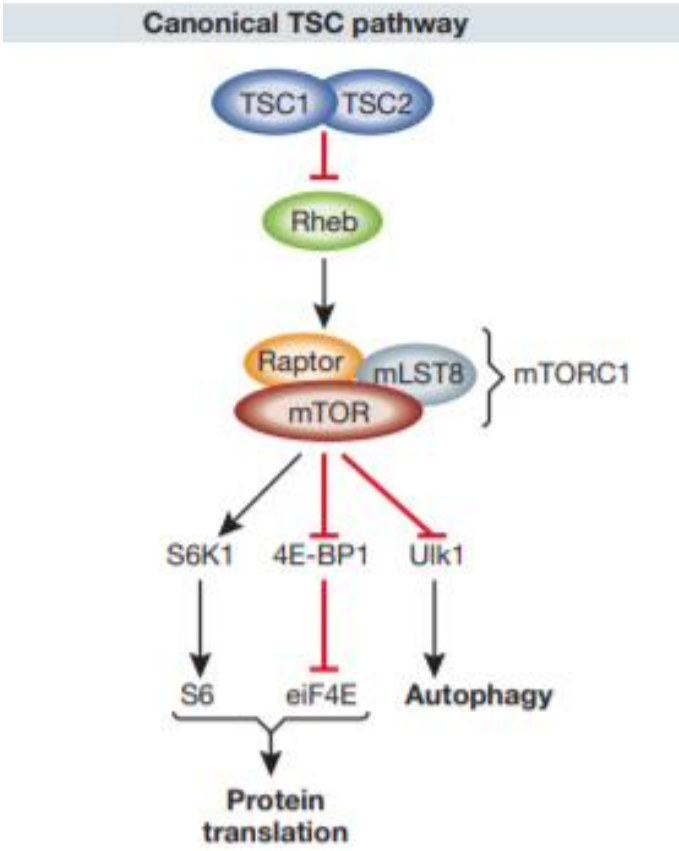
Understanding the gene expression differences between severe and mild forms of Tuberous Sclerosis Complex (TSC) and the role of genetic modifiers to identify novel drug targets

Abigail Howell
Neuroscience 2021

Tuberous Sclerosis Complex (TSC)

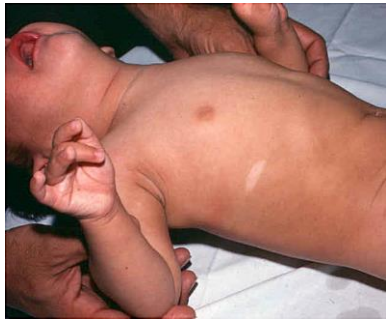
- Tuberous sclerosis complex (TSC) is an autosomal dominant disorder associated with mutations in TSC1 or TSC2
- Protein products of TSC1 and TSC2 inhibit the protein Ras homolog enriched in brain (Rheb), which regulates mTOR Complex 1 (mTORC1), which regulates many cellular functions including growth and survival
- Mutations in TSC1/TSC2 result in constitutive activation of the mTOR signaling pathway which causes increased cell size and cell proliferation
- This leads to the development of numerous noncancerous tumors in many parts of the body (skin, brain, kidneys, and other organs) causing significant health problems



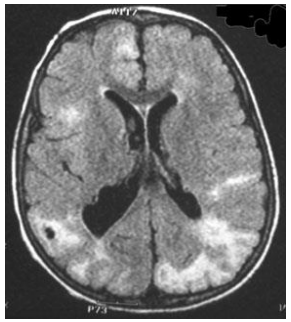
https://pubmed.ncbi.nlm.nih.gov/21412983/

Tuberous Sclerosis Complex (TSC)

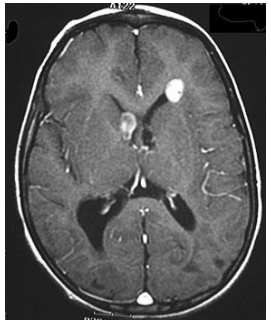
- This leads to the development of numerous noncancerous tumors in many parts of the body (skin, brain, kidneys, and other organs) causing significant health problems



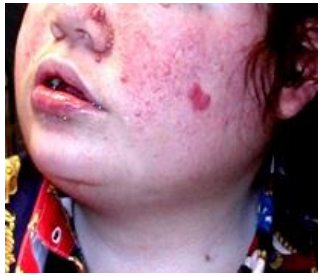
Hypomelanotic macules



Cortical tuber

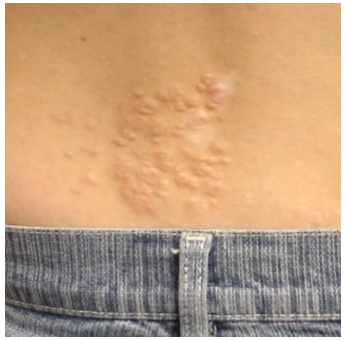


SEGA (subependymal giant cell astrocytoma)

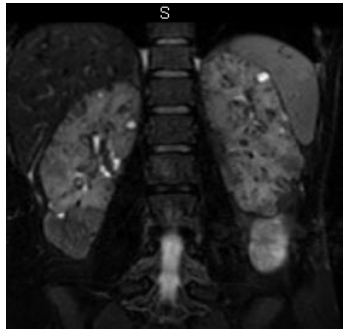


Facial angiofibromas

Periungual fibroma



Shagreen patch



Angiomyolipoma

tgen AN AFFILIATE OF Ci

Tuberous Sclerosis Complex (TSC)

- TSC demonstrates variable severity - affected family members with the same mutation may have mild or severe disease

- Modifier genes could influence disease severity in TSC (mTOR variants)

- **Hypothesis:** Genetic variants falling within a network, or variants that affect expression of genes in such a network, may modify the phenotype of a disease caused by one of the genes in the network

- Goals of this project are:
  - To identify genetic modifiers in Whole Exome Sequencing (WES) data that may be correlated with expression differences in severe vs mildly afflicted patients
  - To use RNAseq data to identify **differential gene expression** in **severe vs mildly** afflicted patients

- In order to develop a method of molecular profiling to predict disease severity in patients to allow for early initiation of disease modifying therapies

Variant summary WES data

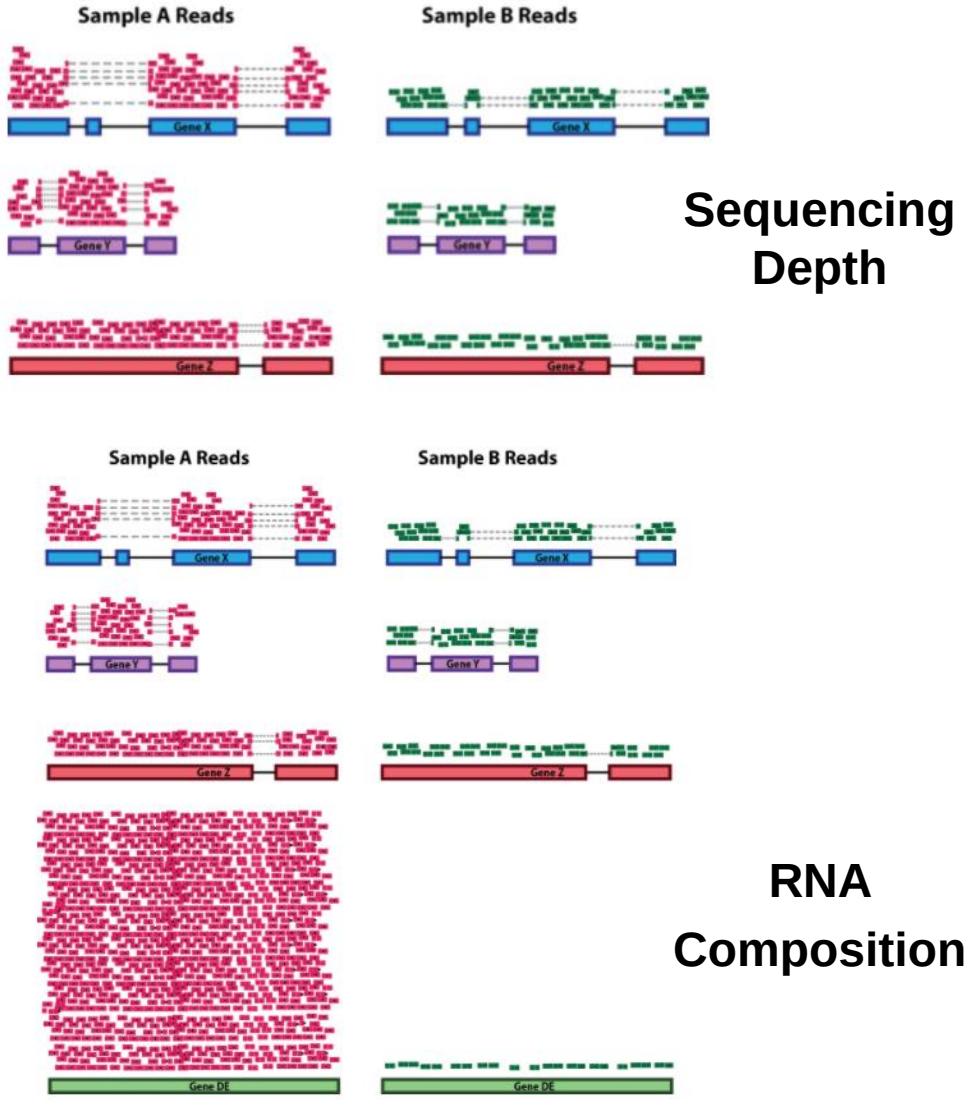
- Illumina FASTQ files aligned to build 37 of the human reference genome using BWA
- BAM files were deduplicated and base quality scores were recalibrated using GATK and Picard
- GATK HaplotypeCaller was used to detect variants in the exome data
- Filtering was used to identify private variants or rare variants with a maximum population frequency of 0.05

	DoD															
	N007		N013		N015		N014		N022		N030		N033		N035	
	proband (M)	affected mother	proband (F)	affected mother	proband (F)	affected father	proband (F)	affected mother	proband (F)	affected mother	proband (M)	affected father	proband (M)	affected mother	proband (M)	affected father)
TSC1	splice variant, c.2626-3C>T, (0/1)	splice variant, c.2626-3C>T, (0/1); M209V (0/1)	splice variant, c.2626-3C>T, (0/1); nonsense (0/1)	splice variant, c.2626-3C>T, (0/1); nonsense (0/1)											FS (0/1)	FS (0/1)
TSC2	FS (0/1)	FS (0/1)			L105R (0/1)	L105R (0/1)	A583T (0/1)	A583T (0/1)	splice variant, c.4663-1G>T(), (0/1)	splice variant, c.4663-1G>T(), (0/1)	insertion (0/1)	insertion (0/1)	splice variant, c.1882C>G (0/1)	splice variant, c.1882C>G (0/1)		
CLIP1						P161R (0/1)										
DDIT4L												K180fs (0/1)				
DEPDC5										G1134R (0/1)						
DVL1		R150C (0/1)														
EIF4G1					L290F (0/1)						P1236A (0/1)				S15P (0/1)	
FZD1	P93dup (0/1)	P93dup (1/1)					P93dup (0/1)	P93dup (1/1)						P93dup (0/1)		P93dup (0/1)
FZD7																G196E (0/1)
FZD8										C118F (0/1)						
IRS1								A512P (0/1)							A512P (0/1)	
LRP5														A1525V (0/1)		
LRP6												V1382F (0/1)				
MIOS												K101R (0/1)				
PIK3C2A	T204I (0/1)					T204I (0/1)										
PIK3C2B										F1029L (0/1)						
PIK3R3											N390S (0/1)					
SLC38A9			S182T (0/1)	S182T (1/1)												
SOS1						R497Q (0/1)										
SREBF1							R673C (0/1)									
TNFRSF11A			A192V (0/1)													
TNFRSF14				SV (0/1)												
TNFRSF19											A169V (0/1)					
TNFRSF4		R10H (0/1)														
WNT2B										P21R (0/1)						
WNT6			P155R (0/1)	P155R (1/1)												
WNT8B						F313S (0/1)										

tgen® AN AFFILIATE OF

Differential Expression Between Severe and Mild TSC Patients

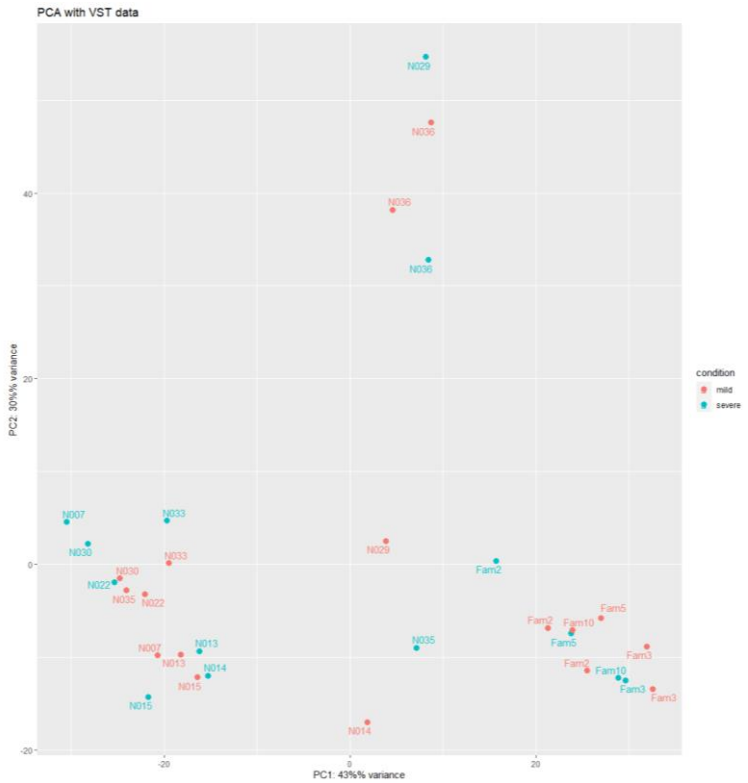
- **Differential Expression** - testing is to determine which genes are expressed at different levels between conditions
- **DESeq2** normalization – controlling for sequencing depth and RNA composition
- Normalized counts are fit to a GLM using a negative binomial – and then a Wald test (modified t-test) is performed to see if model coefficients differ from zero
- **Model = condition + family**



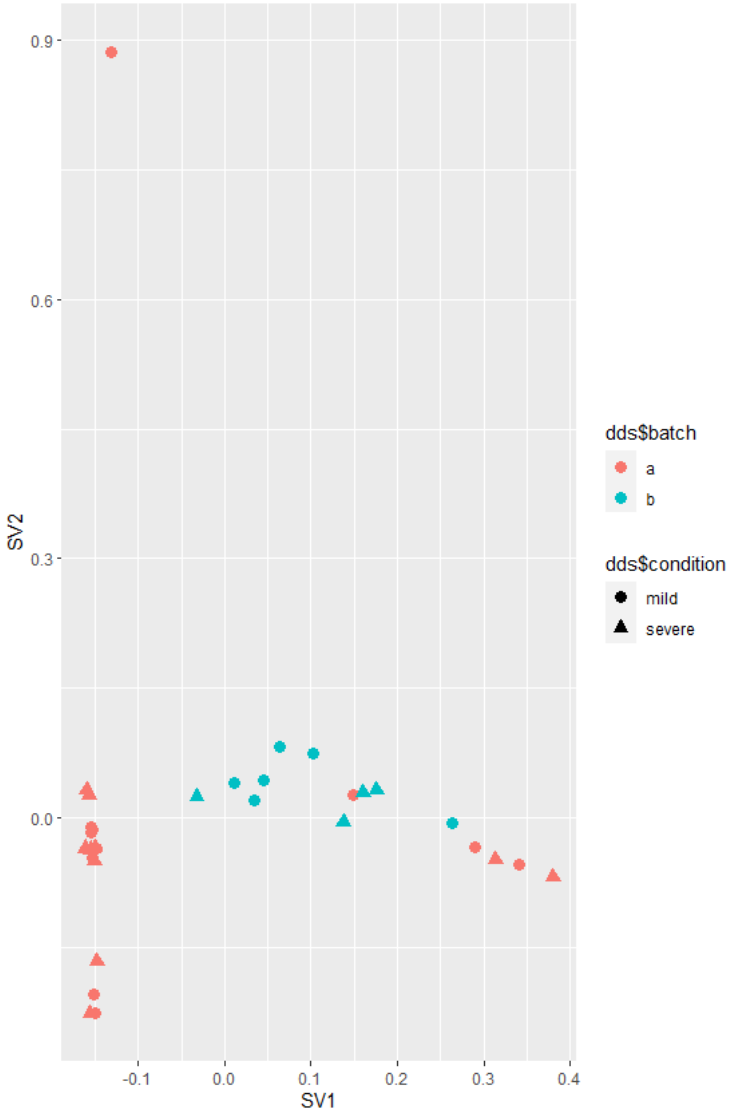
https://hbctraining.github.io/DGE_workshop/lessons/02_DGE_count_normalization.html

Exploratory Analysis and Visualization

- Principal Component Analysis revealed clustering by families rather than by severe vs. mild TSC phenotype – highlighting the importance of the unique genetic background of each family



- Surrogate variable analysis identified hidden batch effects of collection timepoints and were corrected for in the differential expression model



tgen AN AFFILIATE OF City

Differential Gene Expression Analysis Between Parent and Child with Mild and Severe Phenotype

9 genes upregulated in the Severe condition
FBLN2*
ZNF860
PLPP3
IGHG1*
DNTT
FSIP2
ZNF683
FCRL3
NRCAM*

20 genes downregulated in the Severe condition	
UPP1	MBNL3
YOD1	RN7SL752P*
ZNF185	XK*
F13A1*	ITGA2B*
SPARC*	PRKG1
SH3BGRL2	ITGB5
RHOBTB1	CAT
MTURN	SELP
TREML1*	FOXO3
BMP6*	RAB6B

*log fold change > 1

Functional Enrichment Analysis of Gene Expression Data

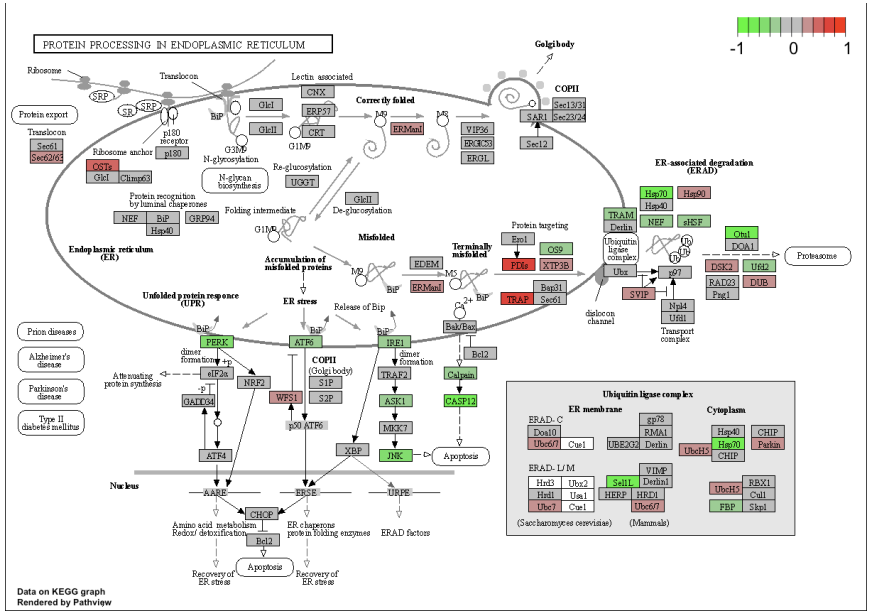
- **Gene sets** - lists of usually functionally related genes
  - **Overrepresentation analysis (ORA):** do your DE genes contribute significantly to a particular set of functionally related genes
- **Databases used:**
- ***KEGG*:** The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway maps representing molecular interaction and reaction networks of 7 categories: metabolism, genetic and environmental information processing, cellular processes, organismal systems, human diseases, and drug development.

Overrepresentation Analysis

- 23 significant gene sets identified with the KEGG database
- Including gene sets related to:
  - Protein processing in endoplasmic reticulum
  - Cardiomyopathy
  - PI3K-Akt signaling pathway
  - T-cell Receptor Signaling
  - Cell Adhesion Molecules
  - FoxO signaling pathway

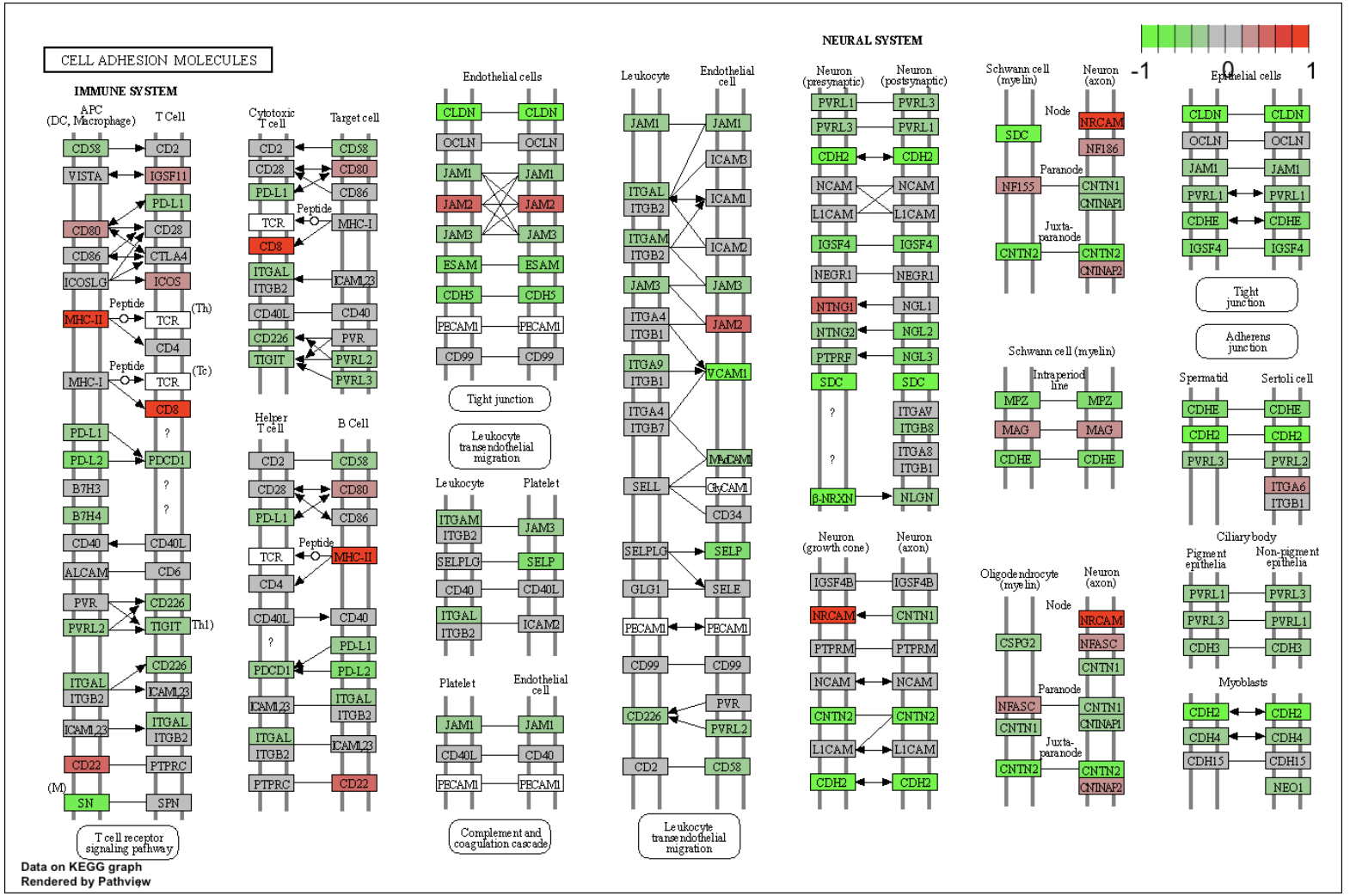
Protein processing in endoplasmic reticulum ORA KEGG

- Protein processing in endoplasmic reticulum
- **Human TSC-associated renal angiomyolipoma cells are hypersensitive to ER stress:** *“Endoplasmic reticulum (ER) stress can develop in TSC-associated cells due to mTORC1-driven protein translation.”*



Cell adhesion molecules ORA KEGG

- Cell adhesion molecules
- **TSC2 modulates cell adhesion and migration via integrin-α1β1:** *"Here we show that loss of TSC2 increased both the attachment and spreading of mouse embryonic fibroblasts to the extracellular matrix proteins collagen type I and fibronectin"*



T-cell Receptor Signaling ORA KEGG

- mTOR-mediated signaling has been implicated in both adaptive and innate immune cell development and function
  - mTORC2 has been found to promote Th2 immune response in T-cells through phosphorylation of the protein kinase Cθ
  - mTOR signaling is also involved in the regulation of effector T cell differentiation, regulatory T cell generation and function, memory T cell responses, and T cell trafficking

Summary

- Identified genetic variants in mTOR pathway genes (genetic modifiers) that may be contributing to differential gene expression
- Identified differentially expressed genes between severe and mild forms of TSC that can contribute to phenotypic variability
- Functional enrichment demonstrated these differentially expressed genes contribute mainly to gene sets related to the PI3K-Akt signaling pathway, T-cell Receptor Signaling, Cell Adhesion Molecules, and the FoxO signaling pathway