Circulating small non-coding RNAs associate with kidney function in people with type 2 diabetes

LM 't Hart1,2,3,4, GA Bouland1, JA de Klerk1, M.T. Blom3,5, SJ Cramer1, JWJ Beulens2,3, RC Sliker1,2,3

1 Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands
2 Department of Epidemiology and Data Sciences, Amsterdam University Medical Center, VU University Amsterdam, Amsterdam, The Netherlands
3 Amsterdam Public Health, Amsterdam, The Netherlands
4 Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
5 Department of General Practice and Elderly Care Medicine, Amsterdam University Medical Center location VUmc, Amsterdam, The Netherlands

Author for correspondence:
LM 't Hart, PhD
Leiden University Medical Center
Department of Cell and Chemical Biology
Albinusdreef 2
2333ZA Leiden, The Netherlands
+31 715269796
lmthart@lumc.nl

Word count:
Abstract: 248 words
Body: 3725 words (excl references)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Aims/hypothesis Micro- and macrovascular complications are common among persons with type 2 diabetes. Interest into the potential of circulating small non-coding RNAs (sRNAs) as biomarkers or drivers of the development of diabetic complications is growing. In this study we investigated if circulating sRNAs levels associate with prevalent chronic kidney disease (CKD) in persons with type 2 diabetes.

Methods Plasma sRNAs levels were determined using sRNA-seq, allowing detection of miRNAs, snoRNAs, piRNAs, tRNA-fragments and various other sRNA classes, in persons with type 2 diabetes, with CKD (n=69) or without CKD (n=405). Multiple regression analyses were used to test for associations between the sRNAs and primary endpoint CKD. In secondary analyses, we also tested the association with eGFR and albuminuria (UACR).

Results In total, twelve sRNA were associated with prevalent CKD (log FC = -0.32 to -1.28, all P≤6.24x10^-4). Although miRNAs represent the majority of the sRNAs measured (64%) only four miRNA were significantly associated with prevalent CKD. Interestingly, the majority of the significant sRNA belonged to the snoRNAs (58%). Similar results were observed for eGFR and UACR. Only one of the twelve sRNAs showed its highest expression in kidney whereas several others showed high expression in liver or colon.

Conclusions/interpretation Small RNAs present in the circulation associate with CKD in persons with type 2 diabetes. Further studies are warranted to elucidate the biological role of sRNA in diabetic CKD and in particular snoRNAs. High expression of these circulating sRNAs in tissues other than kidney suggest a role in inter-organ communication.

Research in context

• What is already known about this subject? (maximum of 3 bullet points)
  miRNAs present in the circulation have been associated with CKD in diabetes. The role of other small non-coding RNAs, for example snoRNAs, have not been studied.

• What is the key question? (one bullet point only; formatted as a question)
  Are circulating small RNAs, such as miRNA, snoRNAs, tRNA fragments and piRNAs associated with CKD in persons with type 2 diabetes (n=475)?

• What are the new findings? (maximum of 3 bullet points)
  Twelve small RNAs associate with prevalent CKD in type 2 diabetes. Especially snoRNAs are more frequently associated with CKD. Identified circulating small RNAs are often expressed in tissues other than kidney, which suggest a role for these small non-coding RNAs in inter-organ communication.

• How might this impact on clinical practice in the foreseeable future? (one bullet point only)
  Circulating small RNAs and especially snoRNAs might represent novel mechanisms and therapeutic targets for CKD in diabetes.
Introduction

Chronic kidney disease (CKD) in diabetes is one of the most common complications in type 2 diabetes. Nonetheless, while some develop complications early in the disease, others will never. The mechanisms underlying CKD are not fully understood (1). Many groups including our own (2), have investigated biomarkers such as genomics, metabolomics, transcriptomics and proteomics data (3; 4). This research has improved our knowledge about the mechanisms underlying CKD development in T2D and have yielded novel biomarkers and drug targets. Some of these biomarkers such as a proteomic score are currently tested for clinical use (3; 4). A novel class of potential biomarkers are the circulating small non-coding RNAs (sRNA) which are mostly contained in extracellular vesicles or exosomes or are bound to lipoproteins and originate from various tissues including tissues important in kidney function (5-7). Whereas it was previously thought that sRNAs merely reflect degradation by-products, it has now been shown that they exert important functions in cell-to-cell communication, the regulation of gene expression and ribosomal RNA maturation (5-7). The composition of sRNAs in the circulation is dependent on tissue homeostasis and can be altered by disease development. Irrespective of the mechanism involved, sRNAs may have an impact on the development and progression of CKD when deregulated. Importantly, each tissue releases its own combination of tissue-specific and more ubiquitous sRNAs into the bloodstream. These circulating sRNAs can be taken up by other tissues and affect target tissue function (8).

Previous studies, focused only on miRNAs, identified several miRNAs associated with kidney functions (9-11). In this study we examined, for the first time, all circulating sRNAs, including miRNAs, piRNAs, tRNA fragments (tRFs) and snoRNAs in relation to prevalent CKD. We use next-generation sequencing to capture the whole sRNA transcriptome allowing for a hypothesis-free, comprehensive analysis to study the role of sRNAs in diabetic kidney disease (12).

Materials and Methods

The study has been approved by the ethics committee of the Amsterdam University Medical Center, location VUmc. All participants provided informed consent prior to this study. A schematic overview of the study design can be found in Supplemental Figure 1.

Study sample

For this project we used data from the Hoorn Diabetes Care System cohort (DCS), a large prospective study comprised of over 14,000 individuals with type 2 diabetes who receive structured diabetes care. (13) Participants visited the Hoorn diabetes research center annually for diabetes care including routine biochemistry, medication review and clinical care. Patients were also invited to participate in DCS biobanks in which we combine the annual phenotypic data with biobanking of liquid biopsies which includes serum, plasma and urine. Blood samples were collected during two phases of biobanking in 2008/2009 and 2012-2014. For the current study we used data collected for a previous case-cohort study.
investigating sRNAs in relation to insulin initiation (‘t Hart et al. in preparation). In short, the study included 145 cases of incident insulin use and 330 random controls all with a plasma sample collected within four years after diagnosis of diabetes and no-insulin use at the time of blood sampling. All laboratory measurements were on samples taken in a fasted state on the same day as the sample for sRNA-seq. Routine anthropometric, biochemistry methods and definitions have been described previously in detail in van der Heijden et al. (13). Kidney function was assessed based on eGFR (CKD-EPI) and urinary albumin to creatinine ratio (UACR) and included as continuous variables in our analyses. Stages of chronic kidney disease (CKD) were defined according the KDIGO criteria as previously described (14; 15). Cases were defined as having CKD stage ≥ 1 at the time of blood sampling for sRNA analysis which resulted in 69 cases.

Circulating small RNA sequencing

Circulating sRNAs were isolated from 900ul of citrate plasma using a commercial cell-free RNA isolation kit (Quick-cfRNA Serum & Plasma Kit, Zymo Research, Irvine, CA, USA). After quality control 475 samples were submitted for sRNA sequencing. Library preps were constructed using the NEBNext® Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs, Ipswich, MA, USA). A Pippin Prep instrument (Sage Science, Beverly, MA, USA) was used for size selection of fragments ranging from 120 to 200 nucleotides to further enrich for sRNAs with a size of ~20 to 100 nucleotides. Sequencing was performed using the Illumina NextSeq500 sequencer with v2.5 sequencing reagent kits and 35 cycle paired-end reads (PE35). On average 11.3±2.8 million reads per sample were generated. The exceRpt pipeline (16), developed by the NIH Extracellular RNA Communication Consortium (ERCC), was used to map the obtained reads, for quality control, to process the data and generate RNA abundance estimates for miRNA, piRNA, IncRNA, tRNA fragments, Y_RNA (fragments), snoRNA, snRNA, scaRNA and various other sRNA species. Using respectively miRBase v22(17), piRNABank v1(18), Gencode (version 38)(19), circBase(20) and GtRNAdb(21). One sample did not pass the QC thresholds.

Statistical analysis

The R-package edgeR was used to identify differentially expressed sRNAs (22). A negative binomial generalized log-linear model was applied using edgeR’s function estimateGLMRobustDisp. The base model was adjusted for age and sex. A fully adjusted model was additionally adjusted for BMI, diabetes duration, HbA1c, HDL, SBP and smoking unless otherwise stated. UACR was log10-transformed before analysis. In sensitivity analyses we additionally included use of antihypertensive medication, eGFR or UACR and case-control status. Where possible missing data were imputed from previous measurements taken within one year of the blood sample for RNA isolation (n_{max}=17). The Benjamini-Hochberg procedure was used to adjust for multiple testing and an FDR below 0.05 was considered significant.
Expression patterns of small RNAs in tissues

Publicly available raw sRNA-seq data for various metabolic tissues were downloaded from the Gene Expression Omnibus (GEO) (23). The tissues for which sRNA-seq data was used included kidney, thyroid, brain, liver, muscle, heart, colon, pancreas, subcutaneous white adipose tissue and whole blood and urine. Accession numbers are given in Supplemental Table 1. Data were processed using the exceRpt pipeline as described above. The tissue data were used to estimate the abundance of the sRNAs in these tissues. Data was z-scaled in figures.

Results

The main clinical characteristics of our study group are shown in Table 1. The characteristics resemble those of a typical T2D population with slightly more males than females, a mean age-at-diagnosis of 59.3 ± 9.0 years, BMI of 30.7 ± 6.0 kg/m2 and an average duration of diabetes at time of blood sampling of 2.2 ± 1.1 years. Glycemic control was good with an average HbA1c of 46 (42, 51) mmol/mol. Prevalence of CKD at the time of blood sampling was 15%.

After quality control 938 sRNAs were expressed above the threshold of a median ≥5 copies in our data, which included miRNAs, piRNA, Y_RNA fragments, lncRNAs, tRNA_fragments, snoRNAs, snRNAs and various other sRNA species. (Table 2). Twelve sRNAs were significantly associated with prevalent CKD (P_{FDR}≤0.05, Table 3, Figure 1A). All identified sRNAs were down-regulated in CKD cases. The top hit was SNORD118 (log2 fold change=-0.873, P_{FDR}=2.35x10^{-3}, Figure 1B). Respectively, 73 and 209 sRNAs were significantly associated with eGFR and UACR, on which the CKD classification is based (Table 3, Supplemental Table 2, Supplemental Figures 2 and 3). The overlap between the sRNAs associated with CKD and its components eGFR and UACR resulted both in unique and overlapping sRNAs (Table 3, Figure 1C). Three sRNAs overlapped between all three endpoints used (U8, SNORD107 and SNORD25), whereas one overlapped between eGFR and CKD (miR-145-5p) and six between UACR and CKD (SNORD118, SNORD24, SNORD87, miR-143-3p, SNORD15B, miR-4685-3p).

In the full model, including adjustment for age, sex, BMI, diabetes duration, HbA1c, HDL, SBP and smoking, respectively 3, 53 and 183 sRNAs remained significantly associated with CKD, eGFR and UACR (Supplemental Table 3, Supplemental Figure 4; 100%; 91% and 86% overlap for CKD, eGFR and UACR respectively). In a sensitivity analysis, additional adjustment for use of antihypertensive medication and baseline eGFR or UACR (only for eGFR and UACR) did not have a major impact on the effect sizes but reduced the number of significant associations (r²>=0.94, Supplemental Figure 5). In addition adding case-control status for incident insulin use only marginally affected the results (data not shown).

Comparing the frequency of different types of sRNAs identified showed that while snoRNAs represent 10% of the total sRNA transcriptome in our RNA-seq study they are the most abundant class
associated with kidney function (≥31%, Table 2). A similar picture is observed in the full models (Supplemental Table 4).

The pairwise correlation between the 12 sRNAs associated with diabetic CKD was very modest (Figure 1D). The only exception were U8 and SNORD118, which is likely due to the fact that they are encoded in the same host gene (TMEM107) on chromosome 17. Results of these analyses for eGFR and UACR are shown in Supplemental Figure 6.

In a next step we examined the tissue-specific expression patterns of the identified sRNAs using publicly available expression data in 11 tissues or body fluids (Figure 1E). This illustrated that many sRNAs show tissue-specific expression patterns. For the 12 sRNAs associated with CKD only mir-143-3p showed its highest expression in kidney tissue suggestive of a direct role in kidney function. piRNA-019825 was the only sRNA that showed high expression in urine. Interestingly, several sRNAs were highly expressed in other tissues such as liver and colon. This suggest that these sRNAs might be involved in inter-organ communication in the regulation of kidney function. It is further interesting that several sRNA show their highest expression in liver or colon. Similar observations were made for eGFR and UACR (Supplemental Figure 7).

Discussion

In this study we show that circulating sRNAs are associated with prevalent CKD in people with type 2 diabetes. In addition, we show that circulating sRNAs are associated with the two components of CKD: eGFR and UACR. Importantly these sRNAs do not only represent miRNAs but also other classes of small non-coding RNAs, particularly snoRNAs. Using publicly available data we show that several of these are highly expressed in colon or liver but not kidney. This suggests that circulating sRNAs may be involved inter-organ crosstalk to regulate kidney function or reflect differences in clearance of sRNAs/vesicles from the circulation.

sRNA tissue specific expression

Interestingly, only one of the significant sRNAs was found at high levels in urine showing that the circulating sRNA transcriptome and the sRNA urine transcriptome are not similar or interchangeable. Besides one sRNA highly expressed in kidney (miRNA-143-3p) we found that many significant sRNAs are overexpressed in one or a few other tissues we tested, where especially liver and colon seem to be overrepresented among significant sRNAs. We speculate that these sRNAs represent interesting candidates for further studies as they are potentially involved in organ crosstalk in the regulation of and or response to kidney dysfunction. On the other hand it cannot be excluded that clearance of (specific) sRNAs or vesicles is affected by the disease.

Associations between kidney function and sRNAs
Although a substantial number of studies have examined circulating sRNAs associated with kidney function in persons with diabetes, most studies are relatively small (n<100) and focused only on microRNAs. From these studies a large number of miRNAs have emerged that are putatively associated with kidney function. However, probably due to small sample sizes, heterogeneity in endpoints and methodology used there is limited consistency between studies as shown in various meta-analyses (11; 25; 26). We have performed a large untargeted sRNA-seq study in a large well defined sample of persons with T2D with and without CKD from the Hoorn DCS study. Although miRNAs represent 64% of all sRNAs detected in our study we found that for the outcomes studied only between 33 and 44% of the significant sRNAs belong to the miRNAs. Highlighting the potential of the other classes of sRNAs in the etiology of CKD in diabetes. Of the twelve sRNAs significantly associated with CKD four were miRNAs, seven belonged to the snoRNAs and we also identified a single piRNA.

**miRNAs**

MicroRNAs are involved in the post-transcriptional regulation of gene expression and are known to be involved in crosstalk between (distant) cells and tissues (5; 6; 10). As such, there is an ever growing interest to study their role in disease development and as biomarkers of disease and disease progression.

In our study one miRNA, miR-145-5p, was associated with both CKD and eGFR. Levels of this miRNA are lower in both CKD and persons with low eGFR. Analysis of the expression pattern of this miRNA showed its highest expression in colon and moderate expression in thyroid but much less in other tissues, including kidney or urine. In a previous meta-analyses of this miR it was downregulated in those with T2D and diabetic nephropathy (26). In addition, it was shown that this miRNA was lower in those with faster eGFR decline (27). Functional studies support a role for this miR in CKD in persons with diabetes and in animal studies (28; 29). miR-143-3p, associated with CKD and UACR, showed its highest expression in kidney suggesting a direct role in kidney function. In previous miRNA studies it was shown that this miR was down-regulated in renal tissue of animal models of (diabetic nephropathy) DN (26). This was further supported by functional studies in an animal model showing impairment of the glomerular filtration barrier (30). miR-143-5p was the only miRNA that remained significant after further adjustment for baseline BMI, diabetes duration, HbA1c, HDL, SBP, smoking, eGFR and UACR in the full models, showing its robustness as a biomarker for CKD in diabetes. To the best of our knowledge this miRNA has not been linked to CKD previously. These three miRNAs are all encoded in the same region on chromosome 5q32. In a GWAS aiming to identify blood miRNA QTLs both miR-143-3p and miR-145-5p but not miR-143-5p or miR-145-3p were shown to be under control of the same SNPs (24). It was further shown that they exert a similar function in regulation of genes involved in thrombosis and angiogenesis. However, we observed quite a different expression pattern in kidney, and other tissues, suggestive that their function in kidney might not be identical. The fourth significant miRNA, miR-4685-3p also associated with UACR and was expressed at very low levels in all tissues studied. In addition, we could not find any other studies investigating this miR or its role in kidney function.
function. Further studies are thus warranted to elucidate its function. In addition to the association with CKD we also found many miRNAs associated with eGFR and or UACR. A large number of which remained significant in the full models. This includes miRNAs that were also found in previous studies, such as let-7a-5p, miR-192-5p, miR-194-5p, miR-214-3p, miR-374a-3p, miR-451a (11; 25; 26), as well as novel ones. Interestingly, of the 123 miRNAs found in the base models for eGFR and UACR none of them were associated with both phenotypes suggesting that the associated miRNAs are specific to either defect in kidney function but not both. If confirmed in independent studies this provides interesting starting points for further functional studies investigating their role in either condition.

**piRNAs**

PIWI-interacting RNAs are involved in silencing of transposons but they can also be involved in the regulation of gene expression in a miRNA-like manner (31). Only one piRNA, piRNA-019825 (DQ597218), was significantly associated with CKD. Previously it was shown that this piRNA is one of the most predominant piRNA species in urine and plasma (32). It was the only of the twelve significant sRNA that showed its highest expression in urine. Whether piRNA-019825 is also associated with CKD in urine requires further investigation, if so, it may prove useful as a non-invasive biomarker of CKD in diabetes. To the best of our knowledge this piRNA has not been studied previously in relation to kidney function. Interestingly, we also find various other piRNAs associated with either eGFR and or UACR (11-14%). Several of which remained significant in the full models. Suggesting that they may be useful early markers of CKD in diabetes. Interestingly, a recent study showed that DNA methylation near the *PIWIL1* gene was associated with diabetic CKD (33). PIWI proteins, like *PIWIL1*, use piRNAs to form the piRNA-induced silencing complex, resulting in RNA degradation and epigenetic silencing. Based on these observations it will be interesting to elucidate whether the piRNAs are causally involved in CKD or merely biomarkers of kidney dysfunction.

**snoRNAs**

Small Nucleolar RNAs (snoRNA) are predominantly localized to the nucleolus where they have a function in rRNA biogenesis, structure and activity (34). There are two major types of snoRNAs containing either box C/D or box H/ACA motifs. In our study we only detected box C/D snoRNAs associated with CKD. Box C/D snoRNAs are involved in 2’-O-methylation of specific sequences in target RNA sequences, mainly rRNA. In addition other functions have been ascribed to the snoRNAs such as mRNA splicing and editing. SnoRNA derived fragments have also been described and it has been speculated that they exert miRNA like functions (34-37).

Although snoRNAs represent only ten percent of the total circulating sRNA transcriptome in our study they represent fifty-eight percent of the sRNAs found associated with CKD suggesting that they may function as early biomarkers of CKD in diabetes. Whether they are causally involved in regulation of kidney function needs to be shown. Of the seven snoRNAs that were associated with CKD
several showed their highest expression in liver or colon whilst in general kidney and urine levels of these snoRNAs are low. Three snoRNAs, U8, SNORD107 and SNORD25 were associated with all three phenotypes but neither had previously been linked to kidney function. SnoRNAs SNORD118, SNORD24, SNORD87 and SNORD15B where associated with CKD and UACR but not eGFR. At present there are no studies that have investigated the role of snoRNAs in diabetic CKD. There are, however, several studies showing the association of snoRNAs with cardiometabolic and other traits (34; 38). This confirms that functional defects in snoRNA function can have a causal role in diseases. snoRNAs U8 and SNORD118 both remained significant in the full models. These snoRNAs are encoded within the TMEM107 gene on chromosome 17p13.1, where U8 is transcribed from SNORD118 locus. Mutations in SNORD118 cause a rare genetic cerebral microangiopathy called leukoencephalopathy with brain calcifications and cysts (LCC; also known as Labrune syndrome) (39). However, based on current evidence this a strictly neurological disease with no evidence for involvement of other tissues or the kidneys. Mutations in TMEM107 though have been linked to various syndromes that often also involve the kidney, mostly through cyst development (40; 41). Further replication and functional studies are thus warranted to elucidate the role of U8/SNORD118 and other snoRNAs in diabetic CKD and the putative mechanism behind our observations. Given their expression pattern in tissues other than kidney it will also be of interest to elucidate the origin of the circulating snoRNAs and their potential role in communication between organs. In addition to the snoRNAs described above various other snoRNAs are associated with eGFR and UACR (31-40% of total number of significant sRNA). Almost all remained significant in the full models and thus represent interesting starting points for further replication and functional studies.

Members of other sRNA classes

Although other classes of sRNAs, such as Y_RNA derived fragments, tRNA fragments, IncRNAs and snRNAs represent roughly twenty percent of all circulating sRNA detected we did not find any associations with CKD. We did find a small number of associations with eGFR and UACR though. The IncRNA GAS5 was associated with both eGFR and UACR which persisted in the full model. GAS5- IncRNA is a snoRNA host gene which encodes a large number of snoRNAs. In addition, a number of studies have linked this IncRNA to kidney function (42-44). Interestingly, this IncRNA is also a binding partner for SNORD118 described above (45).

Because there are to our best knowledge no GWASes for sRNAs other than for miRNAs we performed a GWAS on our own samples. However, we did not identify any genome wide significant cis QTLs (data not shown). Larger, better powered, studies will be needed to elucidate the role of genetics in the regulation of sRNA expression levels.

Strengths and limitations
Our study has several strengths. It is one of the largest circulating sRNA profiling studies studying the role of circulating sRNAs in diabetic CKD. This improves power and reduces the chance of false positives. In addition, it uses untargeted RNA sequencing allowing not only the study of miRNAs but also several other classes of sRNA showing that especially a large number of these other sRNAs associated with CKD and its components eGFR and UACR. The use of early, mild cases of CKD facilitates the identification of early markers of disease and may help to introduce early interventions aiming to preserve kidney function in those at risk. The use of cases with mild CKD is on the other hand also a limitation as it has to be established if our results can also be extrapolated to those with more advanced stages of CKD. We used a total RNA isolation procedure and as such we were not able to investigate if the sRNAs are carried by extracellular vesicles or are associated with carriers such as HDL and Ago-2. Such detailed studies are also needed to establish the cells/tissue of origin of the sRNAs and elucidation of the mechanism and potential crosstalk between tissues via the sRNAs as has already been well documented for miRNAs (5). The use of plasma for RNA isolation is another limitation. Although from a biomarker perspective this isn’t a limitation, from a functional perspective the role of the sRNAs we identified needs to be established in the relevant cells and tissues. Another limitation of our study is its cross-sectional design. Prospective studies are needed to investigate if these circulating sRNAs can be used in the early prediction of those at risk of diabetic CKD before onset.

In this study we have shown that several circulating sRNAs are associated with CKD in subjects with diabetes. Interestingly, only a relatively small fraction belong to the well-studied class of non-coding RNAs, the microRNAs. This highlights the potential of other classes of non-coding sRNAs and especially the snoRNAs in diabetic kidney disease development. Interestingly, many of the sRNAs identified are not highly expressed in kidney and may thus be used in between tissue crosstalk in the regulation of kidney function. Future prospective and functional studies are needed to establish the role of circulating sRNAs in subjects with diabetes and chronic kidney disease.

Acknowledgements
We would like to thank all staff and participants of the Hoorn Diabetes Care System for their support and participation.

Funding
This work was supported by a diabetes breakthrough grant from the Dutch Diabetes Research Foundation and ZonMw (grant number 459001015) and a grant from the EFSD/Boehringer Ingelheim European Research Programme on “Multi-System Challenges in Diabetes” 2021. The study sponsors/funders were not involved in the design of the study; the collection, analysis, and interpretation of data; writing the report; and did not impose any restrictions regarding the publication of the report.
Contribution statement

LMtH, JWJB and RCS designed the study. JAdK and SJC isolated RNA and prepared the samples for RNAseq. MTB and JWJB were responsible for phenotypic data acquisition and preparation. LMtH, GAB and RCS prepared, analyzed and interpreted the data. LMtH and RCS wrote the first draft of the manuscript. LMtH and RCS are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors critically read and revised the manuscript and approved the final version of the manuscript.

References


34. Deogharia M, Majumder M: Guide snoRNAs: Drivers or Passengers in Human Disease? Biology (Basel) 2018;8
Table 1. Clinical characteristics of the study population.

<table>
<thead>
<tr>
<th></th>
<th>all</th>
<th>No CKD</th>
<th>CKD</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>475</td>
<td>406</td>
<td>69</td>
</tr>
<tr>
<td>Gender (%Female)</td>
<td>45</td>
<td>46</td>
<td>42</td>
</tr>
<tr>
<td>Age (years)</td>
<td>61.4 ± 9.1</td>
<td>61.0 ± 8.8</td>
<td>63.8 ± 10.1*</td>
</tr>
<tr>
<td>BMI (Kg/m²)</td>
<td>30.7 ± 6.0</td>
<td>30.3 ± 5.6</td>
<td>32.7 ± 7.3*</td>
</tr>
<tr>
<td>Age-at-Diagnosis (years)</td>
<td>59.3 ± 9.0</td>
<td>58.8 ± 8.7</td>
<td>61.8 ± 9.8*</td>
</tr>
<tr>
<td>Diabetes duration (years)</td>
<td>2.2 ± 1.1</td>
<td>2.2 ± 1.1</td>
<td>2.0 ± 1.1</td>
</tr>
<tr>
<td>Fasting glucose (mmol/L)</td>
<td>8.0 ± 1.6</td>
<td>8.0 ± 1.7</td>
<td>8.1 ± 1.4</td>
</tr>
<tr>
<td>HbA1c (mmol/mol)</td>
<td>46 (42, 51)</td>
<td>46 (42, 51)</td>
<td>46 (43, 54)</td>
</tr>
<tr>
<td>HDL (mmol/L)</td>
<td>1.19 ± 0.32</td>
<td>1.20 ± 0.33</td>
<td>1.11 ± 0.25</td>
</tr>
<tr>
<td>eGFR (ml/min/1.73 m²)</td>
<td>84 ± 16</td>
<td>87 ± 13</td>
<td>70 ± 22*</td>
</tr>
<tr>
<td>eGFR ≤ 60 (%)</td>
<td>34 (7)</td>
<td>0</td>
<td>34 (50)</td>
</tr>
<tr>
<td>Urine Albumin-to-Creatinine ratio (mg/mmol)</td>
<td>0.48 (0.20, 1.00)</td>
<td>0.41 (0.18, 0.79)</td>
<td>3.99 (0.79, 10.5)*</td>
</tr>
<tr>
<td>Albuminuria (%)</td>
<td>43 (9)</td>
<td>0</td>
<td>43 (62)</td>
</tr>
<tr>
<td>Chronic Kidney Disease (%)</td>
<td>69 (15)</td>
<td>0</td>
<td>69 (100)</td>
</tr>
<tr>
<td>Diastolic blood pressure (mm Hg)</td>
<td>78 ± 9</td>
<td>78 ± 9</td>
<td>79 ± 10</td>
</tr>
<tr>
<td>Systolic blood pressure (mm Hg)</td>
<td>139 ± 18</td>
<td>139 ± 18</td>
<td>142 ± 20</td>
</tr>
<tr>
<td>Smoking (current/ever/never)</td>
<td>92/130/250</td>
<td>76/107/220</td>
<td>16/23/30</td>
</tr>
<tr>
<td>Medication use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metformin (%)</td>
<td>340 (72)</td>
<td>287 (71)</td>
<td>53 (77)</td>
</tr>
<tr>
<td>SU (%)</td>
<td>130 (27)</td>
<td>105 (26)</td>
<td>25 (36)</td>
</tr>
<tr>
<td>Other glucose-lowering (%)</td>
<td>13 (3)</td>
<td>12 (3)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Antihypertensive medication (%)</td>
<td>319 (67)</td>
<td>260 (64)</td>
<td>59 (86)*</td>
</tr>
<tr>
<td>Lipid lowering medication (%)</td>
<td>320 (67)</td>
<td>271 (67)</td>
<td>49 (71)</td>
</tr>
</tbody>
</table>

Data are mean ± SD, n (%) or median (IQR). All measures taken at the time of blood sampling.
* P≤0.05 versus those without CKD.
Table 2. Frequency of different types of sRNAs associated with the three endpoints.

<table>
<thead>
<tr>
<th></th>
<th>Total # of sRNA detected</th>
<th>CKD</th>
<th>eGFR</th>
<th>UACR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Total # sRNAs</td>
<td>938</td>
<td>100</td>
<td>12</td>
<td>73</td>
</tr>
<tr>
<td>miRNA</td>
<td>602</td>
<td>64</td>
<td>4</td>
<td>33</td>
</tr>
<tr>
<td>snoRNA</td>
<td>96</td>
<td>10</td>
<td>7</td>
<td>58</td>
</tr>
<tr>
<td>lincRNA</td>
<td>93</td>
<td>10</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>piRNA</td>
<td>45</td>
<td>5</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Y_RNA fragments</td>
<td>33</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>others</td>
<td>26</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tRNA fragments</td>
<td>24</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>snRNA</td>
<td>12</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>scaRNA</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Overview of twelve sRNAs associated with prevalent CKD

<table>
<thead>
<tr>
<th>sRNA</th>
<th>Prevalent CKD</th>
<th>eGFR</th>
<th>UACR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LogFC</td>
<td>P$_{FDR}$</td>
<td>logFC</td>
</tr>
<tr>
<td>SNORD118</td>
<td>-0.873</td>
<td>2.35x10^-3</td>
<td>0.013</td>
</tr>
<tr>
<td>U8</td>
<td>-0.851</td>
<td>2.35x10^-3</td>
<td>0.014</td>
</tr>
<tr>
<td>hsa-miR-143-5p</td>
<td>-0.479</td>
<td>8.04x10^-3</td>
<td>0.007</td>
</tr>
<tr>
<td>SNORD107</td>
<td>-1.283</td>
<td>2.59x10^-2</td>
<td>0.028</td>
</tr>
<tr>
<td>SNORD24</td>
<td>-0.794</td>
<td>2.59x10^-2</td>
<td>0.015</td>
</tr>
<tr>
<td>SNORD87</td>
<td>-1.188</td>
<td>3.03x10^-2</td>
<td>0.018</td>
</tr>
<tr>
<td>hsa-miR-145-5p</td>
<td>-0.321</td>
<td>3.03x10^-2</td>
<td>0.008</td>
</tr>
<tr>
<td>hsa-miR-143-3p</td>
<td>-0.328</td>
<td>3.40x10^-2</td>
<td>0.002</td>
</tr>
<tr>
<td>SNORD15B</td>
<td>-0.756</td>
<td>4.88x10^-2</td>
<td>0.014</td>
</tr>
<tr>
<td>hsa_piR_019825</td>
<td>gb</td>
<td>DQ597218</td>
<td>-0.463</td>
</tr>
<tr>
<td>hsa-miR-4685-3p</td>
<td>-0.430</td>
<td>4.88x10^-2</td>
<td>0.003</td>
</tr>
<tr>
<td>SNORD25</td>
<td>-0.905</td>
<td>4.88x10^-2</td>
<td>0.025</td>
</tr>
</tbody>
</table>

LogFC is the log2 of the fold change. P$_{FDR}$ represents the P value after correction for multiple hypothesis testing using the Benjamini-Hochberg procedure.
Figure 1. sRNAs significantly associated with CKD.

Panel A. Volcano plot showing the association between sRNAs and CKD in the base model. sRNAs in blue are significantly lower in cases with CKD after multiple hypothesis testing (FDR≤0.05). logFC is the natural logarithm of the fold change between cases and controls. Panel B. SNORD118 levels in those with or without CKD. CPM = counts per million reads. Panel C. Overlap between sRNAs significantly associated with CKD, eGFR or UACR in the base model. Panel D. Pairwise correlations for the twelve sRNA associated with CKD in the base model. Colors represent Spearman’s rho. Panel E. Expression pattern for sRNAs associated with CKD in publicly available datasets for 11 tissues of interest. Colors represent the Z-scaled expression of a sRNA where blue indicates relative high expression and red relative low expression.