



Published in final edited form as:

*Sci Transl Med.* 2015 December 2; 7(316): 316ra193. doi:10.1126/scitranslmed.aac7071.

## Tissue transcriptome-driven identification of epidermal growth factor as a chronic kidney disease biomarker

Wenjun Ju<sup>#1,2,\*</sup>, Viji Nair<sup>#1</sup>, Shahaan Smith<sup>1,‡</sup>, Li Zhu<sup>3</sup>, Kerby Shedden<sup>2,4</sup>, Peter X. K. Song<sup>5</sup>, Laura H. Mariani<sup>1,6</sup>, Felix H. Eichinger<sup>1</sup>, Celine C. Berthier<sup>1</sup>, Ann Randolph<sup>1</sup>, Jennifer Yi-Chun Lai<sup>1</sup>, Yan Zhou<sup>5,§</sup>, Jennifer J. Hawkins<sup>1</sup>, Markus Bitzer<sup>1</sup>, Matthew G. Sampson<sup>7</sup>, Martina Thier<sup>8</sup>, Corinne Solier<sup>8</sup>, Gonzalo C. Duran-Pacheco<sup>8</sup>, Guillemette Duchateau-Nguyen<sup>8</sup>, Laurent Essioux<sup>8</sup>, Brigitte Schott<sup>8</sup>, Ivan Formentini<sup>8</sup>, Maria C. Magnone<sup>8,¶</sup>, Maria Bobadilla<sup>8</sup>, Clemens D. Cohen<sup>9,||</sup>, Serena M. Bagnasco<sup>10</sup>, Laura Barisoni<sup>11</sup>, Jicheng Lv<sup>3</sup>,

\*Corresponding author. wenjunj@med.umich.edu (W.J.); kretzler@umich.edu (M.K.).

‡Present address: Department of Medical Oncology, University of Washington, Seattle, WA 98109, USA.

§Present address: Merck Research Laboratories, North Wales, PA 19454, USA.

¶Present address: Cardiovascular and Metabolic Diseases Innovative Medicines, AstraZeneca R&D, Mölndal SE-431 83, Sweden.

||Present address: Department for Kidney Disease and Hypertension and Clinical Immunology, Harlaching Hospital, Munich 81545, Germany.

**Author contributions:** M.K., W.J., and V.N. participated in the study design. W.J., S.S., L.Z., A.R., M.T., C.S., and B.S. (for C-PROBE), W.J., S.S., S.M.B., and L.B. (for NEPTUNE), L.Z. (for PKU-IgAN), and C.D.C. (for ERCB) participated in data generation. K.S., V.N., L.Z., W.J., S.S., F.H.E., C.C.B., J.Y.-C.L., Y.Z., P.X.K.S., L.H.M., L.E., I.F., G.C.D.-P., G.D.-N., B.S., M.C.M., M. Bobadilla, H.-Y.W., H.Z., J.L., L.Z., and M.K. participated in data analysis. W.J., V.N., L.Z., H.Z., L.H.M., K.S., M. Bitzer, S.M.B., L.B., M.G.S., F.C.B., and M.K. participated in data interpretation; J.J.H., C.A.G., H.-Y.W., H.Z., and C.D.C. provided study materials; W.J. and C.C.B. participated in figure preparation; and W.J., V.N., L.H.M., P.X.K.S., F.C.B., K.S., and M.K. participated in writing the paper. The contribution of V.N. in this study is, in part, to fulfill the thesis requirement at the Ludwig Maximilian University of Munich. W.J. and V.N. share first authorship. W.J. and M.K. share correspondence.

**Competing interests:** M. Bobadilla, G.C.D.-P., G.D.-N., L.E., M.C.M., M.T., I.F., C.S., M.K., V.N., and W.J. hold a patent PCT/EP2014/073413 “Biomarkers and methods for progression prediction for chronic kidney disease” related to this work. M.K. reports grants from Hoffman–La Roche during the conduct of the study; research support from AbbVie, AstraZeneca, Boehringer Ingelheim, and Eli Lilly outside the submitted work. M.K. is on the Board Advisory Committee of AbbVie, Eli Lilly, and Pfizer (honoraria paid to institution). C.D.C. received speaker honoraria from Hoffman–La Roche.

**Data and materials availability:** The gene expression Cel files are available at Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) under reference nos. GSE32591, GSE37455, GSE35488, GSE47185, and GSE69438.

### SUPPLEMENTARY MATERIALS

[www.sciencetranslationalmedicine.org/cgi/content/full/7/316/316ra193/DC1](http://www.sciencetranslationalmedicine.org/cgi/content/full/7/316/316ra193/DC1)

#### Materials and Methods

Fig. S1. Baseline eGFR prediction by a three-marker panel in C-PROBE.

Fig. S2. EGF and two other candidate mRNAs' expression patterns in a panel of 84 human organs, tissues, and cell lines.

Fig. S3. Correlation of EGF mRNA and uEGF/Cr with eGFR and eGFR slope in DN patients and CKD patients with diabetes.

Fig. S4. Correlation of uEGF/Cr with ACR.

Fig. S5. EGF as the top upstream regulator of genes correlated with eGFR slope.

Fig. S6. Correlation of uEGF/Cr with eGFR slope.

Fig. S7. ROC curve and corresponding AUC statistics for models with and without uEGF.

Fig. S8. mRNA localization by in situ hybridization: Negative and positive control images.

Table S1. Significantly enriched canonical pathways in intrarenal marker set.

Table S2. qRT-PCR assays used to validate expression of the intrarenal transcripts.

Table S3. Correlations of identified intrarenal transcripts with log<sub>2</sub> eGFR of patients from the discovery and two validation cohorts.

Table S4. Demographic characteristics of NEPTUNE patients with intrarenal EGF expression data available.

Table S5. Top 10 upstream regulators of transcripts correlated with eGFR change over time (eGFR slope).

References (45-53)

**Citation:** W. Ju, V. Nair, S. Smith, L. Zhu, K. Shedden, P. X. K. Song, L. H. Mariani, F. H. Eichinger, C. C. Berthier, A. Randolph, J. Y.-C. Lai, Y. Zhou, J. J. Hawkins, M. Bitzer, M. G. Sampson, M. Thier, C. Solier, G. C. Duran-Pacheco, G. Duchateau-Nguyen, L. Essioux, B. Schott, I. Formentini, M. C. Magnone, M. Bobadilla, C. D. Cohen, S. M. Bagnasco, L. Barisoni, J. Lv, H. Zhang, H.-Y. Wang, F. C. Brosius, C. A. Gadegebeku, M. Kretzler, for the ERCB, C-PROBE, NEPTUNE, and PKU-IgAN Consortium, Tissue transcriptome-driven identification of epidermal growth factor as a chronic kidney disease biomarker. *Sci. Transl. Med.* 7, 316ra193 (2015).

Hong Zhang<sup>3</sup>, Hai-Yan Wang<sup>3</sup>, Frank C. Brosius<sup>1,12</sup>, Crystal A. Gadegbeku<sup>13</sup>, Matthias Kretzler<sup>1,2,\*</sup>, and the ERCB, C-PROBE, NEPTUNE, and PKU-IgAN Consortium

<sup>1</sup>Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109, USA.

<sup>2</sup>Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA. <sup>3</sup>Renal Division, Department of Internal Medicine, Peking University First Hospital, Peking University Institute of Nephrology, Beijing 100034, China. <sup>4</sup>Department of Statistics, University of Michigan, Ann Arbor, MI 48109, USA. <sup>5</sup>Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA. <sup>6</sup>Arbor Research Collaborative for Health, Ann Arbor, MI 48104, USA. <sup>7</sup>Department of Pediatrics, University of Michigan, Ann Arbor, MI 48109, USA. <sup>8</sup>Roche Pharmaceutical Research and Early Development–Roche Innovation Center, 4070 Basel, Switzerland. <sup>9</sup>Division of Nephrology, Institute of Physiology, University of Zurich, CH-8006 Zürich, Switzerland. <sup>10</sup>Department of Pathology, Johns Hopkins School of Medicine, Baltimore, MD 21287, USA. <sup>11</sup>Department of Pathology, Miller School of Medicine, University of Miami, Miami, FL 33136, USA. <sup>12</sup>Department of Molecular and Integrative Physiology, University of Michigan, Ann Arbor, MI 48109, USA. <sup>13</sup>Temple Clinical Research Institute, Temple University School of Medicine, Philadelphia, PA 19140, USA.

# These authors contributed equally to this work.

## Abstract

Chronic kidney disease (CKD) affects 8 to 16% people worldwide, with an increasing incidence and prevalence of end-stage kidney disease (ESKD). The effective management of CKD is confounded by the inability to identify patients at high risk of progression while in early stages of CKD. To address this challenge, a renal biopsy transcriptome-driven approach was applied to develop noninvasive prognostic biomarkers for CKD progression. Expression of intrarenal transcripts was correlated with the baseline estimated glomerular filtration rate (eGFR) in 261 patients. Proteins encoded by eGFR-associated transcripts were tested in urine for association with renal tissue injury and baseline eGFR. The ability to predict CKD progression, defined as the composite of ESKD or 40% reduction of baseline eGFR, was then determined in three independent CKD cohorts. A panel of intrarenal transcripts, including epidermal growth factor (*EGF*), a tubule-specific protein critical for cell differentiation and regeneration, predicted eGFR. The amount of EGF protein in urine (uEGF) showed significant correlation ( $P < 0.001$ ) with intrarenal *EGF* mRNA, interstitial fibrosis/tubular atrophy, eGFR, and rate of eGFR loss. Prediction of the composite renal end point by age, gender, eGFR, and albuminuria was significantly ( $P < 0.001$ ) improved by addition of uEGF, with an increase of the C-statistic from 0.75 to 0.87. Outcome predictions were replicated in two independent CKD cohorts. Our approach identified uEGF as an independent risk predictor of CKD progression. Addition of uEGF to standard clinical parameters improved the prediction of disease events in diverse CKD populations with a wide spectrum of causes and stages.

---

## INTRODUCTION

Chronic kidney disease (CKD) presents a major public health challenge and is an important cause of global mortality (1). Eight to 16% of the population worldwide is affected by CKD

(2–4) with increased risk for end-stage kidney disease (ESKD), cardiovascular disease (CVD), and death (5). A limited set of therapeutic approaches have been shown to slow progression of CKD and reduce its associated CVD risk (2). The main challenges to improve outcomes in patients with CKD are the current inability to identify patients at high risk of loss of renal function in early CKD and then to provide the patients targeted, effective treatments. Identifying pathobiologically important biomarkers that can address both of these challenges would have a major impact on outcomes of patients with progressive CKD.

Multiple lines of evidence have emerged over the last four decades, supporting the concept of a common pathway activated in CKD progression (6–8). This concept suggests that once kidney damage reaches a threshold, the subsequent progression is largely independent of the initial kidney injury and uses a shared molecular mechanism for disease progression(8). Combining information on genetic risk and intrarenal gene regulation allows identification of the functional underpinnings of CKD progression, not only in model systems but also in human disease, and strongly supports the notion of a common progression pathway (9).

Here, we define the shared molecular mechanism of CKD, using four cohorts from across three continents (Fig. 1). Our approach is rooted in the hypothesis that both clinical phenotypes, such as baseline estimated glomerular filtration rate (eGFR) and proteinuria, and renal tissue alterations, such as tubular atrophy (TA), interstitial fibrosis (IF), and glomerulosclerosis, seen in CKD are associated with, and a consequence of, the dynamic molecular mechanisms reflected in the transcriptional programs of the diseased kidney tissues. By implementing this strategy, we identified epidermal growth factor (EGF) as a biomarker, with intrarenal mRNA and urinary EGF protein concentrations tightly correlated with eGFR at the time of biopsy and with the change in renal function over time, independent of eGFR and proteinuria.

## RESULTS

### Identification and validation of intrarenal candidate markers for prediction of kidney function

As the first step in the sequential strategy for CKD predictor identification (Fig. 1), we quantified intrarenal transcripts and evaluated their association with clinical phenotypes. Affymetrix GeneChip–based steady state transcript expression was derived from the tubulointerstitial compartment of kidney biopsies of 164 patients from the European Renal cDNA Bank (ERCB) (Table 1) and correlated with eGFR at the time of biopsy. A total of 72 candidate genes passed the selection criteria of significant correlation [false discovery rate (FDR) < 0.01] with eGFR and differential expression compared to living donors. These transcripts are enriched in the key pathogenic pathways that have been implicated in driving CKD progression (table S1). The expression of these transcripts was measured in 55 CKD patients from the ERCB (Table 1) by quantitative real-time polymerase chain reaction (qRT-PCR) (table S2) and assessed for their ability to predict patients' eGFR at time of biopsy (Fig. 2A). A panel of six genes [nicotinamide *N*-methyltransferase(*NNMT*), *EGF*, thymosin  $\beta$  10 (*TMSB10*), tissue inhibitor of metalloproteinases (*TIMPI*), tubulin  $\alpha$  1a (*TUBA1A*), and annexin A1 (*ANXA1*)] showed the best predictive performance (correlation between predicted and observed eGFR values,  $r = 0.77$ ,  $P < 0.0001$ ; Fig. 2B). Candidate transcripts

were then assessed in a second, independent North American cohort, the Clinical Phenotyping Resource and Biobank Core (C-PROBE), consisting of 42 patients with CKD (Table 1), with the *EGF*, *NNMT*, and *TSMB10* panel showing the best eGFR prediction ( $r=0.62$ ,  $P < 0.0001$ ; fig. S1).

### Intrarenal transcript-based selection of urinary biomarkers

Next, noninvasive biomarker candidates were selected from the intrarenal transcript predictor panel. Markers for noninvasive evaluation were prioritized on the basis of (i) correlation with eGFR in multiple cohorts (table S3), (ii) kidney-specific transcript and protein expression, and (iii) compelling biology for a mechanistic role in CKD. These selection criteria converged on EGF as the top candidate. First, intrarenal *EGF* mRNA expression consistently showed significant correlation with eGFR (FDR  $\leq 0.05$ ;  $r=0.66$ ,  $0.81$ , and  $0.42$  in the discovery and the two validation cohorts, respectively) and contributed to the best performing gene expression predictor panel in both ERCB and C-PROBE (Fig. 2C). Second, the *EGF* transcript showed highly kidney-specific expression compared to other candidates in the human tissue panel (Fig. 2D and fig. S2). *EGF* mRNA was enriched in renal cortex and medulla compared to glomeruli, papillary tips, and renal pelvis (Fig. 2E) (10). In situ hybridization showed *EGF* mRNA expression in tubular cross sections in medulla and cortex (Fig. 2F, I and II), consistent with published immunohistochemical data demonstrating that EGF expression is restricted to the thick ascending limb of Henle and distal tubules (11). Third, in situ hybridization showed diminished *EGF* transcript expression in tubular cells of the cortex and medulla in CKD (Fig. 2F, III and IV versus I and II). Finally, EGF is a known mediator of differentiated epithelial cell function and regeneration. Both in vivo and in vitro studies have shown that EGF enhances renal tubule cell regeneration and repair and accelerates the recovery of renal function after injury (12–14) through its pro-proliferative and antiapoptotic action.

To test whether urinary EGF protein reflected intrarenal *EGF* transcript expression, we measured EGF protein concentrations and normalized them to urine creatinine concentration (uEGF/Cr) in patients with paired urine and kidney biopsy samples. uEGF/Cr correlated significantly with the intrarenal *EGFRNA* expression in 34 C-PROBE ( $r=0.70$ ,  $P < 0.0001$ ; Fig. 3A) and 85 Nephrotic Syndrome Study Network (NEPTUNE) patients ( $r=0.69$ ,  $P < 0.0001$ ; Fig. 3B; for cohort details, see table S4).

Next, we used urine samples of 349 C-PROBE patients (Table 2) to evaluate the association between uEGF/Cr and eGFR, independent of the 34 C-PROBE patients described above. uEGF/Cr correlated significantly with eGFR ( $r=0.77$ ,  $P < 0.0001$ ) at the time of sample procurement (Fig. 3C). The uEGF/Cr-eGFR correlation was significant in C-PROBE subgroups of patients with DN ( $n=70$ ) or with diabetes ( $n=135$ ) (fig. S3). The correlation was further replicated in 141 NEPTUNE patients with biopsy-proven FSGS, MCD, and MGN (Table 2) ( $r=0.82$ ,  $P < 0.0001$ , Fig. 3D) and in 452 patients from a prospective cohort study enrolling patients with biopsy-proven primary immunoglobulin A (IgA) nephropathy at Peking University First Hospital, China (PKU-IgAN) (Table 2) ( $r=0.78$ ,  $P < 0.0001$ ; Fig. 3E). In all three cohorts, uEGF/Cr showed only a weak negative correlation ( $r=-0.14$ ,  $-0.17$ , and  $-0.31$  in C-PROBE, NEPTUNE, and PKU-IgAN, respectively) with urinary

albumin-to-creatinine ratio (ACR, fig. S4), suggesting that uEGF/Cr could represent a different pathophysiologic mechanism than that detected by albuminuria.

### Association of EGF with renal biopsy outcome predictors

The correlation of uEGF/Cr with histological scores of IF and TA was measured to assess the association of this biomarker with tubulointerstitial lesions (15). IF/TA scores have been reported to be associated with cross-sectional eGFR (16) and can predict long-term kidney function (17). IF and TA scores were obtained from 102 NEPTUNE patients, and a significant inverse correlation was observed between uEGF/Cr and IF and TA [ $\rho = -0.75$  and  $-0.74$ , respectively ( $P < 0.001$ ); Fig. 4]. The correlations remained significant after adjusting for eGFR and ACR ( $\rho = -0.33$  for both,  $P < 0.001$ ).

### EGF and CKD progression

To determine whether EGF can predict CKD progression, we assessed the association between EGF and loss of renal function using eGFR slope as a continuous variable. Both intrarenal gene expression data and three or more eGFR values (over  $4.0 \pm 1.8$  years) were available for 29 C-PROBE patients, allowing correlation of renal transcript amounts with eGFR slope. Four hundred thirty-one transcripts were found to be correlated with eGFR slope ( $FDR < 0.05$ ). A functional network analysis predicted EGF to be the top upstream regulator of 37 eGFR slope-correlated transcripts ( $P = 2.09 \times 10^{-12}$ ; table S5 and fig. S5).

Both intrarenal *EGF* mRNA and urinary EGF protein/Cr showed a significant correlation with eGFR slope in the C-PROBE cohort [ $r = 0.58$ ,  $P < 0.001$ , for *EGF* tissue mRNA (Fig. 5A) and  $r = 0.64$ ,  $P < 0.001$ , for uEGF/Cr (fig. S6)]. We then used uEGF/Cr to predict eGFR slope via a multivariable regression model. Adjusted for age and gender, the eGFR slope predicted by uEGF/Cr showed a higher correlation with observed eGFR slope ( $r = 0.65$ ,  $P < 0.001$ ; Fig. 5B), compared to a model using ACR instead of uEGF/Cr ( $r = 0.29$ ,  $P < 0.001$ ; Fig. 5C).

The association of baseline uEGF with renal outcome (composite end point of ESKD or 40% reduction from baseline eGFR) was evaluated by classification power [comparing the area under receiver operating characteristic (ROC) curves of the corresponding models] and survival analysis (time-to-event analysis). ROC analysis showed an area under the curve (AUC) for the base model (including eGFR and ACR, adjusted for age and gender) of 0.84 (95% CI, 0.76 to 0.91) in C-PROBE. Addition of uEGF to the base model increased AUC to 0.90 (95% CI, 0.84 to 0.95) and resulted in a statistically improved model as evaluated by the likelihood test ( $P = 0.001$ ). The classification value of uEGF added to the base model was replicated in the NEPTUNE and PKU-IgAN cohorts (fig. S7).

In survival analysis, addition of uEGF/Cr [model 2 (M2)] to the base prediction model M1 (including eGFR and ACR, adjusted for age and gender) demonstrated a significant ( $P < 0.0001$ ) covariate effect on the likelihood ratio test, as well as an improvement in model fit, as  $R^2$  increased from 0.15 (M1) to 0.22 (M2), and the C-statistic increased from 0.75 (M1) to 0.87 (M2) (Table 3). Thus, addition of uEGF/Cr substantially improved the ability to predict renal outcomes when compared to ACR and baseline eGFR in combination. The improved prediction ability of the composite model M2 (M1 plus uEGF/Cr) was replicated

in both NEPTUNE and PKU-IgAN cohorts (Table 3). Multivariable-adjusted associations of uEGF/Cr with the hazard of progression to composite end point are shown in Fig. 6. Higher uEGF/Cr was associated with decreased risk of ESKD or 40% reduction of baseline eGFR in three independent cohorts: the estimated hazard ratios (HRs, 95% CI) were 0.27 (0.13 to 0.54), 0.29 (0.15 to 0.58), and 0.53 (0.37 to 0.69) for C-PROBE, NEPTUNE, and PKU-IgAN, respectively, adjusted for age, gender, eGFR, and ACR. Thus, with one unit increase of uEGF/Cr (in  $\log_2$  scale) and the other covariates withheld, the hazard of disease progression would decrease by 73, 71, and 47%, respectively, in these three cohorts. This is equivalent to a one-unit decrease of uEGF/Cr ( $\log_2$ ) being associated with an increased risk of CKD progression of 3.73 (1.85 to 7.69)-fold, 3.43 (1.72 to 6.67)-fold, and 1.96 (1.45 to 2.70)-fold in these three cohorts, respectively.

## DISCUSSION

By applying a renal biopsy transcriptome-driven sequential marker discovery approach, we demonstrated that (i) a set of intrarenal transcripts correlates with kidney function across a broad range of renal diseases; (ii) uEGF concentration showed a strong correlation with tissue *EGF* transcript expression and maintained the ability to predict eGFR; (iii) uEGF concentration was associated with eGFR slope; (iv) uEGF added predictive power to traditional clinical prognostic markers of CKD progression end points across cohorts.

The GFR-correlated transcripts are enriched in the key pathogenic pathways that have been implicated in driving CKD progression, ranging from chronic inflammation and extracellular matrix modulation to tubular cell differentiation (table S1). Because these pathways are involved in CKD progression independent of etiology, prognostic biomarkers that best represent them should be applicable to a wide spectrum of causes of CKD.

Estimates of GFR are the best overall indices of kidney function in health and disease, according to the National Kidney Foundation Kidney Disease Outcomes Quality Initiative (NKF KDOQI) Clinical Practice Guidelines (18). Baseline eGFR and proteinuria are established predictors of CKD progression and capture primary elements of glomerular function, with eGFR reflecting glomerular hemodynamics and ultrafiltration capacity and proteinuria reflecting changes in glomerular permselectivity and its impact on CKD progression. Multiple potential biomarkers for renal inflammation and fibrosis, a dominant signature in the CKD pathway analysis presented here, have been shown to be associated with kidney impairment. These include tumor necrosis factor receptors 1 and 2 (TNFR1 and TNFR2) (19, 20), monocyte chemoattractant protein 1 (MCP1) (21), and matrix metalloproteinases (22). Because these molecules are found in plasma as well, their urinary concentration is a function of both glomerular filtration and intrarenal production in proteinuric disease, potentially affecting their utility as urinary biomarkers of tubulointerstitial disease. Biomarkers of tubular function and reserve are not available for clinical use in CKD. As described in this study, uEGF has the potential to serve this role, given its highly restricted intrarenal expression and strong correlation between renal transcripts and urinary protein concentration.

EGF function provides an independent rationale for its evaluation as a CKD progression biomarker. EGF has been reported to be responsible for modulation of tissue response to injury in the kidney after tubulointerstitial damage (23, 24). Exogenous EGF enhanced the tubular cell repair process and accelerated tubular regeneration and recovery of renal function in an animal model of acute renal injury (13), suggesting that uEGF concentrations are not only associated with concurrent renal function but also regulate signaling pathways of tubular recovery after injury. This offers the intriguing possibility that uEGF might serve as a surrogate marker for regenerative functional reserve of the renal tubules, reflecting their ability to respond to future acute or chronic injury, which are key factors in CKD progression. This concept is further supported by the unbiased functional analysis of gene expression data sets, which predicted that EGF was as an upstream regulator of eGFR slope-associated genes. uEGF inversely correlated with IF and TA, established independent morphometric predictors of renal outcomes (25–27). These associations not only provide strong support for our hypothesis but also shed light on the pathogenic role of EGF in CKD progression.

Reduced concentrations of EGF protein in the urine have been previously observed in DN, IgAN, adult polycystic kidney disease, and children with chronic renal failure (28–31). For example, Torres *et al.* showed that in a single center cohort of 132 IgAN patients, the ratio of EGF to MCP1 in the urine was negatively associated with the composite outcome of ESKD and/or doubling of the baseline serum creatinine (23), but it was not compared to conventional predictors of ESKD in predicting renal survival.

A critical factor for the uEGF-eGFR correlation is the high degree of tissue specificity, making uEGF robust to extrarenal events that may affect the accuracy of other nonspecific biomarkers. EGF is absent or only minimally detectable in plasma (32). Urinary EGF is reported to be derived from the ascending portion of Henle's loop and the distal convoluted tubule (33, 34), and in situ hybridization confirmed local synthesis of *EGF* in tubular compartments. The tissue specificity might be responsible for the robustness of the correlation of uEGF with renal function and renal functional decline across a wide spectrum of systemic and primary renal diseases in diverse environments.

Our study has several limitations. The discovery approach used in this study critically depends on renal biopsy tissue, raising concerns about the ability to extrapolate the renal tissue findings to CKD patients not undergoing biopsy. However, uEGF correlated as tightly with eGFR across all C-PROBE patients as it did in the biopsy patients from the same cohort ( $r = 0.77$  for both,  $P < 0.001$ ), supporting the utility of uEGF across CKD. A second potential limitation is that the biopsy cohorts had a relatively small proportion of patients with DN (19 of 261 across the three cohorts). To determine whether the results of our study were generalizable to CKD patients with diabetes, who form the majority of CKD in the United States, we specifically investigated patients with DN. In the ERCB discovery cohort, we saw a correlation between intrarenal *EGF* mRNA expression and eGFR in the 17 DN patients. In addition, of the 349 C-PROBE patients analyzed for noninvasive marker studies, about 39% had diabetes and 20% of patients were diagnosed with DN. Among these patients, we found correlations of uEGF with eGFR and eGFR slope. These data suggest

that the results of our study may be generalizable to CKD patients with diabetes, but further validation in cohorts of patients with DN will be required.

To assess association with and prediction of CKD progression by uEGF, we used two principal approaches. First, eGFR slope was derived by a mixed-effect model to capture the trajectory of the disease over time. eGFR slope estimates assume linearity of progression and are vulnerable to acute kidney injury and treatment effects, particularly in patients with glomerular disease at the time of presentation (35, 36). Given these constraints, we focused the assessment of EGF and its association with eGFR slope in the chronic renal disease cohort (C-PROBE). The second approach used progression to ESKD or 40% eGFR reduction as composite end points, reflecting current concepts of clinical trial end points in CKD (37). The three cohorts used in this study had relatively short prospective follow-up time periods (1.8 to 3.6 years/mean), enriching this analysis for patients with rapid progression of CKD. The ability of uEGF to identify CKD patients with rapid progression might be particularly useful to enrich clinical trials with patients who are more likely to reach end points during the limited trial observation period. To assess the utility of uEGF for informing CKD management in routine clinical settings, further comprehensive studies with structured long-term follow-up will be needed.

In summary, our study used an unbiased screening strategy to identify uEGF as a predictor of CKD progression in patients with glomerular disease, in three independent cohorts with diverse ethnic and geographic backgrounds. uEGF was linked to tubular differentiation and regeneration, a mechanism essential to retain renal function in CKD, but not well reflected by the conventional predictors (proteinuria or baseline GFR). Addition of uEGF into a CKD biomarker panel will likely improve risk stratification of CKD patients and thereby enhance the ability to target clinical care and limited health care resources to those in most need, as well as to optimize clinical trial design. In addition, the importance of EGF-dependent mechanisms in CKD progression suggests potential therapeutic targets for patients with impaired regenerative tubular functional reserve.

## MATERIALS AND METHODS

### Study design

The objective of this study was to identify and validate noninvasive markers for CKD progression. We applied a transcriptome-driven sequential strategy. For tissue transcriptome analysis, gene expression profiles from 164 consecutive biopsies from the ERCB (38) were used as a discovery cohort. Subsequent profiles of 55 patients from ERCB served as the first validation cohort, and 42 patients from C-PROBE whose samples were available for tissue gene expression analysis served as the second validation cohort (Table 1). Thirty-two living donor transplant biopsies obtained at the time of transplantation were used as controls. Urinary biomarker analysis was performed on all samples available at the time of the study from baseline visits from three independent cohort studies (C-PROBE, NEPTUNE, and PKU-IgAN), representing 349, 141, and 452 CKD patients, respectively (Table 2). In the above cohorts, only patients aged 18 years or older were included in this study. All bio-specimens were procured after informed consent and with approval of the local ethics



committees. The details of cohort designs and sample procurement are provided in Supplementary Materials and Methods.

### Gene expression analysis

Transcriptome analysis was performed on microdissected tubulointerstitial components of human renal biopsies prospectively procured for molecular analysis, using Affymetrix GeneChip and TaqMan Low Density Arrays as previously published (39). Data processing details are provided in Supplementary Materials and Methods. Normalized expression data were  $\log_2$ -transformed and batch-corrected. FDR was applied to account for multiple testing. The Cel files are available at Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) under reference nos. GSE32591, GSE37455 (5), GSE35488 (6), GSE47185 (7), and GSE69438.

### In situ hybridization

In situ hybridization was performed by branched-DNA signal amplification with QuantiGene ViewRNA ISH Tissue Assay Kit (Affymetrix/Panomics Solutions) on formalin-fixed paraffin-embedded human kidney tissue using the specific probe set for human *EGF* mRNA. Further details are provided in Supplementary Materials and Methods.

### Measurements of urinary EGF

uEGF concentration was measured in spot urine samples with Human EGF Immunoassay Quantikine ELISA (R&D Systems). Samples (1:150 dilution) and standards were run in duplicate, absorbance was measured with a VersaMax ELISA plate reader, and results were calculated with SoftMax Pro (Molecular Devices). Detailed assay validation information is provided in Supplementary Materials and Methods. uEGF was normalized to urine creatinine concentration and is referred to as uEGF/Cr.

### GFR estimates

GFR was estimated by the four-variable Modification of Diet in Renal Disease (MDRD) study equation (40) in ERCB, C-PROBE, and NEPTUNE. Both MDRD- and CKD-EPI GFR estimates were determined for C-PROBE subjects, and their  $\log_2$ -transformed values showed a high correlation ( $r = 0.993$ ), indicating limited impact of the GFR estimation on the data presented. In PKU-IgAN cohort, GFR was estimated by a modified MDRD equation (c-aGFR<sub>4</sub>) specifically adapted for Chinese CKD patients (41, 42).  $\log_2$ -transformed c-aGFR<sub>4</sub> and MDRD-eGFR were highly correlated ( $r = 0.996$ ).

### Outcome measure

Two outcomes were used to evaluate CKD progression: (i) eGFR slope and (ii) a composite end point of ESKD or 40% reduction of baseline eGFR (37). eGFR slope was calculated for patients with at least three eGFR records over a minimum follow-up of 1.5 years, including eGFRs recorded before the baseline visit, if available. To account for irregularly spaced creatinine measurements, a linear mixed-effects model (43) was applied to calculate subject-specific eGFR slope. More details are provided in Supplementary Materials and Methods.

## Evaluation of tubulointerstitial damage using IF/TA

Histopathology was assessed in NEPTUNE biopsy cohort with digital images obtained from paraffin-embedded kidney tissue stained with hematoxylin and eosin, PAS, silver-based, and trichrome stains (44). Whole slide images of glass slides from 102 cases stored in the NEPTUNE digital pathology repository were assessed for percentage of cortex involved by IF/TA by six pathologists. The percentage of cortex involved by IF/TA was determined in each individual stain and averaged for an overall % value. More details are provided in Supplementary Materials and Methods.

## Statistical analysis

The correlations of  $\log_2$ -transformed transcript expression values with both baseline eGFR and eGFR slope were calculated using Pearson correlation. FDR was applied to account for multiple testing. The selection criteria for candidate transcript markers included (i) significant correlation of transcript levels with baseline eGFR ( $|r| > 0.4$ ,  $FDR < 0.01$ ) and (ii) significant differential expression of transcripts compared to living donor kidney biopsy controls ( $FDR \leq 0.05$ , fold change  $\geq 1.8$  or  $\leq 0.56$ ) in at least three of eight kidney diseases studied in the discovery cohort.

Ridge regression was performed on the discovery cohort to combine transcripts showing associations with baseline eGFR into a multimarker signature for better prediction. In the analysis, the transcripts were ranked by their strength of marginal association with  $\log_2$  eGFR. The top-ranked  $k$  genes were first standardized, then combined as linear predictors with weights of correlation coefficients, and then carried over into two validation cohorts.

Linear regression was applied to investigate the association between eGFR slope and the noninvasive protein marker uEGF, adjusted for age and gender. The overall quality of fit was evaluated by an  $F$  test at the 0.05 level of significance.

To evaluate the classification power of the candidate marker, we used logistic regression and AUC values from the ROC curves to compare nested models ([www.inside-r.org/packages/cran/epicalc/docs/lrtest](http://www.inside-r.org/packages/cran/epicalc/docs/lrtest)). To assess the effect of the candidate marker on the hazard of reaching composite end points, Cox proportional hazards models were fit to the available prospective data. We examined nested models, starting with the base model (M1) including baseline eGFR, ACR, age, and gender. Samples with missing data were excluded from the analysis. In the subsequent model (M2), uEGF/Cr was added, and the goodness of fit and improved prediction ability of the additional parameters were assessed by likelihood ratio tests, C-statistics, and AIC. All statistical analysis was performed using R ([www.R-project.org](http://www.R-project.org)) and SAS software (version 9.3).

Our approach to minimize overfitting included rigorous correction for multiple testing of the initial gene expression analysis, sequential validation of intrarenal markers in two test cohorts with two different mRNA quantification technologies (microarray and qRT-PCR), and replication of the uEGF analysis (association with eGFR and prediction of the composite renal end point) in two independent cohorts. Furthermore, we limited the number of covariates in Cox models with respect to observations. All patients who passed inclusion criteria and had samples available at the time of the study were included in the analysis.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We acknowledge all participating centers of the ERCB–Kröner-Fresenius Biopsy Bank (ERCB-KFB), the C-PROBE, the NEPTUNE, the PKU-IgAN study cohort, and their participants for their cooperation. We thank the support of George M. O'Brien Michigan Kidney Translational Core Center and the Michigan Diabetes Research Center at the University of Michigan. **Funding:** This study was supported by the Else Kröner-Fresenius Foundation (for ERCB); by the European Consortium for High-Throughput Research in Rare Kidney Diseases (EUREnOmics; European Union FP 7:305608); by NIH (R01DK079912, P30DK081943, DK083912, P30DK020572, and UL1RR000433); by Office of Rare Diseases Research, National Center for Advancing Translational Sciences, National Institute of Diabetes and Digestive and Kidney Diseases, University of Michigan and NephCure Kidney International (U54DK083912); and by the University of Michigan Health System and Peking University Health Sciences Center Joint Institute for Translational and Clinical Research. Analysis of urine samples of C-PROBE patients was supported by Hoffman–La Roche.

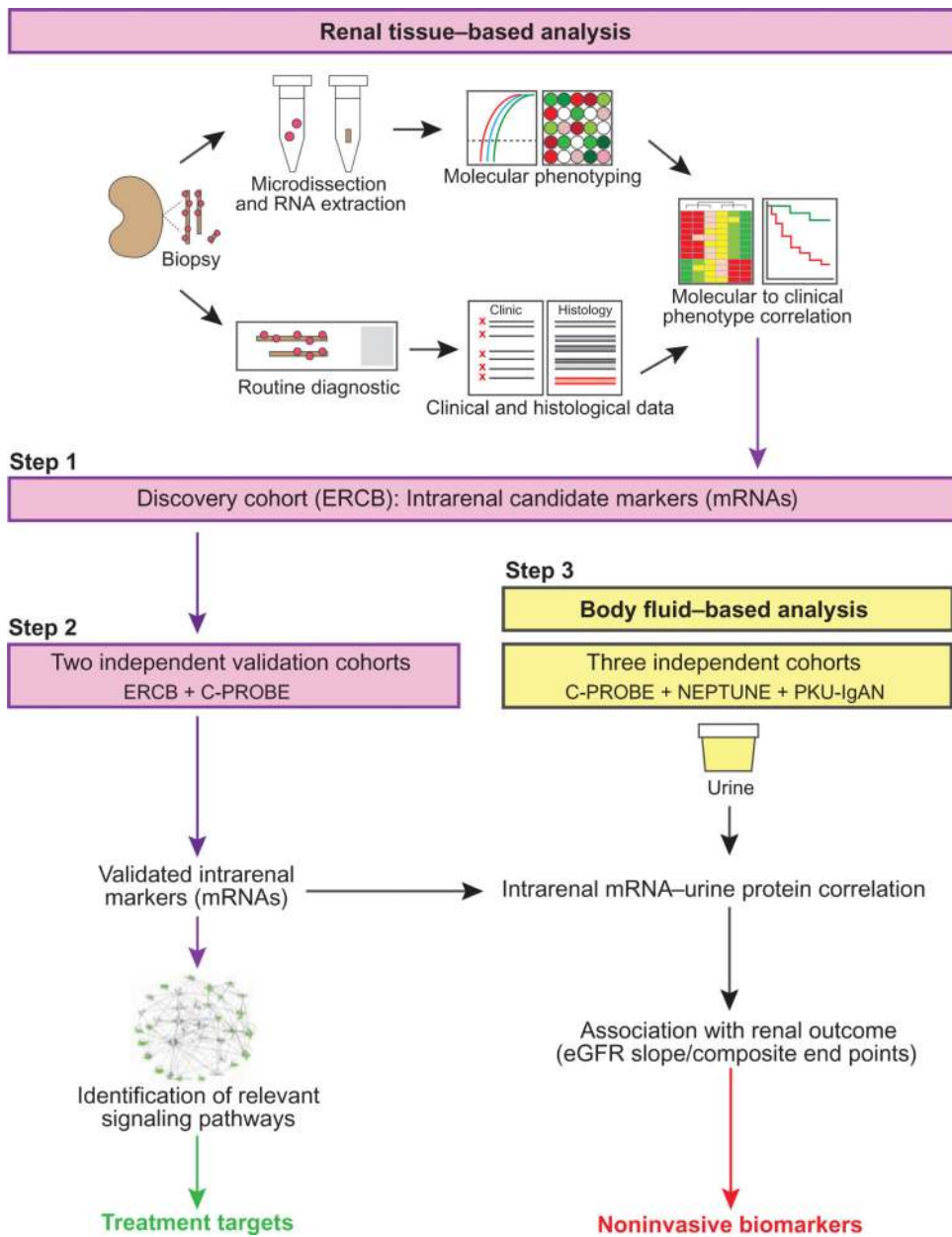
## REFERENCES AND NOTES

- Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Barker-Collo S, Bartels DH, Bell ML, Benjamin EJ, Bennett D, Bhalla K, Bikbov B, Bin Abdulhak A, Birbeck G, Blyth F, Bolliger I, Boufous S, Bucello C, Burch M, Burney P, Carapetis J, Chen H, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahodwala N, De Leo D, Degenhardt L, Delossantos A, Denenberg J, Des Jarlais DC, Dharmaratne SD, Dorsey ER, Driscoll T, Duber H, Ebel B, Erwin PJ, Espindola P, Ezzati M, Feigin V, Flaxman AD, Forouzanfar MH, Fowkes FG, Franklin R, Fransen M, Freeman MK, Gabriel SE, Gakidou E, Gaspari F, Gillum RF, Gonzalez-Medina D, Halasa YA, Haring D, Harrison JE, Havmoeller R, Hay RJ, Hoen B, Hotez PJ, Hoy D, Jacobsen KH, James SL, Jasrasaria R, Jayaraman S, Johns N, Karthikeyan G, Kassebaum N, Keren A, Khoo J-P, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lipnick M, Lipshultz SE, Ohno SL, Mabweijano J, MacIntyre MF, Mallinger L, March L, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGrath J, Mensah GA, Merriman TR, Michaud C, Miller M, Miller TR, Mock C, Mocumbi AO, Mokdad AA, Moran A, Mulholland K, Nair MN, Naldi L, Narayan KM, Nasseri K, Norman P, O'Donnell M, Omer SB, Ortblad K, Osborne R, Ozgediz D, Pahari B, Pandian JD, Rivero AP, Padilla RP, Perez-Ruiz F, Perico N, Phillips D, Pierce K, Pope CA III, Porrini E, Pourmalek F, Raju M, Ranganathan D, Rehm JT, Rein DB, Remuzzi G, Rivara FP, Roberts T, Rodriguez De León F, Rosenfeld LC, Rushton L, Sacco RL, Salomon JA, Sampson U, Sanman E, Schwebel DC, Segui-Gomez M, Shepard DS, Singh D, Singleton J, Sliwa K, Smith E, Steer A, Taylor JA, Thomas B, Tleyjeh IM, Towbin JA, Truelsen T, Undurraga EA, Venketasubramanian N, Vijayakumar L, Vos T, Wagner GR, Wang M, Wang W, Watt K, Weinstock MA, Weintraub R, Wilkinson JD, Woolf AD, Wulf S, Yeh P-H, Yip P, Zabetian A, Zheng Z-J, Lopez AD, Murray CJ, AlMazroa MA, Memish ZA. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012; 380:2095–2128. [PubMed: 23245604]
- Levey AS, Atkins R, Coresh J, Cohen EP, Collins AJ, Eckardt KU, Nahas ME, Jaber BL, Jadoul M, Levin A, Powe NR, Rossert J, Wheeler DC, Lameire N, Eknoyan G. Chronic kidney disease as a global public health problem: Approaches and initiatives—A position statement from Kidney Disease Improving Global Outcomes. *Kidney Int*. 2007; 72:247–259. [PubMed: 17568785]
- Coresh J, Byrd-Holt D, Astor BC, Briggs JP, Eggers PW, Lacher DA, Hostetter TH. Chronic kidney disease awareness, prevalence, and trends among U.S. adults, 1999 to 2000. *J. Am. Soc. Nephrol*. 2005; 16:180–188. [PubMed: 15563563]
- Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, Saran R, Wang AY-M, Yang C-W. Chronic kidney disease: Global dimension and perspectives. *Lancet*. 2013; 382:260–272. [PubMed: 23727169]

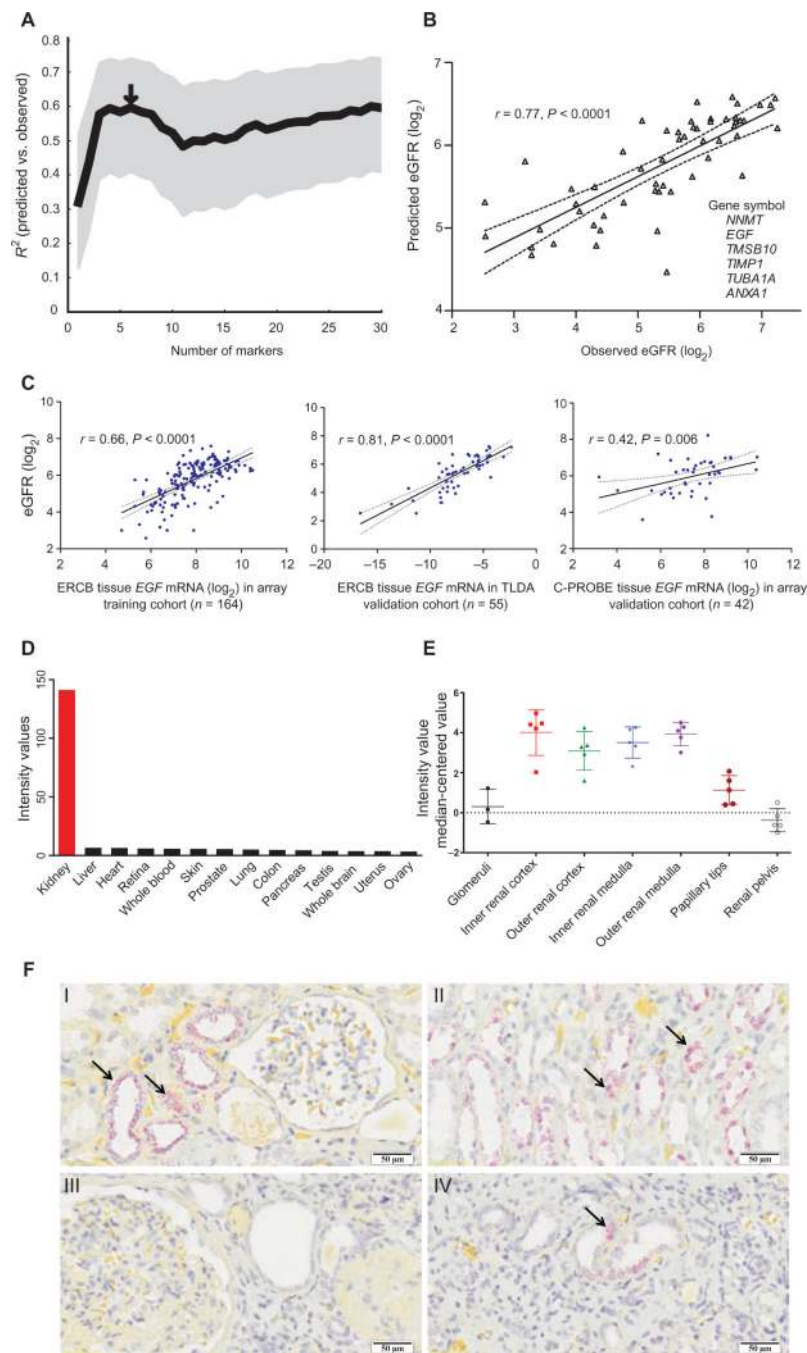
5. Whitman IR, Feldman HI, Deo R. CKD and sudden cardiac death: Epidemiology, mechanisms, and therapeutic approaches. *J. Am. Soc. Nephrol.* 2012; 23:1929–1939. [PubMed: 23100219]
6. Brenner BM, Meyer TW, Hostetter TH. Dietary protein intake and the progressive nature of kidney disease: The role of hemodynamically mediated glomerular injury in the pathogenesis of progressive glomerular sclerosis in aging, renal ablation, and intrinsic renal disease. *N. Engl. J. Med.* 1982; 307:652–659. [PubMed: 7050706]
7. Fogo AB. Mechanisms of progression of chronic kidney disease. *Pediatr. Nephrol.* 2007; 22:2011–2022. [PubMed: 17647026]
8. Hodgkins KS, Schnaper HW. Tubulointerstitial injury and the progression of chronic kidney disease. *Pediatr. Nephrol.* 2012; 27:901–909. [PubMed: 21947270]
9. Martini S, Nair V, Keller BJ, Eichinger F, Hawkins JJ, Randolph A, Böger CA, Gadegbeku CA, Fox CS, Cohen CD, Kretzler M. European Renal cDNA Bank; C-PROBE Cohort; CKDGen Consortium, Integrative biology identifies shared transcriptional networks in CKD. *J. Am. Soc. Nephrol.* 2014; 25:2559–2572. [PubMed: 24925724]
10. Higgins JPT, Wang L, Kambham N, Montgomery K, Mason V, Vogelmann SU, Lemley KV, Brown PO, Brooks JD, van de Rijn M. Gene expression in the normal adult human kidney assessed by complementary DNA microarray. *Mol. Biol. Cell.* 2004; 15:649–656. [PubMed: 14657249]
11. Salido EC, Lakshmanan J, Fisher DA, Shapiro LJ, Barajas L. Expression of epidermal growth factor in the rat kidney. An immunocytochemical and in situ hybridization study. *Histochemistry.* 1991; 96:65–72. [PubMed: 1938482]
12. Humes HD, Cieslinski DA, Coimbra TM, Messana JM, Galvao C. Epidermal growth factor enhances renal tubule cell regeneration and repair and accelerates the recovery of renal function in postischemic acute renal failure. *J. Clin. Invest.* 1989; 84:1757–1761. [PubMed: 2592559]
13. Lechner J, Malloth NA, Jennings P, Heckl D, Pfaller W, Seppi T. Opposing roles of EGF in IFN- $\alpha$ -induced epithelial barrier destabilization and tissue repair. *Am. J. Physiol. Cell Physiol.* 2007; 293:C1843–C1850. [PubMed: 17913840]
14. Safirstein R, Zelent AZ, Price PM. Reduced renal prepro-epidermal growth factor mRNA and decreased EGF excretion in ARF. *Kidney Int.* 1989; 36:810–815. [PubMed: 2615190]
15. Farris AB, Colvin RB. Renal interstitial fibrosis: Mechanisms and evaluation. *Curr. Opin. Nephrol. Hypertens.* 2012; 21:289–300. [PubMed: 22449945]
16. Diaz Encarnacion MM, Griffin MD, Slezak JM, Bergstralh EJ, Stegall MD, Velosa JA, Grande JP. Correlation of quantitative digital image analysis with the glomerular filtration rate in chronic allograft nephropathy. *Am. J. Transplant.* 2004; 4:248–256. [PubMed: 14974947]
17. Nicholson ML, McCulloch TA, Harper SJ, Wheatley TJ, Edwards CM, Feehally J, Furness PN. Early measurement of interstitial fibrosis predicts long-term renal function and graft survival in renal transplantation. *Br. J. Surg.* 1996; 83:1082–1085. [PubMed: 8869307]
18. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: Evaluation, classification, and stratification. *Am. J. Kidney Dis.* 2002; 39(Suppl. 1):S1–S266. [PubMed: 11904577]
19. Kurashina T, Nagasaka S, Watanabe N, Yabe D, Sugi N, Nin K, Hosokawa M, Nomura Y, Fukushima M, Nakai Y, Nishimura F, Taniguchi A. Circulating TNF receptor 2 is closely associated with the kidney function in non-diabetic Japanese subjects. *J. Atheroscler. Thromb.* 2014; 21:730–738. [PubMed: 24717758]
20. Gohda T, Niewczas MA, Ficociello LH, Walker WH, Skupien J, Rosetti F, Cullere X, Johnson AC, Crabtree G, Smiles AM, Mayadas TN, Warram JH, Krolewski AS. Circulating TNF receptors 1 and 2 predict stage 3 CKD in type 1 diabetes. *J. Am. Soc. Nephrol.* 2012; 23:516–524. [PubMed: 22266664]
21. Titan SM, Vieira JM Jr, Dominguez WV, Moreira SRS, Pereira AB, Barros RT, Zatz R. Urinary MCP-1 and RBP: Independent predictors of renal outcome in macroalbuminuric diabetic nephropathy. *J. Diabetes Complications.* 2012; 26:546–553. [PubMed: 22981148]
22. He W, Tan RJ, Li Y, Wang D, Nie J, Hou FF, Liu Y. Matrix metalloproteinase-7 as a surrogate marker predicts renal Wnt/b-catenin activity in CKD. *J. Am. Soc. Nephrol.* 2012; 23:294–304. [PubMed: 22095947]

23. Torres DD, Rossini M, Manno C, Mattace-Raso F, D'Altri C, Ranieri E, Pontrelli P, Grandaliano G, Gesualdo L, Schena FP. The ratio of epidermal growth factor to monocyte chemotactic peptide-1 in the urine predicts renal prognosis in IgA nephropathy. *Kidney Int.* 2008; 73:327–333. [PubMed: 17943082]
24. Grandaliano G, Gesualdo L, Bartoli F, Ranieri E, Monno R, Leggio A, Paradies G, Caldaruolo E, Infante B, Schena FP. MCP-1 and EGF renal expression and urine excretion in human congenital obstructive nephropathy. *Kidney Int.* 2000; 58:182–192. [PubMed: 10886563]
25. Serón D, Moreso F, Bover J, Condom E, Gil-Vernet S, Cañas C, Fulladosa X, Torras J, Carrera M, Grinyó JM, Alsina J. Early protocol renal allograft biopsies and graft outcome. *Kidney Int.* 1997; 51:310–316. [PubMed: 8995748]
26. Serón D. Risk factors associated with the deterioration of renal function: The role of protocol biopsies. *Prilozi.* 2007; 28:291–302. [PubMed: 17932475]
27. An Y, Xu F, Le W, Ge Y, Zhou M, Chen H, Zeng C, Zhang H, Liu Z. Renal histologic changes and the outcome in patients with diabetic nephropathy. *Nephrol. Dial. Transplant.* 2015; 30:257–266. [PubMed: 25063425]
28. Tsau Y-K, Chen C-H. Urinary epidermal growth factor excretion in children with chronic renal failure. *Am. J. Nephrol.* 1999; 19:400–404. [PubMed: 10393378]
29. Mathiesen ER, Nexø E, Hommel E, Parving H-H. Reduced urinary excretion of epidermal growth factor in incipient and overt diabetic nephropathy. *Diabet. Med.* 1989; 6:121–126. [PubMed: 2522852]
30. Ranieri E, Gesualdo L, Petrarulo F, Schena FP. Urinary IL-6/EGF ratio: A useful prognostic marker for the progression of renal damage in IgA nephropathy. *Kidney Int.* 1996; 50:1990–2001. [PubMed: 8943482]
31. Weinstein T, Hwang D, Lev-Ran A, Ori Y, Korzets A, Levi J. Excretion of epidermal growth factor in human adult polycystic kidney disease. *Isr. J. Med. Sci.* 1997; 33:641–642. [PubMed: 9397135]
32. Mattila AL, Viinikka L, Saario I, Perheentupa J. Human epidermal growth factor: Renal production and absence from plasma. *Regul. Pept.* 1988; 23:89–93. [PubMed: 3070644]
33. Jørgensen PE, Hilchey SD, Nexø E, Poulsen SS, Quilley CP. Urinary epidermal growth factor is excreted from the rat isolated perfused kidney in the absence of plasma. *J. Endocrinol.* 1993; 139:227–234. [PubMed: 7508486]
34. Melenhorst WBH, Mulder GM, Xi Q, Hoenderop JG, Kimura K, Eguchi S, van Goor H. Epidermal growth factor receptor signaling in the kidney: Key roles in physiology and disease. *Hypertension.* 2008; 52:987–993. [PubMed: 18981331]
35. Hunsicker LG, Adler S, Caggiula A, England BK, Greene T, Kusek JW, Rogers NL, Teschan PE. Predictors of the progression of renal disease in the Modification of Diet in Renal Disease Study. *Kidney Int.* 1997; 51:1908–1919. [PubMed: 9186882]
36. Li L, Astor BC, Lewis J, Hu B, Appel LJ, Lipkowitz MS, Toto RD, Wang X, Wright JT Jr, Greene TH. Longitudinal progression trajectory of GFR among patients with CKD. *Am. J. Kidney Dis.* 2012; 59:504–512. [PubMed: 22284441]
37. Coresh J, Turin TC, Matsushita K, Sang Y, Ballew SH, Appel LJ, Arima H, Chadban SJ, Cirillo M, Djurdjev O, Green JA, Heine GH, Inker LA, Irie F, Ishani A, Ix JH, Kovesdy CP, Marks A, Ohkubo T, Shalev V, Shankar A, Wen CP, de Jong PE, Iseki K, Stengel B, Gansevoort RT, Levey AS. CKD Prognosis Consortium, Decline in estimated glomerular filtration rate and subsequent risk of end-stage renal disease and mortality. *JAMA.* 2014; 311:2518–2531. [PubMed: 24892770]
38. Yasuda Y, Cohen CD, Henger A, Kretzler M. European Renal cDNA Bank (ERCB) Consortium, Gene expression profiling analysis in nephrology: Towards molecular definition of renal disease. *Clin. Exp. Nephrol.* 2006; 10:91–98. [PubMed: 16791393]
39. Schmid H, Boucherot A, Yasuda Y, Henger A, Brunner B, Eichinger F, Nitsche A, Kiss E, Bleich M, Gröne H-J, Nelson PJ, Schlöndorff D, Cohen CD, Kretzler M. European Renal cDNA Bank (ERCB) Consortium, Modular activation of nuclear factor-kB transcriptional programs in human diabetic nephropathy. *Diabetes.* 2006; 55:2993–3003. [PubMed: 17065335]
40. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, Kusek JW, Van Lente F. Chronic Kidney Disease Epidemiology Collaboration, Using standardized serum creatinine values

- in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann. Intern. Med.* 2006; 145:247–254. [PubMed: 16908915]
41. Ma Y-C, Zuo L, Chen J-H, Luo Q, Yu X-Q, Li Y, Xu J-S, Huang S-M, Wang L-N, Huang W, Wang M, Xu G-B, Wang H-Y. Modified glomerular filtration rate estimating equation for Chinese patients with chronic kidney disease. *J. Am. Soc. Nephrol.* 2006; 17:2937–2944. [PubMed: 16988059]
  42. Zhang L, Wang F, Wang L, Wang W, Liu B, Liu J, Chen M, He Q, Liao Y, Yu X, Chen N, Zhang J.-e. Hu Z, Liu F, Hong D, Ma L, Liu H, Zhou X, Chen J, Pan L, Chen W, Wang W, Li X, Wang H. Prevalence of chronic kidney disease in China: A cross-sectional survey. *Lancet.* 2012; 379:815–822. [PubMed: 22386035]
  43. Ficociello LH, Rosolowsky ET, Niewczas MA, Maselli NJ, Weinberg JM, Aschengrau A, Eckfeldt JH, Stanton RC, Galecki AT, Doria A, Warram JH, Krolewski AS. High-normal serum uric acid increases risk of early progressive renal function loss in type 1 diabetes: Results of a 6-year follow-up. *Diabetes Care.* 2010; 33:1337–1343. [PubMed: 20332356]
  44. Barisoni L, Nast CC, Jennette JC, Hodgins JB, Herzenberg AM, Lemley KV, Conway CM, Kopp JB, Kretzler M, Lienczewski C, Avila-Casado C, Bagnasco S, Sethi S, Tomaszewski J, Gasim AH, Hewitt SM. Digital pathology evaluation in the multicenter Nephrotic Syndrome Study Network (NEPTUNE). *Clin. J. Am. Soc. Nephrol.* 2013; 8:1449–1459. [PubMed: 23393107]
  45. Gadegbeku CA, Gipson DS, Holzman LB, Ojo AO, Song PXX, Barisoni L, Sampson MG, Kopp JB, Lemley KV, Nelson PJ, Lienczewski CC, Adler SG, Appel GB, Cattran DC, Choi MJ, Contreras G, Dell KM, Fervenza FC, Gibson KL, Greenbaum LA, Hernandez JD, Hewitt SM, Hingorani SR, Hladunewich M, Hogan MC, Hogan SL, Kaskel FJ, Lieske JC, Meyers KEC, Nachman PH, Nast CC, Neu AM, Reich HN, Sedor JR, Sethna CB, Trachtman H, Tuttle KR, Zhdanova O, Zilleruelo GE, Kretzler M. Design of the Nephrotic Syndrome Study Network (NEPTUNE) to evaluate primary glomerular nephropathy by a multi-disciplinary approach. *Kidney Int.* 2013; 83:749–756. [PubMed: 23325076]
  46. Cohen CD, Frach K, Schlöndorff D, Kretzler M. Quantitative gene expression analysis in renal biopsies: A novel protocol for a high-throughput multicenter application. *Kidney Int.* 2002; 61:133–140. [PubMed: 11786093]
  47. Berthier CC, Bethunaikkan R, Gonzalez-Rivera T, Nair V, Ramanujam M, Zhang W, Bottinger EP, Segerer S, Lindenmeyer M, Cohen CD, Davidson A, Kretzler M. Cross-species transcriptional network analysis defines shared inflammatory responses in murine and human lupus nephritis. *J. Immunol.* 2012; 189:988–1001. [PubMed: 22723521]
  48. Reich HN, Tritchler D, Cattran DC, Herzenberg AM, Eichinger F, Boucherot A, Henger A, Berthier CC, Nair V, Cohen CD, Scholey JW, Kretzler M. A molecular signature of protein-uria in glomerulonephritis. *PLOS One.* 2010; 5:e13451. [PubMed: 20976140]
  49. Ju W, Greene CS, Eichinger F, Nair V, Hodgins JB, Bitzer M, Lee Y.-s. Zhu Q, Kehata M, Li M, Jiang S, Rastaldi MP, Cohen CD, Troyanskaya OG, Kretzler M. Defining cell-type specificity at the transcriptional level in human disease. *Genome Res.* 2013; 23:1862–1873. [PubMed: 23950145]
  50. Frericks M, Esser C. A toolbox of novel murine house-keeping genes identified by meta-analysis of large scale gene expression profiles. *Biochim. Biophys. Acta.* 2008; 1779:830–837. [PubMed: 18790095]
  51. Matsushita K, Mahmoodi BK, Woodward M, Emberson JR, Jafar TH, Jee SH, Polkinghorne KR, Shankar A, Smith DH, Tonelli M, Warnock DG, Wen C-P, Coresh J, Gansevoort RT, Hemmelgarn BR, Levey AS. Chronic Kidney Disease Prognosis Consortium, Comparison of risk prediction using the CKD-EPI equation and the MDRD study equation for estimated glomerular filtration rate. *JAMA.* 2012; 307:1941–1951. [PubMed: 22570462]
  52. DeSilva B, Smith W, Weiner R, Kelley M, Smolec J, Lee B, Khan M, Tacey R, Hill H, Celniker A. Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharm. Res.* 2003; 20:1885–1900. [PubMed: 14661937]
  53. Zhang P, Song PX-K, Qu A, Greene T. Efficient estimation for patient-specific rates of disease progression using nonnormal linear mixed models. *Biometrics.* 2008; 64:29–38. [PubMed: 17501938]



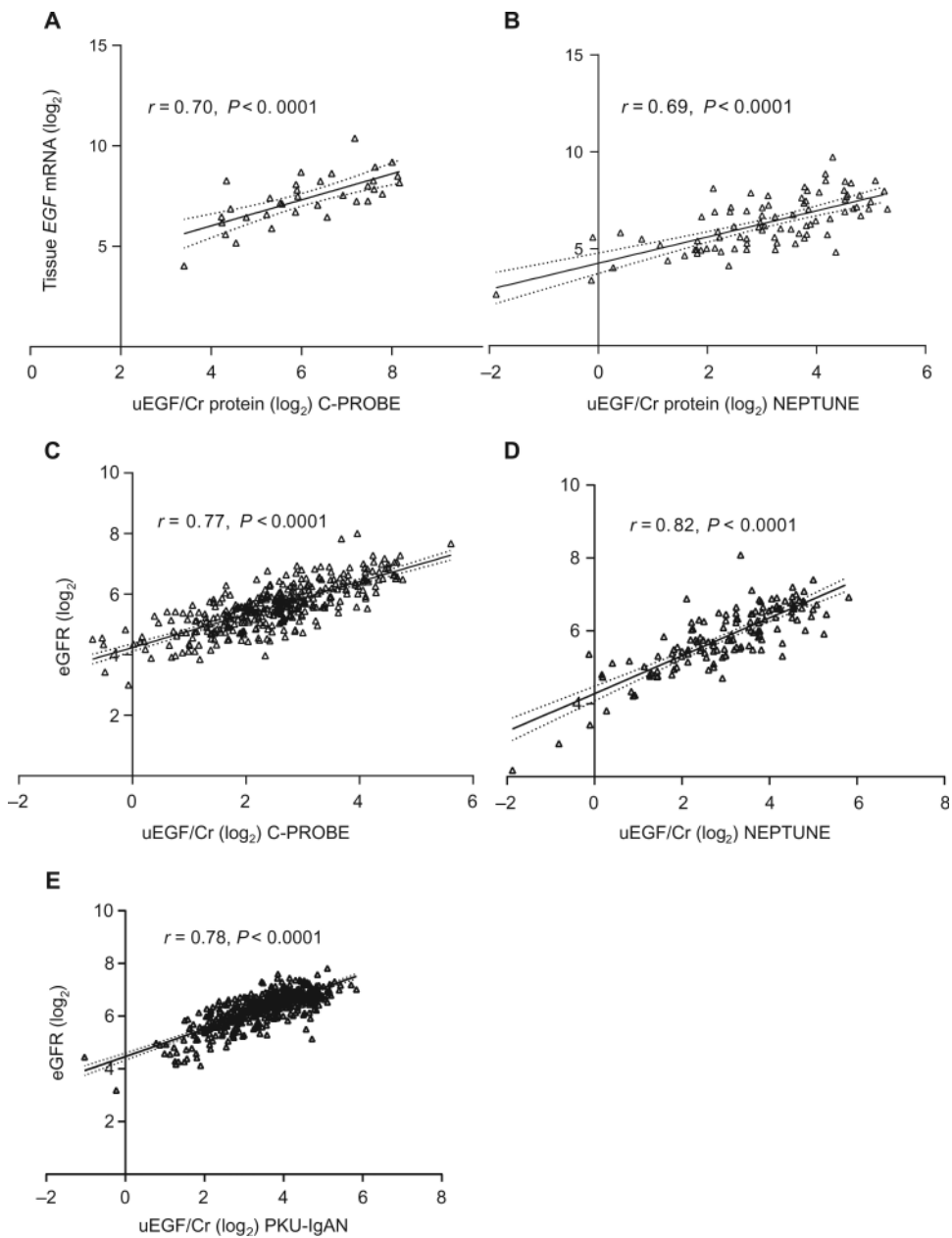
**Fig. 1.** Schematic overview of the tissue transcriptome-driven strategy to identify urinary biomarkers for CKD progression.



**Fig. 2. Transcript-eGFR association and EGF expression in the kidney in patients with CKD** (A) Correlation between observed and predicted eGFR (continuous line) on the basis of intrarenal transcripts integrated by ridge regression analysis (top 1 to 30 transcripts shown), providing eGFR prediction in the validation cohort ( $n = 55$ ). Gray area represents 95% confidence interval (CI). Black arrow indicates the highest correlation provided by a six-marker set, which is depicted in (B). (B) Maximal correlation ( $r = 0.77, P < 0.0001$ ) was demonstrated between observed eGFR and that predicted by gene expression values of six transcripts ( $n = 55$ ). Dotted line represents 95% CI. (C) Intrarenal EGF mRNA showed a

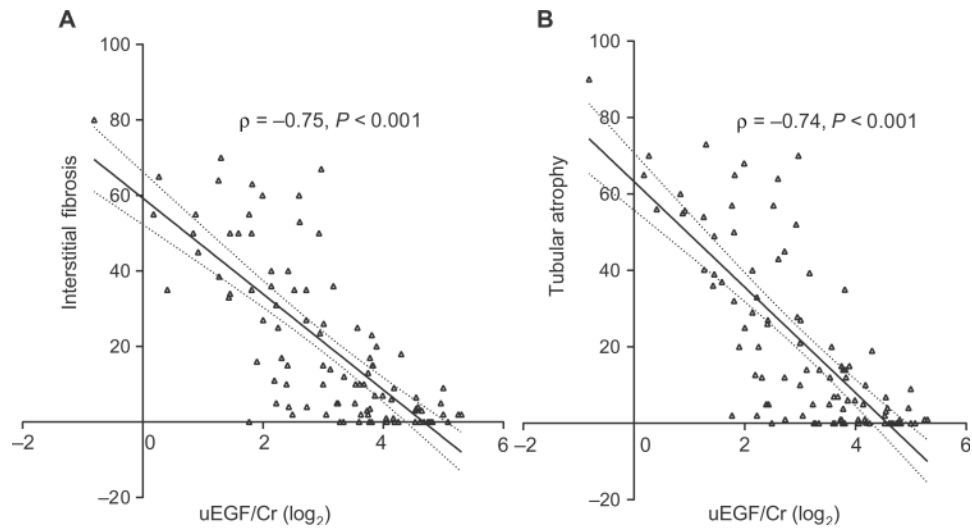


significant correlation with patients' eGFR in discovery ( $n = 164$ ) and validation cohorts ( $n = 55$  and  $42$ ). TLDA, Taq-Man Low Density Array. **(D)** *EGF* mRNA expression in major human organs/tissues (a selection out of a panel of 84 human organs/tissues and cell lines), derived from BioGPS (<http://biogps.org>), indicating a highly kidney-specific expression pattern. For full data set, see fig. S2. **(E)** *EGF* mRNA expression pattern in adult kidney (glomeruli, inner and outer cortex, inner and outer medulla, papillary tips, and renal pelvis); data were extracted from Higgins *et al.* (10). **(F)** In situ hybridization demonstrates tubule-specific *EGF* mRNA in both cortex (I and III) and medulla (II and IV), with reduced expression in CKD (III and IV) compared to healthy controls (I and II). Black arrows indicate positive staining in pink. Scale bars, 50  $\mu$ M.



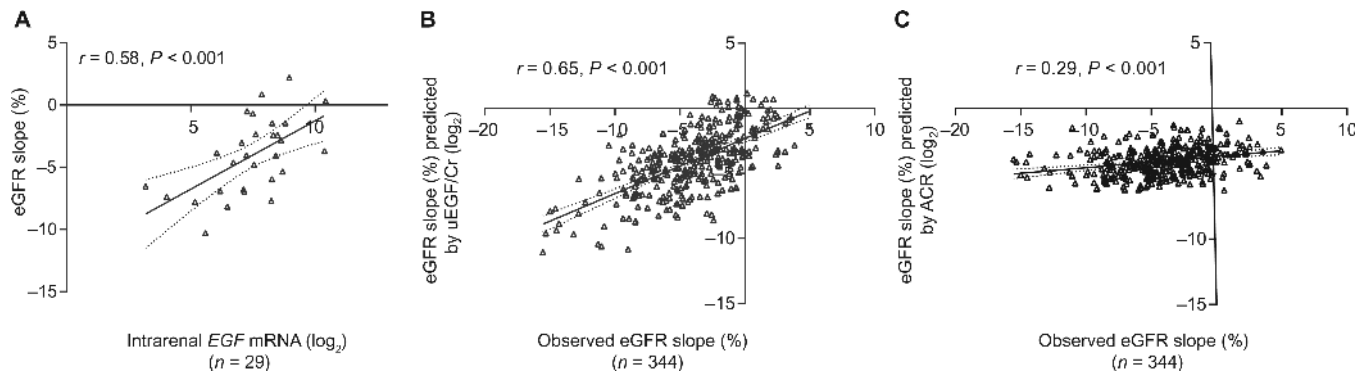
**Fig. 3. Correlation of uEGF/Cr with intrarenal *EGF* mRNA and eGFR**

(**A** and **B**) uEGF/Cr is correlated with intrarenal *EGF* mRNA in patients with matching urine enzyme-linked immunosorbent assay (ELISA) data and tissue mRNA expression data in C-PROBE (**A**) ( $n = 34$ ) and NEPTUNE (**B**) ( $n = 85$ ) cohorts. uEGF/Cr is correlated significantly ( $P < 0.0001$ ) with eGFR at the time of biopsy in patients from C-PROBE [(**C**)  $n = 349$ ], NEPTUNE [(**D**)  $n = 141$ ], and PKU-IgAN [(**E**)  $n = 452$ ].



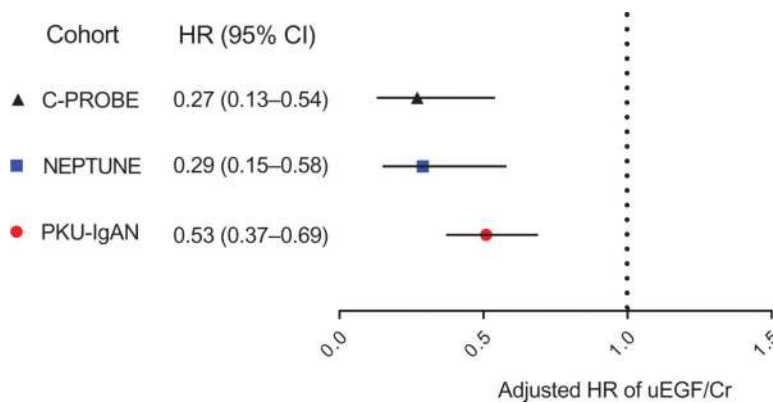
**Fig. 4. Association of uEGF/Cr with tubulointerstitial damage**

(A and B) Tubulointerstitial damage is reflected by the percentage of cortex affected by IF/TA. Scoring of IF/TA was based on evaluation of silver, periodic acid–Schiff (PAS), and trichrome-stained kidney sections ( $n = 102$ ) by six readers who were blinded to the uEGF/Cr value. For each section, the average scores of the six readers for IF and TA were calculated as indicators of tubulointerstitial damage. uEGF/Cr is significantly correlated with IF (A) and TA (B) scores (Spearman correlation,  $P < 0.001$ ).



**Fig. 5. Association of EGF with eGFR slope**

(A) Intrarenal *EGFR* mRNA expression correlated significantly ( $P < 0.001$ ) with eGFR slope in C-PROBE patients ( $n = 29$ ). (B and C) Correlation of the observed eGFR slope of CKD patients in C-PROBE ( $n = 344$ ) with slope predicted by uEGF/Cr (B) or ACR (C) using a regression model (adjusted for age and gender). eGFR slope predicted by uEGF/Cr (B) exhibited a higher correlation with the observed value than slope predicted by ACR (C).



**Fig. 6. Multivariable-adjusted HRs for predicting the composite end point on the basis of uEGF/Cr**

HRs were adjusted by age, gender, eGFR, and ACR. Adjusted HRs and 95% CIs were obtained by separate Cox regression models in each study cohort. A one-unit decrease of uEGF/Cr (in log scale) was associated with an increased risk of CKD progression of 3.73 (1.85 to 7.69)-fold, 3.43 (1.72 to 6.67)-fold, and 1.96 (1.45 to 2.70)-fold in these three cohorts. The unadjusted HRs for EGF were 0.33 (0.21 to 0.51) (C-PROBE), 0.33 (0.21 to 0.52) (NEPTUNE), and 0.57 (0.46 to 0.70) (PKU-IgAN).

**Table 1**  
**Demographic characteristics of CKD patients in the gene expression discovery and validation cohorts**

Age and eGFR are presented as means±SD. SLE, lupus nephropathy; RPGN, rapidly progressive glomerulonephritis; HTN, hypertensive nephropathy; MGN, membranous glomerulonephritis; DN, diabetic nephropathy; FSGS, focal segmental glomerulosclerosis; MCD, minimal change disease; TMD, thin basement membrane disease.

Disease type	CKD patients	eGFR (ml/min per 1.73 m <sup>2</sup> )	Age (years)	Gender (% female)
<b>Discovery cohort: ERCB</b>				
SLE	30	63.7 ± 29.4	34.7 ± 13.3	76.7
IgAN	24	75.9 ± 37.9	36.4 ± 14.6	25
RPGN	21	46.6 ± 31.5	58.5 ± 14.1	42.9
HTN	20	43.9 ± 25.1	57.2 ± 12.1	25
MGN	18	88.9 ± 41.4	53.4 ± 19.3	44.4
DN	17	44.3 ± 24.9	58.3 ± 10.7	29.4
FSGS	16	73.4 ± 38.4	46.2 ± 17.6	56.2
MCD	12	100.7 ± 33.9	35.8 ± 16.8	33.3
TMD	6	93.4 ± 29.4	46.0 ± 14.5	33.3
Total	164	66.4 ± 37.2	46.9 ± 17.6	43.3
<b>Validation cohort 1: ERCB</b>				
SLE	10	60.1 ± 31.5	37.1 ± 13.8	60
IgAN	17	50.8 ± 34.1	46.0 ± 17.5	29.4
RPGN	5	20.9 ± 15.0	50.4 ± 21.7	20
MGN	4	54.4 ± 32.4	58.7 ± 23.6	50
DN	1	97.3	44.8	100
FSGS	8	55.9 ± 36.0	43.6 ± 18.6	25
MCD	4	86.0 ± 44.5	44.4 ± 19.3	0
Other	6	68.4 ± 60.7	45.2 ± 16.7	33.3
Total	55	55.7 ± 38.3	45.1 ± 17.6	32.7
<b>Validation cohort 2: C-PROBE</b>				
SLE	16	106.4 ± 60.1	31.3 ± 9.9	87.5
IgAN	1	110.4	23	100
HTN	1	146.5	29	0
MGN	2	51.9 ± 21.9	53.0 ± 5.7	50
DN	1	50.0	80.8	100
FSGS	11	51.2 ± 21.0	41.9 ± 14.0	27.2
MCD	1	124.7	22	100
Other	9	41.3 ± 23.1	52.4 ± 18.9	55.6
Total	42	76.3 ± 50.8	39.6 ± 15.7	61.9

**Table 2**

Baseline characteristics of patients whose urine samples were analyzed in urine studies.

Baseline characteristics	C-PROBE ( <i>n</i> = 349)	NEPTUNE ( <i>n</i> = 141)	PKU-IgAN ( <i>n</i> = 452)
Age	54.8 ± 16.1	45.7 ± 17.3	36.0 ± 12.0
Gender (% female)	59.3	35.5	51.3
Race (% black)	47.9	23.4	0
eGFR (ml/min per 1.73 m <sup>2</sup> )	55.4 ± 32.2	67.7 ± 37.1	84.8 ± 34.1
ACR (mg/mg)	0.8 ± 1.2	2.3 ± 2.1	0.6 ± 0.6
Log <sub>2</sub> uEGF/Cr (ng/mg)	2.5 ± 1.1	3.0 ± 1.3	3.5 ± 1.0

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3**  
**The association of uEGF with time to composite end point**

The association was evaluated by a Cox proportional hazards regression analysis in C-PROBE, NEPTUNE, and PKU-IgAN cohorts. Renal event is defined as the presence of composite end point of ESKD or 40% reduction in baseline eGFR. The follow-up lengths (in years) for this analysis for the three cohorts were  $1.8 \pm 0.8$ ,  $2.0 \pm 0.8$ , and  $3.6 \pm 2.2$ , respectively. AIC, Akaike information criterion; LR, likelihood ratio test.

Cohort	Model	AIC	C-statistics	P value (LR test)
<b>C-PROBE (25 events/189 patients)</b>				
M1	eGFR + ACR <sup>*</sup>	206.3	0.75 (0.58-0.91)	
M2	eGFR + ACR <sup>*</sup> + EGF	192.1	0.87 (0.77-0.97)	<0.0001
<b>NEPTUNE (26 events/110 patients)</b>				
M1	eGFR + ACR <sup>*</sup>	213.98	0.74 (0.61 -0.86)	
M2	eGFR + ACR <sup>*</sup> + EGF	204.14	0.80 (0.68-0.92)	0.0006
<b>PKU-IgAN (68 events/428 patients)</b>				
M1	eGFR + ACR <sup>*</sup>	660.18	0.71 (0.59-0.83)	
M2	eGFR + ACR <sup>*</sup> + EGF	647.94	0.75 (0.64-0.86)	0.00016

\* Adjusted for age and gender.