Genetic Analysis of Image-Derived Phenotypes in Kidney Substructures

Lida Wang[1,2], Madeleine Cule[1], Ramprakash Srinivasan[1], Kevin Wright[1], Elena Sorokin[1], Anil Raj[1], Jimmy Bell[3]

1. Calico Life Sciences LLC, South San Francisco, CA, USA;
2. Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, USA
3. Research Centre for Optimal Health, School of Life Sciences, University of Westminster, London, UK

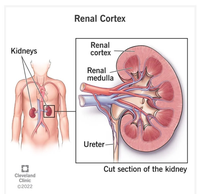
Chronic kidney disease (CKD) is a global health issue causing substantial morbidity and mortality. While factors like diabetes, hypertension, and vascular problems contribute to CKD, its origins remain unclear for many cases. Molecular mechanisms behind declining renal function and age-related structural changes in the kidney are not fully understood. This study employs MRI data from the UK Biobank to explore kidney substructure-disease connections and genetic factors. A 3D kidney segmentation neural network accurately measures substructure volumes, including cortex, medulla, vessels, and parenchyma, revealing their age-related changes and associations with diseases. Integrating kidney images, transcriptomics, and plasma proteomics provides insights into downstream effects of kidney risk genes, uncovers new disease-related biochemical markers, and enhances our grasp of genetic impacts on kidney substructures.

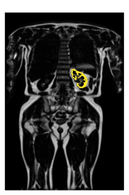
Project Goals:
1.Unravel the connection between kidney substructures, age and disease outcomes.
2.Investigate the genetic factors influencing kidney structural variation.

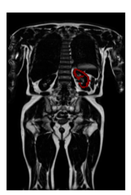
Introduction

Chronic kidney disease (CKD) stands as a significant global health concern, contributing to substantial morbidity and mortality rates worldwide. In developed nations, the prevalence of CKD exceeds 10% among the adult population. The underlying molecular mechanisms triggering the gradual decline in renal function lack comprehensive understanding, hampering the development of effective drug targets.

Aging introduces intricate structural alterations to the kidneys, influencing overall volume, cortical and medullary dimensions, sinus fat distribution, and cyst formation, all of which correlate with diminishing kidney function. The intricate interplay between structural modifications, molecular shifts, and functional changes poses a challenging puzzle to decipher. While the investigation of this interplay is common in the context of chronic kidney disease, it has been less explored in healthy individuals. The exploration of genetic factors contributing to these structural variations remains an uncharted territory.





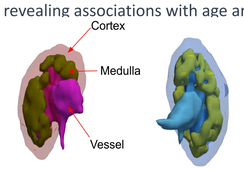


In this study, we harness the imaging data from the UK Biobank to meticulously analyze kidney substructures. Our primary objective is to untangle the intricate relationship between kidney substructures, aging, and disease outcomes. Additionally, we embark on a journey to uncover the genetic determinants that influence structural variations within the kidney. By shedding light on these intricate connections, we aim to advance our understanding of kidney health and contribute to the identification of potential therapeutic targets.

Subjects and methods

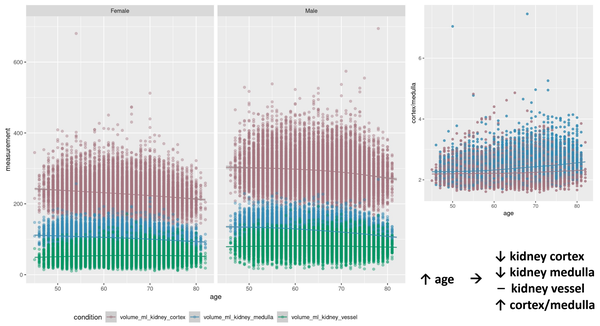
• Image-Derived Phenotypes (IDP)

Leveraging UK Biobank MRI data, we harnessed a convolutional neural network to segment kidneys, creating IDPs for kidney cortex, medulla and vessels. This enabled accurate volume estimation for various substructures and their ratios, revealing associations with age and kidney diseases.



	IDP
Volume	Cortex
	Medulla
	Vessel
	Parenchyma (cortex + medulla)
Ratio	Cortex/medulla

• IDP and Age



↑ age → ↓ kidney cortex
↓ kidney medulla
– kidney vessel
↑ cortex/medulla

• Pheno-wide association analysis

Larger **kidney cortex** volume is associated with increased risk of **diabetes** and **hypertension**

Smaller **kidney medulla** volume is associated with increased risk of **renal failure/ckd**

Larger **kidney vessel** volume is associated with increased risk of **kidney and urine stone**

A greater **ratio of cortex to medulla** is associated with increased risk of **renal failure and hypertension**

• Time-to-event model

COX proportional hazards model: $\lambda(t|X_i) = \lambda_0(t)\exp(\beta_1 X_{i1} + \cdots + \beta_p X_{ip})$

	Chronic renal failure		Hypertension		Kidney stone	
	exp(coef)	p-value	exp(coef)	p-value	exp(coef)	p-value
Kidney cortex (cil)	0.9901	5.74E-02	1.0012	3.24E-01	1.0077	3.34E-01
Kidney medulla (cil)	**0.9727**	**2.85E-02**	**1.0084**	**2.72E-03**	0.9850	4.23E-01
Kidney vessel (cil)	**1.0059**	**2.67E-02**	1.0007	2.82E-01	**1.0144**	**1.30E-04**
Cortex/medulla	0.9386	8.94E-01	**1.8250**	**1.43E-06**	0.6680	6.58E-01
Age (year)	1.0685	< 2e-16	0.9896	1.21E-08	0.9829	4.61E-02
Sex	1.6995	9.41E-05	1.3438	< 2e-16	2.3707	3.07E-05
BMI	1.1070	2.40E-15	0.3513	< 2e-16	1.0688	1.04E-03
Body surface area (cil)	6.6954	2.85E-06	1.1050	< 2e-16	0.2391	1.71E-02

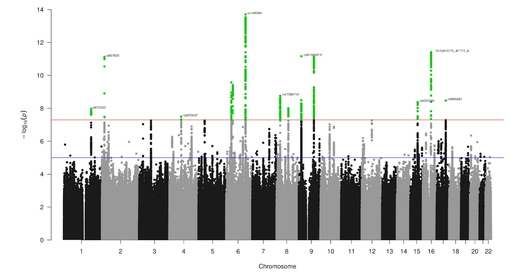
↑ kidney medulla → ↓ renal failure ↑ hypertension
↑ cortex/medulla → ↑ hypertension

↑ kidney vessel → ↓ renal failure ↑ kidney stone

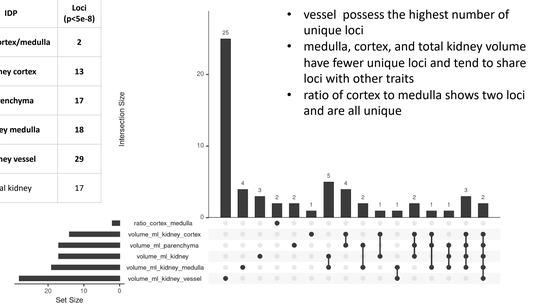
Results

• Genome-Wide Association Studies (GWAS)

Manhattan plot of Kidney cortex

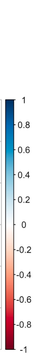


IDP	Loci (p<5e-8)
Ratio cortex/medulla	2
Kidney cortex	13
Parenchyma	17
Kidney medulla	18
Kidney vessel	29
Total kidney	17



- vessel possess the highest number of unique loci
- medulla, cortex, and total kidney volume have fewer unique loci and tend to share loci with other traits
- ratio of cortex to medulla shows two loci and are all unique

• Genetic correlation



Kidney relevant traits:

- eGFR (estimated glomerular filtration rate) to measure how well kidney are working
  high correlation
- Type 2 diabetes
  medium correlation
- Blood pressure(systolic and diastolic blood pressure, pulse pressure)
  low correlation
- Kidney stone
  medium correlation

	egfr	t2d	dbp	sbp	pp	kidney_stone
kidney	0.42	0.25	0.08	0.10	0.08	0.15
kidney_cortex	0.38	0.27	0.09	0.09	0.05	0.17
kidney_medulla	0.42	0.24	0.09	0.10	0.06	0.18
parenchyma	0.40	0.27	0.09	0.09	0.06	0.18
kidney_vessel	0.10	0.22	0.05	0.05	0.04	0.20
ratio_cortex_medulla	-0.17		-0.09	-0.04	-0.02	-0.06

Integration with Multi-omics

Multi-omics refers to the integration of diverse levels of data, including genomic, transcriptomic, epigenomic, proteomic, and other levels, to gain comprehensive insights into complex biological systems. The goal of any multi-omics project is to bridge the biological gap between genotype and phenotype.

• Colocalization

Data:

UKBB PPP: ~3000 protein quantitative trait loci (pQTL)

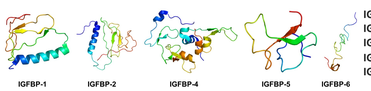
COLOC:

Whether two independent association signals at a locus are consistent with having a shared casual variant.

199 pQTLs could be colocalized with at least one IDP GWAS loci

Insulin-like growth factor-binding protein (IGFBP-family)

The role of IGFBPs in regulating transcription, inducing cell migration and apoptosis is closely related to the occurrence and development of kidney disease.



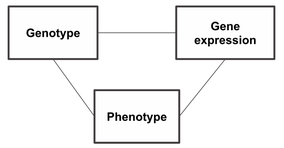
IGFBP-1 – kidney weight – diabetic kidney disease
IGFBP-3 – eGFR/CKD
IGFBP-4 – diabetes mellitus and diabetic nephropathy
IGFBP-6 – proteomic studies of CKD
IGFBP-7 – acute kidney injury

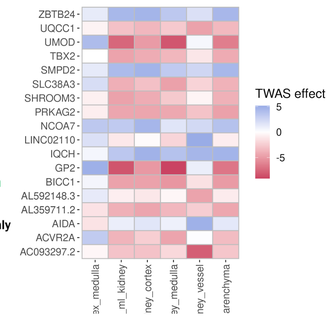
Epidermal growth factor receptors (ErbB receptors)

ErbB signaling is critically involved in renal electrolyte homeostasis and the development of end organ damage, as occurs in hypertension and atherosclerosis.

ERBB2 – renal cyst
ERBB4 – renal fibrosis

• Transcriptome-Wide Association Studies (TWAS)



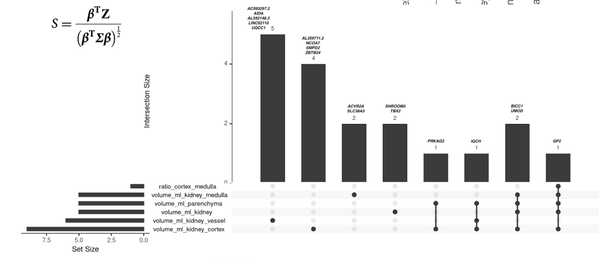


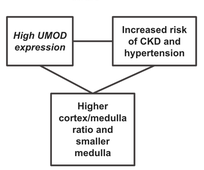
Step 1: build gene expression prediction model (kidney eQTLs atlas)

EXpression PREdiction with Summary Statistics Only (EXPRESSO)

Step 2: associate imputed gene expression with kidney GWAS

$$S = \frac{\beta^T Z}{(\beta^T \Sigma \beta)^{\frac{1}{2}}}$$





UMOD
The gene is coding for uromodulin—the most abundant protein in the kidney and in the normal urine.

GP2
The role of GP2 remains unknown. Show similarities with uromodulin

Novel genes:

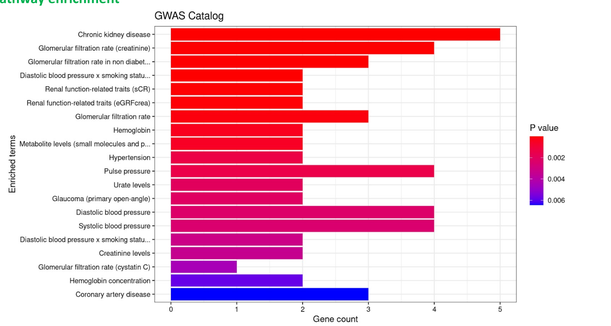
PRKAG2 – Kidney cortex

NCOA7 – Kidney cortex

SLC38A3 – Kidney medulla

AIDA, AL592148.3 – Kidney vessel

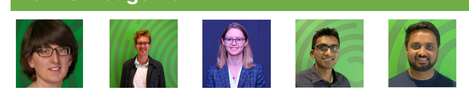
Pathway enrichment



Conclusion

- Kidney substructures are closely related to age and kidney relevant disease.
- Some substructures are useful biomarkers (e.g. cortex/medulla – hypertension, vessel - kidney stone)
- We identified novel genome-wide significant genetic loci for IDPs
- We prioritized potential disease-relevant genes through integrated methods

Acknowledgement



We thank the Statistical Genetics and Machine Learning teams for their assistant and thanks to the collaborators at University of Westminster.

Furthermore, heartfelt gratitude goes out to the dedicated ***intern team*** for their invaluable companionship, making this summer truly exceptional.