Diabetic nephropathy (DNP) is by far the most common cause of ESRD (1). Black Americans, Mexican Americans, and Native Americans are disproportionately affected by DNP compared with white Americans (2). Approximately one third of individuals with diabetes develop DNP with a high likelihood of progression to ESRD. In addition, DNP is associated with considerably increased cardiovascular disease risk and mortality. Thus, the public health burden from DNP is enormous (3). Current evidence suggests that both genetic and environmental factors determine susceptibility to develop DNP and the risk for and rate of progression of DNP (4–6). Hypertension, poor glycemic and lipid control, and smoking increase the risk for development of DNP (7). For example, the Diabetes Control and Complications Trial and the United Kingdom Prospective Diabetes Study showed the importance of strict glucose and BP control in delaying diabetic complications (8,9). Epidemiologic studies have shown that DNP is strongly clustered in families and that race has a major effect on DNP susceptibility and rate of progression, firmly establishing the importance of genetic risk factors in the development of DNP (2,4,10).

Currently available therapeutic approaches are focused on blockade of the renin-angiotensin system (RAS). Thus, angiotensin-converting enzyme inhibitors and angiotensin receptor blockers are able to slow the rate of progression but do not arrest or reverse the disease (11,12). Moreover, RAS blockade is usually initiated only after DNP manifests itself clinically with persistent proteinuria in both type 1 and type 2 diabetes. However, we now can postulate that the initiating pathomechanisms of DNP precede the clinical onset considerably. Indeed, it is possible that molecular and cellular changes that eventually lead to clinical DNP are present in kidneys of individuals who already are at risk shortly after the onset of diabetes. For example, several studies suggest that reduction of podocyte numbers per glomerulus is detectable early in the course of both type 1 and type 2 diabetes and is a strong predictor of subsequent proteinuria (13–15). Thus, although DNP is characterized by a prolonged clinical latency period that lasts for years, a major barrier in DNP research is the absence of targeted investigation of molecular and cellular changes that occur in kidneys of individuals with new onset of diabetes.

We propose that modern genomic and proteomic technologies now can be applied to investigate patterns of gene and protein regulation in kidney at the onset of diabetes with the goal to develop characteristic molecular profiles that predict reliably whether a patient with newly diagnosed diabetes is at risk for DNP, even many years before its first clinical manifes-
tation. The major rationale for such “personalization” of DNP risk at the onset of diabetes would be to allocate resources to individuals who are most likely to develop microvascular disease. This may include intensification of care parameters such as increased support personnel to achieve maximally tight blood glucose and BP control and early initiation of future preventive treatment modalities. Equally important will be the identification of individuals who have diabetes and carry little or no DNP risk, presumably need less assiduous follow-up, and can be spared cost and risk exposure inherent of future chronic treatments. However, considerable barriers that limit rapid progress toward this goal currently exist. The most important limitation is undoubtedly the general lack of renal biopsy material from individuals with diabetes. Here we review existing literature and new approaches that highlight the extraordinary promise of genomic and proteomic strategies for the development of molecular markers and diagnostic tools for early and reliable prediction of DNP in individuals with diabetes.

Proteomics

Proteomics of Kidney Tissue in DNP

Proteomics measures protein level at a global scale. Proteomics studies are becoming increasingly popular because they provide direct information about protein levels, which are essential to execute gene functions (16). Proteomics has been applied successfully to rodent models of diabetic nephropathy (17), but it requires large amounts of kidney material; therefore, it is less suitable for the analysis of human kidney biopsy materials. Recently, Thongboonkerd et al. (17) performed proteomic analysis on kidneys of 120-d-old OVE26 transgenic mice, which display many characteristics of early-onset human type 1 diabetes (18). They used two-dimensional gel electrophoresis with SYPRO Ruby staining, quantitative intensity analysis, and matrix assisted laser desorption ionization–time of flight mass spectrometry. Thirty proteins were identified as differentially expressed in the diabetic kidney, including proteases, protease inhibitors, apoptosis-associated proteins, regulators of oxidative tolerance, calcium-binding proteins, transport regulators, cell signaling proteins, and smooth muscle contractile elements. Nineteen of the altered proteins had previously been shown to be regulated during diabetes, whereas roles for the other 11 altered proteins had not previously been established, suggesting that they may function by novel mechanisms in diabetic nephropathy.

Proteomics to Analyze Urinary Protein Profiles

Application of proteomics tools to analyze urinary protein profiles is rapidly emerging as a promising research area for marker discovery (19). The most important advantage of urine proteomics is the prospect of a noninvasive and easy sampling method. Urinary proteins are a mixture of proteins filtered from plasma or secreted by kidney cells. Normal urinary proteins generally reflect normal tubular physiology. Information on changes in urinary protein excretion by various interventions is essential for a better understanding of tubular and glomerular responses to physiologic stimuli. Rossing et al. (20) performed urinary proteomics analysis in four groups of patients with type 2 diabetes, matched for age gender and diabetes duration, including 20 normoalbuminuric patients, 20 microalbuminuric patients with retinopathy, and 18 macroalbuminuric patients with retinopathy. Furthermore, changes in urinary polypeptide patterns during treatment with the angiotensin receptor blocker candesartan were evaluated in the macroalbuminuric patients in a randomized, double-blinded, cross-over trial in which each patient received treatment with placebo and with candesartan 8, 16, and 32 mg/d each for 2 mo. They used a combination of capillary electrophoresis and mass spectrometry to identify differentially regulated proteins and were able to identify 4551 different polypeptides in the urine samples. Urinary polypeptide patterns were not significantly different in normo- and macroalbuminuric patients, whereas distinct differences were found in macroalbuminuric patients. Differences in urinary polypeptide patterns between normo- and macroalbuminuric patients permitted the establishment of a “diabetic renal damage” pattern that consisted of 113 polypeptides. Eleven of these polypeptides had been sequenced and identified. Candesartan treatment in macroalbuminuric patients significantly changed 15 of the 113 polypeptides in the diabetic renal damage pattern toward levels in normoalbuminuric patients.Change in the diabetic renal damage pattern was not candesartan dose dependent, but individual changes correlated with changes in urinary albumin excretion at each dose level.

Similar studies have been conducted by Mischak et al. (21) on patients with type 2 diabetes and DNP. These studies indicate that urine proteomics could be adapted for pharmacoproteomics approaches, such as identification of responders and nonresponders of RAS blockade, etc. Sharma et al. (22) used two-dimensional differential in-gel electrophoretic analysis to examine protein patterns in urine samples in patients with DNP. All patients had longstanding diabetes, impaired renal function, and overt proteinuria. They identified 99 differentially regulated spots in the urine proteome of the diabetic samples, with 63 up- and 36 downregulated. One spot was consistently upregulated by 19-fold across individuals in the diabetic group, and it was identified as α 1 antitrypsin. ELISA of urine samples from a separate group of patients and control subjects confirmed a marked increase of α 1 antitrypsin in patients with diabetes.

Barriers for Proteomics Applications in Marker Discovery

Although proteomics strategies are rapidly emerging in renal research, considerable technical limitations remain to be overcome. First, current technologies lack sensitivity to allow detection of moderate- and low-abundance proteins. Second, reliable algorithms to facilitate standardization of proteomics data sets to enable reliable and large-scale comparisons between many samples or even between laboratories are not readily available and need to be developed. Third, proteolytic cleavage of proteins into peptides occurs rapidly and uncontrollably and presents a major problem for the analysis of proteomics data sets. New approaches are under development using gel-free technology and sample fractionation before analysis to improve sensitivity for detection of membrane and low-abundance proteins.
Gene Expression Profiling: Technologies and Barriers for Translational Research in DNP  

Technology and Applications

The development of microarray technology has revolutionized functional genomics by providing tools that allow the parallel measurement of gene expression in whole genomes with thousands of genes (23–25). Microarrays are artificially constructed grids of probes, such that each element of the grid probes for a specific RNA sequence; that is, each holds a DNA sequence that is a reverse complement to the target RNA sequence. In gene expression microarrays, either synthetic oligonucleotides or cDNA fragments are used as probes (26,27). Although, there are many protocols and platforms available, the basic technique involves extraction of RNA from biologic samples, followed by complementary RNA copying with incorporation of either fluorescence nucleotides or tags that can be labeled subsequently. Thus, samples contain a complete complement of labeled RNA or cDNA species that can be hybridized to microarrayed specific complementary sequence probes. This procedure typically generates thousands of measurements of gene expression per biologic sample. The cost of microarray experimentation continues to decrease rapidly, allowing for much broader use of this powerful technology in many translational research disciplines, including biomarker discovery, drug discovery, pharmacogenomics, toxicogenomics, systems biology, and molecular pathology (28,29).

Technical Barriers for Application of Gene Expression Profiling in Renal Biopsy Samples

Although clearly very promising, the application of microarray technology to analyze gene expression profiles in renal biopsy tissue is also extremely challenging (30–32). Renal biopsy tissue is typically very heterogeneous, with proportions of glomerular and tubular segments often varying dramatically, depending on the path of the biopsy needle in kidney tissue during the procedure. The considerable sample heterogeneity of undissected kidney biopsy cores precludes reliable sample-to-sample comparisons. Two different solutions to this problem are being pursued by various investigators. Pioneering studies by Kretzler et al. (32,33) have established that manual microdissection of fresh renal biopsy cores is a productive approach that allows efficient separation of glomeruli and tubuli. It is critical that biopsy cores be placed immediately in storage solutions that contain inhibitors of RNase to prevent RNA degradation and to conserve an in vivo transcriptome profile. Similarly, our group has developed a microdissection-based method to prepare glomerular and tubular segments of fresh kidney biopsy cores for microarray analysis (K.S. and E.P.B., unpublished observations). Although this approach provides consistently high-quality RNA samples, it is very labor-intensive and requires considerable skills in microdissection, limiting its current application to clinical and translational research projects.

Laser capture microdissection is an alternative technology that can be applied to retrieve glomeruli from fixed kidney tissue sections under microscopic guidance (34). The procedure allows accurate identification and procurement of glomerular or tubular cells from tissue samples under direct microscopic visualization. Disadvantages of laser capture microdissection approaches are the considerable inconsistency of RNA quality caused by degradation induced during tissue fixation, sectioning, and laser-induced degradation. Both manual and laser capture microdissection methods typically provide very small amounts of tissue. Until recently, it was not possible to isolate and quantify very small amounts of RNA (typically picogram quantities) reliably. In addition, quality control of such small samples was not possible. Several research tool companies recently introduced new technologies to overcome these technical limitations, including resin-based RNA extraction and lab-on-chip–based microfluidic electrophoresis methods. Thus, these technical challenges now have been solved through the development of refined protocols and improved reagents and tools.

Paucity of Kidney Biopsies in Individuals Who Have Diabetes without and with Clinical DNP

In general, the diagnosis of DNP is established on the basis of clinical findings, including patient history, signs and symptoms, and urinary protein excretion. Because analysis of kidney tissue by renal pathologists is not required to diagnose DNP, kidney biopsies are only rarely performed in individuals with diabetes. As a consequence, kidney tissue is typically not available for research studies. The paucity of new kidney biopsy tissue is severely restricting translational molecular research in DNP and prevents the application of microarray analysis for gene expression profiling in human DNP. Thus, gene expression profiling in DNP research is currently largely restricted to experimental animal models. In addition, it has proved very difficult to obtain approval for protocol research kidney biopsies from institutional review boards at US medical centers, with few exceptions, largely because of the risk associated with this invasive procedure. Although considerable progress has been achieved in our understanding of the pathomechanisms of established DNP with clinical manifestations in both type 1 and type 2 diabetes, the early molecular and cellular changes that are induced in kidney with the onset of diabetes in individuals who are at risk for developing DNP remain unclear. Therefore, the paucity of kidney biopsies in individuals who have diabetes without and with clinical DNP is to a large extent responsible for the relative standstill in translational research of early DNP.

Challenges for Biologically or Clinically Meaningful Analysis of Gene Expression Data

Microarrays deliver large amounts of data on tens of thousands of genes. The result is an immense quantity of biologic information that needs to be analyzed, presented, and archived in a meaningful way. Therefore, functional genomic studies should be combined with advanced computational and biostatistical approaches (35). For potential identification of biomarkers, the gene expression data have to be analyzed in conjunction with patient and sample variables. The most basic question that one can ask in a transcriptional profiling experiment is which genes’ expression levels changed significantly when gene expression levels in two different groups are compared.
Although the statistical methods to identify lists of differentially expressed genes are very powerful, it is considerably more challenging to determine meaningful correlations between gene expression patterns and clinical parameters.

In general, pattern discovery methods such as clustering provide a high-level overview of a data set and may be the first analysis step in a study that ultimately involves other analytical methods (36). Such techniques include dimension-reduction methods, as well as various “clustering” techniques designed for finding groups within the data. What these methods have in common is that they simplify the data set, ideally in ways that impart additional information about its structure, and that they are considered “unsupervised,” meaning that the reduction is derived solely from the data rather than reflecting any previous knowledge or classification scheme (37).

In contrast to pattern discovery, class prediction methods are techniques specifically designed to classify objects into known groups (36). Numerous reports describe machine-learning algorithms as computational techniques for classifying multidimensional data. Most methods include a training phase, run-on samples whose classes are already known (“training set”), and a testing phase in which the algorithm generalizes from the training data to predict classifications of previously unseen samples (“test or validation set”). Because direction is provided in the training phase, prediction methods are referred to as “supervised” classification methods. For microarray data derived from clinical studies, prediction generally refers to the classification of patients’ samples by characteristics such as disease subtype or response to treatment (38–40). These reports demonstrate convincingly that microarray data and class prediction methods provide a powerful approach to define disease classifications and to predict outcome or treatment response. Once highly predictive classifier gene sets are identified, expression patterns of these genes must be verified by methods other than microarray. Typically, quantitative real-time PCR analysis is performed on the genes of interest in the same RNA samples that were subjected to microarray analysis. Finally, the performance of prediction and classification algorithms should be verified in new, unknown groups of patients/samples (41).

**Gene Expression Profiling Studies of Human DNP**

Using cadaver kidneys, Baelde et al. (42) characterized gene expression patterns in two normal kidneys and two kidneys with DNP obtained postmortem. Glomerular RNA was hybridized in duplicate on Human Genome U95Av2 Arrays (Affymetrix, Santa Clara, CA). Ninety-six genes were increased in diabetic glomeruli, whereas 519 genes were decreased. The list of genes with increased expression levels in DNP included aquaporin 1, calpain 3, hyaluronoglucosidase, and platelet/endothelial cell adhesion molecule. The list of genes with decreased expression levels included bone morphogenetic protein 2, vascular endothelial growth factor, fibroblast growth factor 1, IGF binding protein 2, and nephrin. A decrease in vascular endothelial growth factor and nephrin was validated at the protein level and also at the RNA level in renal biopsy specimens from five additional patients with diabetes. However, major limitations of this study are the small sample number and the use of cadaveric kidney tissue without confirmation of the integrity of the RNA.

**Gene Expression Profiling Studies in Kidneys of Murine Models of DNP**

Because human kidney tissue is not readily available, mouse models of DNP have been subjected to microarray analysis. Wada et al. (43) analyzed gene expression in kidneys of streptozotocin (STZ)-induced diabetic mice using high-density DNA filter arrays. They studied four experimental groups: Control mice, mice that were subjected to unilateral nephrectomy, STZ-induced diabetic mice, and STZ-induced diabetic mice with unilateral nephrectomy. Histopathologic changes were examined at 24 wk after the induction of diabetes. Gene expression profiles were compared between control and STZ-treated mice with or without nephrectomy, respectively. Both the STZ-treated and the uninephrectomized STZ-treated mice manifested similar degrees of glomerular hypertrophy and glomerulosclerosis. Sixteen transcripts with increased expression and 65 with decreased expression were identified in diabetic kidneys. The identified genes were enriched in functional categories related to glucose and lipid metabolism, ion transport, transcription factors, signaling molecules, and extracellular matrix–related molecules. Mishra et al. (44) performed gene expression profiling of whole kidneys that were obtained from db/db mice with new onset of type 2 diabetes and with long-standing type 2 diabetes with albuminuria and mesangial matrix expansion. A total of 639 RNA transcripts were differentially expressed in kidneys between groups, including new genes that are usually present in adipocytes, such as adipocyte differentiation-regulated protein (ADRP; or adipophilin in humans). ADRP is a perilipin family protein that forms lipid storage vesicles and controls triglyceride utilization. The authors showed that accumulation of lipid storage droplets correlated with the magnitude and localization of ADRP in db/db kidneys. Additional genes that are involved in lipid transport, oxidation, and storage were differentially expressed in db/db kidneys. All are known to be regulated directly by peroxisome proliferator–activated receptor-α (PPAR-α). It is interesting that peroxisome PPAR-α protein was found to be upregulated in glomeruli, cortical tubules, and renal arterial vessels of db/db mice. Wilson et al. (45) analyzed gene expression in kidneys of NOD mice with distinct phenotypic profiles, including prediabetic, new-onset diabetes, and long-term type 1 diabetes. Whereas 27 genes were decreased, only the glutathione peroxidase 3 (Gpx3) gene showed increased expression in the new-onset diabetic mice compared with nondiabetic control NOD mice. Conversely, 19 of the 27 genes that were initially decreased and seven additional genes were increased, whereas Gpx3 was decreased in NOD mice with long-term diabetes compared with controls. The majority of these genes function in four major signaling pathways, including insulin, TGF-β, TNF-α, and PPAR pathways.

Fan et al. (46) combined analysis of quantitative trait loci associated with DNP and gene expression profiling in kidney in a so-called “genomical genetics” approach. They used Af-
fimetryx GeneChip Expression Analysis system to survey gene expression profiles of diabetic KK/Ta mouse kidneys. Profiling was performed in kidneys from three KK/Ta diabetic mice and two BALB/c mice at 20 wk of age. Ninety-eight known genes and 31 expressed sequence tags were found to be differentially expressed between KK/Ta and BALB/c kidneys. The differentially expressed genes were involved in renal function, extracellular matrix expansion and degradation, signal transduction, transcription regulation, ion transport, glucose and lipid metabolism, and protein synthesis and degradation. It is interesting that analysis of genes located on a quantitative trait loci for the development of albuminuria in KK/Ta mice, called UA-1, revealed that five differentially expressed genes resided within the albuminuria quantitative trait loci chromosomal region, including syndecan-4, S-adenosylhomocysteine hydrolase, somatostatin receptor 4, and Kreisler leucine zipper protein.

In contrast with the studies described in the previous paragraphs of this section, we designed gene expression profiling studies in which identification of differentially expressed genes was based on differences in phenotype features instead of predefined study groups of mice with type 1 diabetes (STZ) and type 2 diabetes (db/db) (47). In addition, this study differed from previous studies by a much larger sample size of 65 total animals. Standardized phenotype classification and comparatively large sample size enabled us to use supervised methods to identify classifier genes for mesangial matrix expansion or hyperglycemia/albuminuria, respectively. These studies identified hydroxysteroid dehydrogenase-3ß isotype 4 and osteopontin as top classifier genes for the mesangial matrix expansion phenotype. The expression levels of these genes also allowed the classification of a separate group of animals for the absence or presence of diabetic glomerulopathy with a high degree of precision. A similar analysis identified the scavenger receptor CD36 and kidney androgen regulated protein as top classifiers for hyperglycemia with albuminuria. In a follow-up study, we recently demonstrated that CD36 was increased in proximal tubular cells in human DNP, where it was mediating tubular cell apoptosis induced by glycated and free fatty acids (48).

Conclusions

Genomics and proteomics approaches are now widely used to analyze gene and protein expression profiles that underlie experimental kidney disease, including DNP. These studies have already identified molecular markers that correlate with tubulointerstitial progression of DNP in humans (48). Additional candidate molecular markers for use in molecular characterization of DNP will undoubtedly be discovered by refined, large-scale studies of gene expression at defined stages of diabetes and DNP in mice and other experimental models. In addition, urinary proteome analyses in individuals with diabetes are under way, providing an excellent opportunity for biomarker discovery. In parallel, technological advances and improved protocols enable reliable genome-wide gene expression studies directly in kidney, using extremely small amounts of tissue such as glomeruli and tubular segments obtained from microdissected renal biopsy cores typically obtained in clinical practice. These techniques now are ready to be applied to characterize the molecular events that occur in kidneys of individuals with diabetes many years before the clinical manifestation of DNP and to develop reliable predictors of the risk for DNP already at the onset of diabetes (Figure 1). In addition, the risk for complications that are associated with renal biopsies guided by modern imaging techniques is sufficiently reduced so that it should no longer preclude large-scale studies involving research kidney biopsies for genomic and molecular analysis. Such bold translational studies undoubtedly entail extraordinary potential to transform the clinical management of DNP by individualizing the care of persons with diabetes. In the future, we will be able to identify individuals who have diabetes and are at risk for DNP and to start truly preventive interventions at the onset of diabetes (Figure 1). Equally important, we will be able to avoid unnecessary treatment and exposure of individuals who have diabetes and little or no risk for DNP. DNP is a new frontier for translational and personalized medicine.

Acknowledgments

This work was supported by National Institutes of Health grants R01 DK060043, R01 DK056077, P50 DK064236-010003, and U01DK060995 to E.P.B. K.S. was the recipient of the National Kidney Foundation...
References


Young Investigator Award and the Juvenile Diabetes Research Foundation Career Development Award.


