

From SNPs to Pathways: Interpreting Vitamin D GWAS Results

Kristin Koppelmaa

Supervisors:

Susan Steinbusch-Coort

Martina Summer-Kutmon

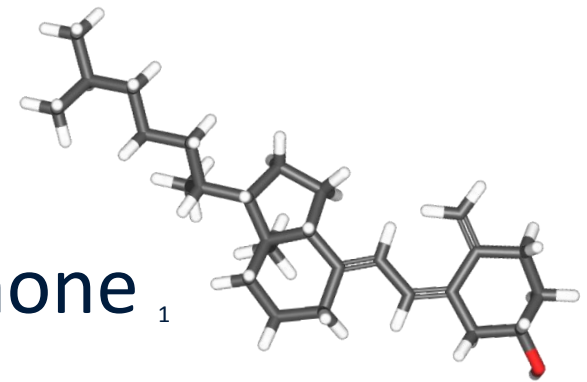
Bachelor thesis presentation

July 1st, 2019



What is vitamin D?

Fat-soluble, secosteroid pro-hormone ¹

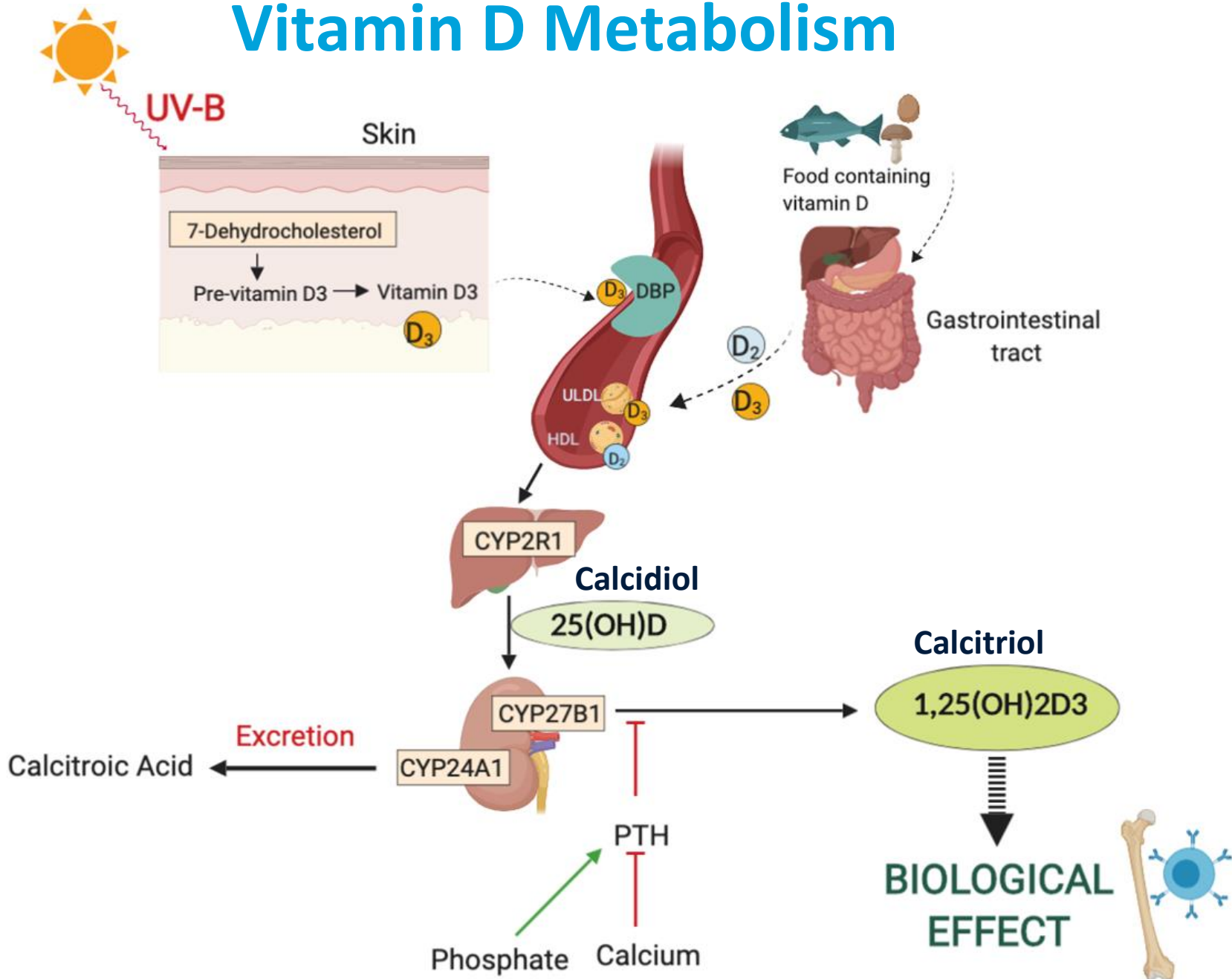


Functions: Bone health, calcium & phosphate homeostasis, immune system ¹

1 billion people are deficient! ²

Health consequences: Osteoporosis, obesity, CVD, cancer, AI diseases ²

Vitamin D Metabolism



Determinants of serum VD concentrations



Environmental factors₃, but also, **GENETICS!**



COMPLEX TRAIT

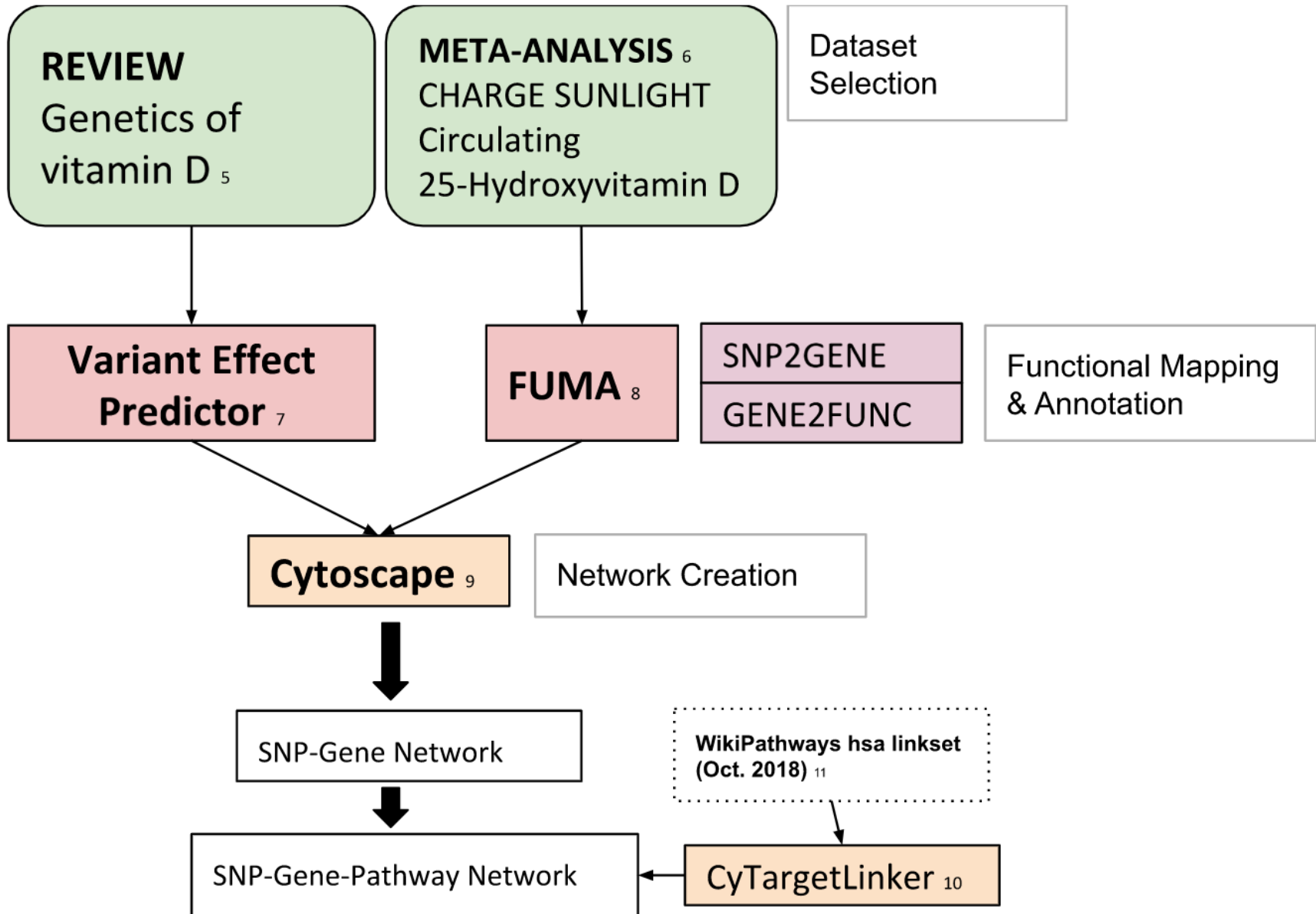
Genome-Wide Association Studies (GWAS)

- Identify **causal variants** of traits, or Single Nucleotide Polymorphisms (SNPs) ⁴
- **Lead SNP** p-value = 5×10^{-8} ⁴
- **Suggestive/Candidate SNP** p-value = 5×10^{-6} to 10^{-4}
- **Vitamin D GWAS:** *GC, CYP2R1, CYP24A1, NADSYN1/DHCR7, AMDHD1, SEC23A* ⁵

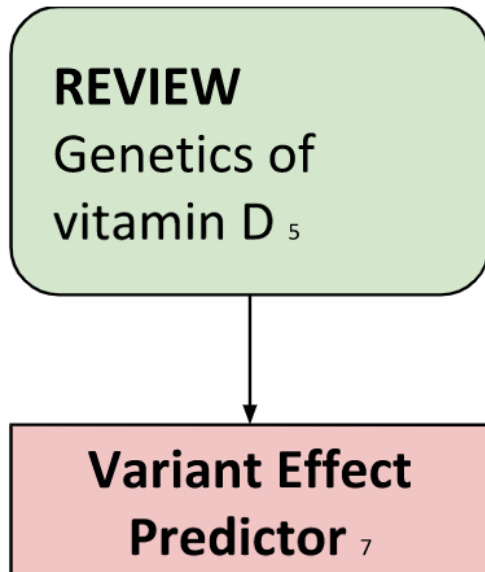
Aim

To investigate the **biological effects** of the **variation in genetic determinants** of **vitamin D serum concentration.**

Materials & Methods



Materials & Methods



Variant set 1₅

22 Significant SNPs

42 Suggestive SNPs

VEP

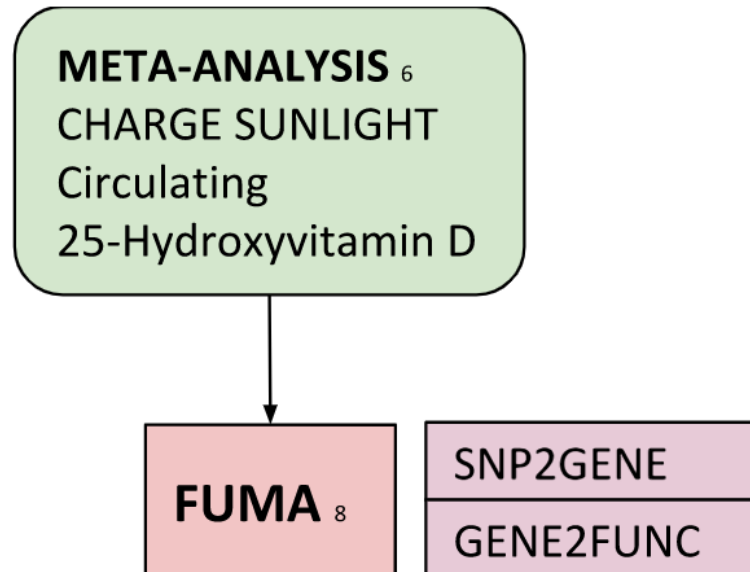
- Genetic mapping
- Location & functional consequence
 - **SIFT** score (**S**orting **I**ntolerant **F**rom **T**olerant)
Scale 0 (deleterious) - 1 (tolerated)

Materials & Methods

Variant set 2

1.048.576
variants

GWAS summary
statistics



Functional Mapping and Annotation of GWAS

1) SNP2GENE

- Lead SNPs
- **CADD** score (threshold **12.37**)
- eQTL mapping

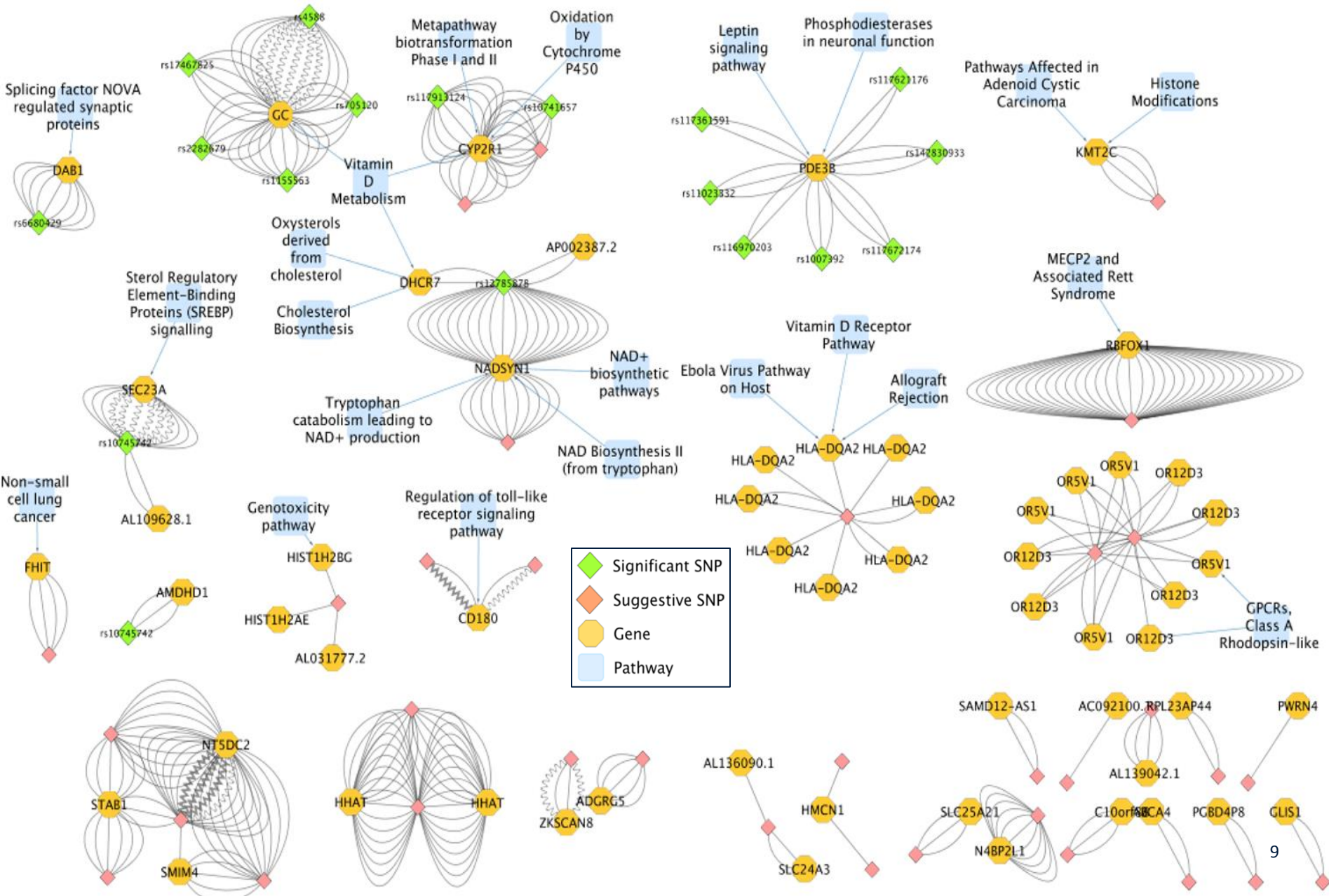
2) GENE2FUNC

- Tissue gene expression heatmap

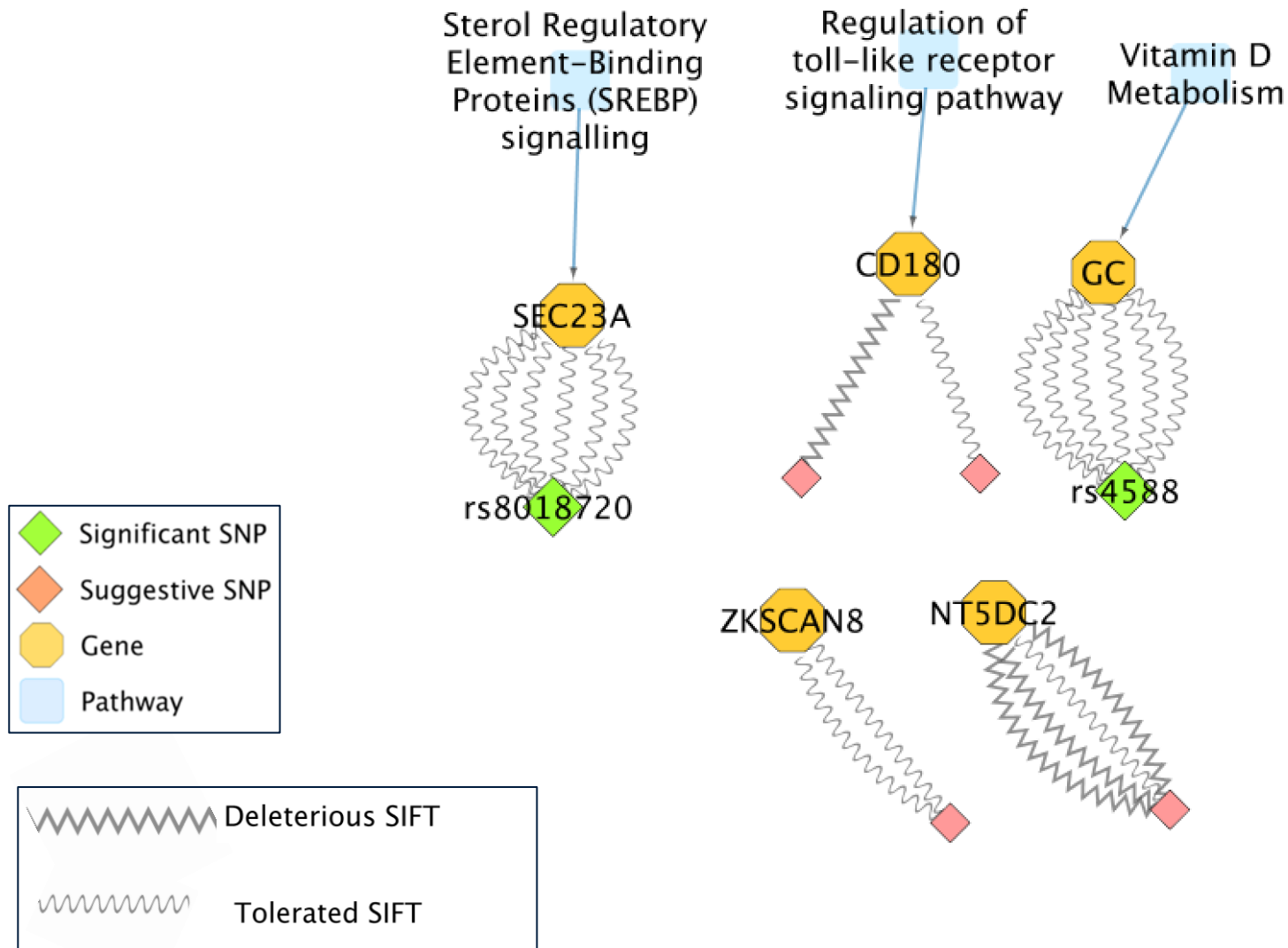
6. Jiang X, O'Reilly PF, Aschard H, Hsu YH, Richards JB, Dupuis J, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nat Commun. 2018;9(1):260.

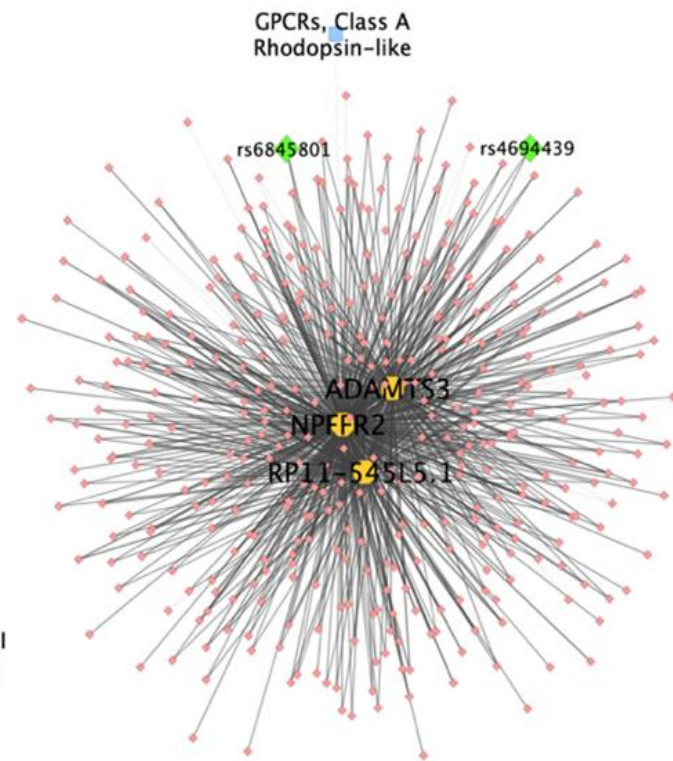
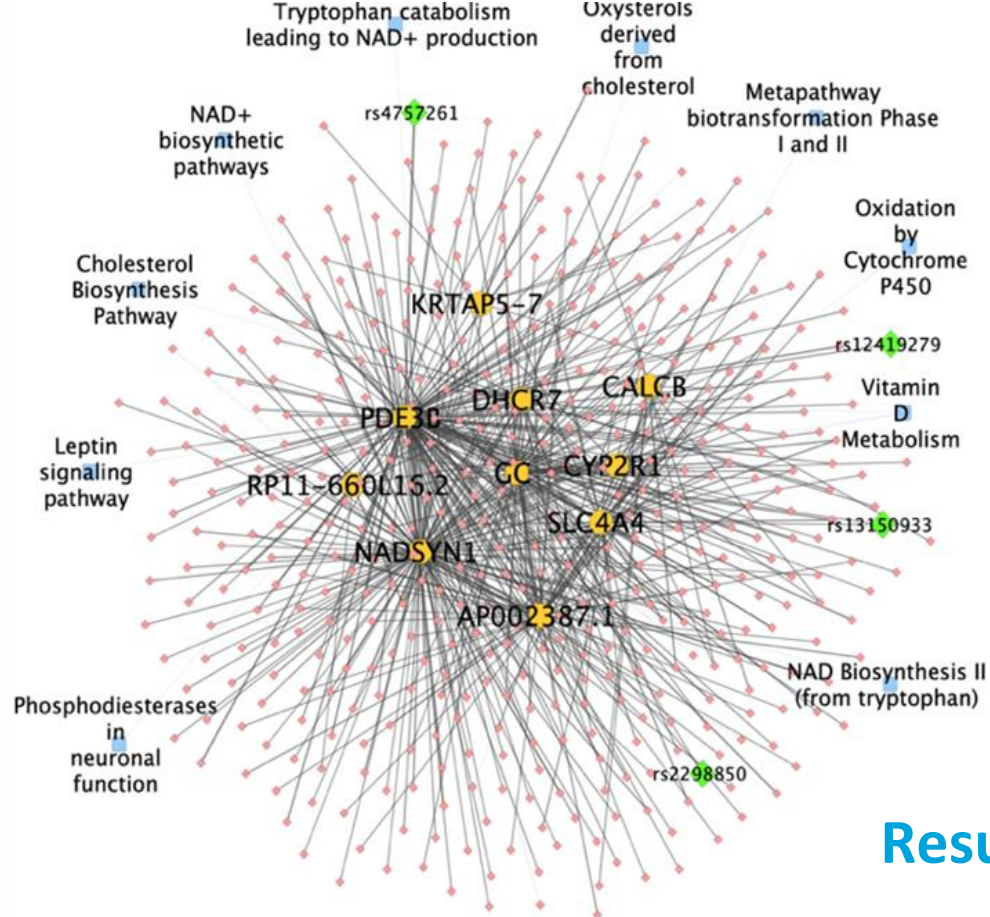
8. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826.

Results: Variant set 1 Network

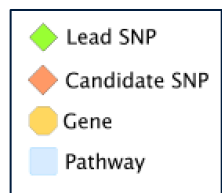
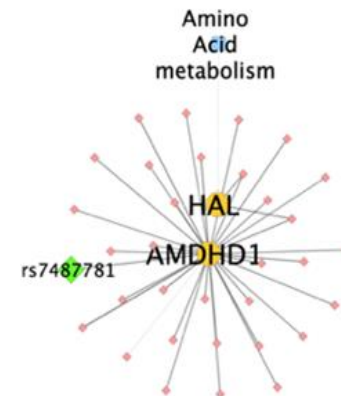
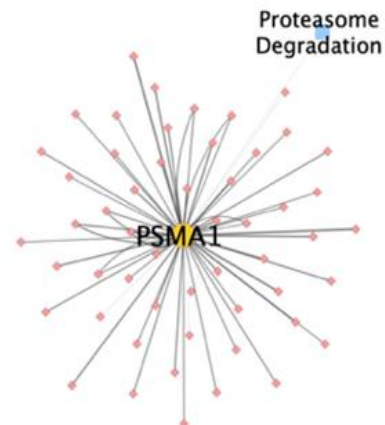
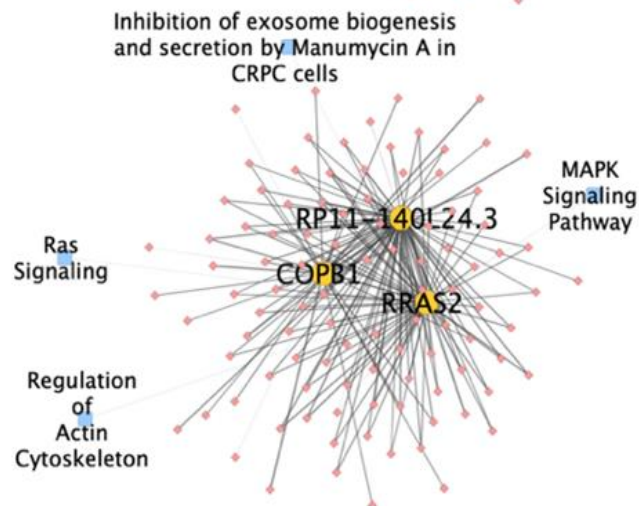


Results: Filtered Variant Set 1 Network w. SIFT score

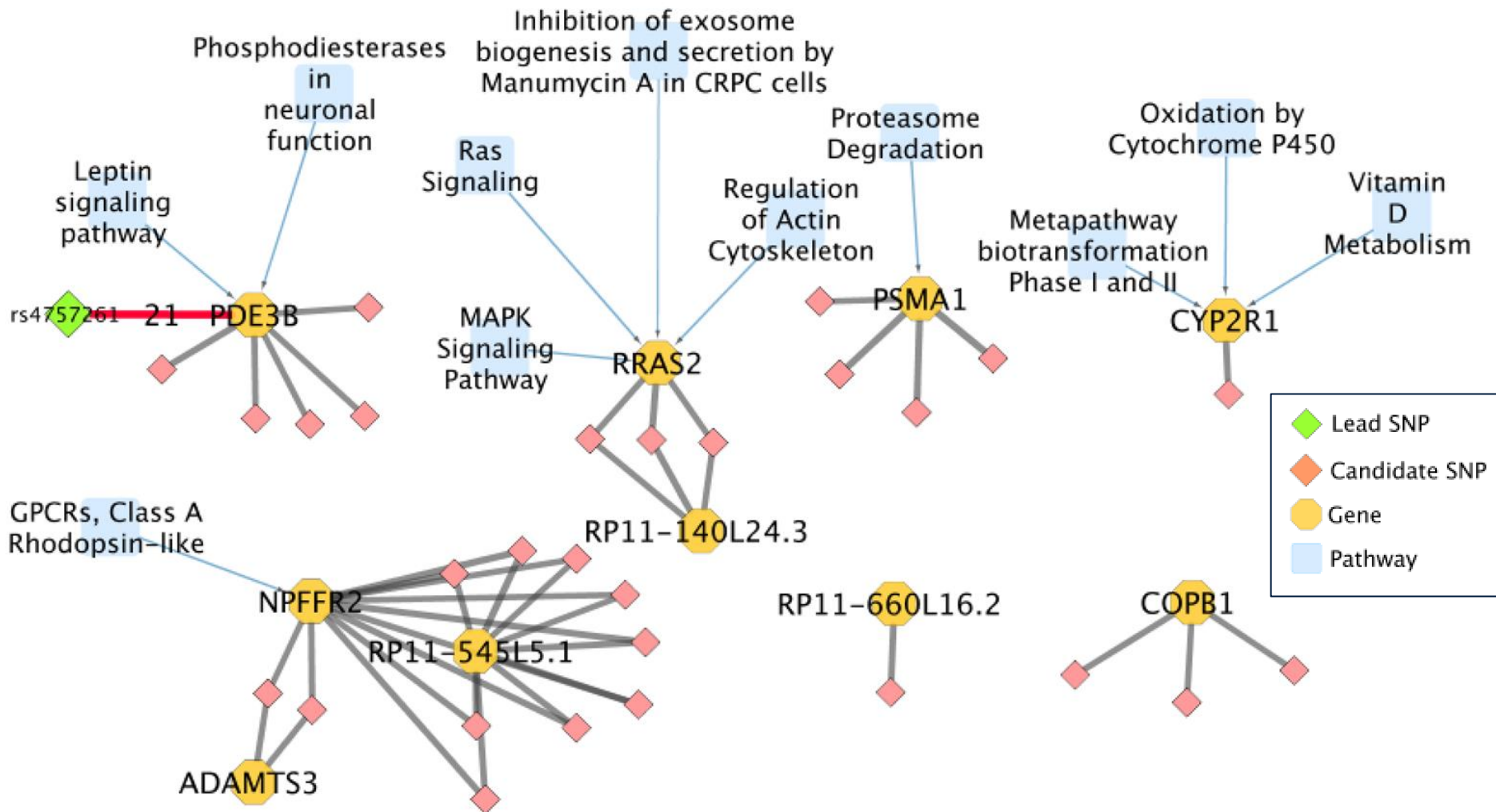




Results: Variant set 2 Network



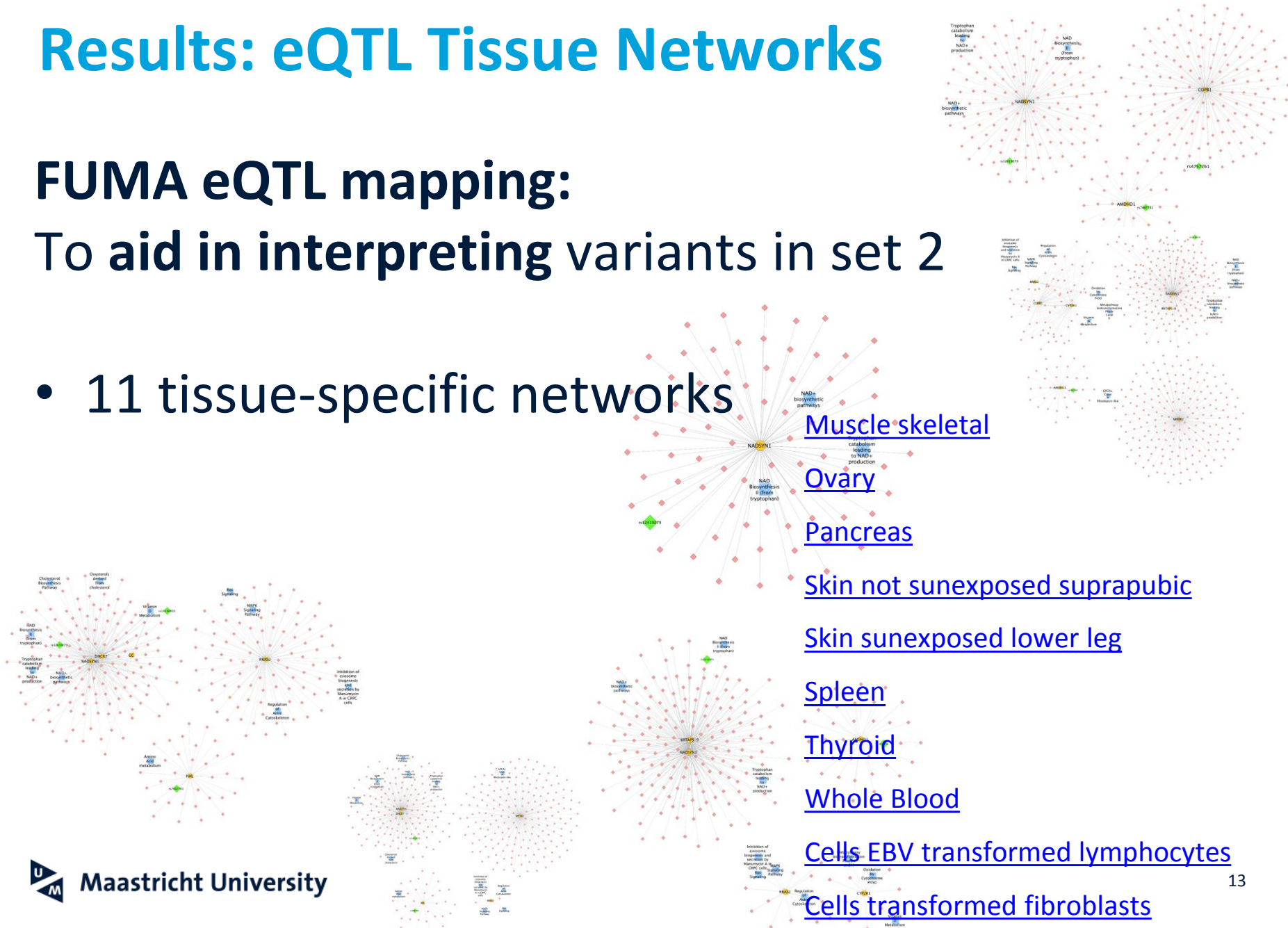
Results: Filtered Variant Set 2 Network w. CADD score >12.37



Results: eQTL Tissue Networks

FUMA eQTL mapping: To aid in interpreting variants in set 2

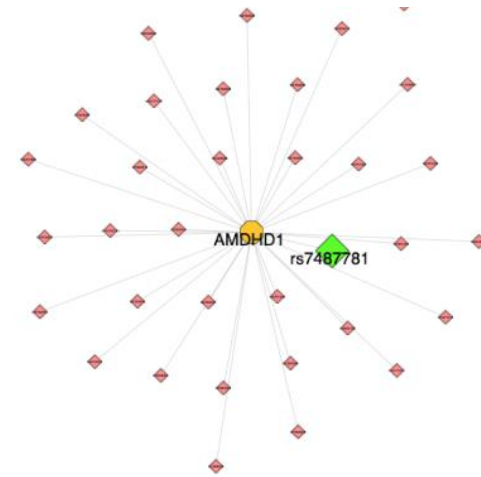
- 11 tissue-specific networks



Discussion

Variant set 2 lead SNP **rs7487781**

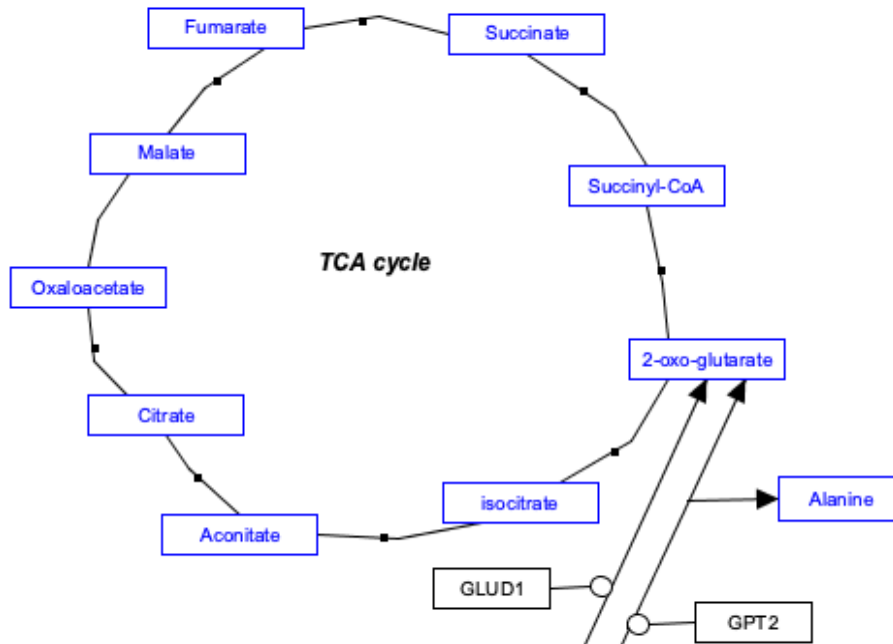
- in *AMDHD1* and *HAL* genes
- Intronic
- Expressed in many eQTL tissues!



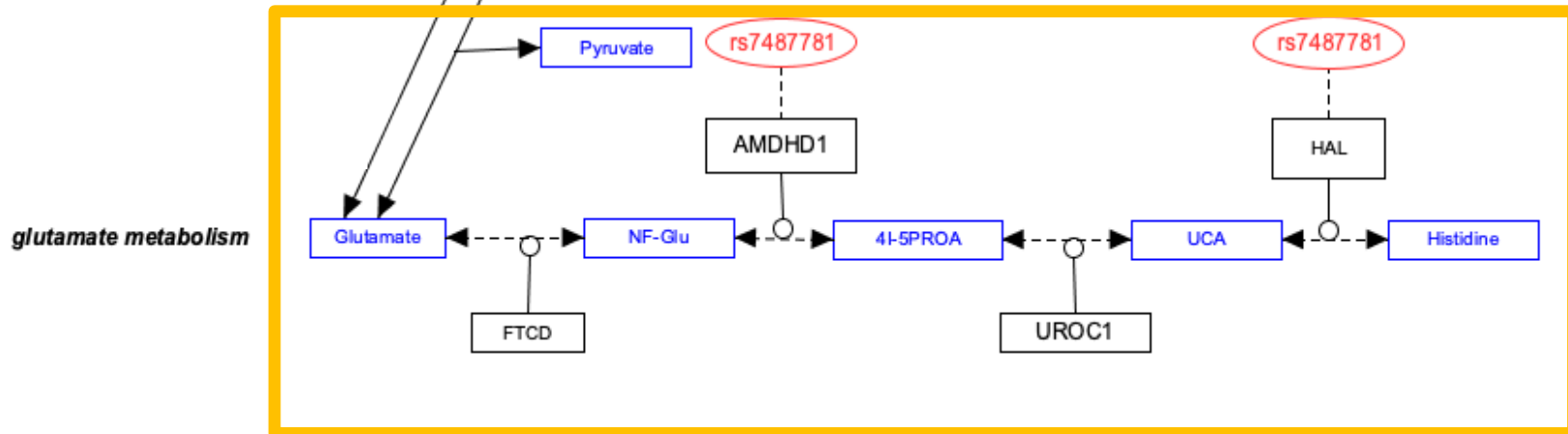
Histidine catabolism

Histidine $\longrightarrow$ Glutamate $\longrightarrow$ TCA cycle

Discussion - Regulation of the TCA Cycle



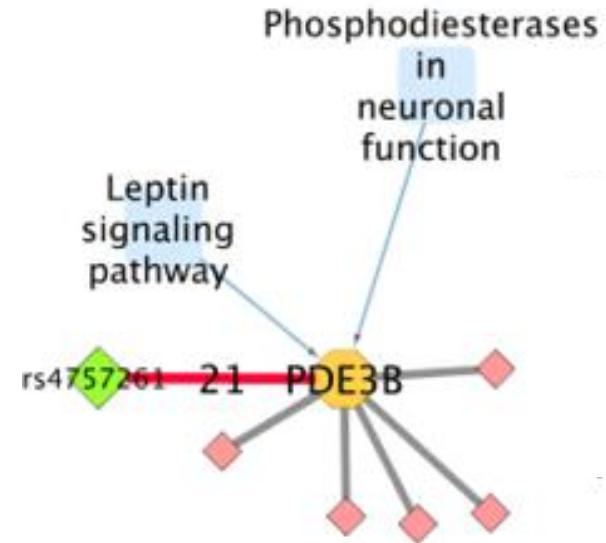
- Vitamin D regulates immunometabolism₉



Discussion

Variant set 2 lead SNP **rs4757261**

- In **PDE3B** gene
- Intronic variant
- Alternative splicing
- Expressed in pancreas



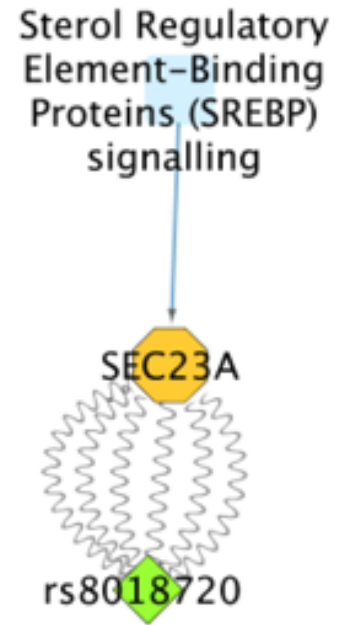
Leptin signaling

- Leptin regulates insulin secretion via PDE3B activation ¹
- Regulation of lipolysis ¹¹

Discussion

Variant set 1 **significant** SNP **rs8018720**

- in **SEC23A** (Sec23 Homolog A, Coat Complex II Component)
- Missense mutation
Leucine → Isoleucine or Valine



Lipid Metabolism

SREBP's = transcription factors
regulate lipid metabolism, cholesterol synthesis ¹²

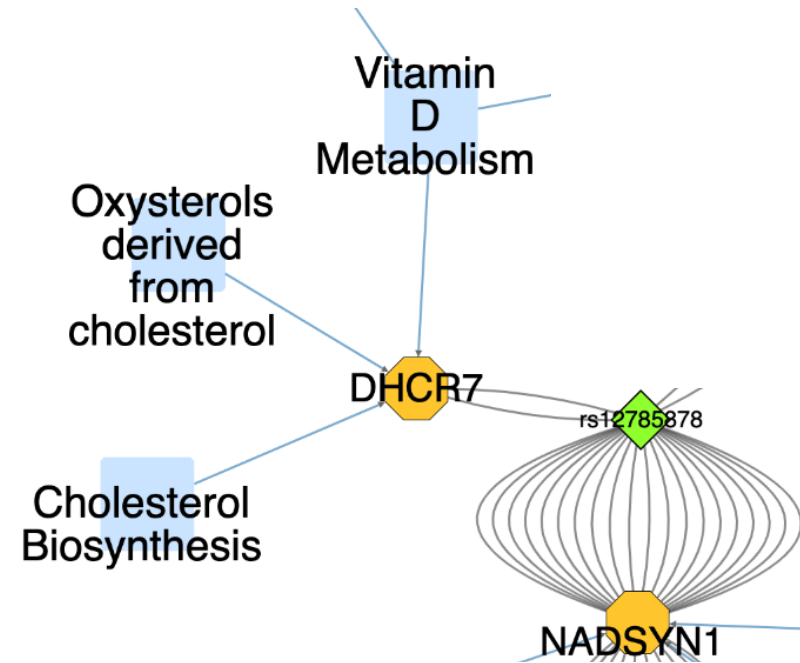
25(OH)D regulates SREBP signalling ¹³

Discussion

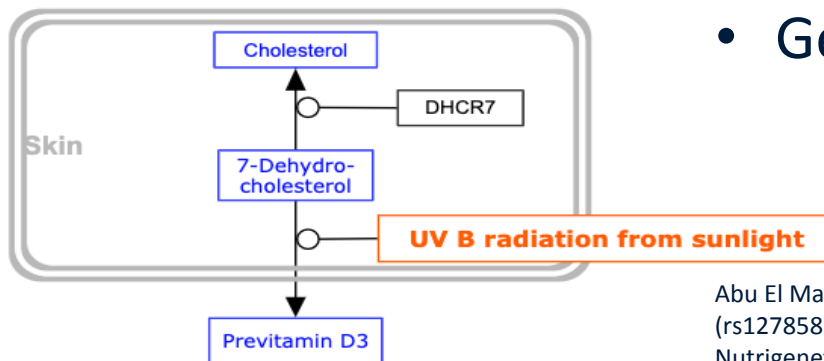
Variant set 1 significant variant

rs12785878

- In ***DHCR7*** and ***NADSYN1*** genes
- Intronic
- Expressed in eQTL skin tissue



Cholesterol biosynthesis



- Genetic marker for CVD ¹⁴

Abu El Maaty MA, Hassanein SI, Sleem HM, Gad MZ. Effect of polymorphisms in the NADSYN1/DHCR7 (rs12785878 and rs1790349) on plasma 25-hydroxyvitamin D levels and coronary artery disease risk. *Nutrigenet Nutrigenomics*. 2013;6(6):327-35.

Conclusion

Biological processes

- Cellular metabolism
- Insulin insensitivity
- Lipolysis
- Lipid (especially cholesterol) synthesis

Possibly related to:

- Obesity, T2D, CVD

However...

- Confirmation of results required
- More diverse studies are necessary



Future Outlook

- Pathway-analysis
- Full understanding of the trait of VD serum levels
- Personalised guidelines e.g. lifestyle changes, supplements
- Prevention

Thank you for your attention!

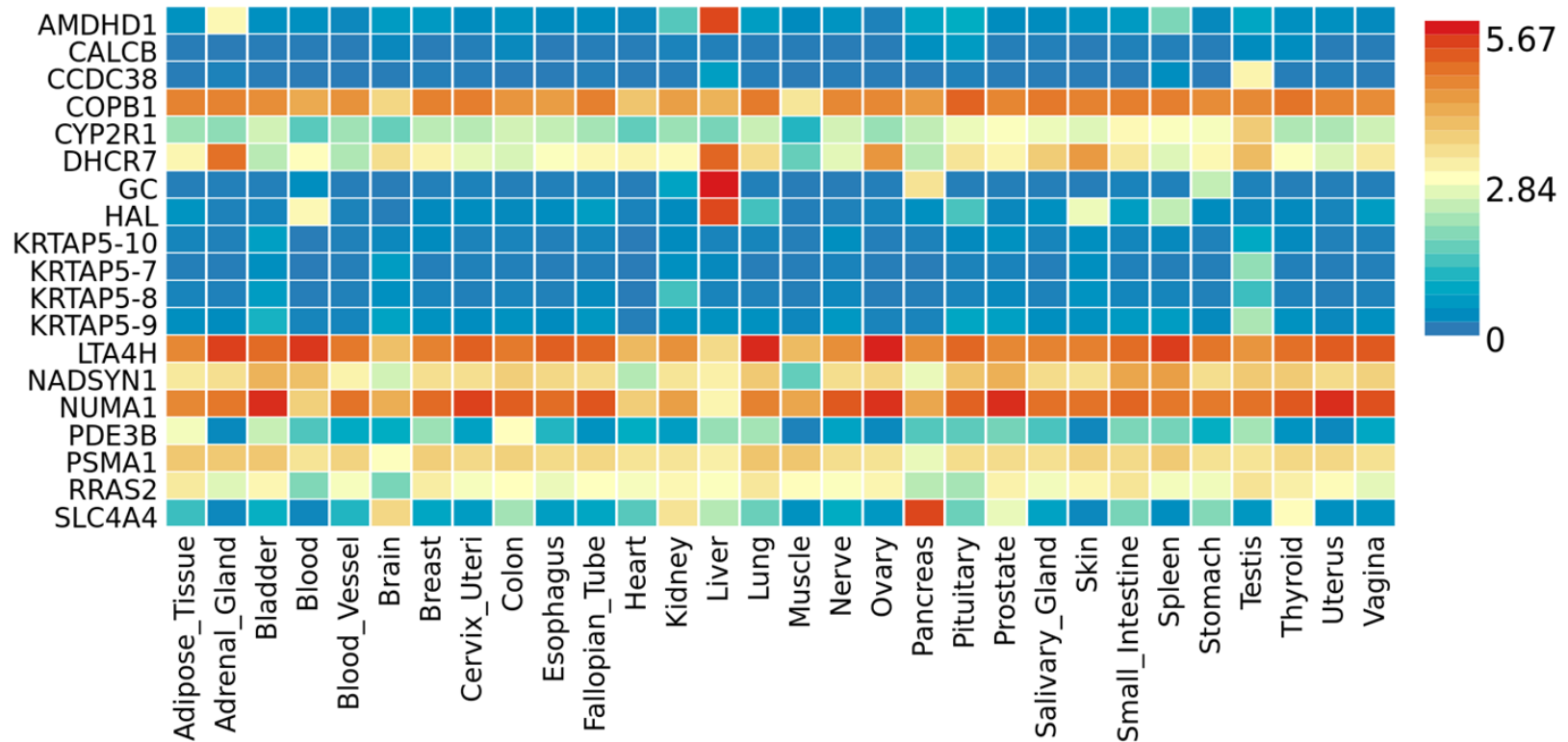


Question time!

Additional slides!

Gene expression heatmap (FUMA GENE2FUNC)

Significant SNPs



Gene expression

Database: GTEx 30 general tissue types

GENE		NADSYN1	AMDHD1	HAL	DHCR7	PDE3B	COPB1
Lead SNP		rs12419279	rs7487781	rs7487781	rs12419279	rs4757261	rs4757261
T I S S U E	Pancreas	✓	✓				
	Skin_Not_Sun_Exposed_Suprapubic	✓		✓	✓	✓	
	Skin_Sun_Exposed_Lower_leg	✓		✓	✓		
	Thyroid	✓	✓				✓
	Whole_Blood	✓	✓				✓
	Spleen	✓					
	Adrenal_Gland	✓					
	Cells_EBV-transformed_lymphocytes	✓					✓
	Cells_Transformed_fibroblasts	✓	✓				
	Muscle_Skeletal	✓	✓				
	Ovary	✓					

Representation of genes (and variant 2 lead SNPs) expressed in Cytoscape networks of 11 different GTEx v7 tissues.

Lead SNP rs12419279 in NADSYN 1 is expressed in all selected GTEx tissues, AMDHD1 in 5 GTEx tissues, COPB1 in 3, HAL and DHCR7 in 2, and PDE3B in 1 GTEx tissue.

Additional slides!

Variant set 2 lead SNPs

rsID	uniqID	Chromosome	Position	P-value
rs13150933	4:72531863:C:T	4	72531863	1.26E-19
rs2298850	4:72614267:C:G	4	72614267	0
rs6845801	4:72730996:A:G	4	72730996	5.20E-34
rs4694439	4:72756696:A:G	4	72756696	4.69E-35
rs4757261	11:14669802:A:C	11	14669802	1.86E-27
rs12419279	11:71139061:A:T	11	71139061	2.58E-50
rs7487781	12:96355376:G:T	12	96355376	1.41E-10