Research Article

Various types of circulating small non-coding RNAs associate with kidney function in people with type 2 diabetes

Running title: Small non-coding RNA and kidney function in type 2 diabetes

LM ‘t Hart¹,²,³,⁴ (ORCID 0000-0003-4401-2938), JA de Klerk¹,⁵,⁶, GA Bouland⁸, JHD Peerlings¹, M.T. Blom⁴,⁵, S J Cramer², R Bijkerk⁵,⁶, JWJ Beulens¹,², G A Bouland¹, M.T. Blom⁵, S J Cramer², R Bijkerk⁵,⁶, J H D Pe e rlin gs¹, M.T. Blom⁵, S J Cramer², R Bijkerk⁵,⁶, J W J Beulens¹,², R C Slieker¹,²,⁵

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, 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 Internal Medicine, Division of Nephrology, Leiden University Medical Center, Leiden, The Netherlands
⁶ Einthoven Laboratory for Vascular and Regenerative Medicine, Leiden University Medical Center, Leiden, The Netherlands
⁷ Department of General Practice and Elderly Care Medicine, Amsterdam University Medical Center, Amsterdam, The Netherlands

Author for correspondence:

Leen M ‘t Hart, PhD
Leiden University Medical Center
Department of Cell and Chemical Biology
Albinusdreef 2
2333ZA Leiden, The Netherlands
+31 715269796
lmtthart@lumc.nl

Keywords: chronic kidney disease, type 2 diabetes, small non-coding RNA, microRNA, snoRNA

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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

Introduction Micro- and macrovascular complications are common among persons with type 2 diabetes. Recently there has been growing interest to investigate the potential of circulating small non-coding RNAs (sncRNAs) as biomarkers and drivers of the development of diabetic complications. In this study we investigate to what extent circulating sncRNAs levels associate with prevalent diabetic kidney disease (DKD) in persons with type 2 diabetes.

Methods Plasma sncRNAs levels were determined using sRNA-seq, allowing detection of miRNAs, snoRNAs, piRNAs, tRNA-fragments and various other sRNA classes. We tested for differentially expressed sncRNAs in persons with type 2 diabetes, with DKD (n=69) or without DKD (n=405). In secondary analyses, we also tested the association with eGFR, albuminuria (UACR) and the plasma proteome.

Results In total seven sncRNAs were associated with prevalent DKD (P_{FDR}≤0.05). Although miRNAs represent the majority of the sncRNAs measured (64%) only one miRNA was significantly associated while the majority of the significant sncRNA belonged to the snoRNA class (71%). Similar results were observed for eGFR and UACR. In addition, the seven sncRNAs, and especially the piRNA piR-019825, were associated with plasma levels of 80 proteins of which several have known associations with kidney function.

Conclusion We have identified novel small non-coding RNAs, primarily from classes other than microRNAs, that are associated with DKD in persons with type 2 diabetes. Further studies, in particular for snoRNAs and the piRNA piR-019825 that have never been studied before, are warranted to explore their biomarker potential and elucidate their biological role in kidney function.
Introduction

Diabetic kidney disease (DKD) 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 DKD 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 DKD development and have yielded novel biomarkers and drug targets.

Circulating small non-coding RNAs (sncRNA) are mostly contained in extracellular vesicles or are bound to RNA-binding proteins and lipoproteins and originate from various tissues including tissues important in kidney function [5]. The composition of sncRNAs in the circulation is dependent on tissue homeostasis and can be altered by disease development. Whereas previously it was thought that sncRNAs 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, splicing and ribosomal RNA maturation [5]. Irrespective of the mechanism involved, sncRNAs may have an impact on the development and progression of DKD when deregulated. Importantly, each tissue releases its own combination of tissue-specific and more ubiquitous sncRNAs into the bloodstream. These circulating sncRNAs can be taken up by other tissues and affect target tissue function [5].

Although a substantial number of studies have examined circulating microRNAs associated with kidney function in persons with diabetes, most studies are relatively small (n<100). 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 [6-8].

To the best of our knowledge there are currently no studies that have systematically examined the other classes (biotypes) of small ncRNA. Therefore, we have performed a large untargeted small non-coding RNA-seq study in a large well defined sample of persons with T2D with and without DKD from the Hoorn DCS study. Importantly, we go beyond the current state-of-the-art by not only investigating microRNAs but also other sncRNAs classes, including besides miRNAs, piRNAs, tRNA fragments (tRFs), lncRNA fragments, Y-RNAs and snoRNAs. For this we used next-generation sequencing to capture the whole sncRNA transcriptome allowing for a hypothesis-free, comprehensive analysis to study the role of all types of sncRNAs in diabetic kidney disease [9].
Methods

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 14,000+ individuals with type 2 diabetes who receive structured diabetes care [10]. Participants visit 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 a subsample from the whole DCS biobank (n=475) using the following inclusion criteria: 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. [10]. 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 [11]. 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 snRNAs 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 to further enrich for snRNAs 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 excRpt pipeline [12], 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), snRNA, snoRNA, scaRNA and various
other sncRNA species. Using respectively miRBase v22 [13], piRNABank v1 [14], Gencode (version 38) [15], circBase [16] and GtRNAdb [17]. One sample did not pass the QC thresholds.

**Taqman validation of the RNA-seq results**

We used Taqman assays (Fisher Scientific, NL) to verify the validity and reproducibility of the small RNA-seq results. We choose four miRNAs with a range from low to relative high expression in our study sample (miR-542-3p, miR-323b-3p, miR-186-5p and miR-423-5p) which were measured in 12 plasma samples for which sncRNA-seq data were available. The Taqman assays were run according to the manufacturer’s instructions. miR-191-3p expression was used as the reference (housekeeping gene) in both the Taqman and small RNA-seq datasets as suggested by the manufacturer.

**Expression patterns of small RNAs in tissues**

Publicly available raw sncRNA-seq data for various metabolic tissues were downloaded from the Gene Expression Omnibus (GEO) [18]. The tissues for which sncRNA-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 sncRNAs in these tissues. Data was z-scaled in figures.

**Plasma proteomics.**

Plasma proteomic data was available for 589 participants of the Hoorn DCS study and generated using the SomaLogic SOMAscan platform (Boulder, CO) as previously described [19]. After quality control 1195 protein levels were available for analysis. For 199 subjects we also had the corresponding sncRNA-seq data from the same sampling date.

**Statistical analysis**

The R-package edgeR (v3.42.4) was used to identify differentially expressed sncRNAs [20]. A negative binomial generalized log-linear model was applied using edgeR’s function estimateGLMRobustDisp. The base model was unadjusted. A fully adjusted model was adjusted for BMI, diabetes duration, HbA1c, HDL, SBP and smoking unless otherwise stated. Age and sex were excluded from the model as these are part of the definition of eGFR. UACR was log10-transformed before analysis. In sensitivity analyses, we additionally included use of antihypertensive medication and or eGFR or UACR. The Benjamini-Hochberg procedure was used to adjust for multiple testing and an FDR below 0.05 was considered significant. The correlation between the sncRNA transcriptome and plasma proteome
was tested using a linear model with the log transformed proteomic measurements as outcome and
the sncRNA transcriptome as predictors with adjustment for age, BMI, sex and technical covariates.
In this exploratory analysis we used an arbitrary significance threshold of \( p \leq 1.00 \times 10^{-3} \).

**Results**

*Clinical characteristics*

The clinical characteristics of our study group are shown in Supplemental Table 2. 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 DKD at the time of blood sampling was 15%.

*sncRNA biotype abundance in our samples*

After quality control, 938 sncRNAs were expressed above the threshold of a median \( \geq 5 \) copies in our
data, which included miRNAs, piRNA, Y RNA fragments, lncRNA fragments, tRNA fragments,
snoRNAs, snRNAs and various other sRNA species. miRNAs, Y-RNAs and tRFs were the most abundant
sncRNAs biotypes comprising respectively 57.6, 34.9 and 4.8 percent of all reads (Table 1). Other
sncRNA biotypes were much less abundant (\( \leq 1.4\% \)). There were no significant differences in the total
number of reads per biotype between cases and controls (all \( P > 0.4 \)).

*Taqman validation shows good reproducibility*

We next sought to verify validity and reproducibility of the sncRNA-seq. Taqman assays for sncRNAs
with high, medium and low expression were used and these showed a good correlation with the
RNA-seq results (\( r^2 = 0.86 \), Supplemental Fig. 2).

*sncRNAs associate with DKD, eGFR and UACR*

Seven sncRNAs were significantly associated with prevalent DKD in the unadjusted model (\( P_{\text{FDR}} \leq 0.05 \),
Table 2, Fig. 1A). All identified sncRNAs were downregulated in DKD cases. The top hit was snoRNA
U8 (log2 fold change = -0.859, \( P_{\text{FDR}} = 1.90 \times 10^{-3} \), Fig. 1B). Five of these sncRNAs also showed a significant
trend for lower levels in those with more advanced stages of CKD (\( P_{\text{FDR}} \leq 0.05 \), Supplemental Fig. 3).
However, due to the small sample size this was non-significant when analyzed only in the cases.
Respectively, 107 and 220 sncRNAs were significantly associated with eGFR and UACR, which are
used for the DKD classification (Table 3, Supplemental Table 3, Supplemental Fig. 4 and 5). The
overlap between the sncRNAs associated with DKD and its components eGFR and UACR is shown in Fig. 1C and Table 2. snoRNA U8 overlapped between all three endpoints used and piR_019825 overlapped between eGFR and DKD. SNORD118, SNORD24, SNORD107, SNORD87 and miR-143-5p all overlapped between UACR and DKD.

In the full model, including adjustment for BMI, diabetes duration, HbA1c, HDL, SBP and smoking, respectively 7, 153 and 205 sncRNAs remained significantly associated with DKD, eGFR and UACR (Supplemental Table 4, Supplemental Fig. 6; 57%; 95% and 84% overlap with the unadjusted model for DKD, eGFR and UACR respectively). U8, SNORD118, piR_019825 and miR-143-5p were significant in both the base and full models. 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 ($r^2$>0.92, Supplemental Fig. 7).

Certain types of sncRNAs are enriched in DKD

Comparing the frequency of different types of sncRNAs identified showed that while snoRNAs represent 10% of the total sncRNA transcriptome in our RNA-seq study they are the most abundant class associated with prevalent DKD (71%, OR=21.9, $P=2.4\times10^{-4}$, Table 1). Enrichment for snoRNAs is also seen for UACR (OR=3.8, $P=8.4\times10^{-13}$). Associations with miRNAs on the other hand are less frequent whilst piRNAs also seem enriched (Table 1). A similar pattern is again seen for UACR and to a lesser extend for eGFR. In the full model a largely similar picture is observed (Supplemental Table 5).

The pairwise correlation in expression between the seven sncRNAs associated with diabetic DKD was very modest (Fig. 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 Fig. 8.

sncRNAs are tissue-enriched and correlate with plasma protein levels

In a next step, to elucidate the putative tissues of origin, we examined the tissue expression patterns of the identified sncRNAs. We used publicly available expression data in 11 tissues or body fluids (Fig. 1E). This illustrated that some sncRNAs show tissue-enriched expression patterns. For the seven sncRNAs associated with DKD none had its highest expression in kidney tissue. piRNA-019825 was the only sncRNA that showed high expression in urine. Similar observations were made for eGFR and UACR associated sncRNAs (Supplemental Fig. 9).
Finally, since the function of most of the sncRNAs we found in our study is unknown we compared in an exploratory study the expression levels of the sncRNAs with expression of 1195 proteins measured in plasma from the same date to identify potential targets or (co-)regulated proteins or pathways. As shown in Supplemental Table 6 we found for the seven sncRNAs 105 nominal significant pairwise correlations (p≤1x10^{-3}) involving 80 proteins. The strongest association was observed between the piRNA piR-019825 and TNF sR-1 (TNFRSF1A, B=\beta=-0.080 (-0.112-0.040), p=4.29x10^{-6}). piRNA-019825 also showed the largest number of correlations (n=32, Fig. 2A) whereas miR-143-5p showed only four. Interestingly, among the proteins associated with piR-019825 was Cystatin C, a known marker of kidney function. Next we assessed whether the 80 proteins were also significantly associated with eGFR or UACR (FDR≤0.05). Fourteen were associated with eGFR and two with UACR. Interestingly, ten of the fourteen eGFR associated and both UACR associated proteins were also associated with piR-019825 expression (Fig. 2A). Protein-protein interactions (PPI) and pathway enrichment was subsequently used for each set of sncRNA associated proteins. This showed that the 32 proteins associated with piR-019825 show a strong PPI enrichment (P=7.0x10^{-4}, Fig. 2B). Pairwise correlation between the proteins is shown in Fig. 2C. Enrichment analysis showed enrichment for KEGG pathways such as Cytokine-cytokine receptor interaction (P=6.7x10^{-6}, NF-kappa B signaling pathway (P=4.6x10^{-3}) and PI3K-Akt signaling pathway (P=3.2x10^{-2}). For the other smaller sncRNA associated protein sets we note that several proteins have previously been implicated in DKD. Fifteen proteins were correlated with two or more sncRNAs. Especially the snoRNAs seem to have multiple overlapping protein associations (Supplemental Fig. 10). Several of these proteins have already previously been implicated in kidney function (including CTAck (CCL27), Chemerin (RARRES2), visfatin (NAMPT), PDE4D, PDGFβ, SDF-1 (CXCL12), sICAM-3, SMAD2) whereas the first two were also in our study significantly associated with eGFR.

**Discussion**

In this study we show that circulating sncRNAs are associated with prevalent DKD in people with type 2 diabetes. Importantly these sncRNAs do not only represent miRNAs but also other classes of small non-coding RNAs. In addition the sncRNAs correlate with proteins that function in pathways known to be involved in DKD/kidney function.

**Associations between kidney function and sncRNAs 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]. 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-143-5p, was associated with both DKD and UACR. Levels of this miRNA are lower in both DKD and persons with high UACR. Analysis of the expression pattern of this miRNA showed its highest expression in thyroid and heart but less in other tissues, including kidney or urine. miR-143-5p remained significant after further adjustment for baseline BMI, diabetes duration, HbA1c, HDL, SBP and smoking in the full models, suggesting it might have biomarker potential. Analysis of the co-expressed proteins showed enrichment for the ovarian steriodogenesis and GnRH signalling pathway but no clear relationship to DKD. However, there is quite some functional evidence from in vitro studies showing that this miRNA is associated with vascular pathology [21], an impaired glomerular filtration barrier [22] and with mesangial cell proliferation and fibrosis [23, 24].

In the full model three additional microRNAs were significant, miR-455-3p, miR-122-3p and miR-193b-3p. Previously it was shown that miR-455-3p was down regulated in human mesangial cells (HMC) and human proximal tubule epithelial cells (HK-2) under DKD stimulating conditions [25]. Interestingly, it was also shown that overexpression of this miR could alleviate the pathological changes such as renal fibrosis in a rat DN model [25]. In a mouse mesangial cell line miR-455-3p improved cell proliferation, inflammation and ECM accumulation induced by high glucose treatment [26]. Involvement of this miR in fibrosis in other tissues like liver and lung has also been reported. A previous report indicated that miR-122, a liver specific miR associated with liver injury, was reduced in patients with end stage renal disease but not in CKD when compared to healthy controls [27]. miR-193b-3p was identified as a potential biomarker for pre-diabetes and also for CKD after nephrectomy for renal cell carcinoma [28, 29].

In addition to the association with DKD we also found many miRNAs associated with eGFR and or UACR. A large number of which remained significant in the full models. Interestingly, of the 154 miRNAs found in the base models for eGFR and UACR only two 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 [30]. One piRNA, piRNA-019825 (DQ597218),
was significantly associated with DKD. Previously it was shown that this piRNA is one of the most abundant piRNA species in urine and plasma [31]. In our exploratory proteomics analysis we found 32 proteins whose expression was correlated with the levels of piR-019825 including cystatin C, a known marker for kidney function. Twelve of the proteins showed an association with eGFR or UACR. In addition, we found that many of the proteins have previously been found in studies of kidney function or function in disease relevant pathways such as the enriched cytokine-cytokine receptor pathway. This includes among the most strong associations proteins such as TNF sR-I (TNFRSF1A), DAN (NBL1), CCL16 and CFD which have all previously been associated DKD or with progression to ESRD [32-34]. Interestingly, we also find various other piRNAs associated with either eGFR and or UACR (10-14%). Most of which remained significant in the full models. Suggesting that they may be useful early markers of DKD. Of note, eGFR-associated piRNAs often show their highest expression in urine, whereas UACR associated piRNAs show low expression in urine, but often high levels in tissues like pancreas, liver and colon.

snoRNAs

Small Nucleolar RNAs (snoRNA) are predominantly localized to the nucleolus where they have a function in rRNA biogenesis, structure and activity [35]. 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 DKD. 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 piRNA or miRNA like functions [36]. Although snoRNAs represent only ten percent of the total circulating sRNA transcriptome in our study they represent seventy-one percent of the sncRNAs found associated with DKD suggesting that they may function as early biomarkers of DKD. The five snoRNAs that were associated with DKD are expressed in most tissues but two showed their highest expression in liver or colon whilst in general kidney and urine levels of these snoRNAs are much lower. SnoRNA U8 was associated with all three phenotypes but has not been linked to kidney function previously. SnoRNAs SNORD118, SNORD24, SNORD107 and SNORD87 where associated with DKD and UACR but not eGFR. Interestingly, especially the snoRNAs seem to be associated with the same set of proteins suggestive of coregulation of the snoRNAs and these proteins. Two of them (CTACK, RARE52) are also associated with eGFR in our study. Several other proteins associated with these snoRNAs have been implicated in kidney function previously. SnoRNAs U8 and SNORD118 both remained significant in the full models. In addition to the snoRNAs described above various other snoRNAs are associated with eGFR
and UACR (13-31% of total number of significant sRNA). Almost all remained significant in the full models.

At present there are no studies that have investigated the role of snoRNAs in diabetic DKD. Recently, it was shown that four snoRNAs were differentially expressed in both plasma and urine samples from 48 hypertensive persons with or without albuminuria [37]. SNORD116 and SNORD127, two highly correlated snoRNAs ($R^2=0.81$), were also measured in our study and both were significantly and negatively associated with UACR which is in contrast to this small previous study which showed an upregulation in plasma but down-regulation in urine.

**Members of other sncRNA classes**

Although other classes of sncRNAs, such as YRNA derived fragments, tRNA fragments (TRF), IncRNA fragments and sncRNAs represent roughly twenty percent of all circulating sRNA detected we did not find any significant associations with DKD. We did find a small number of associations with eGFR and UACR though.

**Strengths and limitations**

Our study has several strengths. It is one of the largest circulating sncRNA profiling studies studying the role of circulating sncRNAs in DKD. 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 sncRNA showing that especially these other types of sncRNAs associated with DKD and its components eGFR and UACR. The use of early, mainly mild cases of DKD facilitates the identification of early markers of disease and may help to introduce early interventions aiming to preserve kidney function in those at risk.

A limitation of our study is the use of a total RNA isolation procedure and as such we were not able to investigate if the sncRNAs are carried by extracellular vesicles or are associated with carriers such as HDL and Ago-2 [5]. The use of plasma for RNA isolation is another limitation. Although from a biomarker perspective this is not a limitation, from a functional perspective the role of the sncRNAs we identified needs to be established in the relevant cells and tissues. Another limitation of our study is its cross-sectional, observational design. As such we cannot exclude that (part of) the observed changes and associations are a reflection of the reduced kidney function in the participants.

In conclusion, in this study we have shown that several circulating sncRNAs are associated with DKD. Interestingly, only a relatively small fraction belong to the well-studied microRNAs. This highlights the potential of other classes of sncRNAs, such as snoRNAs and piRNAs, that have so far not been...
investigated in relation to kidney function before. Future prospective and functional studies are needed to establish the role of these circulating sncRNAs in subjects with diabetes and chronic kidney disease.
Acknowledgement (optional)

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

Statement of Ethics

Study approval statement: The study has been approved by the ethics committee of the Amsterdam University Medical Center, location VUmc (approval number 2007/57; NL14692.029.07)

Consent to participate statement: All participants provided informed consent prior to participating in this study.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

Funding Sources

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 proteomic analyses were funded by a grant from the Innovative Medicines Initiative 2 Joint Undertaking, under grant agreement no. 115881 (RHAPSODY). This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation program and EFPIA and is supported by the Swiss State Secretariat for Education, Research and Innovation (SERI), under contract no. 16.0097. The funders had no role in the design, data collection, data analysis, and reporting of this study.

Author Contributions

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. RB contributed to the critical interpretation of the results. 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.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are not publicly available
due restrictions in the informed consent but are available from the corresponding author on
reasonable request.
References


Figure Legends

Fig. 1. Panel A. Volcano plot showing the association between sncRNAs and DKD in the base model. sncRNAs in blue are significantly lower in cases with DKD after multiple hypothesis testing (FDR≤0.05). logFC is the Log2 of the fold change between cases and controls. Panel B. SNORD118 levels in those with or without DKD. CPM = counts per million reads. Panel C. Overlap between sncRNAs significantly associated with DKD, eGFR or UACR in the base model. Panel D. Pairwise correlations for the seven sncRNAs associated with DKD. Colors represent Spearman’s rho. Panel E. Expression pattern for sncRNAs associated with DKD in publicly available datasets for 11 tissues of interest. Colors represent the Z-scaled expression of a sncRNA where blue indicates relative high expression and red relative low expression.

Fig. 2. Panel A. Overview of the 32 proteins that associate with piR-019825 expression levels in plasma from persons with T2D (P≤1x10^-3). Proteins in green are also significantly associated with eGFR and those in blue are associated with UACR (PFDR≤0.05). Panel B. Protein-protein interaction network from the proteins correlated with piR-019825 levels as predicted by String-db.org (accessed Nov-2023). Panel C. Pairwise correlations between the 32 piR-019825 associated proteins.