defined nephron segments. Nonetheless, SAGE remains tedious, and its use limited to too few laboratories for two reasons: firstly, it requires relatively large capacities for DNA sequencing for a physiology laboratory, but this can be palliated by recouring to national large sequencing facilities, and secondly, tag identification remains a fastidious step. However, this last limit is toning down along with the increasing number of identified tags and the increasing information about genomes. When SAGE allows drawing up of the complete catalogue of transcripts expressed in the kidney, a foreseeable evolution will be to generate dedicated DNA arrays, inasmuch the sensitivity of array methodology will have been increased to become compatible with microdissected tissue.

Microdissection of renal tissue does not circumvent totally the kidney heterogeneity, since glomeruli and distal nephron segments are made of several cell types. To circumvent this restriction, it is conceivable, however, to generate SAGE libraries from subpopulations of cells isolated from these structures by immunodissection [16]. Alternately, cell localisation of transcripts identified by SAGE in these renal structures can be determined *a posteriori* by in situ hybridisation on kidney sections (when expression is high enough to be detectable) or by single-cell RT-PCR [39].

Transcriptome comparison between human and mouse [7] reveals quantitative differences that might be either of biological origin (which raises the problem of the generalised use of mouse models for human physiology and pathologies) or of methodological origin

(inherent to the condition of material collection in humans: ethical questions, pathology of donors, latency in treatment of samples). To make possible reliable studies on human kidney tissue, a solution might be to couple the recourse to biopsy samples, as managed for example by the European Renal cDNA Bank network [50], with laser microdissection to separate different structures or cells.

### Fields of application of renal transcriptome analysis

Discovery of new genes and/or transcripts is the most obvious application of renal transcriptomes. Preliminary reports in human as well as in mouse revealed an important number of transcripts displaying a high specificity of expression along the nephron, and therefore of putative physiological relevance, but of unknown molecular or functional identity. In particular, a special attention should be paid to the tags corresponding to putative antisense transcripts which display the same specificity of expression as their cognate transcript. The role of these naturally occurring antisense transcripts in the regulation of transcript expression is an open new field of investigation of major interest in regulatory physiology.

One of the most original observations of human kidney transcriptome analysis was the axial specificity of expression of genes involved in cell proliferation/death/differentiation. It will be important to define the relevance of this specificity and in particular, its relationship with terminal cell differentiation.

Finally, one of the most exciting application concerns the establishment of relationships between molecular and functional phenotypes, with the physiological objective of defining the whole network of molecules and interactions underlying specific functions. In this regard, the kidney appears as an outstanding system, owing to its high cell plasticity. SAGE is also a powerful approach in that field because, as it provides absolute tag counts, it allows *in silico* comparison to previously published transcriptomes.

**Acknowledgements** Works from our laboratory reported in this review were initiated when we were settled in the Department of Cell and Molecular Biology at the Commissariat à l'énergie atomique (CEA Saclay). They are based in part on the SADE method, which was developed under the leadership of JM Elalouf at CEA Saclay.

---

## References

1. Bonvalet JP, Doignon I, Blot-Chabaud M, Pradelles P, Farman N (1990) Distribution of 11 beta-hydroxysteroid dehydrogenase along the rabbit nephron. *J Clin Invest* 86:832-837
2. Bonventre JV (2003) Dedifferentiation and proliferation of surviving epithelial cells in acute renal failure. *J Am Soc Nephrol* [14 Suppl] 1:S55-S61

3. Boulkroun S, Fay M, Zennaro MC, Escoubet B, Jaisser F, Blot-Chabaud M, Farman N, Courtois-Coutry N (2002) Characterization of rat NDRG2 (N-Myc downstream regulated gene 2), a novel early mineralocorticoid-specific induced gene. *J Biol Chem* 277:31506–31515
4. Cai L, Huang H, Blackshaw S, Liu JS, Cepko C, Wong WH (2004) Clustering analysis of SAGE data using a Poisson approach. *Genome Biol* 5:R51
5. Chabardes-Garonne D, Mejean A, Aude JC, Cheval L, Di Stefano A, Gaillard MC, Imbert-Teboul M, Wittner M, Balian C, Anthouard V, Robert C, Segurens B, Wincker P, Weissenbach J, Doucet A, Elalouf JM (2003) A panoramic view of gene expression in the human kidney. *Proc Natl Acad Sci U S A* 100:13710–13715
6. Chen SY, Bhargava A, Mastroberardino L, Meijer OC, Wang J, Buse P, Firestone GL, Verrey F, Pearce D (1999) Epithelial sodium channel regulated by aldosterone-induced protein sgk. *Proc Natl Acad Sci USA* 96:2514–2519
7. Cheval L, Duong Van Huyen JP, Bruneval P, Verbavatz JM, Elalouf JM, Doucet A (2004) Plasticity of mouse renal collecting duct in response to potassium depletion. *Physiol Genomics* 19:61–73
8. Clarke L, Carbon J (1976) A colony bank containing synthetic Col El hybrid plasmids representative of the entire *E. coli* genome. *Cell* 9:91–99
9. Datson NA, van der Perk-de Jong J, van den Berg MP, de Kloet ER, Vreugdenhil E (1999) MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. *Nucleic Acids Res* 27:1300–1307
10. Deschenes G, Doucet A (2000) Collecting duct (Na<sup>+</sup>/K<sup>+</sup>)-ATPase activity is correlated with urinary sodium excretion in rat nephrotic syndromes. *J Am Soc Nephrol* 11:604–615
11. Duriez C, Falette N, Audouyraud C, Moyret-Lalle C, Bensaad K, Courtois S, Wang Q, Soussi T, Puisieux A (2002) The human BTG2/TIS21/PC3 gene: genomic structure, transcriptional regulation and evaluation as a candidate tumor suppressor gene. *Gene* 282:207–214
12. Ecelbarger CA, Terris J, Frindt G, Echevarria M, Marples D, Nielsen S, Knepper MA (1995) Aquaporin-3 water channel localization and regulation in rat kidney. *Am J Physiol* 269:F663–F672
13. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863–14868
14. Eladari D, Cheval L, Quentin F, Bertrand O, Mouro I, Cherif-Zahar B, Cartron JP, Paillard M, Doucet A, Chambrey R (2002) Expression of RhCG, a new putative NH(3)/NH(4)(<sup>+</sup>) transporter, along the rat nephron. *J Am Soc Nephrol* 13:1999–2008
15. Escoubet B, Coureau C, Bonvalet JP, Farman N (1997) Non-coordinate regulation of epithelial Na channel and Na pump subunit mRNAs in kidney and colon by aldosterone. *Am J Physiol* 272:C1482–C1491
16. Fejes-Toth G, Naray-Fejes-Toth A (1989) Isolated principal and intercalated cells: hormone responsiveness and Na<sup>+</sup>-K<sup>+</sup>-ATPase activity. *Am J Physiol* 256:F742–F750
17. Francki A, Sage EH (2001) SPARC and the kidney glomerulus: matricellular proteins exhibit diverse functions under normal and pathological conditions. *Trends Cardiovasc Med* 11:32–37
18. Garbers DL, Dubois SK (1999) The molecular basis of hypertension. *Annu Rev Biochem* 68:127–155
19. Gibbs RA, Weinstock GM, Metzker ML, Muzny DM, Sodergren EJ, Scherer S, Scott G, Steffen D, Worley KC, Burch PE, Okwuonu G, Hines S, Lewis L, DeRamo C, Delgado O, Dugan-Rocha S, Miner G, Morgan M, Hawes A, Gill R, Celera, Holt RA, Adams MD, Amanatides PG, Baden-Tillson H, Barnstead M, Chin S, Evans CA, Ferriera S, Fosler C, Glodek A, Gu Z, Jennings D, Kraft CL, Nguyen T, Pfannkoch CM, Sitter C, Sutton GG, Venter JC, Woodage T, Smith D, Lee HM, Gustafson E, Cahill P, Kana A, Doucette-Stamm L, Weinstock K, Fectel K, Weiss RB, Dunn DM, Green ED, Blakesley RW, Bouffard GG, De Jong PJ, Osoegawa K, Zhu B, Marra M, Schein J, Bosdet I, Fjell C, Jones S, Krzywinski M, Mathewson C, Siddiqui A, Wye N, McPherson J, Zhao S, Fraser CM, Shetty J, Shatsman S, Geer K, Chen Y, Abramzon S, Nierman WC, Havlak PH, Chen R, Durbin KJ, Egan A, Ren Y, Song XZ, Li B, Liu Y, Qin X, Cawley S, Cooney AJ, D'Souza LM, Martin K, Wu JQ, Gonzalez-Garay ML, Jackson AR, Kalafus KJ, McLeod MP, Milosavljevic A, Virk D, Volkov A, Wheeler DA, Zhang Z, Bailey JA, Eichler EE, Tuzun E, et al (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* 428:493–521
20. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6,000 genes. *Science* 274:546, 563–547
21. Griffin SV, Pichler R, Wada T, Vaughan M, Durvasula R, Shankland SJ (2003) The role of cell cycle proteins in glomerular disease. *Semin Nephrol* 23:569–582
22. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17:994–999
23. Harris PC, Germino G, Klinger K, Landes G, van Adelsberg J (1995) The PKD1 gene product. *Nat Med* 1:493
24. Hastie ND, Bishop JO (1976) The expression of three abundance classes of messenger RNA in mouse tissues. *Cell* 9:761–774
25. Heidenblut AM, Luttes J, Buchholz M, Heinitz C, Emmersen J, Nielsen KL, Schreiter P, Souquet M, Nowacki S, Herbrand U, Kloppel G, Schmiegel W, Gress T, Hahn SA (2004) aRNA-longSAGE: a new approach to generate SAGE libraries from microdissected cells. *Nucleic Acids Res* 32:e131
26. Hollander MC, Fornace AJ Jr (2002) Genomic instability, centrosome amplification, cell cycle checkpoints and Gadd45a. *Oncogene* 21:6228–6233
27. Katz AI, Doucet A, Morel F (1979) Na-K-ATPase activity along the rabbit, rat, and mouse nephron. *Am J Physiol* 237:F114–F120
28. Kershaw DB, Thomas PE, Wharram BL, Goyal M, Wiggins JE, Whiteside CI, Wiggins RC (1995) Molecular cloning, expression, and characterization of podocalyxin-like protein 1 from rabbit as a transmembrane protein of glomerular podocytes and vascular endothelium. *J Biol Chem* 270:29439–29446
29. Kim GH, Masilamani S, Turner R, Mitchell C, Wade JB, Knepper MA (1998) The thiazide-sensitive Na-Cl cotransporter is an aldosterone-induced protein. *Proc Natl Acad Sci USA* 95:14552–14557
30. Kojima R, Sekine T, Kawachi M, Cha SH, Suzuki Y, Endou H (2002) Immunolocalization of multispecific organic anion transporters, OAT1, OAT2, and OAT3, in rat kidney. *J Am Soc Nephrol* 13:848–857
31. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Showkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissole SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, et al (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921

32. Lavorgna G, Dahary D, Lehner B, Sorek R, Sanderson CM, Casari G (2004) In search of antisense. *Trends Biochem Sci* 29:88–94
33. Liebermann DA, Hoffman B (2003) Myeloid differentiation (MyD) primary response genes in hematopoiesis. *Blood Cells Mol Dis* 31:213–228
34. Liu Z, Chen Y, Mo R, Hui C, Cheng JF, Mohandas N, Huang CH (2000) Characterization of human RhCG and mouse Rhcg as novel nonerythroid Rh glycoprotein homologues predominantly expressed in kidney and testis. *J Biol Chem* 275:25641–25651
35. Liu Z, Peng J, Mo R, Hui C, Huang CH (2001) Rh type B glycoprotein is a new member of the Rh superfamily and a putative ammonia transporter in mammals. *J Biol Chem* 276:1424–1433
36. Loffing J, Loffing-Cueni D, Valderrabano V, Klausli L, Hebert SC, Rossier BC, Hoenderop JG, Bindels RJ, Kaissling B (2001) Distribution of transcellular calcium and sodium transport pathways along mouse distal nephron. *Am J Physiol* 281:F1021–F1027
37. Mene P, Polci R, Festuccia F (2003) Mechanisms of repair after kidney injury. *J Nephrol* 16:186–195
38. Morel F, Doucet A (1986) Hormonal control of kidney functions at the cell level. *Physiol Rev* 66:377–468
39. Nissant A, Lourdel S, Baillet S, Paulais M, Marvaux P, Teulon J, Imbert-Teboul M (2004) Heterogeneous distribution of chloride channels along the distal convoluted tubule probed by single-cell RT-PCR and patch-clamp. *Am J Physiol Renal Physiol* 287:F1233–F1243
40. Nony PA, Schnellmann RG (2003) Mechanisms of renal cell repair and regeneration after acute renal failure. *J Pharmacol Exp Ther* 304:905–912
41. Nurse P (2000) A long twentieth century of the cell cycle and beyond. *Cell* 100:71–78
42. Peters DG, Kassam AB, Yonas H, O'Hare EH, Ferrell RE, Brufsky AM (1999) Comprehensive transcript analysis in small quantities of mRNA by SAGE-lite. *Nucleic Acids Res* 27:e39
43. Price PM, Megyesi J, Safirstein RL (2003) Cell cycle regulation: repair and regeneration in acute renal failure. *Semin Nephrol* 23:449–459
44. Quentin F, Eladari D, Cheval L, Lopez C, Goossens D, Colin Y, Cartron JP, Paillard M, Chambrey R (2003) RhBG and RhCG, the putative ammonia transporters, are expressed in the same cells in the distal nephron. *J Am Soc Nephrol* 14:545–554
45. Robert-Nicoud M, Flahaut M, Elalouf JM, Nicod M, Salinas M, Bens M, Doucet A, Wincker P, Artiguenave F, Horisberger JD, Vandewalle A, Rossier BC, Firsov D (2001) Transcriptome of a mouse kidney cortical collecting duct cell line: effects of aldosterone and vasopressin. *Proc Natl Acad Sci USA* 98:2712–2716
46. Royaux IE, Wall SM, Karniski LP, Everett LA, Suzuki K, Knepper MA, Green ED (2001) Pendrin, encoded by the Pendred syndrome gene, resides in the apical region of renal intercalated cells and mediates bicarbonate secretion. *Proc Natl Acad Sci USA* 98:4221–4226
47. Sabolic I, Katsura T, Verbavatz JM, Brown D (1995) The AQP2 water channel: effect of vasopressin treatment, microtubule disruption, and distribution in neonatal rats. *J Membr Biol* 143:165–175
48. Saha S, Sparks AB, Rago C, Akmaev V, Wang CJ, Vogelstein B, Kinzler KW, Velculescu VE (2002) Using the transcriptome to annotate the genome. *Nat Biotechnol* 20:508–512
49. Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467–470
50. Schmid H, Cohen CD, Henger A, Schlondorff D, Kretzler M (2004) Gene expression analysis in renal biopsies. *Nephrol Dial Transplant* 19:1347–1351
51. Spindler B, Mastroberardino L, Custer M, Verrey F (1997) Characterization of early aldosterone-induced RNAs identified in A6 kidney epithelia. *Pflügers Arch* 434:323–331
52. Templin MF, Stoll D, Schrenk M, Traub PC, Vohringer CF, Joos TO (2002) Protein microarray technology. *Trends Biotechnol* 20:160–166
53. Tessema M, Lehmann U, Kreipe H (2004) Cell cycle and no end. *Virchows Arch* 444:313–323
54. Tirone F (2001) The gene PC3(TIS21/BTG2), prototype member of the PC3/BTG/TOB family: regulator in control of cell growth, differentiation, and DNA repair? *J Cell Physiol* 187:155–165
55. Trougakos IP, Gonos ES (2002) Clusterin/apolipoprotein J in human aging and cancer. *Int J Biochem Cell Biol* 34:1430–1448
56. Tsuchiya K, Giebisch G, Welling PA (1996) Aldosterone-dependent regulation of Na-K-ATPase subunit mRNA in the rat CCD: competitive PCR analysis. *Am J Physiol* 271:F7–F15
57. Vairapandi M, Balliet AG, Hoffman B, Liebermann DA (2002) GADD45b and GADD45 g are cdc2/cyclinB1 kinase inhibitors with a role in S and G2/M cell cycle checkpoints induced by genotoxic stress. *J Cell Physiol* 192:327–338
58. Van den Berg A, van der Leij J, Poppema S (1999) Serial analysis of gene expression: rapid RT-PCR analysis of unknown SAGE tags. *Nucleic Acids Res* 27:e17
59. Vande Walle JG, Donckerwolcke RA (2001) Pathogenesis of edema formation in the nephrotic syndrome. *Pediatr Nephrol* 16:283–293
60. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. *Science* 270:484–487
61. Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE, Jr., Hieter P, Vogelstein B, Kinzler KW (1997) Characterization of the yeast transcriptome. *Cell* 88:243–251
62. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gano W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, et al (2001) The sequence of the human genome. *Science* 291:1304–1351
63. Vilain C, Libert F, Venet D, Costagliola S, Vassart G (2003) Small amplified RNA-SAGE: an alternative approach to study transcriptome from limiting amount of mRNA. *Nucleic Acids Res* 31:e24
64. Virlon B, Cheval L, Buhler JM, Billon E, Doucet A, Elalouf JM (1999) Serial microanalysis of renal transcriptomes. *Proc Natl Acad Sci USA* 96:15286–15291
65. Wagner SN, Atkinson MJ, Wagner C, Hofler H, Schmitt M, Wilhelm O (1996) Sites of urokinase-type plasminogen activator expression and distribution of its receptor in the normal human kidney. *Histochem Cell Biol* 105:53–60
66. Wang J, Chen X, Zhu H, Peng L, Hong Q (2003) Relationship between aging and renal high-affinity sodium-dependent dicarboxylate cotransporter-3 expression characterized with antifusion protein antibody. *J Gerontol A Biol Sci Med Sci* 58:B879–888
67. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS,



- Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Esvara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, et al (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520–562
68. Weinberg WC, Denning MF (2002) P21Waf1 control of epithelial cell cycle and cell fate. *Crit Rev Oral Biol Med* 13:453–464
69. Wu MX (2003) Roles of the stress-induced gene IEX-1 in regulation of cell death and oncogenesis. *Apoptosis* 8:11–18
70. Yang T, Huang YG, Singh I, Schnerrmann J, Briggs JP (1996) Localization of bumetanide- and thiazide-sensitive Na-K-Cl cotransporters along the rat nephron. *Am J Physiol* 271:F931–F939
71. Ye SQ, Zhang LQ, Zheng F, Virgil D, Kwiterovich PO (2000) miniSAGE: gene expression profiling using serial analysis of gene expression from 1 microg total RNA. *Anal Biochem* 287:144–152
72. Zhang R, Skach W, Hasegawa H, van Hoek AN, Verkman AS (1993) Cloning, functional analysis and cell localization of a kidney proximal tubule water transporter homologous to CHIP28. *J Cell Biol* 120:359–369