The transcriptome as a window into pathogenesis of type 1 diabetes

Nat Goodman
Institute for Systems Biology
ICSB 2006

Type 1 diabetes research at ISB

- Projects in this talk
  1) Beta cell gene atlas
  2) Gene regulatory network driving beta cell pathology
- Other projects
  3) T1DBase http://T1DBase.org
  4) Fine scale genetic map of human MHC
     Roach JC et al. Am J Hum Genet. 2006 Oct.
     PMID 16960798
  5) Regeneration of beta cell function
  6) Biomarker discovery

Outline of project 1. Beta cell gene atlas

Rationale Basal gene expression in beta cells defines normal function and provides baseline for abnormal states.

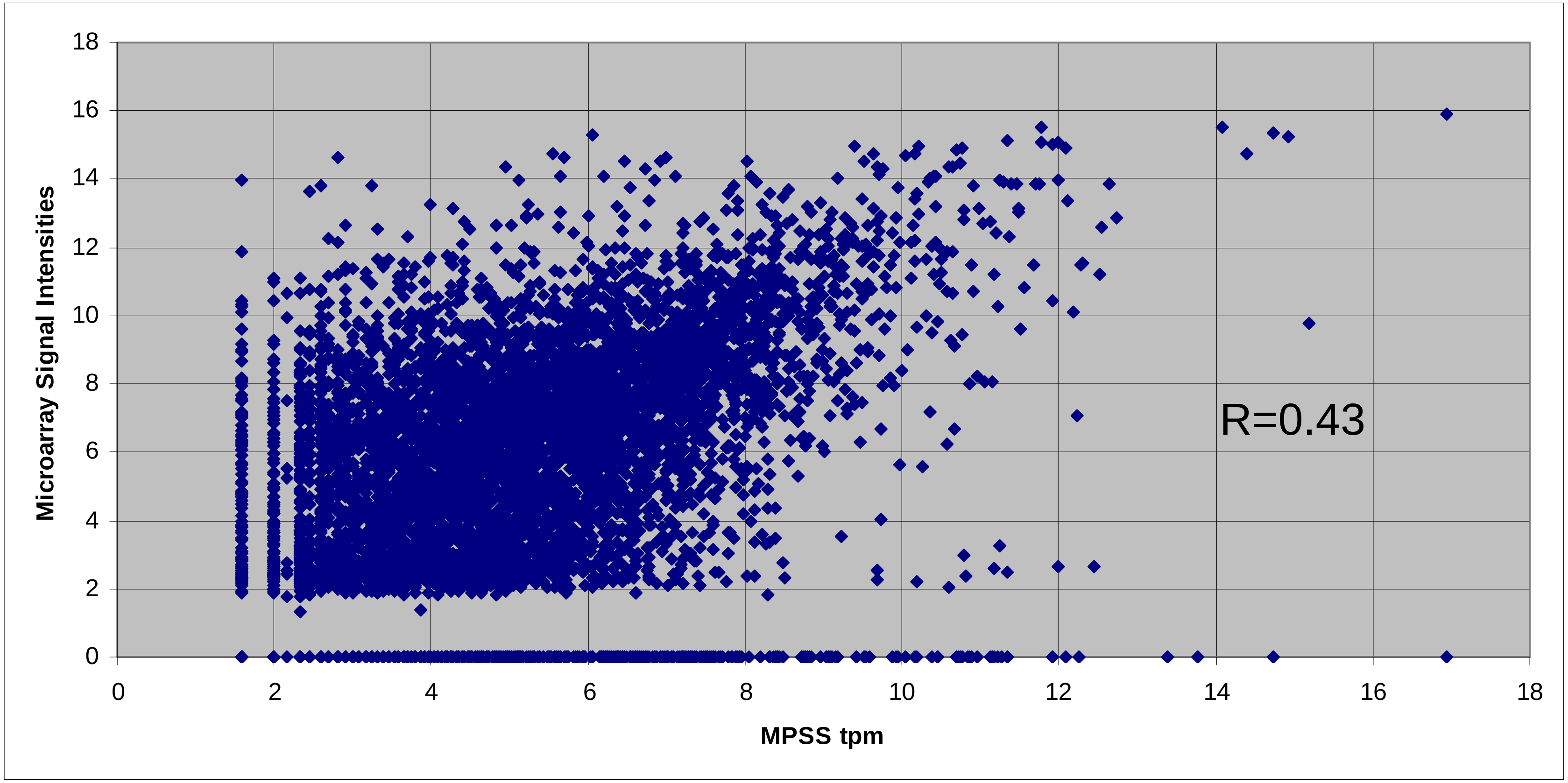
Aims

A. Deep transcriptome analysis of human islets by Massively Parallel Signature Sequencing (MPSS)

B. Integrate MPSS and microarray data (mostly previously published)

C. Generate atlas: comprehensive list of genes basally expressed in beta cells, islets, and cell lines across different species

Comparison of MPSS vs Microarray Data from Untreated Control Islets



Advantages of MPSS:
- Unbiased : The method does not depend upon prior knowledge of the genome
- Sensitive : Library size at least 1,000,000
- Quantitative : No prior distribution

Project 1.
β cell transcriptome

Aim A. Performance of two MPSS analysis of untreated human pancreatic islets

Two islet samples - **1,051,000** and **1,313,239** signatures

Distinct signatures identified: **20,796**

Signature mapped to a unique sequence: **75%**

Signatures high quality: **63.8%**

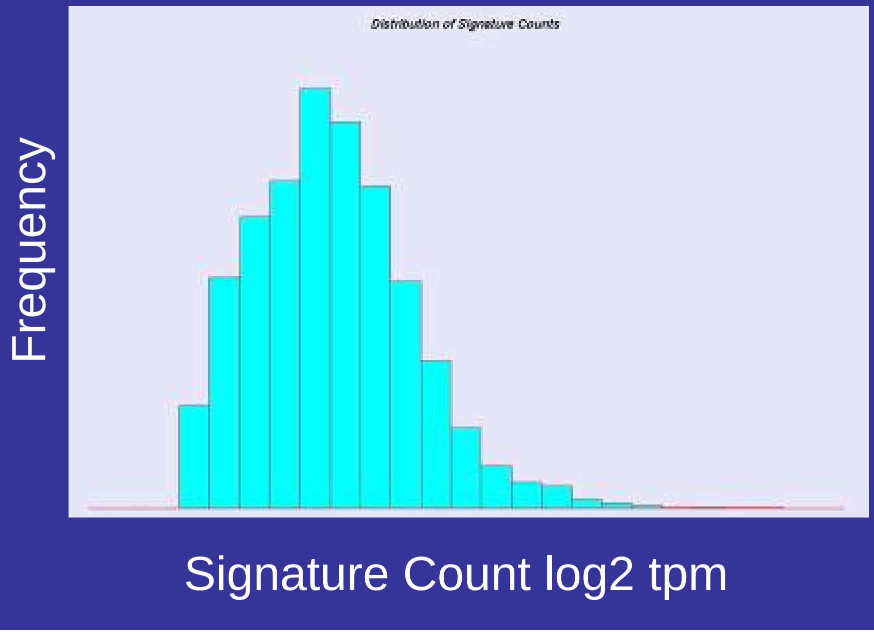
Signatures > 5 tpm
Islet 1: **5,855**
Islet 2: **6,221**

Total # of genes detected: **7,640**

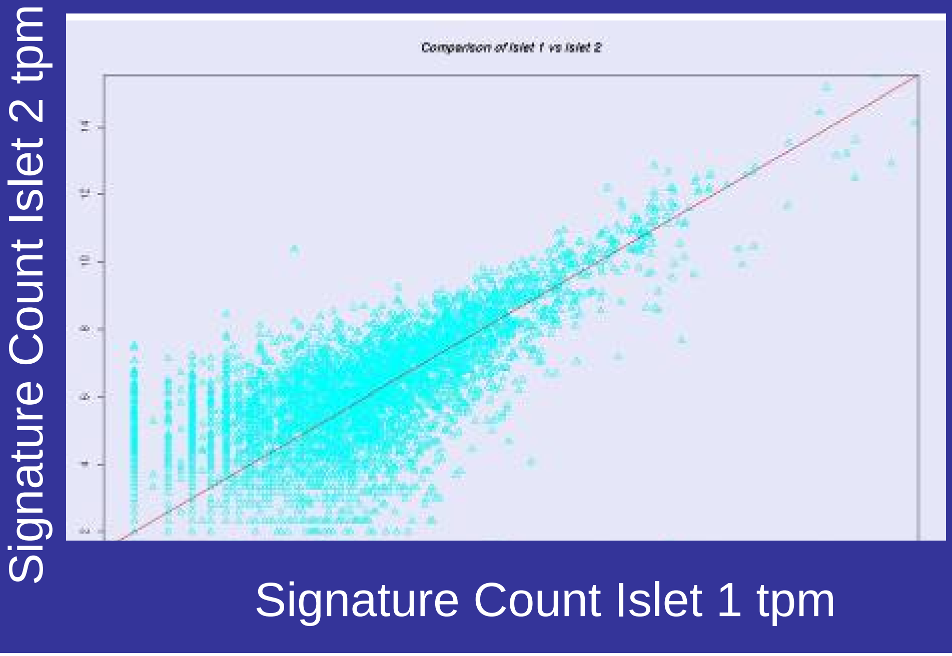
Total # of transcription factors**: 409**

Project 1.
β cell transcriptome

Distribution of MPSS Signature Counts



Scatter Plot of Sample1 vs Sample2



Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
5.0	18.0	45.0	210.7	110.0	142700.0

Project 1.
β cell transcriptome

Signature Counts of the Most Abundant Transcripts

Gene Id	Symbol	Gene Description	Islet 1 (tpm)	Islet 2 (tpm)	Average (tpm)
3630	INS	insulin	119627	144213	131920
5068	REG3A	regenerating islet-derived 3 alpha	28560	47369	37964
5967	REG1A	regenerating islet-derived 1 alpha	44875	17861	31368
2641	GCG	glucagon	15634	37024	26329
5645	TRYP2	protease, serine, 2 (trypsin 2)	34149	7784	20966
2778	GNAS	GNAS complex locus	14405	22141	18273
6750	SST	somatostatin	22176	12652	17414
2495	FTH1	ferritin, heavy polypeptide 1	19839	9480	14659
1504	CTRB1	chymotrypsinogen B1	21986	5727	13856
4535	NADPH	Mitochondrial NADH dehydrogenase 1	17556	9210	13383
1915	EEF1A	eukaryotic translation elongation factor 1a	9884	11889	10886
6137	RPL13	ribosomal protein L13	6656	7261	6958
5968	REG1B	regenerating islet-derived 1 beta	9803	3269	6536
23521	RPL13A	ribosomal protein L13a	6471	6539	6505

Project 1.
β cell transcriptome

Over-represented Biological Processes GO terms
In Top 10 % of MPSS results

GO term	Description	P value*	n**
GO:0008104	protein localization	0.016322	56
GO:0043069	negative regulation of programmed cell death	0.016322	22
GO:0006091	generation of precursor metabolites and energy	0.016322	53
GO:0042775	ATP synthesis coupled electron transport	0.018646	9
GO:0051234	establishment of localization	0.019127	179
GO:0006457	protein folding	0.020346	27
GO:0050875	cellular physiological process	0.022832	514
GO:0009145	purine nucleoside triphosphate biosynthesis	0.022832	13
GO:0009201	ribonucleoside triphosphate biosynthesis	0.022832	13
GO:0009206	purine ribonucleoside triphosphate biosynthesis	0.022832	13
GO:0009142	nucleoside triphosphate biosynthesis	0.029385	13
GO:0045184	establishment of protein localization	0.029385	53

Project 1.
β cell transcriptome

*P value = Fisher's Exact Test corrected by "Benjamini-Yekuteli" procedure
**n = number of genes with the assigned GO term

Under-represented Biological Processes GO terms
In Top 10 % of MPSS results

GO term	Description	P value*	n**
GO:0006351	transcription, DNA-dependent	2.17E-06	46
GO:0006355	regulation of transcription, DNA-dependent	2.95E-06	45
GO:0006350	transcription	3.25E-05	56
GO:0019219	regulation of nucleotide and nucleic acid metabolism	5.17E-05	55
GO:0045449	regulation of transcription	5.23E-05	54
GO:0031323	regulation of cellular metabolism	0.000748	65
GO:0019222	regulation of metabolism	0.002163	70
GO:0006139	nucleotide and nucleic acid metabolism	0.022669	114
GO:0006468	protein amino acid phosphorylation	0.048571	10

*P value = Fisher's Exact Test corrected by "Benjamini-Yekuteli" procedure
**n = number of genes with the assigned GO term

Project 1.
β cell transcriptome

Signature Counts of Some Beta Cell Specific Transcription Factors

	Islet 1 (tpm)	Islet 2 (tpm)
MafA	163	32
IPF1	38	48
NKX6.1	67	90
PAX6	418	703

Pancreatic Islet Enriched Genes in MPSS datasets

Comparison of our MPSS data with 32 human tissues:
An atlas of human gene expression from massively parallel signature sequencing. Jongeneel et al Genome Res 2005 15:1007-1014

Highly-enriched in pancreatic islets:
950 genes

Genes-depleted in islets:
255 genes

Definition of enrichment: 4 standard deviations from mean

Aim B. Integrate MPSS and array data

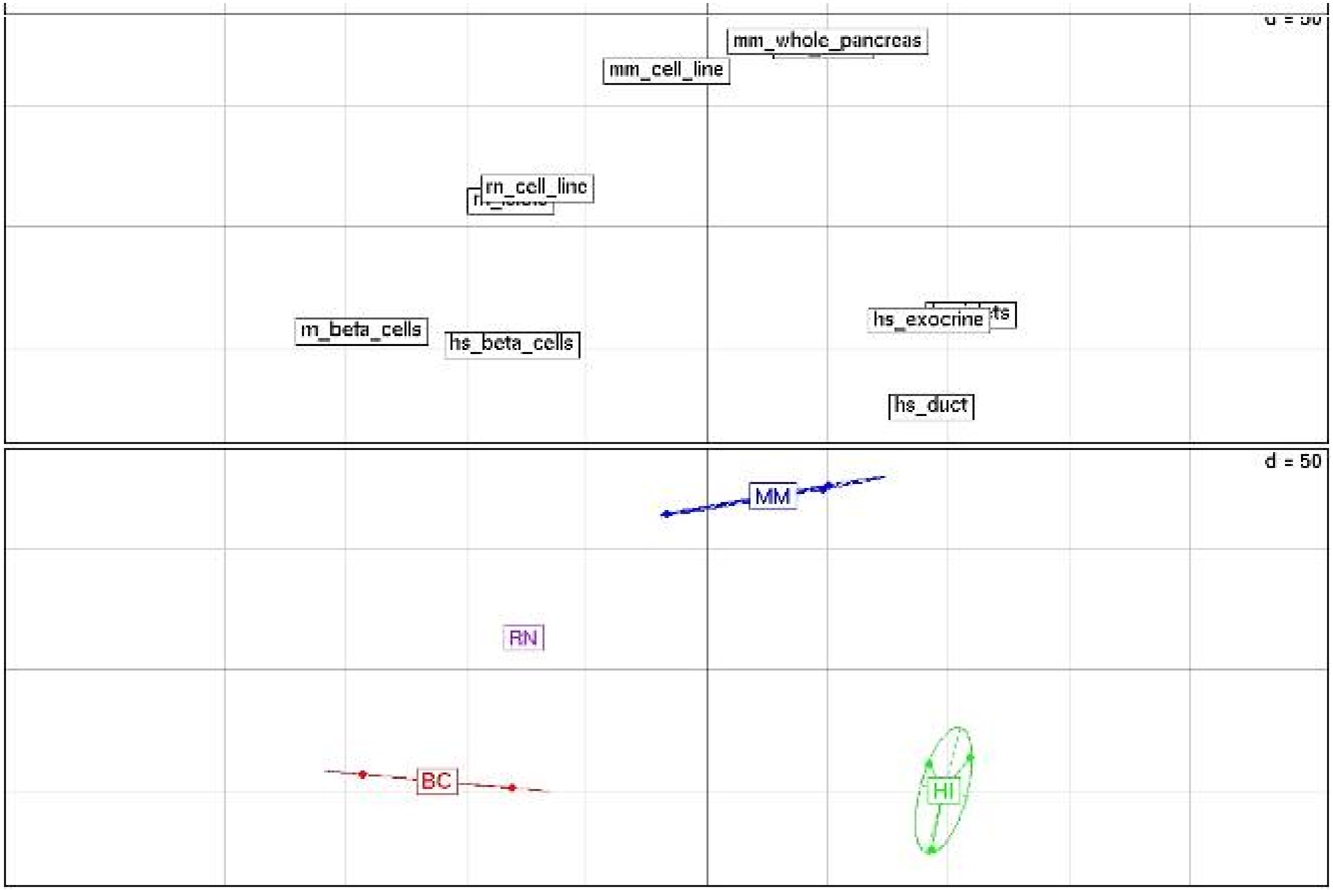
Number of microarrays: **402**
Number of studies: **38**
Number of different platform types: **21**
Species represented = Human, Rat, Mouse
Tissue types = Primary beta cells, pancreatic islets, exocrine pancreas, duct cells, whole pancreas, cell lines

New study: FACS-purified rat beta cells and insulin-producing INS-1 cells using latest Affy chip

Note: Little data from human beta cells
No data from mouse beta cells

Project 1.
β cell transcriptome

Principal Component Analysis of Tissue Profiles



Observations:

1) Mouse tissues distinct
2) Human and rat beta cells are more related to each other than the rest of the tissues. But remember: little human and no mouse data for beta cells

Project 1.
β cell transcriptome

Aim C. Generate atlas

Gene expression in pancreas:

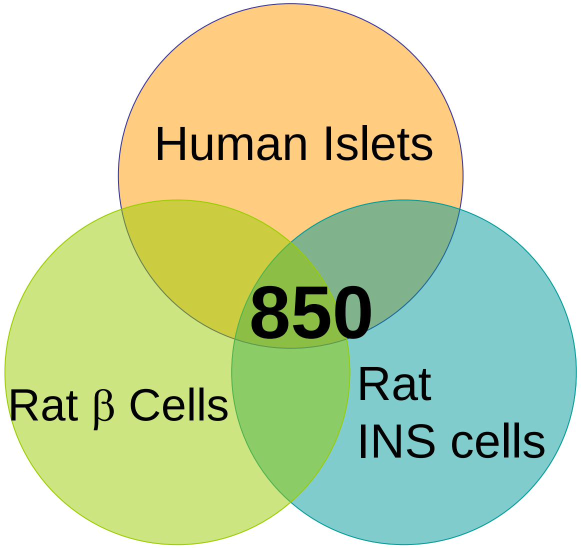
Endocrine	Exocrine	Duct cells	Rat Beta cells	Rat INS cells
9779	11261	4347	7993	7594

Algorithm to obtain the beta cell specific and enriched genes

1. Normalize data to express genes in rank percentiles
2. Islet Enriched Genes – 5-10% difference in ranks
3. Get the intersection of #2 and the primary beta cells and cell lines

Project 1.
β cell transcriptome

Genes enriched in islets vs exocrine and duct cells: 2025



Localization and Secretion Predictions (Secretome 3.0 – TMHMM 2.0)

TransMembrane	114
Secreted	228

Project 1.
β cell transcriptome

Access via T1DBase

http://T1DBase.org/page/AtlasHome

Affymetrix Results - Expression values for each gene were ranked on a scale of 0-100. Cutoffs for expression level are as follows:	
No expression	0-25
Low	25-50
Moderate	50-75
Enriched	75-100

MPSS Results - Signature count for a given gene was averaged over two runs. Cutoffs for expression level are as follows:	
No data (nd)	0
Low	0-20
Moderate	20-100
Enriched	100+

	Human					Mouse		Rat		
	primary beta cells	pancreatic islets	Pancreatic Islets MPSS	exocrine pancreas	ductal cells	pancreatic islets	whole pancreas	primary beta cells	pancreatic islets	beta cell line
INS	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
IPF1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
NOS2A	Y	Y	nd	Y	Y	Y	Y	N	Y	Y
TAP1	N	Y	Y	Y	Y	Y	Y	Y	N	Y
TAP2	Y	Y	Y	Y	Y	Y	Y	Y	N	Y

Conclusions

- Islets have distinct profile compared to other tissues
- Globally mouse tissues are close to each other, human and rat beta cells have similar profiles. Note: we have no data for mouse beta cells
- 850 human beta cell enriched genes contain candidate biomarkers for human beta cells
- Atlas can be used for many purposes including prediction of beta cell function, disease progress etc.

Project 2. Gene regulatory network driving beta cell pathology

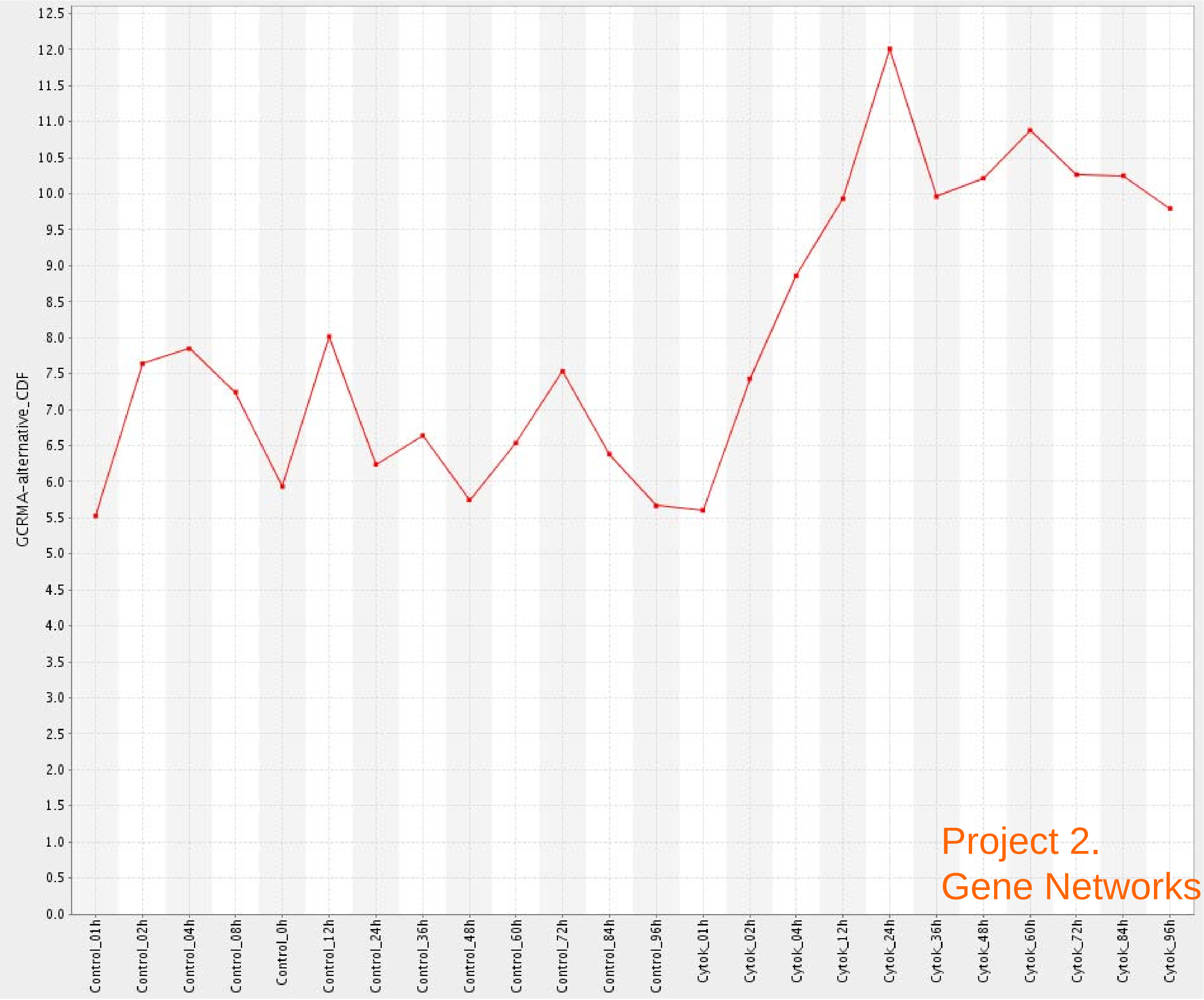
- Time course gene expression study of cytokine-induced beta cell dysfunction and death
- Human islets. *n*=1 today
- Proinflammatory cytokines IL-1 β , IFN-γ known to be synergistically cytotoxic to beta cells
  - Other known killing agents not studied: TNF-α, IL-6, viral infection
- Cytokines activate many genes and transcription factors

Protocol

	1h	2h	4h	8h	12h	24h	36h	48h	72h	84h	96h
Control	X	X	X	X	X	X	X	X	X	X	X
IL1+IFN	X	X	X	X	X	X	X	X	X	X	X

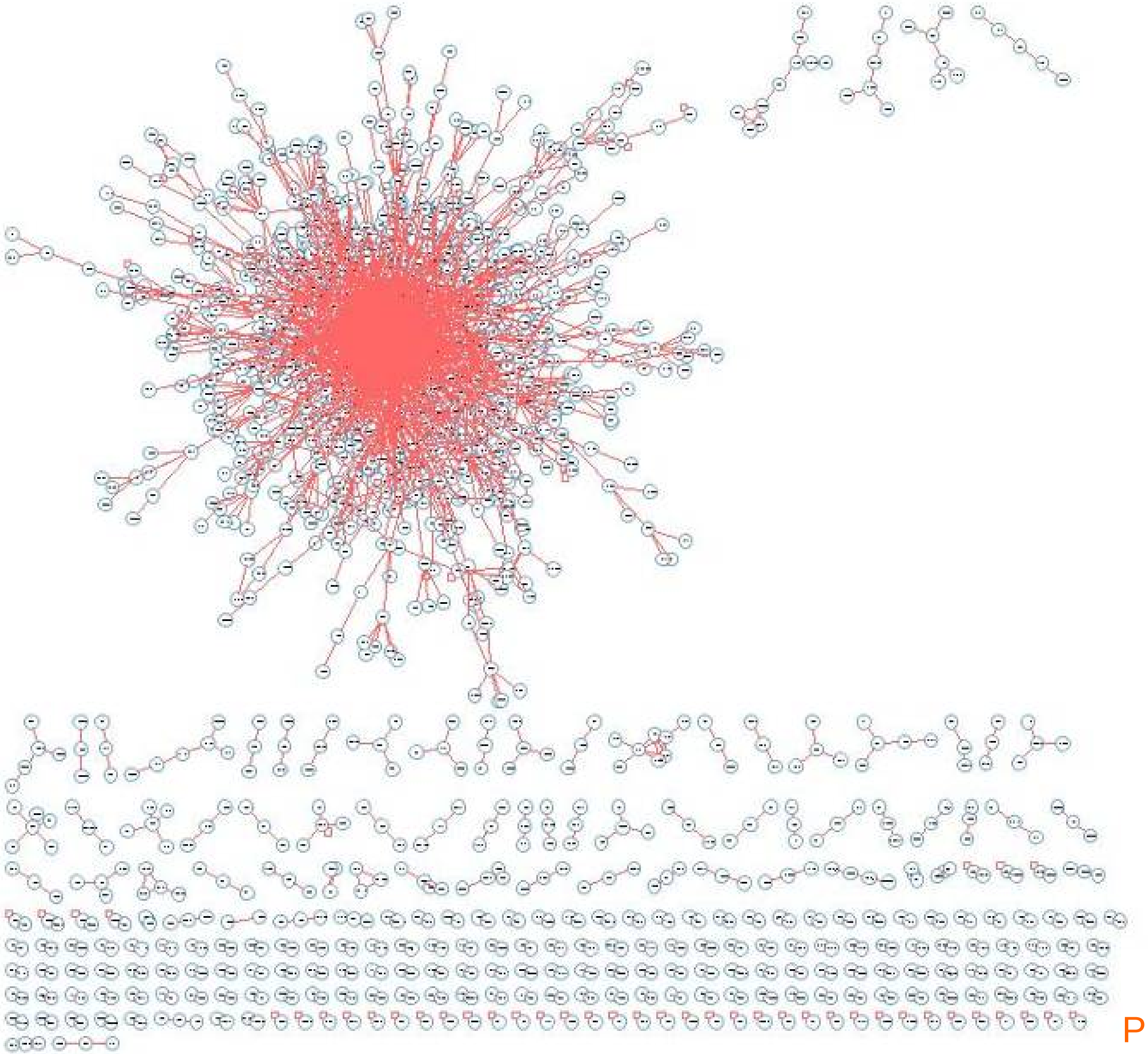
Project 2.
Gene Networks

STAT1 gene expression in response to IL-1β+IFN-γ



Project 2.
Gene Networks

Protein-Protein Interaction Networks



Project 2.
Gene Networks

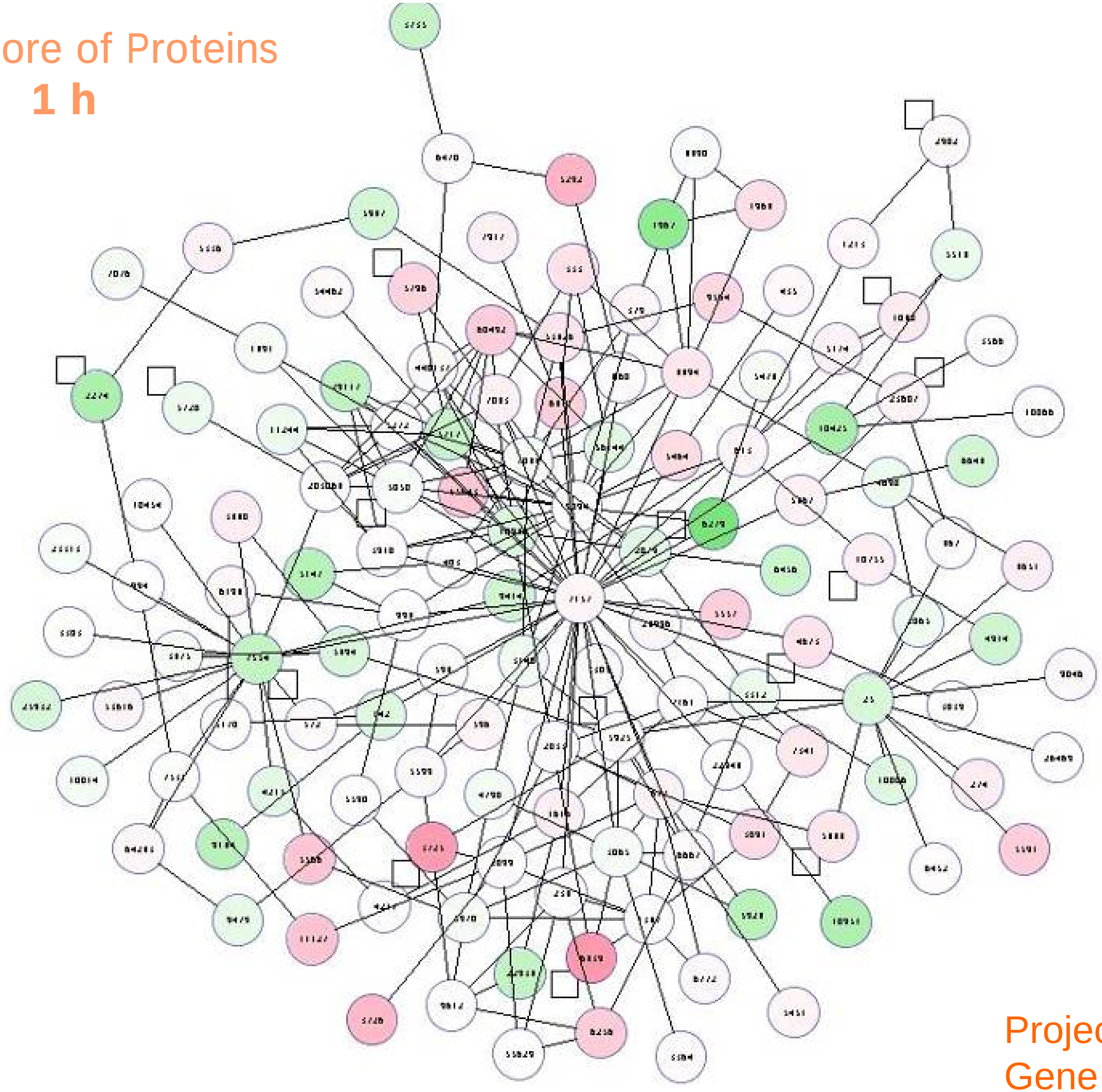
Active modules in beta cell dysfunction and death

Conditions vs. Pathways

Network	Size	Score	il24	cytok6	cytoknf	cytok24	ifn24	ifnpic24	pic24	ifn6	pic6	ifnpic6
1	63	10.058		🟥		🟥		🟥		🟥		🟥
2	29	9.736	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
3	46	9.387	🟥	🟥		🟥	🟥	🟥		🟥		🟥
4	83	9.28	🟥	🟥		🟥	🟥	🟥		🟥	🟥	🟥
5	45	9.273		🟥		🟥		🟥		🟥		🟥
6	36	9.2	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
7	32	9.094	🟥	🟥		🟥		🟥		🟥		🟥
8	37	9.076	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
9	38	9.018	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
10	47	8.946	🟥	🟥		🟥		🟥		🟥		🟥
11	32	8.857		🟥		🟥		🟥		🟥		🟥
12	54	8.83		🟥		🟥		🟥		🟥		
13	49	8.773	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
14	40	8.708	🟥	🟥		🟥		🟥		🟥		🟥
15	43	8.672	🟥	🟥		🟥		🟥		🟥		🟥
16	30	8.648		🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
17	34	8.636	🟥	🟥		🟥		🟥		🟥		🟥
18	76	8.589	🟥	🟥		🟥	🟥	🟥		🟥		🟥
19	41	8.589		🟥		🟥	🟥	🟥		🟥	🟥	🟥
20	32	8.519	🟥	🟥		🟥	🟥	🟥	🟥	🟥		🟥
21	36	8.509		🟥		🟥		🟥		🟥		🟥
22	50	8.506		🟥		🟥	🟥	🟥	🟥	🟥		🟥
23	56	8.48	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
24	28	8.476	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
25	40	8.336	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
26	62	8.332	🟥	🟥		🟥	🟥	🟥		🟥		🟥
27	47	8.313		🟥		🟥		🟥		🟥		🟥
28	37	8.31	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
29	50	8.307		🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
30	35	8.26	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
31	72	8.239		🟥		🟥		🟥		🟥		🟥
32	53	8.229		🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
33	36	8.221		🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
34	51	8.204	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
35	33	8.183	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
36	40	8.181	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
37	25	8.175	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
38	22	8.171		🟥		🟥	🟥	🟥	🟥			🟥
39	36	8.124	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥		🟥
40	55	8.042	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
41	26	8.038		🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
42	59	8.005		🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
43	50	7.991	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
44	48	7.987		🟥		🟥	🟥	🟥	🟥	🟥		🟥
45	55	7.978	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥
46	28	7.954	🟥	🟥		🟥		🟥		🟥		🟥
47	30	7.939		🟥		🟥	🟥	🟥	🟥	🟥		🟥
48	31	7.929	🟥	🟥	🟥	🟥		🟥		🟥		🟥

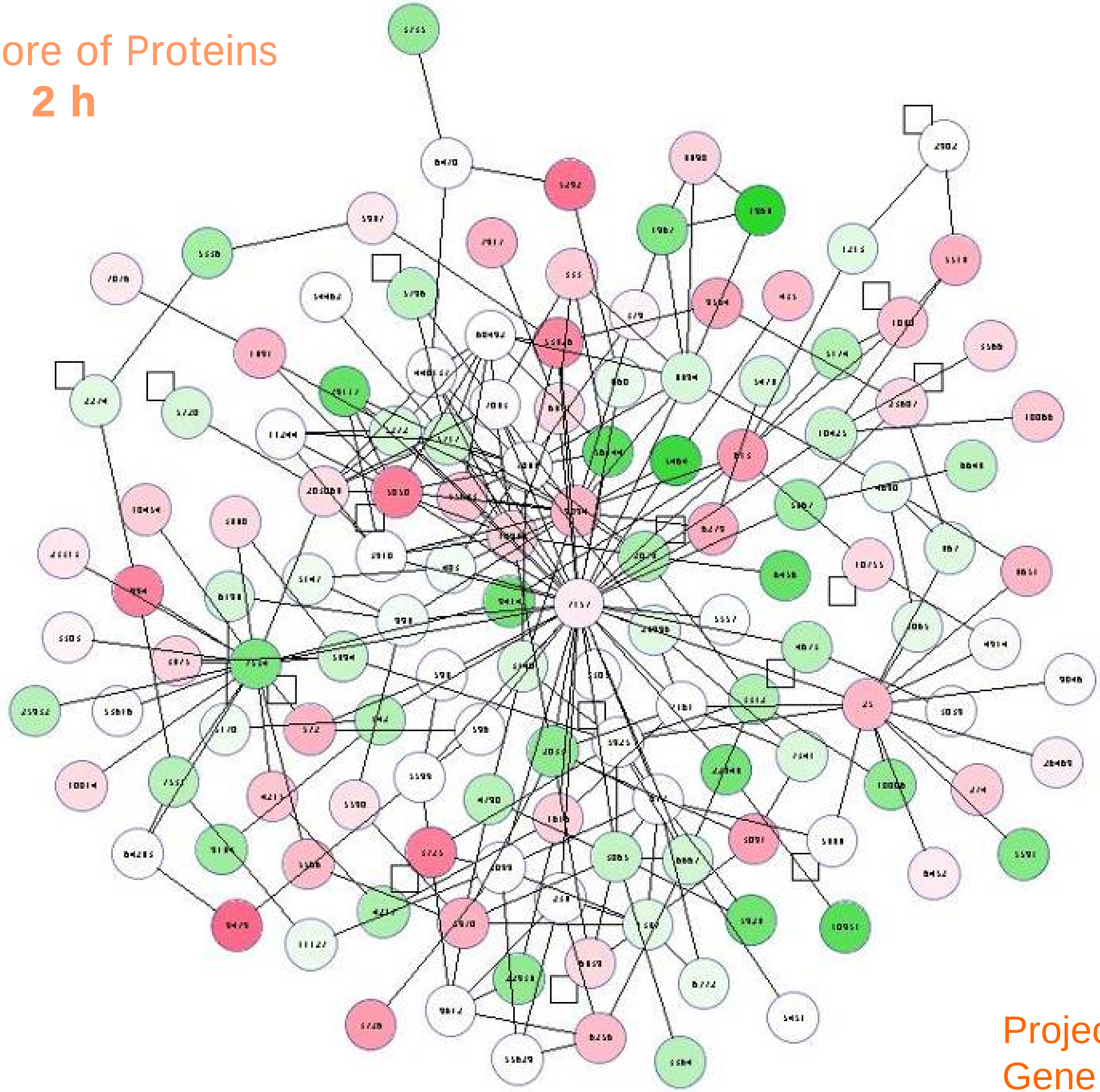
Project 2.
Gene Networks

Active Core of Proteins
1 h



Project 2.
Gene Networks

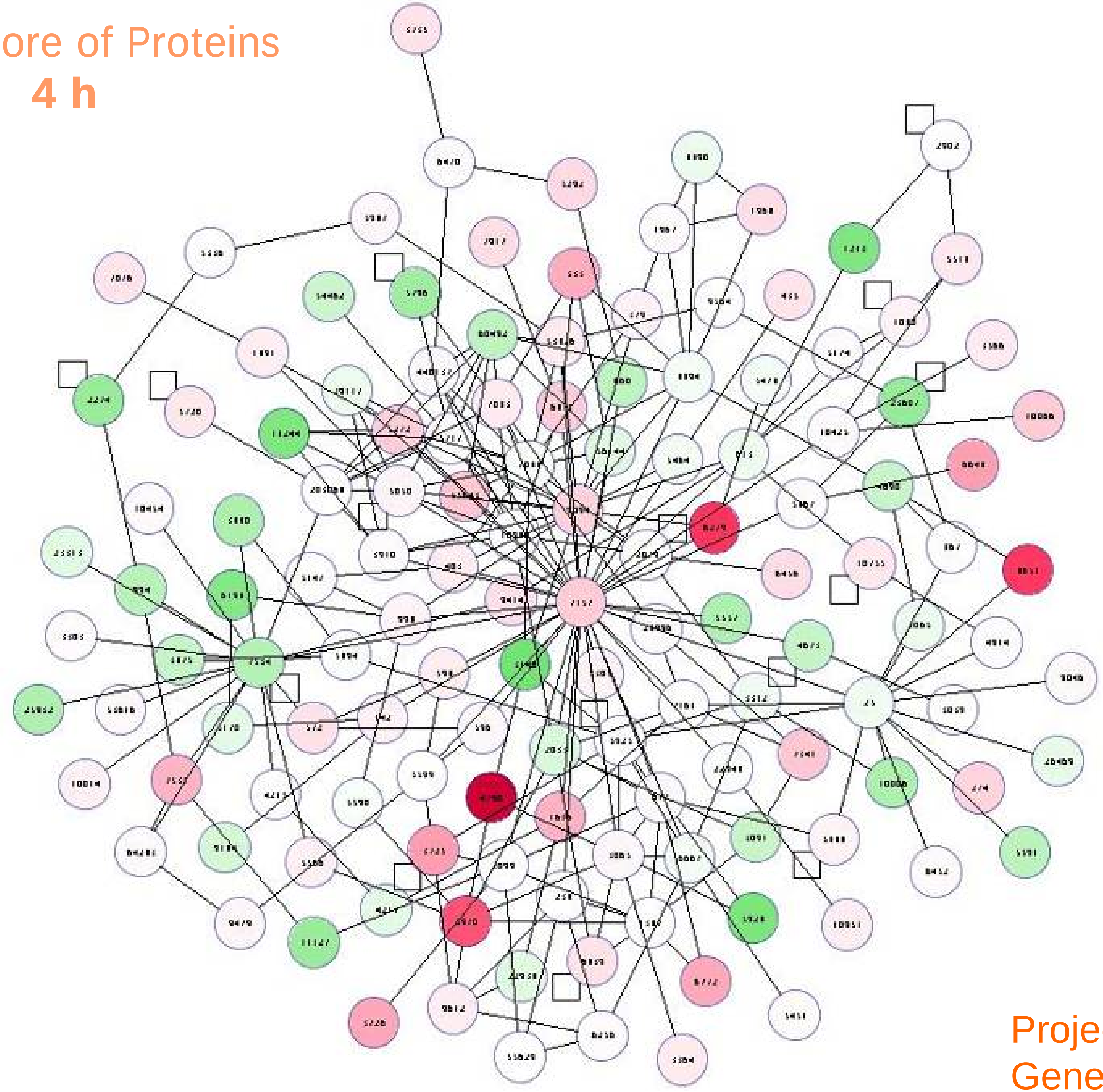
Active Core of Proteins
2 h



Project 2.
Gene Networks

Active Core of Proteins

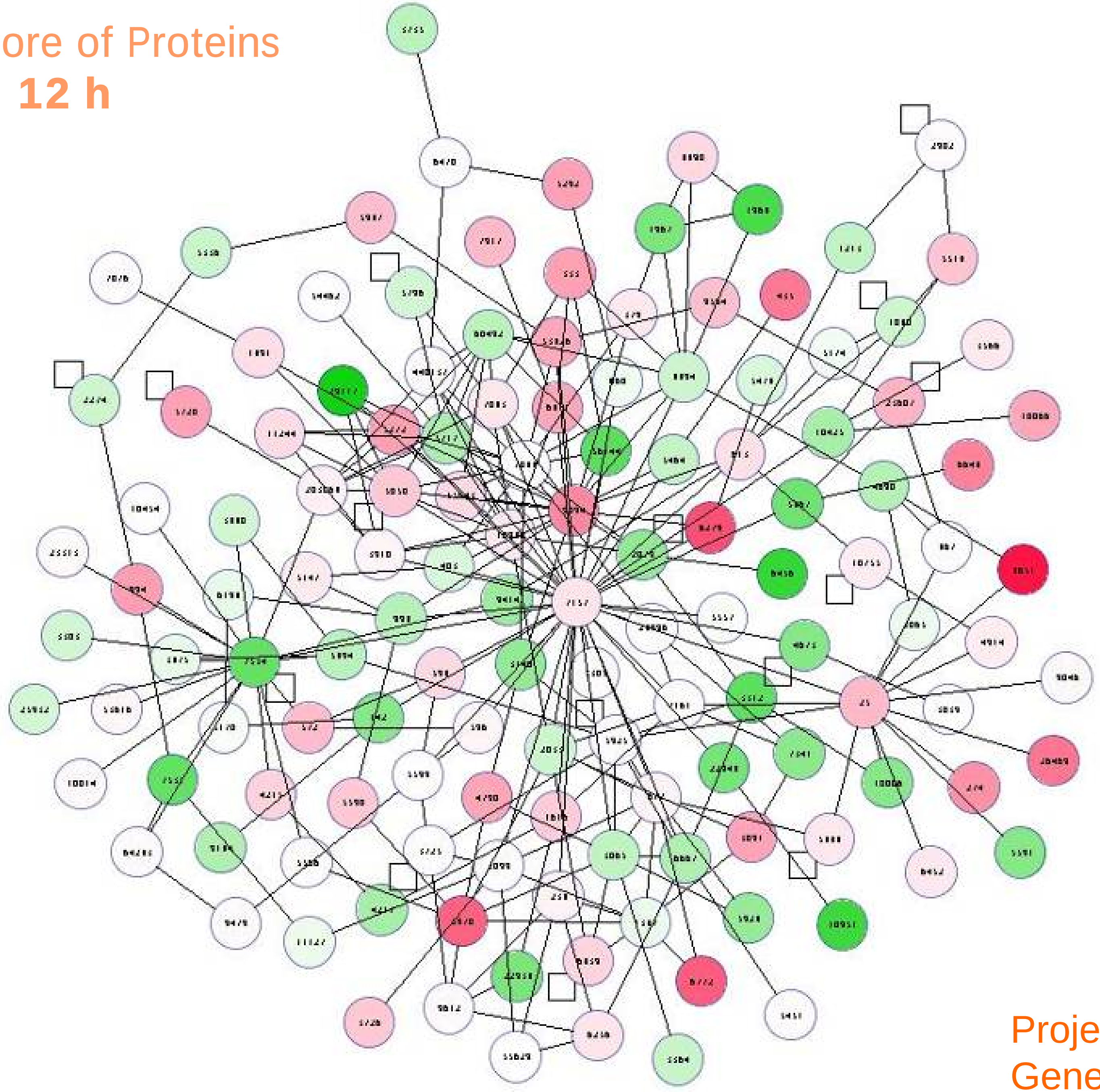
4 h



Project 2.
Gene Networks

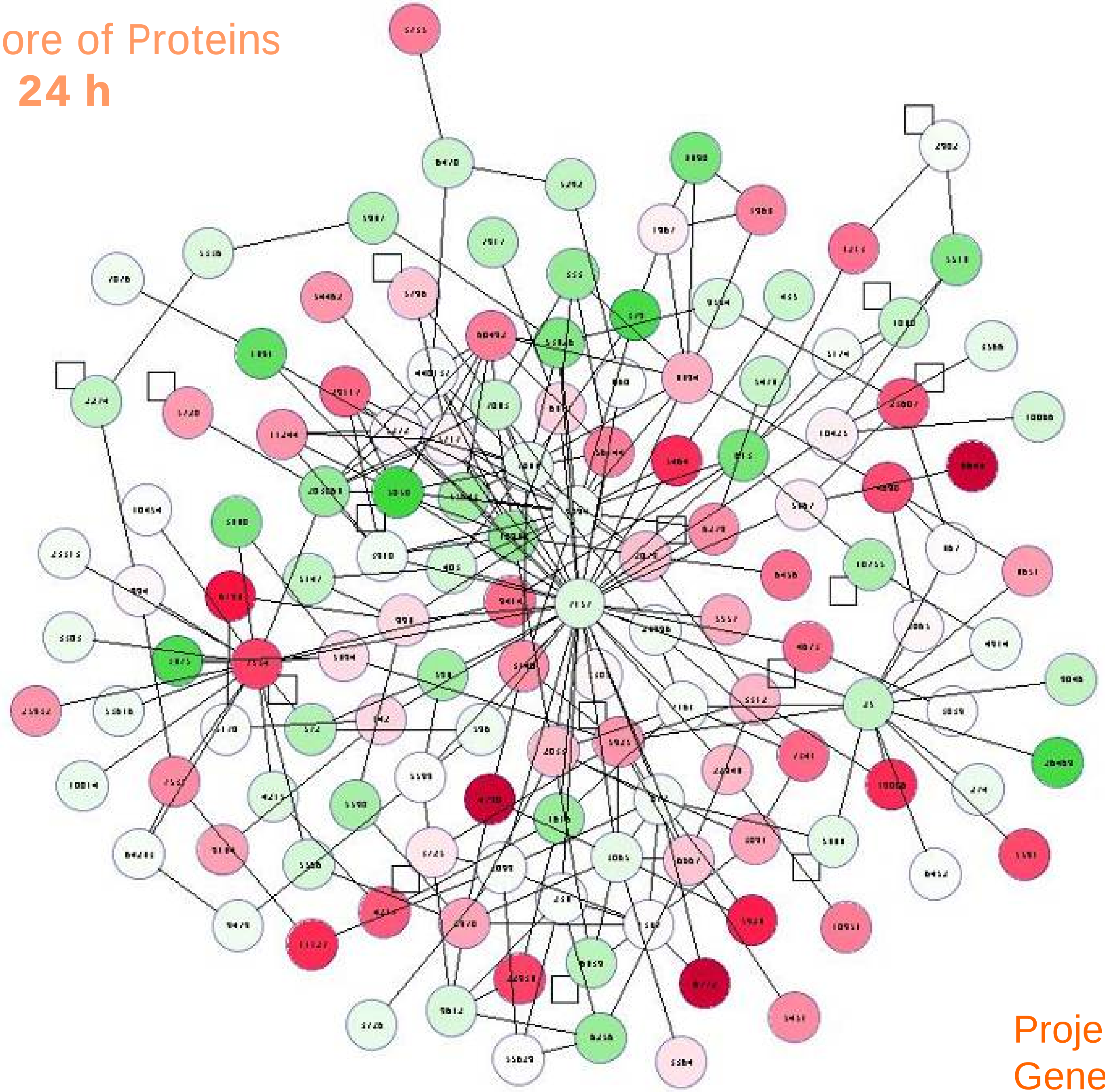
Active Core of Proteins

12 h



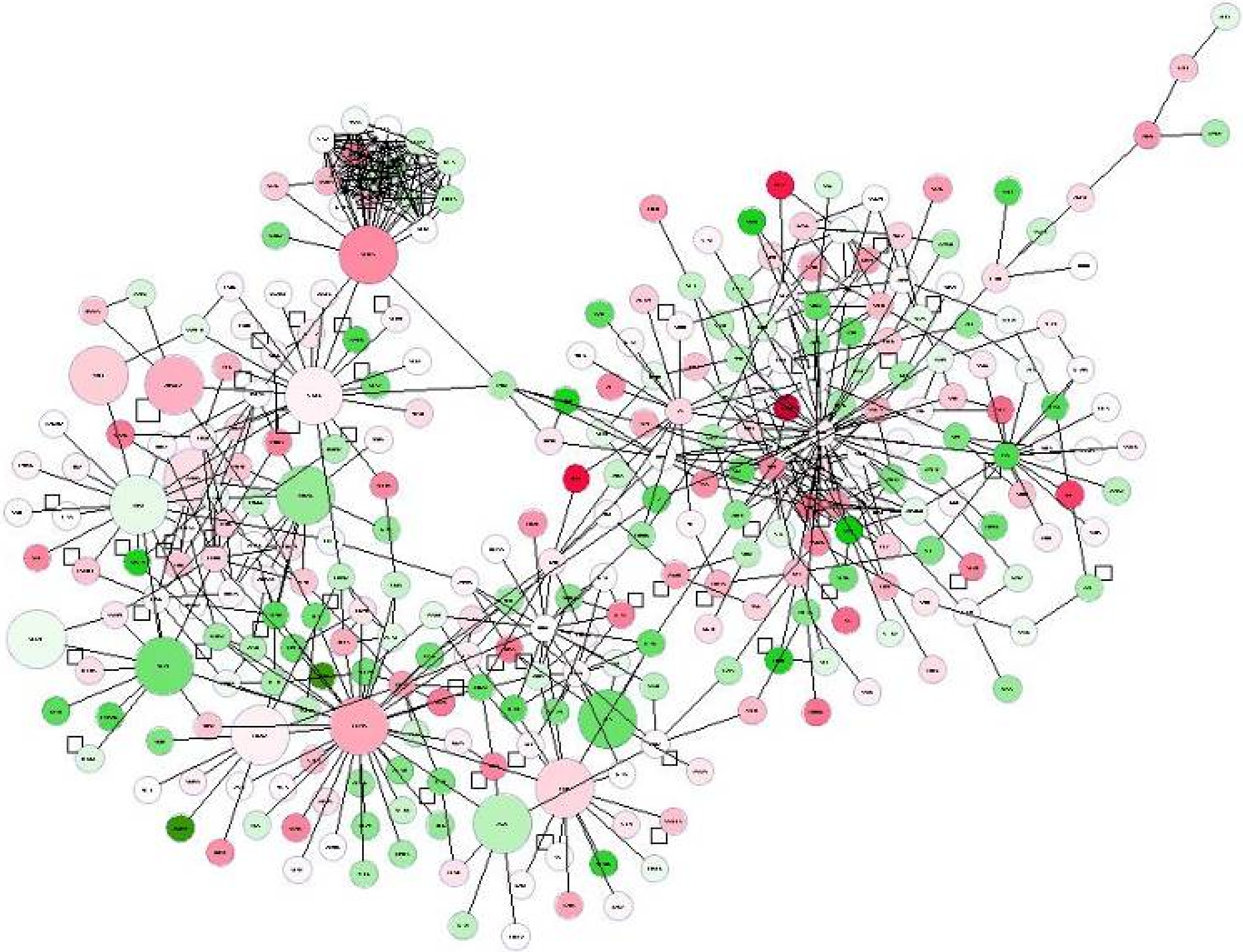
Project 2.
Gene Networks

Active Core of Proteins
24 h



Project 2.
Gene Networks

Active Core of Proteins **PLUS** beta cell-specific genes



Project 2.
Gene Networks

Conclusions

- Core network activated in response to all killing agents
- Time series shows progressive activation
- 24h is critical time point for these islets
- Beta cell enriched network connected to core contain genes related to insulin secretion among others
- Role of beta cell enriched network in pathology yet to be determined: cause or effect?

Acknowledgements

- ISB
  - Lee Hood
  - **Burak Kutlu**
  - Bruz Marlzhof
  - Pamela Troisch
  - Paul Shannon
  - T1DBase team
    - Erin Helton
    - David Burdick
    - Victor Cassen
    - Chris Cavnor
    - Karen Friery
    - Sarah Killcoyne
- Universite Libre de Bruxelles
  - Decio Eizirik
  - Pierre Pirot
  - Daisy Flamez
- University of Uppsala
  - Nils Welsh
  - Olle Korsgren

Funding from JDRF Juvenile Diabetes Research Foundation International

dedicated to finding a cure

We are : Incredibly Speedy Biologists !!

Beat the Bridge 2006 Seattle