Materials and methods

Ethics statement

This study was approved by the ethics committees of Kyoto University, Japan, and the Faculty of Medicine, University of Peradeniya, Sri Lanka. All human samples, clinical reports, and questionnaire data were obtained after receiving written informed consent, and the study was performed in accordance with the guidelines of the Declaration of Helsinki.

Study populations for environmental and genetic studies

CKDu cases were defined as cases who developed CKDu in the clinical course of tubulointerstitial damages. Thus, we eliminated CKD that can be a secondary complication of diabetes (history of diabetes mellitus and HbA1C >6.5% at the time of diagnosis of CKD) or hypertension (history of chronic and/or severe hypertension at the time of diagnosis of CKD), or other known renal diseases such as autoimmune diseases, glomerular nephritis, Fanconi syndrome or IgA nephropathy (presence of histopathological and immunoflurocence evidence). Previous studies suggested that areas of residence and male sex are strong risk factors for CKDu[2,3,5]. Therefore, we selected two areas, Medawachchiya and Girandurukotte, where CKDu has been postulated to be more prevalent, and male sex for this study (Figure 1A).

Figure 1B shows a schematic representation of the selection methodology for the cases and controls recruited for this study. We invited male CKDu case-series patients (age range, 16–70 years) from a single ethnic group (Sinhalese), who were registered in Medawachchiya and Girandurukotte renal clinics from January 2005 to December 2010. The response rate was about 90%, and these patients were recruited as cases for this study (n=311). Cases were confirmed to have tubulointerstitial damage by

78 either renal biopsy or clinical data; 93% (*n*=288) were biopsy-proven (interstitial
79 fibrosis with or without interstitial inflammation and negative immunofluorescence for
80 IgG, IgM, IgA, and C3) and 7% (*n*=23) had serum creatinine >1.2 mg/dL and/or A1M
81 >15.5 mg/L. In parallel, we used the clinical records to confirm that these patients
82 developed CKDu in the course of tubulointerstitial damage and the cases did not have
83 uncontrolled hypertension or diabetes at the time of initial diagnosis.

84 With assistance from community administrative leaders, randomly selected
85 apparently healthy Sinhalese males (age range, 16–70 years), who had no past
86 history of hypertension, diabetes mellitus, or renal diseases, were not on treatment for
87 any other disease condition, and had resided in the region for at least 10 years, were
88 invited to participate. The response rate was about 60%. As illustrated in Figure 1B,
89 blood pressure measurement, dipstick test for proteinuria and glycosuria in spot urine
90 samples, HbA1c, serum creatinine, and A1M were used to recruit subjects without
91 undiagnosed hypertension, diabetes, or renal function impairment.

92 All the selected cases and controls were interviewed using a structured
93 questionnaire to collect personal, occupational, lifestyle-related, food habit-related,
94 and clinical information. Blood samples (10 mL) were collected from peripheral veins
95 into K-EDTA tubes. Serum was separated immediately by centrifugation at 3000 rpm
96 for 10 min. Spot urine samples were also collected from all the recruited cases and
97 controls into empty polypropylene tubes.

98 Water samples (1 L each) were collected into empty polypropylene bottles from
99 drinking water sources in two areas of NCR (Medawachchiya, n=15; Girandurukotte,
100 n=16) (Figure 1A). The water samples were randomly selected from common drinking
101 water sources in villages with a high incidence of CKDu.

102 All the drinking water, urine, serum, and whole-blood samples were immediately

103 stored at 4°C and then transferred to −30°C within 6 h of collection. The samples were
104 shipped to Japan at −20°C and stored at −30°C until analysis.
105 The estimated glomerular filtration rate (eGFR) was obtained using the modified
106 diet in renal disease (MDRD) formula, i.e. GFR (mL/min/1.73 m^2) = 186 × serum
107 creatinine$^{-1.154}$ × (age)$^{-0.203}$ × (0.742 if female), based on the serum creatinine levels[8].
108 In addition, we conducted a field survey to investigate food habits of the local
109 residents.
110
111 **Drinking water, serum and urine analyses**
112 Drinking water and urine samples were analyzed for heavy metals using inductively
113 coupled plasma mass spectrometry (ICP-MS). Urinary heavy metals were checked in
114 all subjects recruited for the genome-wide association study (GWAS) (n=597). From
115 the 286 controls, 154 subjects (54%; 77 per area) and from the 311 cases, 176
116 subjects (57%; 88 per area) were randomly selected for As speciation. Urinary As
117 speciation was carried out using Agilent 1260 HPLC BIO inert, Agilent ICP-MS 7700x
118 (HPLC-ICP-MS). Serum fluoride was analyzed using flow injection analysis with a
119 fluoride ion-selective electrode. Urinary creatinine concentrations, Urinary A1M and
120 HbA1c were measured at Special Reference Laboratory (Tokyo, Japan). The
121 detection limits for As species are 0.02 μg/L and standard reference materials
122 (Seronorm Trace Elements Urine L-1 and L-2, SERO AS, Billingstad, Norway) were
123 analyzed in every analytical batch. Coefficients of variance (n=26) ranged from 4.2%
124 (Cs) to 42.3% (Ni) in the Urine L-1 analysis and from 3.8% (Tl) to 22.6% (Fe) in the
125 Urine L-2 analysis.
126
127 **Genome-wide association study and direct sequencing**

DNA was extracted from the whole-blood samples using a QIAamp DNA Blood Mini Kit (Qiagen, Chatsworth, CA) for the GWAS. The GWAS was performed on 597 subjects (n=311 for cases; n=286 for controls) using Illumina 500K Chips (Human 610-Quad-Custom v1.0 DNA Analysis Bead Chip; Illumina, San Diego, CA) according to standard protocols. Prior to the analysis, the data sets were filtered on the basis of the single nucleotide polymorphism (SNP) genotyping call rates (≥99% completeness), sample call rates (≥95%), non-deviation from the Hardy–Weinberg equilibrium (P<0.001), and evidence of cryptic family relationships (Figure 1C).

Principal component analysis showed no outliers and all subjects were confirmed to be of a single ethnicity (Figure S1). A quantile-quantile plot revealed that the distribution of observed P values followed the expected distribution (Figure S2). The inflation factor was calculated to be 1.0. The analyses were based on an additive genetic model. Values of $P<5.0\times10^{-8}$ were considered to show genome-wide significance and values of $5\times10^{-8}<P<5\times10^{-6}$ were considered to show suggestive evidence of an association.

Direct sequencing of the *SLC13A3* gene and *TP53RK* gene including the 5'-UTR, 3'-UTR, exons and 100-bp covering splicing donor and acceptor sites was performed using the Sanger method in eight randomly selected cases and eight randomly selected controls. Information on the primer sequences and PCR conditions will be provided upon request to the corresponding author.

Determination of the numbers of CAA repeats

We determined the numbers of CAA repeats in the 5' promoter region of *SLC13A3* in all cases and controls. The regions around the CAA repeats were amplified by PCR using fluorescently tagged primers (forward, 5'-FAM-AGC CTG GGT AAC AGA GTG

AGA-3'; reverse, 5'-TCC CTT TAA GAC CTC ATC ACC-3'). The expected size of PCR products was 215 + 3n (where n=CAA repeat number). The sizes of amplification products were analyzed by DNA sequencer.

Statistics

Univariate and multiple logistic analyses for demographic factors past events and life-style related factors were conducted to identify the risk factors. To show the distribution of urinary creatinine levels and each metal parameter, we presented means with standard deviations, medians, minimums, maximums and percentiles (25%, 75%). Some urinary parameters were not normally distributed and we computed the statistical significance of differences between cases and controls using both parametric (Student's t-test) and non parametric (Kolmogorov-Smirnov) methods. A comparison of serum fluoride levels among different CKD stages (stages 1–5) was made using the Student–Newman–Keuls (SNK) multiple range test. A P value of <0.05 following adjustment by Bonferroni correction, using the number of urinary metal parameters in multiple comparisons, was considered significant. All statistical procedures were performed using STATISTICA 64 (supplied by StatSoft, OK, USA).

Results

Population characteristics and characterization of risk factors

As shown in Figure 1B, we recruited 504 candidate control subjects. Blood pressure measurements eliminated 154 subjects owing to hypertension. A further 24 subjects were eliminated because of proteinuria and/or glycosuria. Seventeen subjects were eliminated owing to high HbA1c levels and 23 subjects were eliminated because of elevated serum creatinine or urinary A1M levels. In total, 43% of the 504 participants

178 were eliminated.

179 The demographic and clinical characteristics of the selected cases and controls are
180 summarized in Table 1. The CKDu cases were older and had stayed longer in the
181 region than the controls. On average, the subjects in both groups were neither
182 overweight nor underweight. We identified several risk factors (Table S1). The
183 percentage of farmers was significantly higher in the cases (97.1%) than in the
184 controls (69.9%). Some lifestyle-related factors that are commonly associated with
185 farming occupation, such as smoking, betel chewing, tobacco chewing, and history of
186 snake bites were also significantly more common in the cases than in the controls.
187 The presence of CKDu family history was significantly higher in the cases than in the
188 controls. Then, we evaluated factors which were significant ($p<0.05$) by univariate
189 analyses, by multiple logistic analysis. The analysis revealed that farming, family
190 history of CKDu, tobacco chewing and history of snake bites were independent risk
191 factors (Table 2).

192

193 **Metal concentrations in drinking water**

194 Metal concentrations in drinking water samples collected from Medawachchiya and
195 Girandurukotte are summarized in Table S2 together with the maximum allowable
196 limits in drinking water recommended by the WHO
197 (http://www.who.int/water_sanitation_health/publications/2011/dwq_guidelines/en/ind
198 ex.html) and/or United States Environmental Protection Agency
199 (http://water.epa.gov/drink/contaminants/index.cfm). We further compared the metal
200 concentrations with the Japanese water quality standards (Ministry of Environment,
201 Japan, 2004; http://www.env.go.jp/en/water/wq/wp.pdf). According to the available
202 recommendations, none of the metals, which have guideline values including

nephrotoxic metals such as Cd, As, and Pb, were present at toxic concentrations in the drinking water consumed daily by the CKDu cases in Medawachchiya and Girandurukotte.

Metal concentrations in urine

In Table S3, we presented metal concentrations (μg/L) in urine in cases and controls. We could not conduct urine analyses for 10 cases and 10 controls due to inadequate sample volume. To decrease false positive rates, a significant level ($p=0.003=0.05/18$) was adjusted by numbers of metals examined (n=18). None of the 18 metal concentrations was higher in urines in cases than in controls, including nephrotoxic metals such as Pb, Cd, As, Ni, V, and Al. Reversely, several metals were higher in controls than in cases ($p<0.001$); Pb, Tl, Cs, Cd, Mo, Sr, Rb, Se, As, Ni, Co, V, and Al. Cases had significantly lower creatinine concentrations in urine (n=301: 0.72±0.53 g/L) than controls (n=276: 1.18±0.70 g/L) ($p<0.0001$), suggesting impaired urine concentration ability due to CKDu. Thus, we also investigated the metal concentrations in urine adjusted by urinary creatinine levels (Table S3). Creatinine-adjusted As, Zn, Cu, Co, Mn, and Al concentrations in urine were significantly higher in cases than in controls ($p<0.05/18=0.003$) in both parametric and non-parametric analyses (Table 3). However, it should be addressed that irrespective of the adjustment for creatinine level, urinary Cd concentration was significantly higher in controls than in cases (Table 3 and S3), indicating that Cd is not a causative factor for CKDu.

In terms of As, our results were much lower than the reported concentrations in As-contaminated areas in countries such as Bangladesh and Taiwan[10,11]. However, the total As concentrations were relatively higher than the concentrations observed in

some uncontaminated areas in Europe, the United States, and Canada (http://www.hc-sc.gc.ca/ewh-semt/alt_formats/hecs-sesc/pdf/pubs/contaminants/chms-ecms/report-rapport-eng.pdf)[12,13]. Therefore, we conducted speciation to investigate a possibility that modest contamination with As may be associated with CKDu. A total of 330 urine samples were duplicated and analyzed at two independent laboratories for total As and As speciation. The sum of arsenic speciation components (As-III, As-V, MMA, DMA, AB, Table 4) measured by HPLC ICP-MS nicely matched the total As levels measured by ICP-MS (Table 3 and Table 4, and Figure S3: R^2=0.97, $p<0.0001$). As speciation results revealed that 75–80% of total urinary As in both cases and controls was in the form of arsenobetaine (Table 4), which is non-toxic. Adjustment with creatinine showed that total arsenic concentration (μg/g-Cr) in urine was greater in cases than in controls. However, the difference of the total arsenic concentration was mostly attributable to the difference in arsenobetaine fraction (Table 4). Hata et al (2007) reported that the median total urinary arsenic level of a Japanese healthy male adult is 141.3 μg/L and the arsenobetaine level is 61.3 μg/L; these levels are about three times higher than those of CKDu cases[9]. The modestly elevated levels of arsenobetaine in urine are in accord with the common habit of eating dry fish in these areas, as confirmed by the field study. Typical signs of chronic As exposure, such as skin keratosis and pigmentation, were observed in neither the cases nor the controls.

Fluoride concentration in serum

An inverse relationship was observed between the eGFR and the serum fluoride concentration (Figure S4). However, the eGFR stratified analysis demonstrated that serum fluoride levels were not different between cases and controls (Table S4) for the

253 subjects with eGFR >60 mL/min/1.73m^2. Moreover, the serum fluoride concentrations
254 in the majority of the unaffected and early-stage CKDu patients were below the
255 reported upper normal concentration of 50 μg/L[14]. Thus, we concluded that higher
256 serum fluoride is a result of lowered eGFR (impaired renal function) and not the
257 primary cause of CKDu.
258
259 **GWAS**
260 A total of 543,848 GWAS SNPs in 577 samples were left after the quality control
261 process for the associations test with the eGFR (quantitative trait locus (QTL)
262 analysis) or the affected status (dichotomous analysis) (Figure 1C). A genome-wide
263 view of the single point association data is shown in Figure 1D.
264 A summary of the association results (QTL analysis and dichotomous analysis, with
265 and without age adjustment) for 10 candidate SNPs located in four genes is shown in
266 Table 5 (see also Figure 1D). The top hit that reached the genome-wide significance
267 level ($P<5\times10^{-8}$) in both the QTL analysis and dichotomous analysis was rs6066043
268 (G>A), which is localized on the *SLC13A3* gene encoding sodium-dependent
269 dicarboxylate transporter 3 ($P=5.23\times10^{-9}$ for age-adjusted QTL analysis; $P=3.73\times10^{-8}$
270 for age-adjusted dichotomous analysis). The population attributable fraction and odds
271 ratio for this SNP were 50% and 2.13. The allele frequencies for this SNP vary among
272 populations. The G allele is more common among European populations (88% G
273 allele frequency), while it is found in less than 50% of individuals in Asian populations
274 (HapMap data). In the present study, the G allele frequencies in the cases and
275 controls were 77% and 62%, respectively.
276 Since the SNP rs6066043 is located in the intergene region of *SLC13A3 and*
277 *TP53RK* (Figure S5), it is hard to explain the functional impairment by this SNP. We

thus screened for variants in these two genes. Direct sequencing was conducted for all coding exons and the 5'UTR and 3'UTR for these two genes. The results of direct sequencing are shown in Table S5. We identified variations in the CAA repeat number in the 5' promoter region, 14.6 Kb upstream of exon 1 of *SLC13A3* ($p<0.01$). Minor allele frequencies of other variants did not seem to differ in cases and controls. Thus, the sizes of the CAA repeat expansion were determined using fluorescently tagged primers in all cases and controls and this revealed variation in the CAA repeat expansion from $(CAA)_{10}$ to $(CAA)_{18}$. We investigated the association between repeat numbers and CKD. Considering $(CAA)_{15}$ as the wild-type, an odds ratio was calculated for the repeat polymorphism (Table S6) to be 1.6, being smaller than odds ratio of rs6066043 (2.13) (Table 5). No significant difference in mean eGFR was observed among various CAA repeat combinations (Table S6). Thus, we considered that (CAA) repeats were not a causative variant for CKDu.

The GWAS study revealed three other genes with suggestive evidence ($5x10^{-8}<p<5x10^{-6}$) of an association: *OLFM3, TMEM128* and *LOC157273* (Table 5). However, *OLFM3* can be discarded because it is an olfactory gene.

Discussion

This is the first comprehensive study addressing both the genetic, social, and environmental contributors of CKDu with tubulointerstitial damages in NCR of Sri Lanka. Ecological investigations in the present study confirmed the absence of nephrotoxic heavy metal contamination in drinking water. We found that the SNP rs6066043 of *SLC13A3* loads discernible risk of CKDu in to the local population as indicated by a large population attributable risk of 50%. It should also be addressed that occupation, family history of CKD, tobacco chewing, and history of snakebites

were also identified as independent risk factors.

None of 18 metals had urinary concentrations (μg/L) greater in cases than in controls. Only when adjusted to the urinary creatinine levels (μg/g-cre), it was found that the Zn, Cu, Co, Mn, and Al levels in urine were higher in cases than in controls. However, comparison of these levels with those in other populations is difficult due to the limited availability of published data. We considered that due to the impaired urine concentration ability in cases, those apparent elevations of urine concentrations for some metals might emerge by adjustment with creatinine levels. Further studies are needed to verify the individual and combined effects of these metal parameters on CKDu.

It is worthwhile mentioning for the urinary As levels because urinary concentrations of total As levels in cases and controls were suggestive of modest exposure to As. Thus, we conducted speciation of As compounds in urine. Speciation revealed that arsenobetaine is the major As compound, suggesting that seafood is the major sources of As[13]. The high urinary As concentrations in countries like Japan and Korea were found to result from the high amounts of seafood intake, and As speciation studies showed high proportions of arsenobetaine in urine, which is non toxic[9,15]. A field survey on the food habits of the population in NCR also supports regular intake of dried sea fish which is likely to be the source of high urinary arsenobetaine levels.

It is worth highlighting the fact that some studies have reported a possible relationship between As exposure and CKD[16,17]. Hsueh et al. (2009) reported a positive association between urinary total As concentration and CKD for the first time in Taiwan, and the total As concentrations in their study population closely resemble the observations in the present study (total mean urinary As in the CKD group and healthy controls: 31.95 and 20.71 versus 50.3 and 39.4 μg/gCr, respectively)[17].

However, the dominant component of total As was arsenobetaine in this study, but dimethylarsinic acid in the Taiwan study. Urinary Cd levels in controls were higher than those in cases, suggesting a minimal involvement of Cd in CKDu. The Cd levels reported in this study are comparable to other published data[18,19].

The present study revealed that a common genetic variant close to *SLC13A3* is associated with CKDu. *SLC13A3* encodes the high-affinity sodium dicarboxylate transporter 3 located in the basolateral membranes of human renal proximal tubules, liver, brain, and placenta[20]. On the other hand, *TP53RK*, encodes a protein, TP53-regulated kinase, which is a serine/threonine protein kinase[21] and functions in a *P53* signal pathway. *SLC13A3* has been identified as one of the most sensitive marker genes for predicting the clinical course to end-stage renal disease in type 2 diabetes mellitus and has an association with blood pressure[22,23]. Recently, *SLC13A3* was identified as one of the 43 genes that can be used in protein expression signatures to predict progressive renal fibrosis in mice, and was suggested to be a potentially useful molecular predictor for CKD progression in humans[24]. *SLC13A3* has also been shown to play a role in the accumulation of Hg-thiol conjugates in the basolateral membrane vesicles in proximal tubular cells in the rat[25]. Those pieces of evidence strongly support that *SLC13A3* is more likely mechanistically associated with CKD rather than *TP53RK*. However, further studies are needed to explain a mechanism by which CKDu is associated with rs6066043.

It was interesting to note that 43% of the apparently healthy male population, most of whom were farmers, had, at least, one undiagnosed non communicable disease, mainly high blood pressure. Mendis et al. (1988) reported that the prevalence of high blood pressure among Sri Lankans is one of the highest reported in the world, and Wijewardene *et al* (2005) reported that the area in Sri Lanka with the highest

353 prevalence of high blood pressure is the Uva province[26,27]. One of our study areas
354 (Girandurukotte) is located in this province. Therefore, it is likely that undiagnosed
355 hypertension may accelerate CKDu progression. In one of our previous studies, we
356 found hypertension as a risk factor for CKDu progression[26].

357 Some limitations of this study warrant mention. We used the eGFR based on the
358 MDRD formula rather than using the direct glomerular filtration rate, since it was not
359 feasible to measure the direct glomerular filtration rate in this setting. However, the
360 MDRD formula is not validated for the Sri Lankan population. Owing to the
361 discrepancies present in the validated MDRD formulas for other Asian populations, we
362 could not use these coefficients for this population[29-32]. A smaller sample size without
363 replication evidence may decrease the robustness of the results, and we have
364 planned to incorporate it in the next stage of this study. Although we confirm the
365 farmer preponderance among CKDu patients, this finding needs cautious
366 interpretation. Given that we selected patients by a case-series method to eliminate
367 possible biases, there may be selection bias in selecting controls. We eliminated 43%
368 of controls due to undiagnosed non-communicable diseases, most of them were
369 farmers. Such an elimination of a large portion of the population from the control
370 candidates suggests that local farmers have less chance to access medical care or
371 preventive care than controls with other occupations. Thus, our conclusion on being a
372 farmer as a risk of CKDu is very likely confounded by the accessibility to medical care
373 or public health services. On the other hand, the elimination of such subjects from
374 control candidates strengthened our genetic analysis. Because we primarily aimed to
375 determine genetic factors for tubulointerstitial damages, we selected cases who
376 developed CKDu in the course of primary tubulointerstitial damages without
377 hypertension or diabetes mellitus or other comorbidities. To unify the selection criteria,

controls should be selected with the same backgrounds. Applying the same backgrounds made our genetic analysis for tubulointerstitial damages scientifically sound. If we included controls with hypertension or other morbidities, phenotypic heterogeneity between cases and controls may contaminate our study.

In conclusion, our results do not support the involvement of a founder mutation or single heavy metal in the pathogenesis of CKDu. However, the present study found a significant SNP in *SLC13A3* with an odds ratio of 2.13 and a 50% population attributable fraction, indicating major genetic susceptibility to CKDu. Thus, further study is warranted to elucidate causal link of the SNP rs6066043 with CKDu.

Acknowledgement: We thank Dr Meiko Takahashi (Kyoto University) for support in performing the GWAS, and the staff of the Medawachchiya and Girandurukotte renal clinics for their assistance with the sampling program. The authors would like to acknowledge the Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology in Japan. The funding agency had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Reference

1) Athuraliya TN, Abeysekera DT, Amerasinghe PH, Kumarasiri PV, Dissanayake V. Prevalence of chronic kidney disease in two tertiary care hospitals: high proportion of cases with uncertain etiology. Ceylon Med J 2009; 54: 23-5.
2) Nanayakkara S, Komiya T, Ratnatunga N, et al. Tubulointerstitial damage as the major pathological lesion in endemic chronic kidney disease among farmers in North Central Province of Sri Lanka. Environ Health Prev Med 2012; 17: 213-21.
3) Nanayakkara S, Senevirathna ST, Karunaratne U, et al. Evidence of tubular damage in the very early stage of chronic kidney disease of uncertain etiology in the North Central Province of Sri Lanka: a cross-sectional study. Environ Health Prev Med 2012; 17: 109-17.
4) Bandara JM, Senevirathna DM, Dasanayake DM, Herath V, Abeysekara T, Rajapaksha KH. Chronic renal failure among farm families in cascade irrigation systems in Sri Lanka associated with elevated dietary cadmium levels in rice and freshwater fish (Tilapia). Environ Geochem Health 2008; 30: 465-78.
5) Chandrajith R, Nanayakkara S, Itai K, et al. Chronic kidney diseases of uncertain etiology (CKDue) in Sri Lanka: geographic distribution and environmental implications. Environ Geochem Health 2011; 33: 267-78.
6) Jayasumana MACS, Paranagama PA, Amarasinghe MD, et al. Possible link of Chronic arsenic toxicity with chronic kidney disease of unknown etiology. J Nat Sci Res 2011; 13: 64-73.
7) Athuraliya NTC, Abeysekera TDJ, Amerasinghe PH, et al. Uncertain etiologies of proteinuric-chronic kidney disease in rural Sri Lanka. Kidney Int 2011; 80: 1212–21.

8) Kooman JP. Estimation of renal function in patients with chronic kidney disease. J Magn Reson Imaging 2009; 30: 1341-6.
9) Hata A, Endo Y, Nakajima Y, et al. HPLC-ICP-MS speciation analysis of arsenic in urine of Japanese subjects without occupational exposure. J Occup Health 2007; 49: 217-23.
10) Chen Y, van Geen A, Graziano JH, et al. Reduction in urinary arsenic levels in response to arsenic mitigation efforts in Araihazar, Bangladesh. Environ Health Perspect 2007; 115: 917-23.
11) Chen JW, Chen HY, Li WF, et al. The association between total urinary arsenic concentration and renal dysfunction in a community-based population from central Taiwan. Chemosphere 2011; 84: 17-24.
12) Caldwell KL, Jones RL, Verdon CP, Jarrett JM, Caudill SP, Osterloh JD. Levels of urinary total and speciated arsenic in the US population: National Health and Nutrition Examination Survey 2003-2004. J Expo Sci Environ Epidemiol 2009; 19: 59-68.
13) Saoudi A, Zeghnoun A, Bidondo ML, et al. Urinary arsenic levels in the French adult population: the French National Nutrition and Health Study, 2006-2007. Sci Total Environ 2012; 433: 206-15.
14) Joshi S, Hlaing T, Whitford GM, Compston JE. Skeletal fluorosis due to excessive tea and toothpaste consumption. Osteoporos Int 2011; 22: 2557-60.
15) Lee JW, Lee CK, Moon CS, et al. Korea National Survey for Environmental Pollutants in the Human Body 2008: heavy metals in the blood or urine of the Korean population. Int J Hyg Environ Health 2012; 215: 449-57.
16) Zheng LY, Umans JG, Tellez-Plaza M, et al. Urine arsenic and prevalent albuminuria: evidence from a population-based study. Am J Kidney Dis 2013; 61:

385-94.

17) Hsueh YM, Chung CJ, Shiue HS, et al. Urinary arsenic species and CKD in a Taiwanese population: a case-control study. Am J Kidney Dis 2009; 54: 859-70.

18) Huang M, Choi SJ, Kim DW, et al. Evaluation of factors associated with cadmium exposure and kidney function in the general population. Environ Toxicol 2011: doi: 10.1002/tox.20750.

19) Batariova A, Spevackova V, Benes B, Cejchanova M, Smid J, Cerna M. Blood and urine levels of Pb, Cd and Hg in the general population of the Czech Republic and proposed reference values. Int J Hyg Environ Health 2006; 209: 359-66.

20) Pajor AM. Molecular properties of the SLC13 family of dicarboxylate and sulfate transporters. Pflugers Arch 2006; 451: 597-605.

21) Miyoshi A, Kito K, Aramoto T, Abe Y, Kobayashi N, Ueda N. Identification of CGI-121, a novel PRPK (p53-related protein kinase)-binding protein. Biochem Biophys Res Commun 2003; 303: 399-405.

22) Bento JL, Palmer ND, Zhong M, et al. Heterogeneity in gene loci associated with type 2 diabetes on human chromosome 20q13.1. Genomics 2008; 92: 226-34.

23) Simino J, Shi G, Arnett D, Broeckel U, Hunt SC, Rao DC. Variants on chromosome 6p22.3 associated with blood pressure in the HyperGEN study: follow-up of FBPP quantitative trait loci. Am J Hypertens 2011; 24: 1227-33.

24) Ju W, Eichinger F, Bitzer M, et al. Renal gene and protein expression signatures for prediction of kidney disease progression. Am J Pathol 2009; 174: 2073-85.

25) Lash LH, Hueni SE, Putt DA, Zalups RK. Role of organic anion and amino acid carriers in transport of inorganic mercury in rat renal basolateral membrane vesicles: influence of compensatory renal growth. Toxicol Sci 2005; 88: 630-44.

26) Mendis S, Ranasinghe P, Dharmasena BD. Prevalence of hypertension in Sri

472 Lanka. A large population study in the central province. Public Health 1988; 102:
473 455-62.
474 27) Wijewardene K, Mohideen MR, Mendis S, et al. Prevalence of hypertension,
475 diabetes and obesity: baseline findings of a population based survey in four
476 provinces in Sri Lanka. Ceylon Med J 2005; 50: 62-70.
477 28) Senevirathna L, Abeysekera T, Nanayakkara S, et al. Risk factors associated with
478 disease progression and mortality in chronic kidney disease of uncertain etiology:
479 a cohort study in Medawachchiya, Sri Lanka. Environ Health Prev Med 2012; 17:
480 191-8.
481 29) Praditpornsilpa K, Townamchai N, Chaiwatanarat T, et al. The need for robust
482 validation for MDRD-based glomerular filtration rate estimation in various CKD
483 populations. Nephrol Dial Transplant 2011; 26: 2780-5.
484 30) Imai E, Horio M, Nitta K, et al. Estimation of glomerular filtration rate by the MDRD
485 study equation modified for Japanese patients with chronic kidney disease. Clin
486 Exp Nephrol 2007; 11: 41-50.
487 31) Ma YC, Zuo L, Chen JH, et al. Modified glomerular filtration rate estimating
488 equation for Chinese patients with chronic kidney disease. J Am Soc Nephrol
489 2006; 17: 2937-44.
490 32) Matsuo S, Imai E, Horio M, et al. Revised equations for estimated GFR from serum
491 creatinine in Japan. Am J Kidney Dis 2009; 53: 982-92.
492
493
494

Figure legends

Figure 1. Study area and design. (A) North Central Province in Sri Lanka and sampling locations. The map shows the locations of the CKDu clinics (Medawachchiya and Girandurukotte) from where the human samples were collected. Water samples were collected from two areas in NCR (Medawachchiya and Girandurukotte).

(B) Flow chart summarizing the selection process for the cases and controls. MW, Medawachchiya; GD, Girandurukotte; SCr, serum creatinine; A1M, alpha-1-microglobulin; GWAS, genome-wide association study.

(C) Study design for the genetic analysis. GWAS, genome-wide association study; HWE, Hardy–Weinberg equation; QTL, quantitative trait locus; eGFR, estimated glomerular filtration rate, PCA, principal component analysis.

(D) Manhattan plot. The Manhattan plot shows the significance of associations for all SNPs. The SNPs are plotted on the X-axis according to their position on each chromosome against the association with the estimated glomerular filtration rate on the Y-axis, shown as the $-\log_{10}$(P-value). One SNP showing a genome-wide significant association and all SNPs achieving $P<5x10^{-6}$ are shown with their P-values.

Figure S1. Principal Component Analysis (PCA) of the cases and controls with other populations.

Figure S2. Q-Q plot.

Figure S3. Comparison of total As levels measured by two different methodologies. As speciation and As determination by ICP-MS were conducted for 154 controls and 176 cases. The total As shown in Table 3 (Y) is plotted against total As shown in Table

4 (X). R was significant ($p<0.001$

Figure S4. Variation of serum fluoride with eGFR level and CKD stage.

Scattered plots of serum fluoride levels (X) and eGFR values (Y) for cases (n=301) and controls (N=276) are shown. 265 controls (96%) and 66 cases (22%) had eGFR greater than 60 ml/min/1.73 m^2.

Figure S5. Linkage Disequilibrium block and position of the *SLC13A3* gene (NM_022829.5), GWAS significant SNP and CAA repeat (Chr20: 44630 – 44767 kbp in NCBI36/hg18). D' (D prime) represents strength of linkage disequilibrium. The $(CAA)_n$ repeat was located in the intergene region between *SLC13A3* and *TP53RK.*