



HAL
open science

Overlap of vitamin A and vitamin D target genes with CAKUT-related processes

Ozan Ozisik, Friederike Ehrhart, Chris Evelo, Alberto Mantovani, Anaïs
Baudot

► **To cite this version:**

Ozan Ozisik, Friederike Ehrhart, Chris Evelo, Alberto Mantovani, Anaïs Baudot. Overlap of vitamin A and vitamin D target genes with CAKUT-related processes. *F1000Research*, 2021, 10, pp.395. 10.12688/f1000research.51018.2 . hal-03654274

HAL Id: hal-03654274

<https://hal.science/hal-03654274>

Submitted on 28 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



BRIEF REPORT

REVISED **Overlap of vitamin A and vitamin D target genes with CAKUT-related processes [version 2; peer review: 1 approved, 3 approved with reservations]**Ozan Ozisik¹, Friederike Ehrhart ^{2,3}, Chris T. Evelo ², Alberto Mantovani⁴, Anaïs Baudot ^{1,5}¹Aix Marseille University, Inserm, MMG, Marseille, 13385, France²Department of Bioinformatics - BiGCaT, Maastricht University, Maastricht, 6200 MD, The Netherlands³Department of Bioinformatics, NUTRIM/MHeNs, Maastricht University, Maastricht, 6200 MD, The Netherlands⁴Istituto Superiore di Sanità, Rome, 00161, Italy⁵Barcelona Supercomputing Center (BSC), Barcelona, 08034, Spain**v2** **First published:** 18 May 2021, 10:395
<https://doi.org/10.12688/f1000research.51018.1>**Latest published:** 21 Apr 2022, 10:395
<https://doi.org/10.12688/f1000research.51018.2>**Abstract**

Congenital Anomalies of the Kidney and Urinary Tract (CAKUT) are a group of abnormalities affecting the kidneys and their outflow tracts. CAKUT patients display a large clinical variability as well as a complex aetiology. Only 5% to 20% of the cases have a monogenic origin. It is thereby suspected that interactions of both genetic and environmental factors contribute to the disease. Vitamins are among the environmental factors that are considered for CAKUT aetiology. In this study, we aimed to investigate whether vitamin A or vitamin D could have a role in CAKUT aetiology. For this purpose we collected vitamin A and vitamin D target genes and computed their overlap with CAKUT-related gene sets. We observed limited overlap between vitamin D targets and CAKUT-related gene sets. We however observed that vitamin A target genes significantly overlap with multiple CAKUT-related gene sets, including CAKUT causal and differentially expressed genes, and genes involved in renal system development. Overall, these results indicate that an excess or deficiency of vitamin A might be relevant to a broad range of urogenital abnormalities.

Keywords

CAKUT, vitamin A, vitamin D, nutrients, environmental factor, renal development

Open Peer Review**Approval Status** ? ? ✓ ?

	1	2	3	4
version 2 (revision) 21 Apr 2022			✓ view	
version 1 18 May 2021	? view	? view	? view	? view

- Elena Menegola**, Università degli Studi di Milano, Milan, Italy
- Matias Simons**, Université Paris Descartes-Sorbonne, Paris, France
- Olivia Angelin-Bonnet** , Massey University, Palmerston North, New Zealand
- Sean Eddy**, University of Michigan, Ann Arbor, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding authors: Ozan Ozisik (ozan.ozisik@univ-amu.fr), Anaïs Baudot (anais.baudot@univ-amu.fr)

Author roles: **Ozisik O:** Conceptualization, Data Curation, Formal Analysis, Project Administration, Writing – Original Draft Preparation; **Ehrhart F:** Conceptualization, Writing – Review & Editing; **Evelo CT:** Conceptualization, Supervision, Writing – Review & Editing; **Mantovani A:** Conceptualization, Supervision, Writing – Original Draft Preparation; **Baudot A:** Conceptualization, Data Curation, Project Administration, Supervision, Writing – Original Draft Preparation

Competing interests: No competing interests were disclosed.

Grant information: This study has been conducted in context of the European Joint Programme on Rare Diseases (EJP RD) which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°825575. Ozan Ozisik has received funding from the Excellence Initiative of Aix-Marseille University - A*Midex, a French "Investissements d'Avenir" programme.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2022 Ozisik O *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Ozisik O, Ehrhart F, Evelo CT *et al.* **Overlap of vitamin A and vitamin D target genes with CAKUT-related processes [version 2; peer review: 1 approved, 3 approved with reservations]** F1000Research 2022, **10**:395 <https://doi.org/10.12688/f1000research.51018.2>

First published: 18 May 2021, **10**:395 <https://doi.org/10.12688/f1000research.51018.1>

REVISED Amendments from Version 1

In line with the comments from the reviewers, we made the following new analyses and changes:

In the previous version of the manuscript, we presented the overlap analysis of vitamin A and vitamin D target genes with 13 CAKUT-related gene sets. In this revised version, we added a set of CAKUT differentially expressed genes to the analyses. We hence repeated the hypergeometric test for 14 gene sets. Due to the multiple testing correction, previous adjusted p-values slightly changed.

We added explanations for the statistical test, the null hypothesis, and the background sets used in the analyses.

As a supplementary analysis, we calculated the overlap statistics with randomized sampling (in addition to the hypergeometric test in the main manuscript).

As a supplementary file, we provide the overlap of CAKUT causal genes with other CAKUT-related gene sets without considering being a vitamin target, for the interested readers.

We added descriptions for the roles of BMP4, HNF1B, RET and NRIP1, which are CAKUT causal genes and vitamin A targets according to both CTD and Balmer and Blomhoff (2002).

Any further responses from the reviewers can be found at the end of the article

Introduction

Congenital Anomalies of the Kidney and Urinary Tract (CAKUT) are a group of abnormalities affecting the kidneys and their outflow tracts, including the ureters, the bladder, and the urethra. In the European Union, the overall prevalence of CAKUT (in live plus stillbirths) between 2011 and 2018 was approximately 35:10,000.¹ The main anomalies observed in CAKUT are hydronephrosis (13.02:10,000) and multicystic renal dysplasia (4.33:10,000); these are followed by posterior urethral valves (1.26:10,000), bilateral renal agenesis (1.27:10,000) and bladder exstrophy/epispadia (0.62:10,000); many other anomalies with low prevalence can also be observed.¹ This clinical variability observed in patients is accompanied by a variable severity of the phenotypes.²

Approximately 40 different genes are known to be associated with monogenic causes of CAKUT in humans, but they explain only 5% to 20% of the cases.^{3,4} The clinical variability and the complex aetiology of CAKUT cases suggest a multifactorial origin with complex interactions of both genetic and environmental factors contributing to the disease.^{2,3,5}

The role of different environmental factors in CAKUT pathogenesis has been studied previously. It has been shown, for instance, that the drugs inhibiting the angiotensin-converting enzymes cause a specific form of CAKUT, namely tubular dysgenesis.⁶ Many prenatal and maternal factors were also assessed for their involvement in CAKUT and related anomalies. For example, studies have observed CAKUT associations with pregestational maternal diabetes mellitus,^{7,8,9} gestational maternal diabetes mellitus,^{8,10,11} abnormal volume of amniotic fluid,^{7,10} urogenital infections before the pregnancy,¹² any infection during the pregnancy,¹² maternal overweight and obesity,^{7,8,11,13} and lower infant birth weight.^{7,8,9,10} These maternal and prenatal factors are modulated by environmental risk factors. In particular, maternal diabetes and overweight/obesity as well as low birth weight are related to the nutrition of the maternal-fetal dyad. In this regard, trace nutrients, like vitamins, are likely to play important roles that still await full characterization.¹⁴

Vitamins are known to regulate renal development.^{15,16,17,18} The association of maternal vitamin A deficiency with nephron reduction was studied on a rat model.¹⁹ The inactivation of the retinoic acid nuclear receptors also resulted in renal malformations in mice.²⁰ An additional study on rats showed that vitamin A deficiency downregulates RET expression which is essential for epithelial-mesenchymal interaction during renal development.²¹ Goodyear *et al.* compared a group of pregnant women in Bangalore with a high (55%) prevalence of vitamin A deficiency with a group of pregnant women in Montreal with negligible vitamin A deficiency.²² The authors found that renal volume in newborns, adjusted for body surface area, was significantly lower in the vitamin A-deficient population. While the two populations are different for other genetic and environmental reasons, these results suggest that vitamin A deficiency may have a role in human CAKUT. El-Khashab *et al.* reported a similar finding in their study on a cohort of Egyptian mother-child pairs;²³ children of vitamin A deficient mothers had significantly lower kidney sizes.

The research on the effects of vitamin D deficiency on kidney development has generated interesting results. Rogers *et al.*²⁴ observed that incubation of metanephroi with vitamin D3 prior to implantation into adult rats increased the number of glomeruli. Conversely, Maka *et al.*²⁵ and Nascimento *et al.*²⁶ observed that maternal vitamin D deficiency stimulated nephrogenesis in rats. Following these studies, Boyce *et al.* examined the long term effects of maternal vitamin D

deficiency and observed that adult male rat offspring of vitamin D deficient dams had reduced creatinine clearance, indicating reduced renal functional capacity.²⁷ They also observed a significant upregulation of renal renin mRNA expression in fetuses and adult male offspring, and reported smaller kidneys in the offspring. Miliku *et al.* examined the association of 25-hydroxyvitamin D levels during mid-pregnancy with childhood kidney outcomes among 4212 mother-child pairs.²⁸ They observed that children of mothers who were vitamin D deficient during pregnancy had a larger combined kidney volume compared to children of mothers who had optimal vitamin D levels.

In this study, we examined whether vitamin A and vitamin D target genes are related to gene sets involved in CAKUT and renal system development. Our aim is to generate hypotheses for the involvement of these vitamins in the disease aetiology. We first created a list of vitamin A and vitamin D target genes, extracting information from the Comparative Toxicogenomics Database (CTD)²⁹ as well as from two publications.^{30,31} We then constructed different gene sets relevant to CAKUT, including genes mutated in monogenic forms of CAKUT, genes involved in renal system development, and genes involved in different pathways of interest. Finally, we performed overlap analyses to identify the genes involved in these different gene sets as well as the significant overlaps.

Methods

Vitamin A target genes

We first queried vitamin A and downloaded all gene interactions of vitamin A and its descendants from the [Comparative Toxicogenomics Database \(CTD\)](#).²⁹ We selected only the interactions supported by at least two references. This gave a list of 1086 target genes.

As an independent source, we used the data from the study of Balmer and Blomhoff.³⁰ In this study, the authors reviewed published data from 1191 articles to identify genes regulated by retinoic acid (the active metabolite of vitamin A) in humans and other species. Only a small subset of these articles are part of the CTD curation. The authors provide a list of retinoic acid target genes split into four categories, ranging from strong evidence to indirect regulation through a transcriptional intermediary. We selected the genes from all of these categories, converted gene symbols to human orthologs, and updated gene symbols when necessary using the [HUGO Gene Nomenclature Committee \(HGNC\)](#), [Rat Genome Database \(RGD\)](#), [Mouse Genome Database \(MGD\)](#), [BioMart](#) and [SynGo](#). The final list of genes obtained from Balmer and Blomhoff contains 521 target genes, of which 229 are common with genes from CTD.

Vitamin D target genes

We queried vitamin D and downloaded all gene interactions of vitamin D and its descendants from the CTD. We selected only the interactions supported by at least two references, and obtained a list of 263 target genes.

Ramagopalan *et al.*³¹ identified 230 genes with significant expression changes in response to vitamin D on lymphoblastoid cell lines using microarrays. This publication is not part of the CTD curation, hence it brings additional and independent information. We checked the gene names using HGNC and obtained a list of 210 target genes.

There are only 15 common genes between the list from CTD and the list from Ramagopalan *et al.* The combined list of vitamin D target genes (458 genes) has 134 genes in common with the combined list of vitamin A target genes (1378 genes).

CAKUT-related gene sets

We collected complementary gene sets relevant to CAKUT and renal system development from several sources, outlined below.

CAKUT causal genes set

First, we created a set of genes known to be mutated in the monogenic form of CAKUT. To do so, we combined two lists of genes provided in the studies of van der Ven *et al.*^{3,4} Only six genes are different between these two lists, and their union leads to 42 causal genes (see *Extended data*³²).

CAKUT differentially expressed genes set

Jovanovic *et al.*³³ identified 27 upregulated and 51 downregulated genes as a result of the transcriptome profiling of 15 CAKUT and 7 control ureter samples. We retrieved these differentially expressed genes (DEGs) from the supplementary file they provided. We updated gene names with the help of HUGO Gene Nomenclature Committee (HGNC) and obtained a final list of 74 DEGs (see *Extended data*³²).

Gene Ontology kidney development gene sets

We extracted the genes annotated with the three terms below from [Gene Ontology \(GO\)](#).^{34,35} Please note that these terms have parent-child relationships.

- GO:0072001 Renal system development (315 genes)
- GO:0001822 Kidney development (306 genes)
- GO:0060993 Kidney morphogenesis (96 genes)

Reactome pathways

We selected different pathways of interest for CAKUT. The renin-angiotensin system is essential for kidney development and mutations in the genes of this system result in CAKUT.³⁶ In addition, Davis *et al.* discussed the role of RET signaling in kidney development and CAKUT.³⁷ In multiple studies^{38–43} the roles of WNT, NOTCH, and Hedgehog signaling in kidney development and kidney diseases are discussed. Based on these studies we selected the following Reactome⁴⁴ pathways:

- R-HSA-2022377 Metabolism of angiotensinogen to angiotensins (18 genes)
- R-HSA-8853659 RET signaling (38 genes)
- R-HSA-195721 Signaling by WNT (328 genes)
- R-HSA-157118 Signaling by NOTCH (233 genes)
- R-HSA-5358351 Signaling by Hedgehog (149 genes)

WikiPathways

From [WikiPathways Rare Disease Portal](#)⁴⁵ we obtained the following four CAKUT-related pathways:

- WP5053 Development of ureteric collection system (47 genes)
- WP4823 Genes controlling nephrogenesis (43 genes)
- WP5052 Nephrogenesis (17 genes)
- WP4830 GDNF/RET signaling axis (23 genes)

Overlap analyses

We computed the overlaps between vitamin target gene sets and the CAKUT gene sets of interest defined previously. After obtaining the overlap results, we calculated the significance of the overlaps using the hypergeometric test. Hypergeometric test calculates the statistical significance of getting k successes in n draws without replacement from a finite population of size N that contains a total of K successes. In our context, k is the number of genes overlapping with a specific gene set, K is the size of that gene set, n is the number of vitamin targets, and N is the background gene set (also known as the reference set). We set N to the number of annotated genes in the database from which we obtained the tested gene set. For the CAKUT causal genes set and the differentially expressed genes set, we used the largest N among the different databases, which corresponds to Gene Ontology. The null hypothesis is that the vitamin target genes are not associated with the gene set and the selection of k genes from that gene set is just a result of simple random sampling.

We performed a hypergeometric test for all the gene sets. A consequence of multiple testing is false discoveries in which null hypothesis is rejected erroneously. To control for the false discovery rate, we used the Benjamini-Hochberg correction method (BH adjusted).

Supplementary analyses

We performed two supplementary analyses.

- We calculated the overlap statistics with a randomized sampling approach instead of hypergeometric test as a confirmation
- We calculated the overlaps of the CAKUT causal genes set with the other CAKUT-related gene sets (without considering being a vitamin target)

The details of the supplementary analyses and results are provided in *Extended data*.³²

Results and discussion

The results of the overlap analyses between vitamin A and D target genes and CAKUT-related gene sets are presented in *Tables 1* and *2*, and the corresponding data is available in *Underlying data*.³²

Table 1. Overlap analysis results for vitamin A. Pub: Publications,^{3,4,33} GO: Gene Ontology, Reac: Reactome, WP: WikiPathways.

Source	Term name	BH adjusted p-value and overlap with genes from Comparative Toxicogenomics Database	BH adjusted p-value and overlap with genes from Balmer and Blomhoff
Pub	CAKUT causal genes	(p-value = 1.90E-04) BMP4, EYA1, GATA3, GREB1L, HNF1B, NRIP1, PBX1, RET, SLIT2, WNT4	(p-value = 2.35E-03) AGTR1, BMP4, HNF1B, ITGA8, NRIP1, RET
Pub	CAKUT DEGs	(p-value = 1.39e-05) CCL20, CCN2, DKK1, EEF1A2, GAD1, IL6, INA, LCN2, LGI1, MYC, NUSAP1, PTGS2, SGK1, SOCS2, TOP2A, TRIL	(p-value = 8.56e-02) IL6, KRT13, MYC, PTGS2, TOP2A
GO	Renal system development	(p-value = 3.62E-26) ACTA2, ALDH1A2, ASS1, BAX, BCL2, BMP2, BMP4, BMP7, CASP9, CAT, CDKN1C, CFLAR, COL4A1, CTNNB1, CYP26A1, CYP26B1, DCN, EFNB2, EGR1, EYA1, FGF10, FGF2, FGF8, FOXC1, GATA3, GREB1L, HAS2, HES1, HNF1B, HOXB7, HPGD, ID2, ID3, IRX3, ITGA3, JAG1, LAMA5, LHX1, MMP9, MYC, NFIA, NOTCH1, NOTCH2, ODC1, PBX1, PDGFA, PDGFRA, PDGFRB, PLCE1, RARA, RARB, RBP4, RDH10, RET, RIDA, SDC1, SERPINF1, SHH, SLIT2, SMAD2, SMAD3, SOX9, STAT1, STRA6, SULF2, TACSTD2, TFAP2A, TGFB1, TGFB2, TGFBR1, TIPARP, VEGFA, WNT4, WNT5A, WT1, ZBTB16	(p-value = 6.44E-17) ACTA2, AGTR1, BCL2, BMP2, BMP4, BMP7, CA2, CALB1, COL4A1, CTNNB1, EGR1, FGF1, FGF2, FGFR2, HNF1B, HOXD11, IL6R, ITGA8, LHX1, MME, MMP9, MYC, NOTCH1, ODC1, PDGFA, PDGFRA, PDGFRB, PECAM1, RARA, RARB, RBP4, RET, SHH, SOX9, STAT1, STRA6, TFAP2A, TGFB1, TGFB2, TGFBR1, VEGFA, WNT1, WT1
GO	Kidney development	(p-value = 4.21E-25) ACTA2, ALDH1A2, ASS1, BAX, BCL2, BMP2, BMP4, BMP7, CASP9, CAT, CDKN1C, CFLAR, CTNNB1, CYP26A1, CYP26B1, DCN, EFNB2, EGR1, EYA1, FGF10, FGF2, FGF8, FOXC1, GATA3, GREB1L, HAS2, HES1, HNF1B, HOXB7, HPGD, ID2, ID3, IRX3, ITGA3, JAG1, LAMA5, LHX1, MMP9, MYC, NOTCH1, NOTCH2, ODC1, PBX1, PDGFA, PDGFRA, PDGFRB, PLCE1, RARA, RARB, RDH10, RET, RIDA, SDC1, SERPINF1, SHH, SLIT2, SMAD2, SMAD3, SOX9, STAT1, STRA6, SULF2, TACSTD2, TFAP2A, TGFB1, TGFB2, TGFBR1, TIPARP, VEGFA, WNT4, WNT5A, WT1, ZBTB16	(p-value = 3.90E-16) ACTA2, AGTR1, BCL2, BMP2, BMP4, BMP7, CA2, CALB1, CTNNB1, EGR1, FGF1, FGF2, FGFR2, HNF1B, HOXD11, IL6R, ITGA8, LHX1, MME, MMP9, MYC, NOTCH1, ODC1, PDGFA, PDGFRA, PDGFRB, PECAM1, RARA, RARB, RET, SHH, SOX9, STAT1, STRA6, TFAP2A, TGFB1, TGFB2, TGFBR1, VEGFA, WNT1, WT1
GO	Kidney morphogenesis	(p-value = 2.17E-12) BCL2, BMP2, BMP4, BMP7, CTNNB1, EYA1, FGF10, FGF2, FGF8, GATA3, GREB1L, HES1, HNF1B, HOXB7, IRX3, LAMA5, LHX1, MYC, PBX1, PDGFRB, SHH, SOX9, STAT1, TACSTD2, TGFB1, VEGFA, WNT4, WT1	(p-value = 7.49E-12) BCL2, BMP2, BMP4, BMP7, CALB1, CTNNB1, FGF1, FGF2, HNF1B, HOXD11, LHX1, MYC, PDGFRB, SHH, SOX9, STAT1, TGFB1, VEGFA, WNT1, WT1

Table 1. Continued

Source	Term name	BH adjusted p-value and overlap with genes from Comparative Toxicogenomics Database	BH adjusted p-value and overlap with genes from Balmer and Blomhoff
Reac	Metabolism of Angiotensinogen to Angiotensins	(p-value = 4.30E-01) CTSD, CTSG	(p-value = 6.41E-02) CTSD, CTSG, MME
Reac	RET signaling	(p-value = 1.82E-01) GRB10, IRS2, PIK3R1, PRKCA, RET	(p-value = 2.93E-01) GFRA1, PRKCA, RET
Reac	Signaling by WNT	(p-value = 1.46E-01) AKT1, CLTB, CTBP2, CTNNB1, DACT1, DKK1, FZD2, GNB4, GNG4, GSK3B, H2AC6, H2BC12, H4-16, ITPR3, MYC, PPP3CA, PRICKLE1, PRKCA, PRKCB, PSMB8, PSMB9, PSME2, RAC1, SOX3, SOX7, SOX9, TERT, WNT3, WNT4, WNT5A, XIAP	(p-value = 5.45E-01) CAMK2A, CLTA, CTNNB1, LEF1, MYC, PPP3CA, PPP3CB, PRKCA, RUNX3, SOX9, TERT, WNT1, WNT3A, WNT8A
Reac	Signaling by NOTCH	(p-value = 1.08E-03) ACTA2, ADAM10, AKT1, CCND1, CREB1, DLK1, EGFR, ELF3, FABP7, H2AC6, H2BC12, H4-16, HES1, HEY1, JAG1, JUN, LFNG, MDK, MYC, NOTCH1, NOTCH2, PBX1, PSMB8, PSMB9, PSME2, RUNX1, SMAD3, ST3GAL4, ST3GAL6, STAT1, TFPD2, TP53	(p-value = 5.45E-01) ACTA2, CNTN1, EGFR, FCER2, JUN, MDK, MYC, NOTCH1, STAT1, TP53
Reac	Signaling by Hedgehog	(p-value = 8.84E-01) GSK3B, PSMB8, PSMB9, PSME2, SHH, TUBA4A, TUBB2B, TUBB3	(p-value = 9.87E-01) IHH, SHH
WP	Development of ureteric collection system	(p-value = 1.59E-05) BMP4, CCND1, CTNNB1, FOXC1, GATA3, GREB1L, LHX1, MYCN, RARA, RARB, RARG, RET, SHH, SLIT2, TGFB2, WT1	(p-value = 7.21E-08) BMP4, CTNNB1, FGFR2, GFRA1, HOXD11, ITGA8, LHX1, MYCN, RARA, RARB, RARG, RET, SHH, TGFB2, WT1
WP	Genes controlling nephrogenesis	(p-value = 1.84E-05) CTNNB1, EYA1, FGF8, FOXC1, ITGA3, KDR, LHX1, NOTCH2, PDGFRB, RET, SHH, SLIT2, VEGFA, WNT4, WT1	(p-value = 2.92E-04) CTNNB1, FGFR2, HOXD11, ITGA8, LHX1, PDGFRB, RET, SHH, VEGFA, WT1
WP	Nephrogenesis	(p-value = 1.90E-04) ALDH1A2, BMP7, FGF8, GREB1L, JAG1, MEIS1, NOTCH2, WNT4	(p-value = 3.16E-01) BMP7, MEIS1
WP	GDNF/RET signaling axis	(p-value = 1.57E-03) BMP4, CTNNB1, EYA1, FOXC1, GATA3, LHX1, RET, SLIT2	(p-value = 1.61E-02) BMP4, CTNNB1, GFRA1, LHX1, RET

Vitamin A target genes and CAKUT-related gene sets show significant overlaps

Vitamin A target genes obtained from both CTD and Balmer and Blomhoff are significantly enriched in the set of CAKUT causal genes, with an overlap of ten and six genes, respectively (Table 1). Vitamin A target genes obtained from CTD are also significantly enriched in the set of CAKUT differentially expressed genes (DEGs).

Significant overlaps are observed focusing on GO terms related to renal system development, kidney development, and kidney morphogenesis. These results are expected due to the recognized role of vitamin A in differentiation.^{46,47} It should be noted that all overlapping CAKUT causal genes except NRIP1 are also part of the kidney development GO term.

Four genes (BMP4, HNF1B, RET, NRIP1) appear as particularly interesting as they are CAKUT causal genes and vitamin A targets according to both CTD and Balmer and Blomhoff. BMP4 is a growth factor of the TGF-beta family involved in a wide variety of developmental processes.⁴⁸ BMP4 inhibits ectopic budding of the ureteric tips and promotes elongation of the ureter.⁴⁹ RET is a tyrosine-protein kinase receptor, involved in a large variety of cellular processes. RET signaling is important, in particular, for the terminal growth of the nephric duct, and for initial formation and outgrowth of the ureteric bud.⁵⁰ The Hepatocyte Nuclear Factor 1-beta (HNF1B) is a transcription factor involved in ureteric bud branching and induction of nephrogenesis. It is required for normal S-shaped bodies patterning and subsequent morphogenesis of all nephron segments.⁵¹ Finally, Nuclear receptor interacting protein 1 (NRIP1) modulates the activity of nuclear receptors. It has been shown to play roles in renal malformations via deregulations of retinoic acid signaling.⁵²

Table 2. Overlap analysis results for vitamin D. Pub: Publications,^{3,4,33} GO: Gene Ontology, Reac: Reactome, WP: WikiPathways.

Source	Term name	BH adjusted p-value and overlap with genes from Comparative Toxicogenomics Database	BH adjusted p-value and overlap with genes from Ramagopalan <i>et al.</i>
Pub	CAKUT causal genes	(p-value = 1)	(p-value = 1)
Pub	CAKUT DEGs	(p-value = 6.84e-02) ANGPT2, GEM, MYC, PTGS2	(p-value = 7.64e-01) CD79A, IRF8
GO	Renal system development	(p-value = 1.45E-05) ANGPT2, BAX, BCL2, CA2, CALB1, CD24, CYP26B1, DCN, EFNB2, GREM1, ID2, ITGA3, MYC, OSR2, SMAD7, TACSTD2, TGFB2	(p-value = 1) CFLAR, SULF2
GO	Kidney development	(p-value = 1.45E-05) ANGPT2, BAX, BCL2, CA2, CALB1, CD24, CYP26B1, DCN, EFNB2, GREM1, ID2, ITGA3, MYC, OSR2, SMAD7, TACSTD2, TGFB2	(p-value = 1) CFLAR, SULF2
GO	Kidney morphogenesis	(p-value = 5.07E-02) BCL2, CALB1, GREM1, MYC, TACSTD2	(p-value = 1)
Reac	Metabolism of Angiotensinogen to Angiotensins	(p-value = 1)	(p-value = 7.64E-01) CTSZ
Reac	RET signaling	(p-value = 1)	(p-value = 7.64E-01) FRS2, PIK3CA
Reac	Signaling by WNT	(p-value = 1) ITPR1, MYC, XPO1	(p-value = 1) DAAM1, PIP5K1B, RAC2
Reac	Signaling by NOTCH	(p-value = 7.37E-01) CCND1, ELF3, HIF1A, LFNG, MYC, NCOR2	(p-value = 1) HIF1A, KAT2B, MAML1
Reac	Signaling by Hedgehog	(p-value = 1)	(p-value = 1)
WP	Development of ureteric collection system	(p-value = 7.39E-01) CCND1, TGFB2	(p-value = 1)
WP	Genes controlling nephrogenesis	(p-value = 1) ITGA3	(p-value = 7.64E-01) CD2AP, CXCR4
WP	Nephrogenesis	(p-value = 1)	(p-value = 1)
WP	GDNF/RET signaling axis	(p-value = 8.65E-01) GREM1	(p-value = 1)

Concerning the overlap with Reactome pathways, the vitamin A target genes obtained from CTD are significantly enriched in signaling by NOTCH. The NOTCH signaling pathway is mainly involved in cell-to-cell communication,⁵³ and it plays a role in the development of most organs and tissues.⁵⁴ Other Reactome terms also have overlapping genes, even if not significant. This includes the WNT signaling pathway (which is involved in developmental processes such as patterning, differentiation/proliferation shift, and migration) and the metabolism of angiotensinogen to angiotensins pathway, from which CTSD, CTSG and MME are vitamin A targets and play roles in the generation of different angiotensin peptides.^{55,56,57} The impaired conversion of angiotensin I to angiotensin II is the pharmacodynamic adverse mechanism underlying the fetopathy caused by ACE-inhibiting drugs.⁵⁸ One main feature of this teratogen-induced fetopathy, namely renal tubular dysgenesis, belongs to the CAKUT spectrum. In addition, an overproduction of angiotensin II might have a role in CAKUT anomalies involving renal tubules, as observed in cases of the autosomal recessive polycystic kidney disease.⁵⁹

Finally, vitamin A target genes extracted from both CTD and Balmer and Blomhoff are significantly enriched in the development of ureteric collection system, genes controlling nephrogenesis, and GDNF/RET signaling axis pathways, as defined in WikiPathways. The target genes extracted from CTD are also significantly enriched in nephrogenesis. Signaling by GDNF through the RET receptor is required for normal growth of the ureteric bud.⁶⁰ Localized expression

of GDNF by the metanephric mesenchyme is important to elicit and correctly position the initial budding event from the Wolffian duct and to promote the continued branching of the ureteric bud.

Overall, the significant overlaps observed between vitamin A and the CAKUT gene sets indicate that vitamin A might be relevant to a broad range of urogenital abnormalities.

Vitamin D target genes overlap with renal system development

In the analyses of vitamin D target gene sets, we overall observed very few significant overlaps with CAKUT-related gene sets. Vitamin D target genes obtained from CTD are only significantly enriched in renal system development and kidney development GO terms. It should be noted that in each case 70% of the enriched genes are also vitamin A target genes (according to CTD or Balmer and Blomhoff).

CTSZ, which cleaves angiotensin I to yield angiotensin II,⁶¹ is the only vitamin D target in the metabolism angiotensinogen to angiotensins pathway. Thus, within CAKUT, vitamin D might also have a role in renal tubular defects related to anomalies of the renin-angiotensin system.

WNT and NOTCH signaling pathways have genes that are targets of vitamin D but there is a small, non-significant overlap, and the overlapping genes are not the WNT and NOTCH receptors nor ligands.

Conclusions

In this study we examined the overlaps of vitamin A and vitamin D target gene sets extracted from different databases and publications with CAKUT-related genes and biological processes. While we only observed significant enrichment of vitamin D target genes in GO kidney development terms, there is significant enrichment of vitamin A target genes in most of the gene sets we tested. Indeed, we observed that vitamin A target genes are enriched in CAKUT causal genes set, CAKUT differentially expressed genes set, kidney development pathways, the signaling by NOTCH pathway, and the GDNF/RET signaling axis pathway. There is also an overlap with the signaling by the WNT pathway, although it is not statistically significant. Overall, the significant overlaps observed between vitamin A and CAKUT gene sets indicate that vitamin A might be relevant to a broad range of urogenital abnormalities, pointing out the importance of nutritional management during pregnancy for fetal development. However, it should be clear that our study is dedicated to the generation of hypotheses. Additional studies are needed, using up-to-date datasets obtained with new omics techniques, to investigate the fine-grained effects of vitamin excess and deficiency on CAKUT, and to decipher the pathophysiological mechanisms. The increased availability of multi-omics technologies, such as whole genome sequencing and metabolomics will certainly open new directions, e.g. by revealing new causal genes or identifying metabolic disturbances due to environmental factors.

Data availability

Underlying data

Zenodo: Overlap of vitamin A and vitamin D target genes with CAKUT-related processes. <https://www.doi.org/10.5281/zenodo.4501623>.³²

This project contains the following underlying data in the 'Data' folder:

- VitA-CTD-Genes.txt (list of vitamin A target genes from CTD used for overlap analyses).
- VitA-Balmer2002-Genes.txt (list of vitamin A target genes from Balmer and Blomhoff used for overlap analyses).
- VitD-CTD-Genes.txt (list of vitamin D target genes from CTD used for overlap analyses).
- VitD-Ramagopalan2010.txt (list of vitamin D target genes from Ramagopalan *et al.* used for overlap analyses).
- PathwaysOfInterest.gmt (CAKUT-related gene sets used for overlap analyses).
- PathwaysOfInterestBackground.txt (list indicating which background GMT file to be used for the analysis of each CAKUT-related gene set).
- hsapiens.GO:BP.name.gmt (all gene sets from GO:BP used as background information for overlap analyses).

- hsapiens.REAC.name.gmt (all gene sets from Reactome used as background information for overlap analyses).
- hsapiens.WP.name.gmt (all gene sets from WikiPathways used as background information for overlap analyses).

This project contains the following underlying data in the 'Result' folder:

- VitA-CTD-Genes.csv (overlap analysis results for vitamin A targets from CTD).
- VitA-Balmer2002-Genes.csv (overlap analysis results for vitamin A targets from Balmer and Blomhoff).
- VitA-Merged.csv (merged table for the results presented in VitA-CTD-Genes.csv and VitA-Balmer2002-Genes.csv).
- VitD-CTD-Genes.csv (overlap analysis results for vitamin D targets from CTD).
- VitD-Ramagopalan2010.csv (overlap analysis results for vitamin D targets from Ramagopalan *et al.*).
- VitD-Merged.csv (merged table for the results presented in VitD-CTD-Genes.csv and VitD-Ramagopalan2010.csv).

Extended data

Zenodo: Overlap of vitamin A and vitamin D target genes with CAKUT-related processes. <https://www.doi.org/10.5281/zenodo.4501623>.³²

This project contains the following extended data:

- overlapAnalysis.py (Python script used for all the analyses).

This project contains the following extended data in the 'Supp' folder:

- CAKUT-CausalGenesPMID29079659-PMID30143558.txt (list of genes known to be mutated in the monogenic form of CAKUT).
- DEGsFromPMID30261159.txt (list of differentially expressed genes in CAKUT)
- SupplementaryAnalyses.odt (descriptions of the supplementary analyses).
- targetsFromDifferentSourcesOverlap.odt (contains the information on vitamin A and D targets from different sources (CTD, publications), presents source specific genes and common genes).
- VitA-Balmer2002-GeneMapping.csv (detailed table for the gene list from Balmer and Blomhoff).

This project contains the following extended data in the 'Supp/CAKUTCausalGenesOverlapWithOthers' folder:

- CAKUTCausalGenesOverlap.csv (overlap of CAKUT causal genes with other CAKUT-related gene sets without considering whether genes are vitamin targets or not).

This project contains the following extended data in the 'Supp/RandomizedSampling' folder:

- VitA-CTD-Genes-RS.csv (overlap analysis results for vitamin A targets from CTD using the randomized sampling approach)
- VitA-Balmer2002-Genes-RS.csv (overlap analysis results for vitamin A targets from Balmer and Blomhoff using the randomized sampling approach - please see answer to the reviewers).

- VitD-CTD-Genes-RS.csv (overlap analysis results for vitamin D targets from CTD using the randomized sampling approach).
- VitD-Ramagopalan2010-RS.csv (overlap analysis results for vitamin D targets from Ramagopalan et al. using the randomized sampling approach)

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

Author contributions

Ozisk O: Conceptualization, Data Curation, Formal Analysis, Project Administration, Writing – Original Draft Preparation; Ehrhart F: Conceptualization, Writing – Review & Editing; Evelo C: Conceptualization, Supervision, Writing – Review & Editing; Mantovani A: Conceptualization, Supervision, Writing – Original Draft Preparation; Baudot A: Conceptualization, Data Curation, Project Administration, Supervision, Writing – Original Draft Preparation.

Acknowledgements

We thank the members of European Joint Programme on Rare Diseases, in particular the members of the Environment/AOP Work Focus.

References

1. EU RD Platform prevalence charts and tables: 2020. Accessed: 22.01.2021.
[Reference Source](#)
2. Nicolaou N, Renkema KY, Bongers EM, *et al.*: **Genetic, environmental, and epigenetic factors involved in CAKUT.** *Nat Rev Nephrol.* Dec 2015; **11**(12): 720–731.
[PubMed Abstract](#) | [Publisher Full Text](#)
3. van der Ven AT, Vivante A, Hildebrandt F: **Novel Insights into the Pathogenesis of Monogenic Congenital Anomalies of the Kidney and Urinary Tract.** *J Am Soc Nephrol.* 01 2018; **29**(1): 36–50.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
4. van der Ven AT, Connaughton DM, Ityel H, *et al.*: **Whole-Exome Sequencing Identifies Causative Mutations in Families with Congenital Anomalies of the Kidney and Urinary Tract.** *J Am Soc Nephrol.* 09 2018; **29**(9): 2348–2361.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
5. Murugapopathy V, Gupta IR: **A Primer on Congenital Anomalies of the Kidneys and Urinary Tracts (CAKUT).** *Clin J Am Soc Nephrol.* 05 2020; **15** (5): 723–731.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Boubred F, Vendemmia M, Garcia-Meric P, *et al.*: **Effects of maternally administered drugs on the fetal and neonatal kidney.** *Drug Saf.* 2006; **29**(5): 397–419.
[PubMed Abstract](#) | [Publisher Full Text](#)
7. Parikh CR, McCall D, Engelman C, *et al.*: **Congenital renal agenesis: case-control analysis of birth characteristics.** *Am J Kidney Dis.* Apr 2002; **39**(4): 689–694.
[PubMed Abstract](#) | [Publisher Full Text](#)
8. Hsu CW, Yamamoto KT, Henry RK, *et al.*: **Prenatal risk factors for childhood CKD.** *J Am Soc Nephrol.* Sep 2014; **25**(9): 2105–2111.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Dart AB, Ruth CA, Sellers EA, *et al.*: **Maternal diabetes mellitus and congenital anomalies of the kidney and urinary tract (CAKUT) in the child.** *Am J Kidney Dis.* May 2015; **65**(5): 684–691.
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Tain YL, Luh H, Lin CY, *et al.*: **Incidence and Risks of Congenital Anomalies of Kidney and Urinary Tract in Newborns: A Population-Based Case-Control Study in Taiwan.** *Medicine (Baltimore).* Feb 2016; **95**(5): e2659.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Groen In 't Woud S, Renkema KY, Schreuder MF, *et al.*: **Maternal risk factors involved in specific congenital anomalies of the kidney and urinary tract: A case-control study.** *Birth Defects Res A Clin Mol Teratol.* Jul 2016; **106**(7): 596–603.
[PubMed Abstract](#) | [Publisher Full Text](#)
12. Postoev VA, Grijbovski AM, Kovalenko AA, *et al.*: **Congenital anomalies of the kidney and the urinary tract: A murmanski county birth registry study.** *Birth Defects Res A Clin Mol Teratol.* Mar 2016; **106**(3): 185–193.
[PubMed Abstract](#) | [Publisher Full Text](#)
13. Jadresić L, Au H, Woodhouse C, *et al.*: **Pre-pregnancy obesity and risk of congenital abnormalities of the kidney and urinary tract (CAKUT)-systematic review, meta-analysis and ecological study.** *Pediatr Nephrol.* Jan 2021; **36**(1): 119–132.
[PubMed Abstract](#) | [Publisher Full Text](#)
14. de Souza Mesquita LM, Mennitti LV, de Rosso VV, *et al.*: **The role of vitamin A and its pro-vitamin carotenoids in fetal and neonatal programming: gaps in knowledge and metabolic pathways.** *Nutr Rev.* Jan 2021; **79**(1): 76–87.
[PubMed Abstract](#) | [Publisher Full Text](#)
15. Burrow CR: **Retinoids and renal development.** *Exp Nephrol.* 2000; **8**(4-5): 219–225.
[PubMed Abstract](#) | [Publisher Full Text](#)
16. Charlton JR, Springsteen CH, Carmody JB: **Nephron number and its determinants in early life: a primer.** *Pediatr Nephrol.* Dec 2014; **29**(12): 2299–2308.
[PubMed Abstract](#) | [Publisher Full Text](#)
17. Lee YQ, Collins CE, Gordon A, *et al.*: **The Relationship between Maternal Nutrition during Pregnancy and Offspring Kidney Structure and Function in Humans: A Systematic Review.** *Nutrients.* Feb 2018; **10**(2).
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Lumbers ER, Kandasamy Y, Delforce SJ, *et al.*: **Programming of Renal Development and Chronic Disease in Adult Life.** *Front Physiol.* 2020; **11**: 757.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Lelièvre-Pégurier M, Vilar J, Ferrier ML, *et al.*: **Mild vitamin A deficiency leads to inborn nephron deficit in the rat.** *Kidney Int.* Nov 1998; **54**(5): 1455–1462.
[PubMed Abstract](#) | [Publisher Full Text](#)
20. Mendelsohn C, Batourina E, Fung S, *et al.*: **Stromal cells mediate retinoid-dependent functions essential for renal development.** *Development.* Mar 1999; **126**(6): 1139–1148.
[PubMed Abstract](#)
21. Batourina E, Gim S, Bello N, *et al.*: **Vitamin A controls epithelial/mesenchymal interactions through Ret expression.** *Nat Genet.* Jan 2001; **27**(1): 74–78.
[PubMed Abstract](#) | [Publisher Full Text](#)
22. Goodyer P, Kurpad A, Rekha S, *et al.*: **Effects of maternal vitamin A status on kidney development: a pilot study.** *Pediatr Nephrol.* Feb 2007; **22**(2): 209–214.
[PubMed Abstract](#) | [Publisher Full Text](#)
23. El-Khashab EK, Hamdy AM, Maher KM, *et al.*: **Effect of maternal vitamin A deficiency during pregnancy on neonatal kidney size.** *J Perinat Med.* Mar 2013; **41**(2): 199–203.
[PubMed Abstract](#) | [Publisher Full Text](#)

24. Rogers SA, Droegge D, Dusso A, *et al.*: **Incubation of metanephroi with vitamin d(3) increases numbers of glomeruli.** *Organogenesis.* Oct 2004; **1**(2): 52–54.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. Maka N, Makrakis J, Parkington HC, *et al.*: **Vitamin D deficiency during pregnancy and lactation stimulates nephrogenesis in rat offspring.** *Pediatr Nephrol.* Jan 2008; **23**(1): 55–61.
[PubMed Abstract](#) | [Publisher Full Text](#)
26. Nascimento FA, Ceciliano TC, Aguila MB, *et al.*: **Maternal vitamin D deficiency delays glomerular maturity in F1 and F2 offspring.** *PLoS One.* 2012; **7**(8): e41740.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Boyce AC, Palmer-Aronsten BJ, Kim MY, *et al.*: **Maternal vitamin D deficiency programmes adult renal renin gene expression and renal function.** *J Dev Orig Health Dis.* Oct 2013; **4**(5): 368–376.
[PubMed Abstract](#) | [Publisher Full Text](#)
28. Miliku K, Voortman T, Franco OH, *et al.*: **Vitamin D status during fetal life and childhood kidney outcomes.** *Eur J Clin Nutr.* 05 2016; **70**(5): 629–634.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Davis AP, Grondin CJ, Johnson RJ, *et al.*: **Comparative Toxicogenomics Database (CTD): update 2021.** *Nucleic Acids Res.* Jan 2021; **49**(D1): D1138–D1143.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Balmer JE, Blomhoff R: **Gene expression regulation by retinoic acid.** *J Lipid Res.* Nov 2002; **43**(11): 1773–1808.
[PubMed Abstract](#) | [Publisher Full Text](#)
31. Ramagopalan SV, Heger A, Berlanga AJ, *et al.*: **A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution.** *Genome Res.* Oct 2010; **20**(10): 1352–1360.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Ozisik O, Ehrhart F, Evelo C, *et al.*: **Overlap of vitamin A and vitamin D target genes with CAKUT-related processes.** *Zenodo.* April 2022.
[Publisher Full Text](#)
33. Jovanovic I, Zivkovic M, Kostic M, *et al.*: **Transcriptome-driven integrative exploration of functional state of ureter tissue affected by CAKUT.** *Life Sci.* Nov 2018; **212**: 1–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
34. Ashburner M, Ball CA, Blake JA, *et al.*: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet.* May 2000; **25**(1): 25–29.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
35. The Gene Ontology Consortium: **The Gene Ontology Resource: 20 years and still GOing strong.** *Nucleic Acids Res.* 01 2019; **47**(D1): D330–D338.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. Yosypiv IV: **Renin-angiotensin system in mammalian kidney development.** *Pediatr Nephrol.* Feb 2020.
[PubMed Abstract](#) | [Publisher Full Text](#)
37. Davis TK, Hoshi M, Jain S: **To bud or not to bud: the RET perspective in CAKUT.** *Pediatr Nephrol.* Apr 2014; **29**(4): 597–608.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
38. Kreidberg JA: **Murine models of kidney development.** In: Pollak MR, Mount DB, editor, *Molecular and Genetic Basis of Renal Disease.* 2014 chapter 4, pages 41–48. W.B. Saunders.
39. Pulkkinen K, Murugan S, Vainio S: **Wnt signaling in kidney development and disease.** *Organogenesis.* Apr 2008; **4**(2): 55–59.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
40. Sirin Y, Susztak K: **Notch in the kidney: development and disease.** *J Pathol.* Jan 2012; **226**(2): 394–403.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
41. Edeling M, Ragi G, Huang S, *et al.*: **Developmental signalling pathways in renal fibrosis: the roles of Notch, Wnt and Hedgehog.** *Nat Rev Nephrol.* 07 2016; **12**(7): 426–439.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
42. Wang Y, Zhou CJ, Liu Y: **Wnt Signaling in Kidney Development and Disease.** *Prog Mol Biol Transl Sci.* 01 2018; **153**: 181–207.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
43. Mukherjee M, Fogarty E, Janga M, *et al.*: **Notch Signaling in Kidney Development, Maintenance, and Disease.** *Biomolecules.* 11 2019; **9**(11).
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
44. Jassal B, Matthews L, Viteri G, *et al.*: **The reactome pathway knowledgebase.** *Nucleic Acids Res.* 01 2020; **48**(D1): D498–D503.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
45. Martens M, Ammar A, Riutta A: **WikiPathways: connecting communities.** *Nucleic Acids Res.* 01 2021; **49**(D1): D613–D621.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
46. Ross AC, Gardner EM: **The function of vitamin A in cellular growth and differentiation, and its roles during pregnancy and lactation.** *Adv Exp Med Biol.* 1994; **352**: 187–200.
[PubMed Abstract](#) | [Publisher Full Text](#)
47. Love JM, Gudas LJ: **Vitamin A, differentiation and cancer.** *Curr Opin Cell Biol.* Dec 1994; **6**(6): 825–831.
[PubMed Abstract](#) | [Publisher Full Text](#)
48. Rezzola S, Di Somma M, Corsini M, *et al.*: **VEGFR2 activation mediates the pro-angiogenic activity of BMP4.** *Angiogenesis.* Nov 2019; **22**(4): 521–533.
[PubMed Abstract](#) | [Publisher Full Text](#)
49. Miyazaki Y, Oshima K, Fogo A, *et al.*: **Bone morphogenetic protein 4 regulates the budding site and elongation of the mouse ureter.** *J Clin Invest.* Apr 2000; **105**(7): 863–873.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
50. Costantini F: **RET signaling in ureteric bud formation and branching.** In: Little MH, editor, *Kidney Development, Disease, Repair and Regeneration.* chapter 4, pp. 41–56. Academic Press; 2016.
[Publisher Full Text](#)
51. Heliot C, Desgrange A, Buisson I, *et al.*: **HNF1B controls proximal-intermediate nephron segment identity in vertebrates by regulating Notch signalling components and Irx1/2.** *Development.* Feb 2013; **140**(4): 873–885.
[PubMed Abstract](#) | [Publisher Full Text](#)
52. Vivante A, Mann N, Yonath H, *et al.*: **Dysregulation of Retinoic Acid Signaling.** *J Am Soc Nephrol.* Aug 2017; **28**(8): 2364–2376.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
53. Bigas A, Espinosa L: **Notch signaling in cell-cell communication pathways.** *Current Stem Cell Reports.* Dec 2016; **2**(4): 349–355.
[Publisher Full Text](#)
54. Siebel C, Lendahl U: **Notch Signaling in Development, Tissue Homeostasis, and Disease.** *Physiol Rev.* 10 2017; **97**(4): 1235–1294.
[PubMed Abstract](#) | [Publisher Full Text](#)
55. Belova LA: **Angiotensin II-generating enzymes.** *Biochemistry (Moscow).* Dec 2000; **65**(12): 1337–1345.
[PubMed Abstract](#) | [Publisher Full Text](#)
56. Rice GI, Thomas DA, Grant PJ, *et al.*: **Evaluation of angiotensin-converting enzyme (ACE), its homologue ACE2 and neprilysin in angiotensin peptide metabolism.** *Biochem J.* Oct 2004; **383**(Pt 1): 45–51.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
57. Wu C, Lu H, Cassis LA, *et al.*: **Molecular and Pathophysiological Features of Angiotensinogen: A Mini Review.** *N Am J Med Sci (Boston).* Oct 2011; **4**(4): 183–190.
[PubMed Abstract](#) | [Free Full Text](#)
58. Buttar HS: **An overview of the influence of ACE inhibitors on fetal-placental circulation and perinatal development.** *Mol Cell Biochem.* Nov 1997; **176**(1-2): 61–71.
[PubMed Abstract](#)
59. Loghman-Adham M, Soto CE, Inagami T, *et al.*: **Expression of components of the renin-angiotensin system in autosomal recessive polycystic kidney disease.** *J Histochem Cytochem.* Aug 2005; **53**(8): 979–988.
[PubMed Abstract](#) | [Publisher Full Text](#)
60. Costantini F, Shakya R: **GDNF/Ret signaling and the development of the kidney.** *Bioessays.* Feb 2006; **28**(2): 117–127.
[PubMed Abstract](#) | [Publisher Full Text](#)
61. Nägler DK, Kraus S, Feierler J, *et al.*: **A cysteine-type carboxypeptidase, cathepsin X, generates peptide receptor agonists.** *Int Immunopharmacol.* Jan 2010; **10**(1): 134–139.
[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status: ? ? ✓ ?

Version 2

Reviewer Report 28 April 2022

<https://doi.org/10.5256/f1000research.132373.r135411>

© 2022 **Angelin-Bonnet O.** This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Olivia Angelin-Bonnet 

School of Fundamental Sciences, College of Sciences, Massey University, Palmerston North, New Zealand

The authors answered satisfactorily my comments. The added paragraph in the Methods section about the hypergeometric is clear and will be helpful to readers.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Statistics, Systems Biology, Biological networks analysis and reconstruction

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 17 December 2021

<https://doi.org/10.5256/f1000research.54123.r100885>

© 2021 **Eddy S.** This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Sean Eddy

Division of Nephrology, Michigan Medicine, University of Michigan, Ann Arbor, MI, USA

This study involves generating curated vitamin A and D gene sets from the public domain and assessing the relevance of the genes to CAKUT and CAKUT related pathways. The authors

meticulously report their results and curation efforts so the findings are readily reproducible and are appropriate for F1000Research.

Is the study design appropriate and is the work technically sound?

The addition of targeted vitamin A/retinoic acid and D gene sets is appreciated. Limiting CTD exported gene sets to those with at least 2 references is a reasonable approach to reducing potential false positives. I agree with the authors that this study is currently hypothesis generating. Given that patient CAKUT gene sets already exist and are publicly available (Jovanich *et al.*, 2018, GEO dataset GSE [GSE83946](#))¹, the authors should also assess the enrichment of their gene sets against differentially expressed genes from patients with CAKUT and healthy control reference tissue.

If applicable, is the statistical analysis and its interpretation appropriate?

Because the gene sets are substantially different in size, a direct comparison of enrichment and p-values can be misleading. I would suggest the authors generate a randomized gene set of equivalent size to each of the vitamin A and vitamin D gene sets and perform comparative enrichment analyses on the relevant pathways.

While not necessary, the addition of an odds ratio from 2x2 contingency matrices to assess the strength of enrichment would enhance the findings.

The targeted enrichment analysis of kidney and CAKUT related pathways is a reasonable approach to demonstrating a positive finding, ie. the gene sets identified associated with some of the expected/hypothesized pathways, it would be interesting to see if a more unbiased enrichment approach might also highlight other vitamin A and D relevant pathways, and perhaps additional renal-associated pathways not readily apparent from the targeted approach. This which could also substantiate the authors' vitamin A and D curation efforts.

References

1. Jovanovic I, Zivkovic M, Kostic M, Krstic Z, et al.: Transcriptome-driven integrative exploration of functional state of ureter tissue affected by CAKUT. *Life Sci.* 2018; **212**: 1-8 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Renal systems and computational biology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 15 Apr 2022

Ozan Ozisik, Aix Marseille University, France

We thank the reviewer for the valuable comments. We present the reviewer comments, our answers to them, and quoted texts from the revised manuscript below.

Given that patient CAKUT gene sets already exist and are publicly available (Jovanich et al., 2018, GEO dataset GSE [GSE83946](#))¹, the authors should also assess the enrichment of their gene sets against differentially expressed genes from patients with CAKUT and healthy control reference tissue.

We performed the suggested analysis, described the analysis in the manuscript and presented the results. We observed that vitamin A targets obtained from CTD significantly overlapped with the differentially expressed genes.

“CAKUT differentially expressed genes set Jovanovic et al.³³ identified 27 upregulated and 51 downregulated genes as a result of the transcriptome profiling of 15 CAKUT and 7 control ureter samples. We retrieved these differentially expressed genes (DEGs) from the supplementary file they provided. We updated gene names with the help of HUGO Gene Nomenclature Committee (HGNC) and obtained a final list of 74 DEGs (see Extended data ³²).”

Because the gene sets are substantially different in size, a direct comparison of enrichment and p-values can be misleading. I would suggest the authors generate a randomized gene set of equivalent size to each of the vitamin A and vitamin D gene sets and perform comparative enrichment analyses on the relevant pathways.

We used hypergeometric tests in order to take into account the sizes of gene sets. Indeed, these tests are used in enrichment analyses to assess sets of genes' overrepresentation in annotation terms of different sizes. However, we do agree that the experiment proposed by the reviewer is interesting to confirm our observations. We hence additionally performed a randomized sampling analysis.

We selected a set of random genes having the same size as the target genes set from the

background genes list. We repeated this selection 2000 times. We counted the number of times that the random set had better overlap with the CAKUT-related gene set than the target genes set.

The results are consistent with the hypergeometric tests, except for two comparisons in which the overlaps were not significant according to the hypergeometric test but are significant according to the randomized sampling approach. Indeed, the adjusted p-value for the overlap between vitamin A targets retrieved from CTD and Reactome Signaling by WNT term is 0.146 (>0.05 , insignificant) in the hypergeometric test but it is 0.0426 (<0.05 , significant) in randomized sampling. Additionally, the adjusted p-value for the overlap between vitamin D targets retrieved from CTD and GO kidney morphogenesis term is 0.0507 (>0.05 , insignificant) in the hypergeometric test but it is 0.042 (<0.05 , significant) in randomized sampling. We should note that the results of the randomized sampling approach may change in different runs.

As this is an extra test for confirmation, we provided related documents in Underlying data and Extended data for the interested readers. The analysis process is described in "Supp/SupplementaryAnalyses.odt". The code of the randomized sampling test is available as part of overlapAnalysis.py; the results are in the "Supp/RandomizedSampling" folder.

While not necessary, the addition of an odds ratio from 2x2 contingency matrices to assess the strength of enrichment would enhance the findings.

All the numbers necessary to compute the enrichments (gene set sizes, domain sizes, intersection sizes, p-values) are provided in the tables. We think that providing additional tables and metrics might be confusing.

The targeted enrichment analysis of kidney and CAKUT related pathways is a reasonable approach to demonstrating a positive finding, ie. the gene sets identified associated with some of the expected/hypothesized pathways, it would be interesting to see if a more unbiased enrichment approach might also highlight other vitamin A and D relevant pathways, and perhaps additional renal-associated pathways not readily apparent from the targeted approach. This which could also substantiate the authors' vitamin A and D curation efforts.

We agree that our initial analyses could be extended. As a matter of fact, we are currently extending our approach from CAKUT-associated pathways to the available rare disease pathways. However providing all the enrichment results for all the pathways without any interpretation will not be useful for the audience. As we provide the vitamin target gene lists and the code to reproduce or extend our work, the interested (and expert) researchers with a precise question in mind could perform the analyses.

Competing Interests: No competing interests were disclosed.

<https://doi.org/10.5256/f1000research.54123.r98474>

© 2021 Angelin-Bonnet O. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Olivia Angelin-Bonnet 

School of Fundamental Sciences, College of Sciences, Massey University, Palmerston North, New Zealand

The article "Overlap of vitamin A and vitamin D target genes with CAKUT-related processes" investigates the overlap between vitamin A and D target genes and genes associated with congenital anomalies of the kidney and urinary tract (CAKUT). More specifically, the authors assessed the significance of the intersection between sets of genes identified by previous publications or through databases as targets of either vitamin A or D, and sets of genes identified as causal to CAKUT or associated with pathways of interest with respect to CAKUT. This review focuses on the statistical analysis presented in this paper.

I find the use of a hypergeometric test for this analysis completely appropriate. However, I found that while the Methods section provides a good explanation of the construction of the different gene sets investigated, there is no presentation or discussion of the statistical test used to assess the significance of the overlap between different gene sets. More specifically, I would like to see a brief explanation of the hypergeometric test, in order to justify its use to the user, as well as a statement of its null hypothesis. In addition, there is no mention of the background sets used for the test. I think they should be explicitly mentioned, especially since they differ with the pathway of interest tested (e.g. all GO terms for GO-related gene sets, all Reactome genes for Reactome sets, etc) and are used to constrain the sets of vitamin-related genes (i.e. how many vitamin genes were discarded because they did not appear in the background set?). I also want to note that the availability of the data and the (clearly commented) python script was really appreciated. It ensures that the statistical analysis can be fully reproduced.

Lastly, a minor presentation point: I am not sure that having the p-value between brackets in the same cell as the list of interesting genes in the two tables is very clear, especially when the intersection set is empty and the resulting p-value is 1. I would suggest to add "p-value = " before the actual value to make it clearer.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Statistics, Systems Biology, Biological networks analysis and reconstruction

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 15 Apr 2022

Ozan Ozisik, Aix Marseille University, France

We thank the reviewer for the valuable comments. We present the reviewer comments, our answers to them, and quoted texts from the revised manuscript below.

I find the use of a hypergeometric test for this analysis completely appropriate. However, I found that while the Methods section provides a good explanation of the construction of the different gene sets investigated, there is no presentation or discussion of the statistical test used to assess the significance of the overlap between different gene sets. More specifically, I would like to see a brief explanation of the hypergeometric test, in order to justify its use to the user, as well as a statement of its null hypothesis. In addition, there is no mention of the background sets used for the test. I think they should be explicitly mentioned, especially since they differ with the pathway of interest tested (e.g. all GO terms for GO-related gene sets, all Reactome genes for Reactome sets, etc) and are used to constrain the sets of vitamin-related genes (i.e. how many vitamin genes were discarded because they did not appear in the background set?). I also want to note that the availability of the data and the (clearly commented) python script was really appreciated. It ensures that the statistical analysis can be fully reproduced.

We thank the reviewer for pointing out the lack of necessary information on the applied test. We added the explanations for the statistical test, the null hypothesis, and the background sets used in the tests to the "Methods" section under "Overlap analyses" heading.

"We computed the overlaps between vitamin target gene sets and the CAKUT gene sets of interest defined previously. After obtaining the overlap results, we calculated the significance of the overlaps using the hypergeometric test. Hypergeometric test calculates the statistical significance of getting k successes in n draws without replacement from a finite population of size N that contains a total of K successes. In our context, k is the number of genes overlapping with a specific gene set, K is the size of that gene set, n is the number of vitamin targets, and N is the background gene set (also known as the reference set). We set N to the number of annotated

genes in the database from which we obtained the tested gene set. For the CAKUT causal genes set and the differentially expressed genes set, we used the largest N among the different databases, which corresponds to Gene Ontology. The null hypothesis is that the vitamin target genes are not associated with the gene set and the selection of k genes from that gene set is just a result of simple random sampling.

We performed a hypergeometric test for all the gene sets. A consequence of multiple testing is false discoveries in which null hypothesis is rejected erroneously. To control for the false discovery rate, we used the Benjamini-Hochberg correction method (BH adjusted)."

I would suggest to add "p-value = " before the actual value to make it clearer.

As suggested, we added "p-value = " before the actual value.

Competing Interests: No competing interests were disclosed.

Reviewer Report 27 October 2021

<https://doi.org/10.5256/f1000research.54123.r95476>

© 2021 Simons M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Matias Simons

Laboratory of Epithelial Biology and Disease, Imagine Institute, Université Paris Descartes-Sorbonne, Paris, France

The paper by Ozisik *et al* investigates associations of Vitamin A and D target genes with renal development and CAKUT pathways. They find that a significant overlap for vitamin A target genes but less so for vitamin D target genes. This fits with previous epidemiological studies on vitamin A and renal size. Overall, I find the paper interesting because it provides insights into non-genetic causes of CAKUT. I have only a few minor points:

- Given that a list of 1086 Vit A target genes but only of 263 Vit D target genes, it is not surprising that less overlap is found for Vit D. This should be commented on.
- It is said in abstract and introduction that only 5% to 20% of CAKUT cases have a monogenic origin. However, it has to be noted that these numbers are only based on whole exome sequencing studies. It is to be expected that whole genome sequencing will discover new monogenic causes.
- The discussion could include more perspectives on how new -omic techniques could be used in the future.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Human genetics, developmental biology, Wnt

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 15 Apr 2022

Ozan Ozisik, Aix Marseille University, France

We thank the reviewer for the valuable comments. We present the reviewer comments, our answers to them, and quoted texts from the revised manuscript below.

Given that a list of 1086 Vit A target genes but only of 263 Vit D target genes, it is not surprising that less overlap is found for Vit D. This should be commented on.

Considering the overlaps between gene lists, as stated, with less target genes, less genes are expected to overlap with the CAKUT-related gene sets. However, in the manuscript, we do not focus on the exact number of overlapping genes. Indeed, in order to take into account the variable sizes of the different datasets, we used the hypergeometric test. Our focus is hence on the significance of the overlaps considering the dataset sizes, i.e. the observed overlaps as compared to the expected overlaps given the dataset sizes. This means that a small number of overlapping genes could lead to significance for vitamin D overlap with CAKUT-related genes. However, we did not observe many significant overlaps. The overlaps between vitamin A targets and CAKUT-related gene sets were greater than expectations, which is the reason they are statistically significant. In the revised manuscript we added the explanations for the statistical test.

It is said in abstract and introduction that only 5% to 20% of CAKUT cases have a monogenic origin. However, it has to be noted that these numbers are only based on whole exome sequencing studies. It is to be expected that whole genome sequencing will discover new monogenic causes.

We used a list of causal genes curated from multiple studies. We are not sure if they are all based on whole exome sequencing. However, we agree that whole genome sequencing will help discovering new monogenic causes. We hence added a statement related to the whole genome sequencing to the conclusion:

"The increased availability of multi-omics technologies, such as whole genome sequencing and metabolomics will certainly open new directions, e.g. by revealing new causal genes or identifying metabolic disturbances due to environmental factors."

The discussion could include more perspectives on how new -omic techniques could be used in the future.

We agree that we do expect a lot of developments in this area, including whole genome sequencing but also single-cell or spatial transcriptomics, on human, animal models or organoids. In order to mention all these potential technologies while keeping the text focused, we added the following sentence to the conclusion:

"The increased availability of multi-omics technologies, such as whole genome sequencing and metabolomics will certainly open new directions, e.g. by revealing new causal genes or identifying metabolic disturbances due to environmental factors."

Competing Interests: No competing interests were disclosed.

Reviewer Report 25 June 2021

<https://doi.org/10.5256/f1000research.54123.r85565>

© 2021 Menegola E. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Elena Menegola

Department of Environmental Sciences and Policies, Università degli Studi di Milano, Milan, Italy

The manuscript "Overlap of vitamin A and vitamin D target genes with CAKUT-related processes" is aimed to indirectly evaluate the possible relationship between Vitamin A and D intake imbalance and congenital anomalies of the kidney and urinary tract (CAKUT). The work considers genes reported as dysregulated in CAKUT and key genes involved in different urinary morphogenetic pathways. Both were overlapped with target genes of Vitamin A and Vitamin D (extracted from CTD and from literature: Balmer and Blomhoff, 2002 for Vitamin A, Ramagopalan *et al.*, 2010 for

Vitamin D). As expected considering the biological activity of the Vitamin A morphogenetic derivative (retinoid acid), significant overlaps are reported in the manuscript between Vitamin A target genes and genes involved in several urinary morphogenetic processes and signalling. Interestingly, significant overlaps were observed between Vitamin A target genes and CAKUT-related genes. As far as Vitamin D is concerned, overlappings were limited to some kidney morphogenetic genes but no overlapping was evidenced with CAKUT-related genes.

In general, this work can be considered of interest in order to postulate a possible relationship between Vitamin imbalance (mainly Vitamin A) and CAKUT and to suggest that a nutritional evaluation in pregnancy is essential to ensure a physiological development and maternal/fetal health.

Some suggestions in order to improve the impact of the manuscript:

Abstract:

The abstract should be rewritten in order to better clarify aim and results of the work. In particular, Vitamin D results are lacking.

Introduction:

The aim should be better explained. A brief description for the reason of the choice of literature (Balmer and Blomhoff, 2002 for Vitamin A, Ramagopalan *et al.*, 2010 for Vitamin D) could enhance the comprehension.

Methods:

Indicate retinoic acid as "the active metabolite of vitamin A" instead of "vitamin A acid". Better explain the statistical analysis.

Results:

Tables are not fully self explicative. I suggest:

- a separate column for p-value;
- write in bold the genes both overlapping with CAKUT and urinary pathways/signalling.

Moreover, (1) in table 2 is not explained.

Looking in detail at Vitamin A table, four CAKUT-causal genes (BMP4, HNF1B, NRIP1 and RET, overlapping both with CTD and Balmer and Blomhoff) appear as a leitmotiv in different vitamin-A related morphogenetic pathways. In my personal opinion, this result should be stressed and a brief description of functions of BMP4, HNF1B, NRIP1 and RET in kidney and urinary tract development introduced in Results and/or Conclusions.

Conclusions:

Can be rewritten in order to better focus the relevance of the nutritional management during pregnancy in order to ensure both maternal and fetal health. A general consideration should be introduced, considering also the pleiotropic functions of the involved genes.

References:

Please check references, i.e. reference 54 seems not pertinent.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Developmental toxicology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 15 Apr 2022

Ozan Ozisik, Aix Marseille University, France

We thank the reviewer for the detailed summary and valuable comments. We present the reviewer comments, our answers to them, and quoted texts from the revised manuscript below.

Abstract

The abstract should be rewritten in order to better clarify aim and results of the work. In particular, Vitamin D results are lacking.

We revised the abstract; we clarified our aim and added the results for vitamin D.

"...In this study, we aimed to investigate whether vitamin A or vitamin D could have a role in CAKUT aetiology. For this purpose we collected vitamin A and vitamin D target genes and computed their overlap with CAKUT-related gene sets. We observed limited overlap between vitamin D targets and CAKUT-related gene sets. We however observed that vitamin A target genes significantly overlap with multiple CAKUT-related gene sets, including CAKUT causal and differentially expressed genes, and genes involved in renal system development. Overall, these

results indicate that an excess or deficiency of vitamin A might be relevant to a broad range of urogenital abnormalities."

Introduction

The aim should be better explained. A brief description for the reason of the choice of literature (Balmer and Blomhoff, 2002 for Vitamin A, Ramagopalan et al., 2010 for Vitamin D) could enhance the comprehension.

We clarified the aim of the study in the introduction.

"In this study, we examined whether vitamin A and vitamin D target genes are related to gene sets involved in CAKUT and renal system development. Our aim is to generate hypotheses for the involvement of these vitamins in the disease aetiology."

We used additional datasets from two publications to have independent sources for vitamin target information. Only a small subset of the articles (27 of 1192) reviewed by Balmer and Blomhoff (2002) are reviewed in CTD, and Ramagopalan et al. (2010) is not reviewed in CTD. We clarified this in the methods section.

"As an independent source, we used the data from the study of Balmer and Blomhoff.³⁰ In this study, the authors reviewed published data from 1191 articles to identify genes regulated by retinoic acid (the active metabolite of vitamin A) in humans and other species. Only a small subset of these articles are part of the CTD curation."

"Ramagopalan et al.³¹ identified 230 genes with significant expression changes in response to vitamin D on lymphoblastoid cell lines using microarrays. This publication is not part of the CTD curation, hence it brings additional and independent information."

Methods

Indicate retinoic acid as "the active metabolite of vitamin A" instead of "vitamin A acid". Better explain the statistical analysis.

We changed "vitamin A acid" to "the active metabolite of vitamin A".

"In this study, the authors reviewed published data from 1191 articles to identify genes regulated by retinoic acid (the active metabolite of vitamin A) in humans and other species."

We added the description of the statistical analysis.

"We computed the overlaps between vitamin target gene sets and the CAKUT gene sets of interest defined previously. After obtaining the overlap results, we calculated the significance of the overlaps using the hypergeometric test. Hypergeometric test calculates the statistical significance of getting k successes in n draws without replacement from a finite population of size N that contains a total of K successes. In our context, k is the number of genes overlapping with a specific gene set, K is the size of that gene set, n is the number of vitamin targets, and N is the

background gene set (also known as the reference set). We set N to the number of annotated genes in the database from which we obtained the tested gene set. For the CAKUT causal genes set and the differentially expressed genes set, we used the largest N among the different databases, which corresponds to Gene Ontology. The null hypothesis is that the vitamin target genes are not associated with the gene set and the selection of k genes from that gene set is just a result of simple random sampling.

We performed a hypergeometric test for all the gene sets. A consequence of multiple testing is false discoveries in which null hypothesis is rejected erroneously. To control for the false discovery rate, we used the Benjamini-Hochberg correction method (BH adjusted)."

Results

Tables are not fully self explicative. I suggest:

- **a separate column for p-value;**

Adding a new column to the table for p-values causes difficulty in layout. We hence added "p-value =" before each value in the parentheses to clarify the table.

- **write in bold the genes both overlapping with CAKUT and urinary pathways/signalling.**

The overlap of CAKUT causal genes with other gene sets can be interesting for the reader. However, as the aim of the tables in the manuscript is to present overlap of vitamin target genes and CAKUT related gene sets, presenting this extra information might be confusing for the reader. Additionally, the information of overlap between CAKUT causal genes and other pathways will be incomplete and could be misleading as there might be other genes which are common to these but are not vitamin targets. For these reasons, we now provide the overlap of CAKUT causal genes with all the other pathways (without considering whether genes are vitamin targets or not) in a supplementary dataset. This analysis is described in "Supp/SupplementaryAnalyses.odt", the code for the calculation of this overlap is available as part of overlapAnalysis.py, and the overlap table is presented in "Supp/CAKUTCausalGenesOverlapWithOthers/CAKUTCausalGenesOverlap.csv" in Underlying data and Extended data.

Moreover, (1) in table 2 is not explained.

"(1)"s in table 2 were p-values of 1. We hope that with the addition of "p-value =", these p-values are clearer now.

Looking in detail at Vitamin A table, four CAKUT-causal genes (BMP4, HNF1B, NRIP1 and RET, overlapping both with CTD and Balmer and Blomhoff) appear as a leitmotiv in different vitamin-A related morphogenetic pathways. In my personal opinion, this result should be stressed and a brief description of functions of BMP4, HNF1B, NRIP1 and RET in kidney and urinary tract development introduced in Results and/or Conclusions.

We agree that these are indeed interesting genes to highlight. We stressed that BMP4, HNF1B, NRIP1 and RET, which are CAKUT causal genes, are vitamin A targets according to both CTD and Balmer and Blomhoff. We added a description of their roles.

“Four genes (BMP4, HNF1B, RET, NRIP1) appear as particularly interesting as they are CAKUT causal genes and vitamin A targets according to both CTD and Balmer and Blomhoff. BMP4 is a growth factor of the TGF-beta family involved in a wide variety of developmental processes.⁴⁸ BMP4 inhibits ectopic budding of the ureteric tips and promotes elongation of the ureter.⁴⁹ RET is a tyrosine-protein kinase receptor, involved in a large variety of cellular processes. RET signaling is important, in particular, for the terminal growth of the nephric duct, and for initial formation and outgrowth of the ureteric bud.⁵⁰ The Hepatocyte Nuclear Factor 1-beta (HNF1B) is a transcription factor involved in ureteric bud branching and induction of nephrogenesis. It is required for normal S-shaped bodies patterning and subsequent morphogenesis of all nephron segments.⁵¹ Finally, Nuclear receptor interacting protein 1 (NRIP1) modulates the activity of nuclear receptors. It has been shown to play roles in renal malformations via deregulations of retinoic acid signaling.⁵²”

Conclusions:

Can be rewritten in order to better focus the relevance of the nutritional management during pregnancy in order to ensure both maternal and fetal health. A general consideration should be introduced, considering also the pleiotropic functions of the involved genes.

We added that our findings indicated the importance of nutritional management during pregnancy.

“Overall, the significant overlaps observed between vitamin A and CAKUT gene sets indicate that vitamin A might be relevant to a broad range of urogenital abnormalities, pointing out the importance of nutritional management during pregnancy for fetal development.”

We thank the reviewer for the suggestion for considering the pleiotropic functions of the involved genes, this can be interesting, however we think that getting into this subject may cause loss of the focus in the conclusion, in which we only state the involvement of vitamin targets in CAKUT related gene sets with the note that our study aims at hypothesis generation.

References:

Please check references, i.e. reference 54 seems not pertinent.

We checked the references and we removed the sentence containing this reference.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research