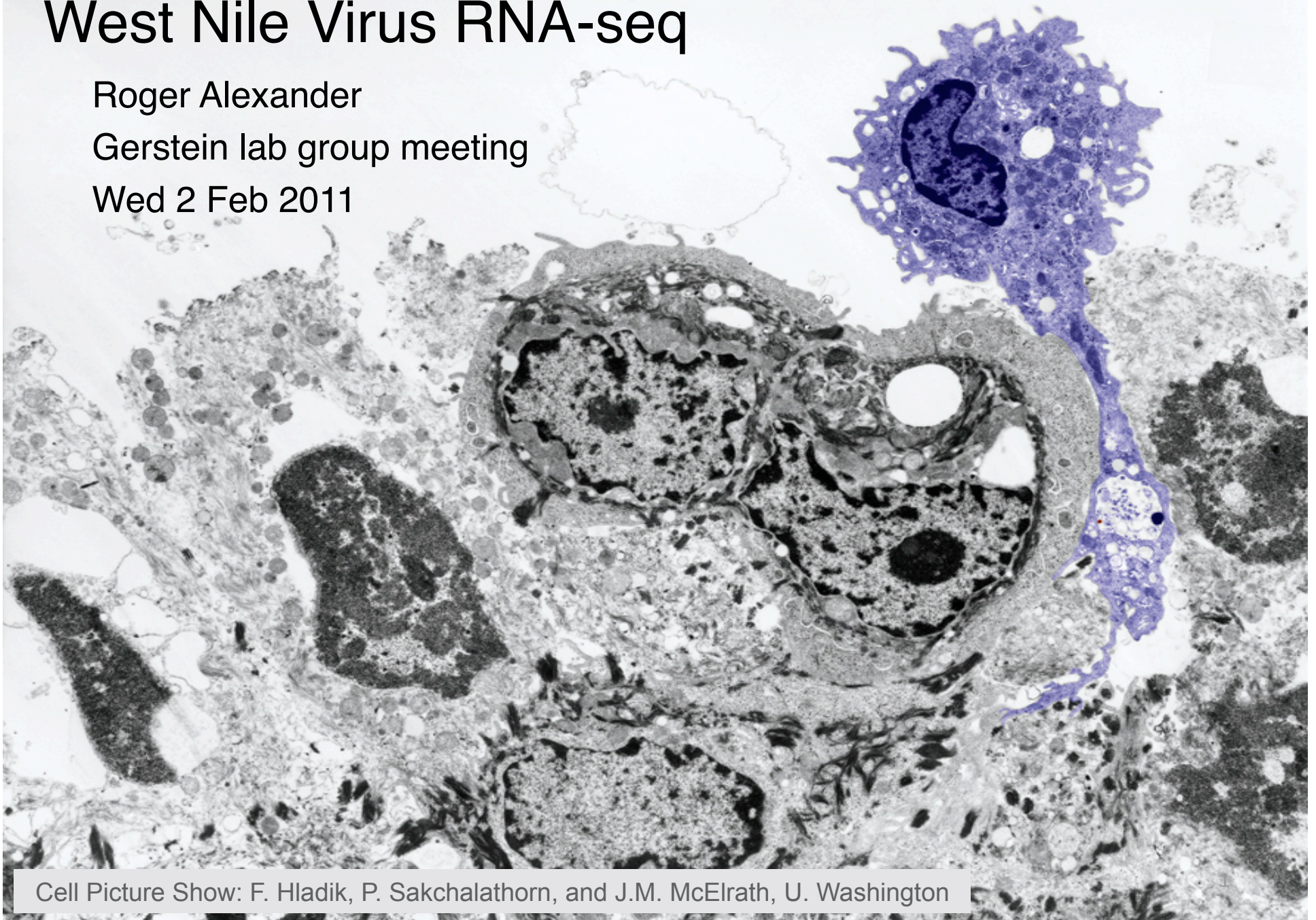


West Nile Virus RNA-seq

Roger Alexander

Gerstein lab group meeting

Wed 2 Feb 2011



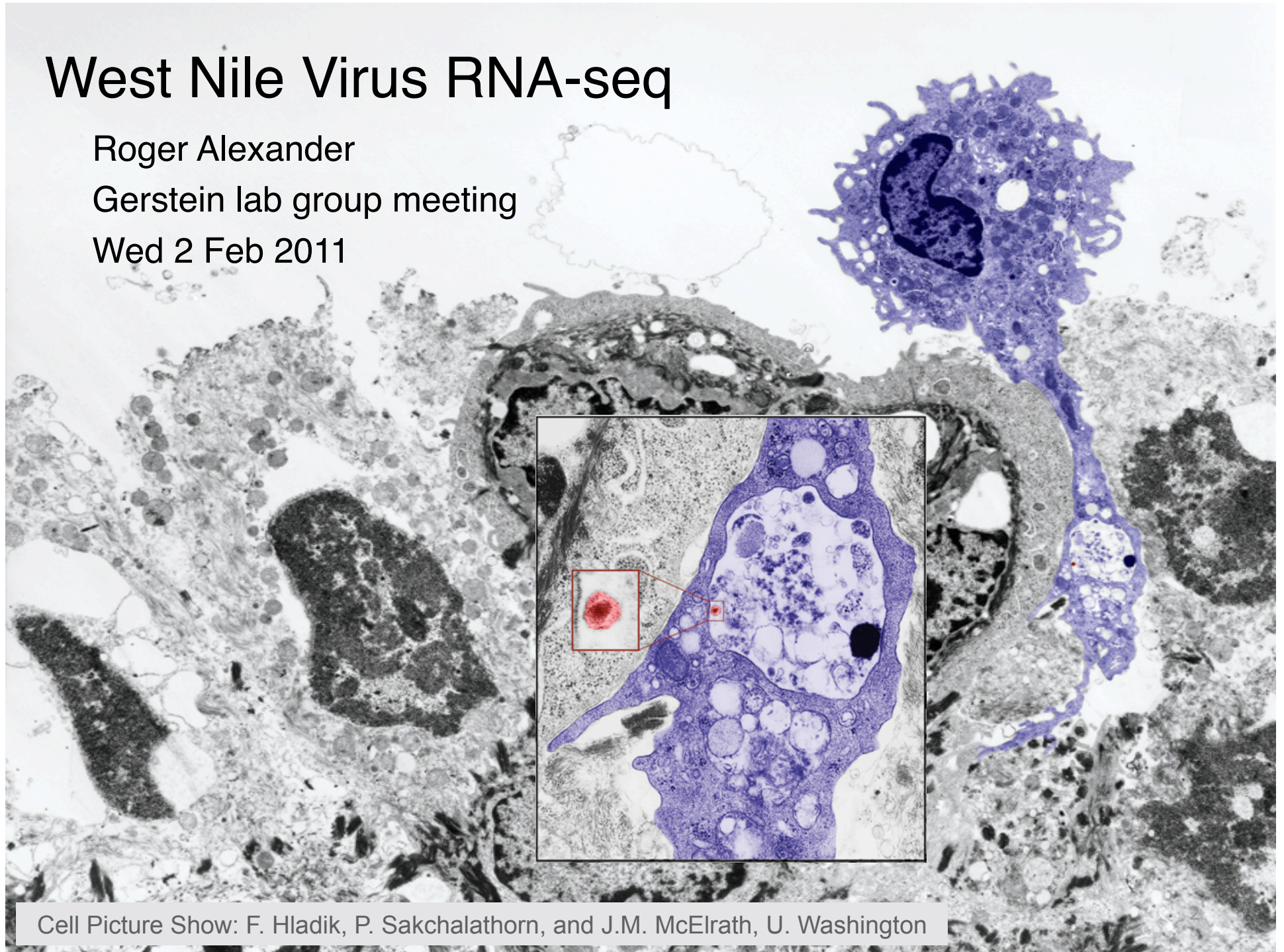
Cell Picture Show: F. Hladik, P. Sakchalathorn, and J.M. McElrath, U. Washington

West Nile Virus RNA-seq

Roger Alexander

Gerstein lab group meeting

Wed 2 Feb 2011



Cell Picture Show: F. Hladik, P. Sakchalathorn, and J.M. McElrath, U. Washington

West Nile Virus RNA-seq



Talk Goal: foster discussion of WNV biology and differential expression (DE) analysis

Talk Outline

- Acknowledgements
- Project Goals
- Biology of West Nile Virus infection life cycle
- Role of host factors identified by RNAi in this dataset
- Mapping RNA-seq reads onto WNV genome
 - good quantitation of viral load?
- Wei Zheng's prior analysis
 - differential expression (DE) at gene level => low power
 - discussion of pathway analysis methods

WNV RNA-seq Acknowledgements

Gerstein Lab

for RSEQtools

Lukas Habegger

Andrea Sboner

for discussion

Joel Rozowsky

Chao Cheng

Montgomery Lab

Ruth Montgomery

Feng Qian (experiments)

Zhao lab

Wei Zheng (first round of analysis)

Modis Lab

Moshe Dessau

West Nile Virus RNA-seq



Goals of this project:

- get funded! (unsuccessful since '07)
- write an interesting paper
- understand the biology of WNV infection of humans
 - cellular vs organismal level
 - organismal:
 - interplay between cell types
 - blood-brain barrier
 - cellular: HeLa vs macrophages: same biology?
- understand the variability of infection response
 - Who in the population is most susceptible?
- implications for understanding Dengue virus

West Nile Virus life cycle

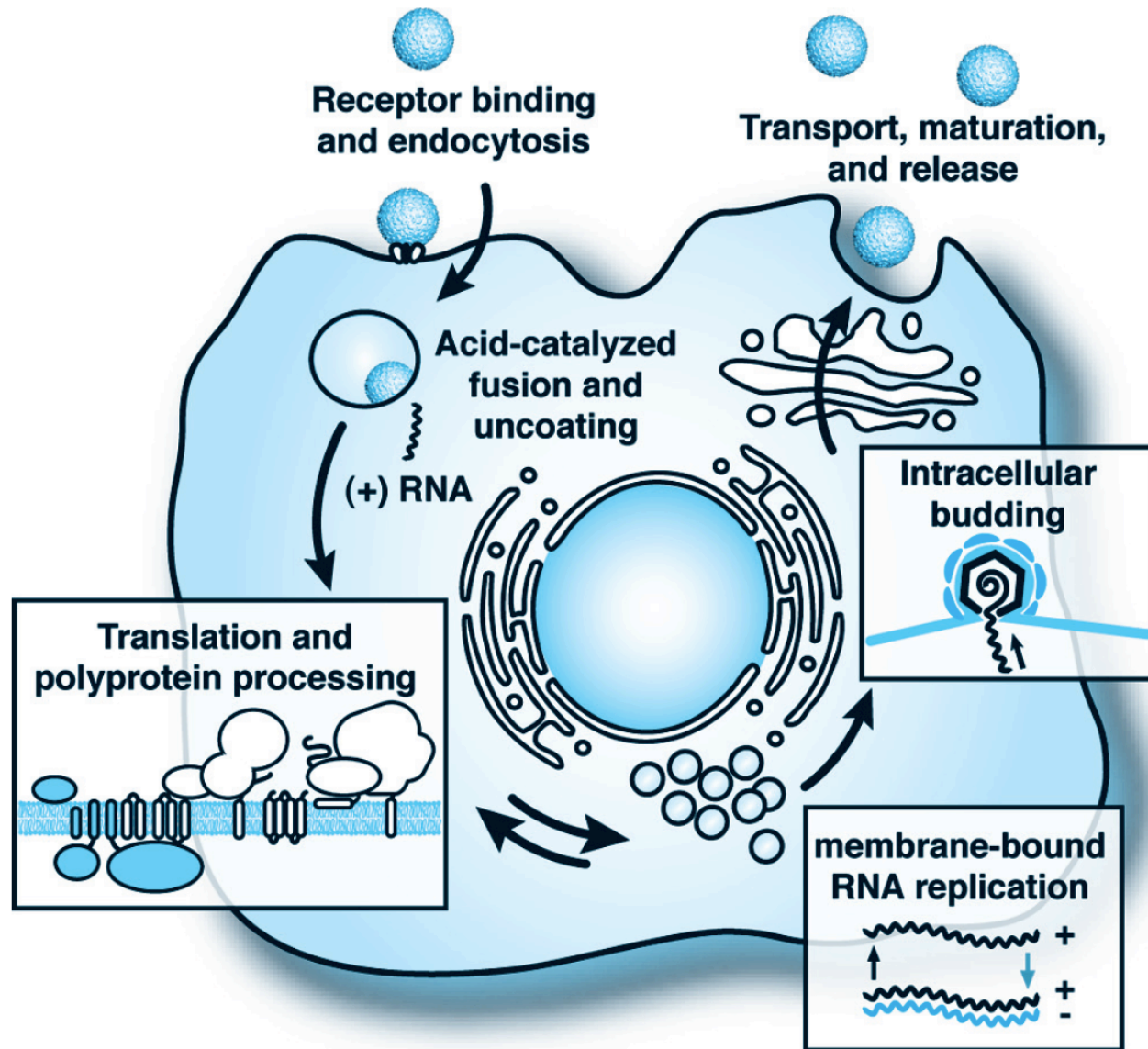


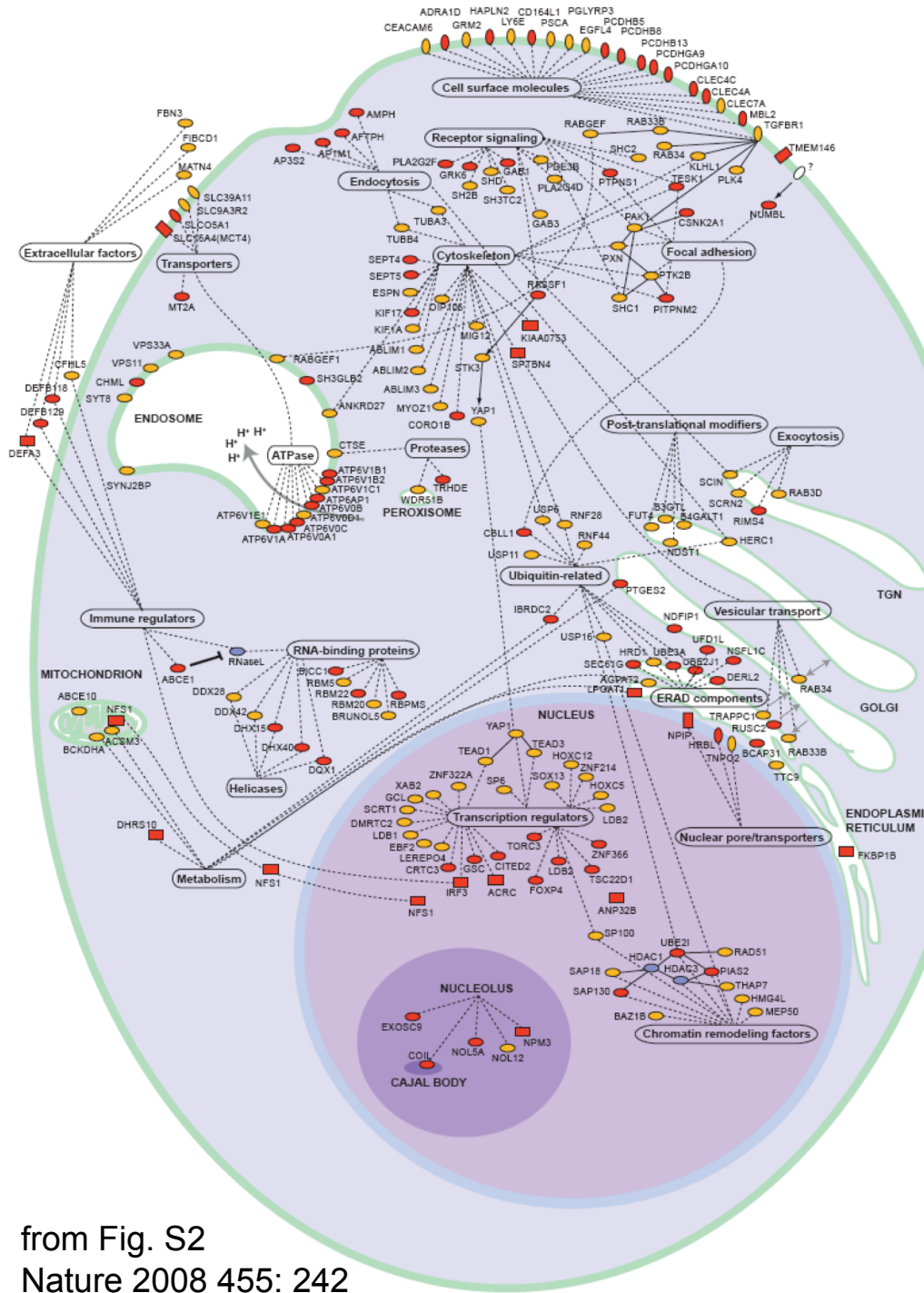
Figure 33.2. Fields Virology 5th ed. (2007)

West Nile Virus life cycle

at gene level according to Fikrig HeLa RNAi paper

Host Susceptibility Factors (HSF)

Host Resistance Factors (HRF)



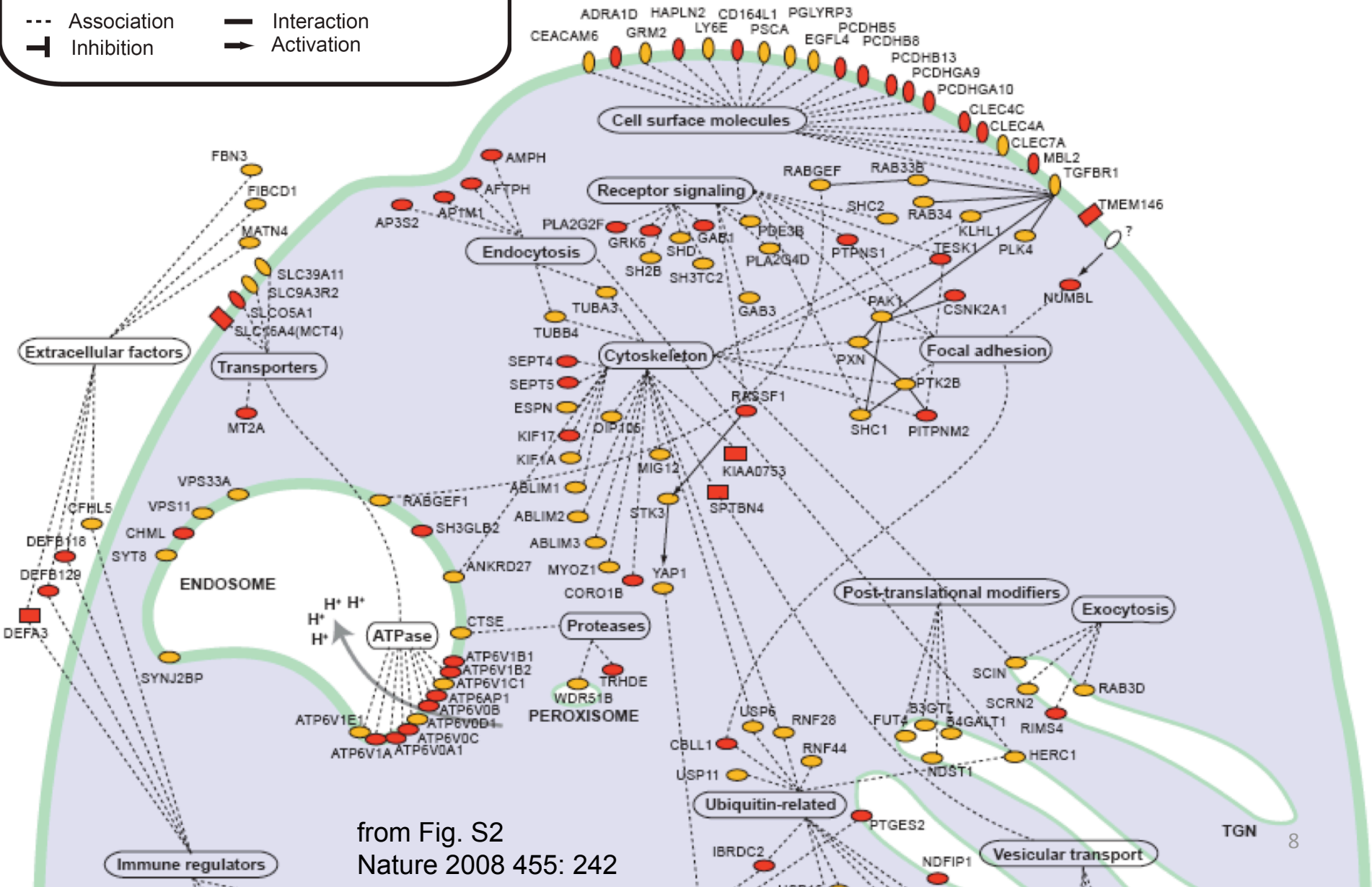
- HRF: Silencing enhances WNV / DENV infection
- HSF: Silencing reduces WNV/DENV infection
- HSF: Silencing reduces WNV, not DENV infection
- Host protein of interest
- Association
- Interaction
- | Inhibition
- Activation

from Fig. S2
Nature 2008 455: 242

West Nile Virus life cycle

■ HRF: Silencing enhances WNV/ DENV infection
● HSF: Silencing reduces WNV/DENV infection
● HSF: Silencing reduces WNV, not DENV infection
● Host protein of interest

- - - Association — Interaction
 ⊣ Inhibition → Activation

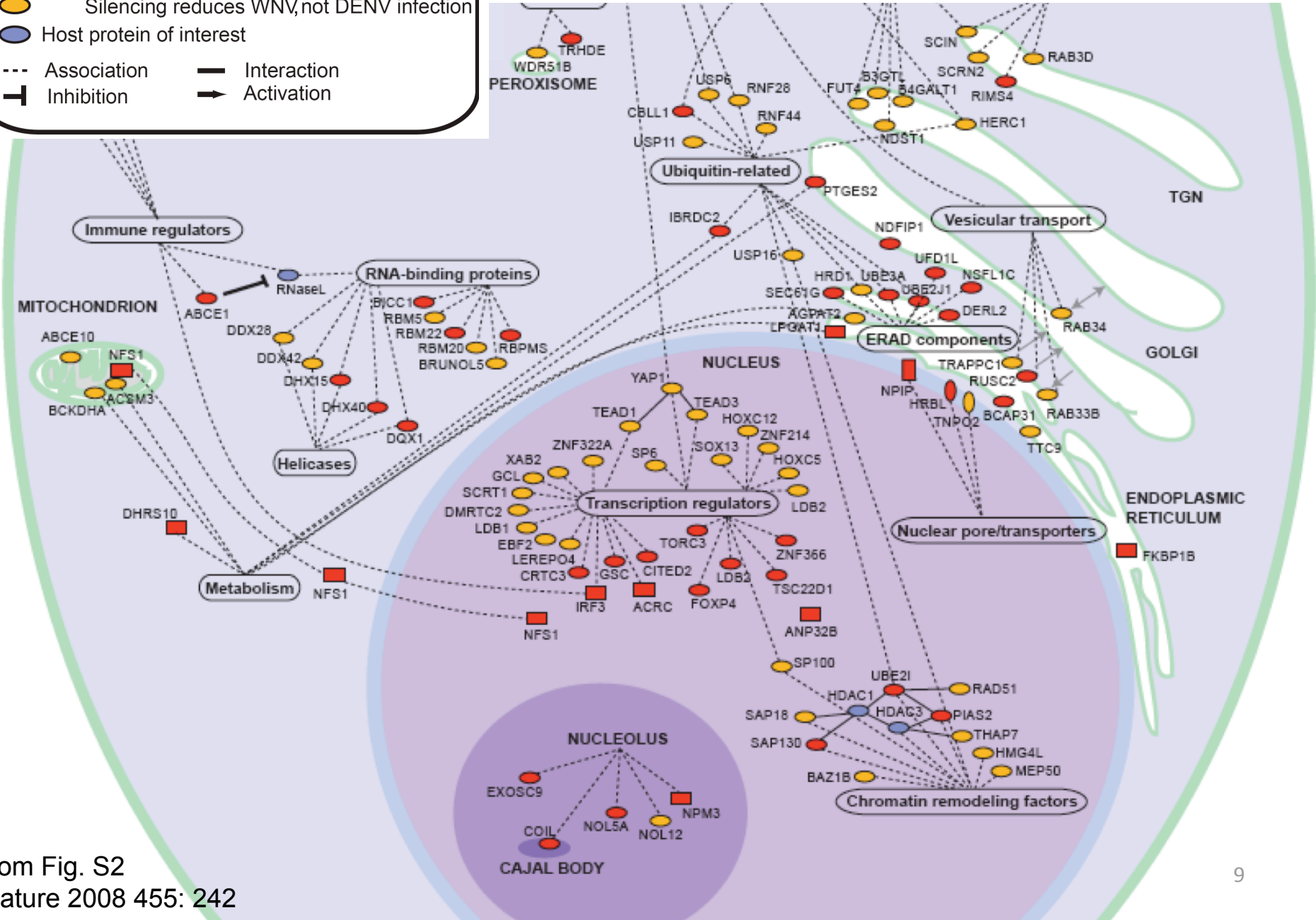


from Fig. S2
Nature 2008 455: 242

West Nile Virus life cycle

■ HRF: Silencing enhances WNV/ DENV infection
● HSF: Silencing reduces WNV/DENV infection
● HSF: Silencing reduces WNV, not DENV infection
● Host protein of interest

--- Association — Interaction
 ⊣ Inhibition ➔ Activation



from Fig. S2
Nature 2008 455: 242

WNV RNA-seq

- prior work: WNV RNAi in HeLa cells (Nature 2008 455: 242)
 - identified ~250 Host Susceptibility Factors (HSF)
 - ~50 Host Resistance Factors (HRF)
 - KO HSF => decreased viral load
 - KO HRF => increased viral load
- this dataset: polyA-selected RNAseq: single-end 36mers
 - primary macrophage cells from blood extracted from 10 healthy volunteers
 - control
 - 24 hrs after WNV infection
 - for validation, qPCR of host IL-8 and virus envelope genes

West Nile Virus RNA-seq

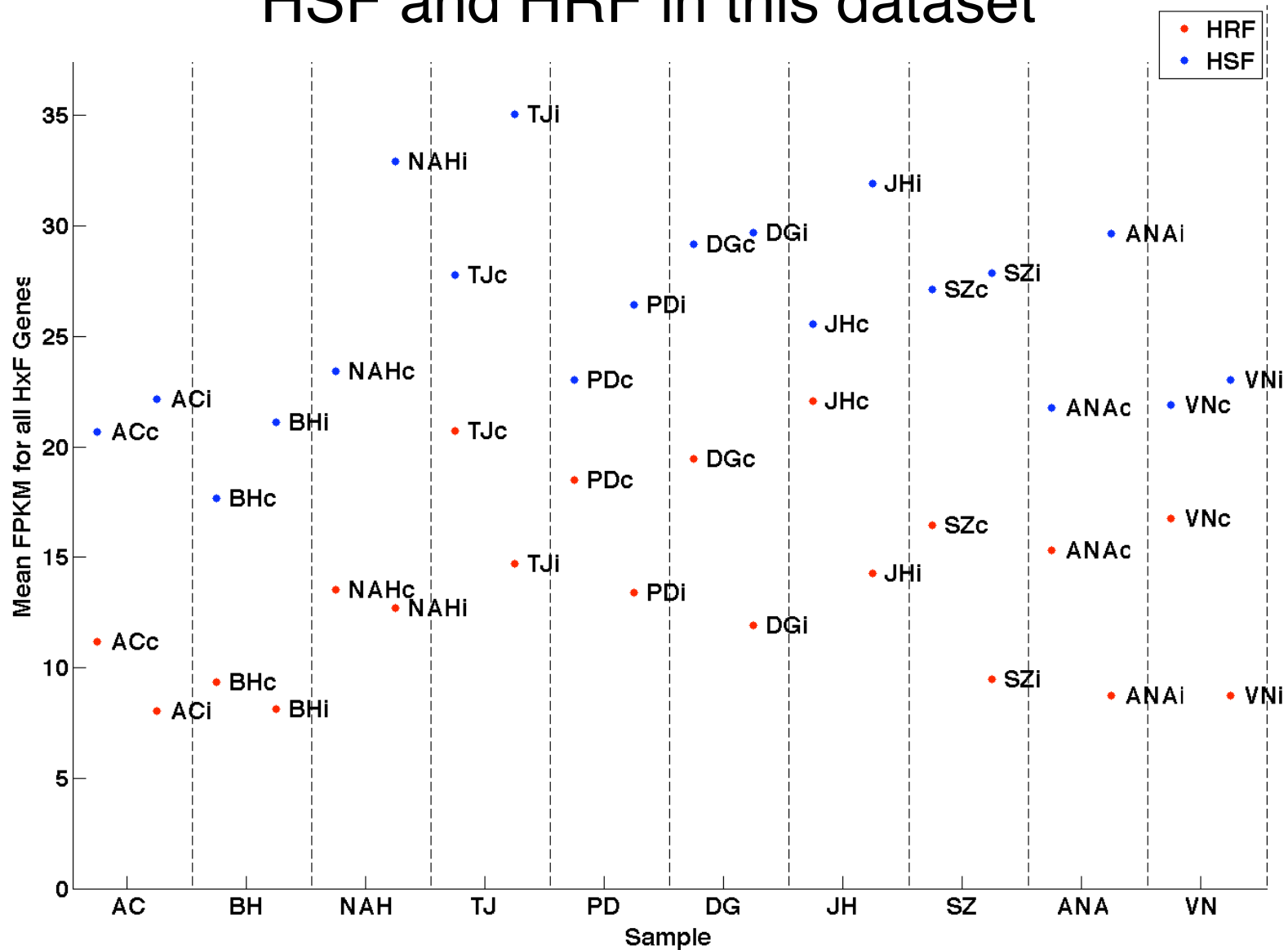


Talk Goal: foster discussion of WNV biology and differential expression (DE) analysis

Talk Outline

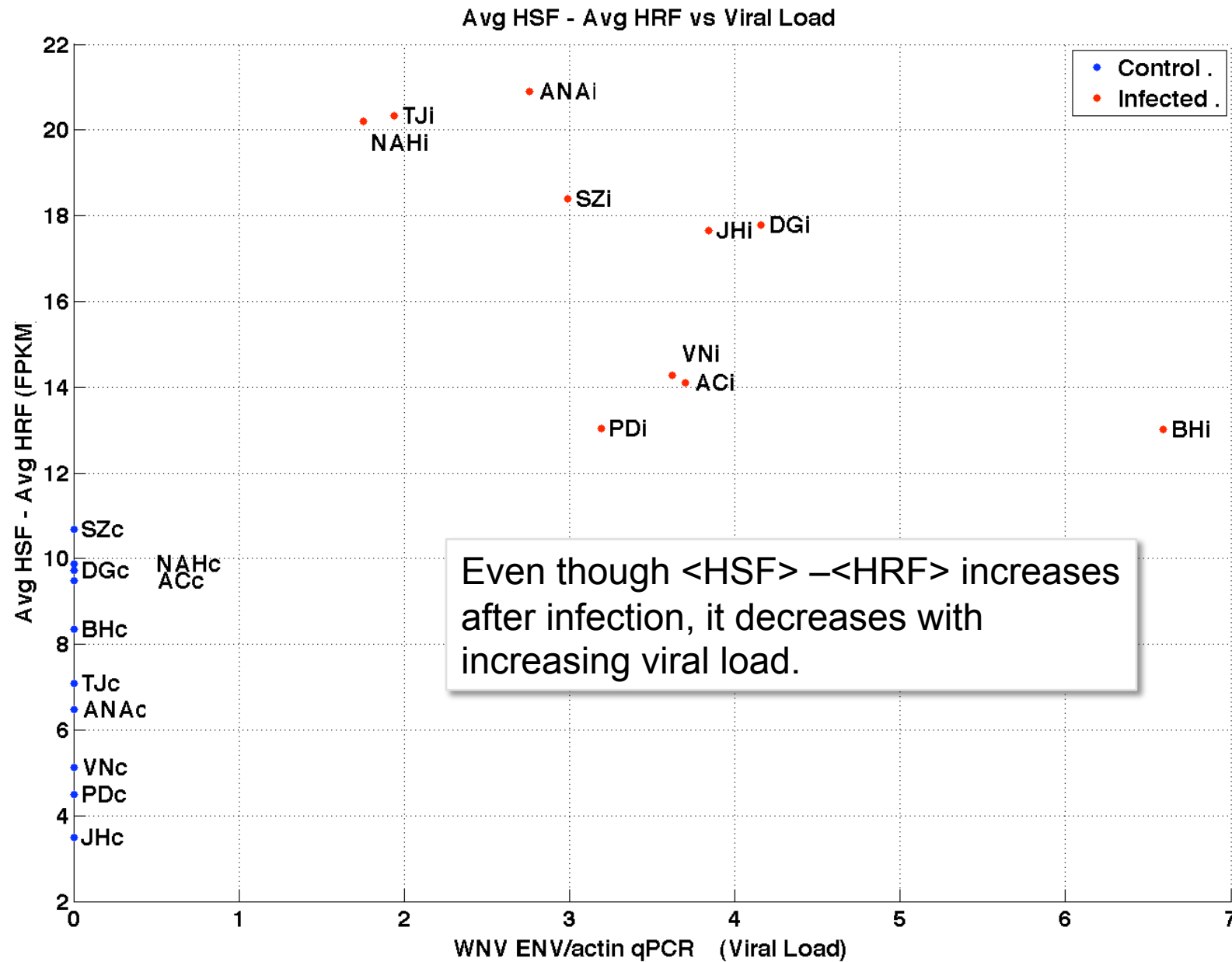
- Acknowledgements
- Project Goals
- Biology of West Nile Virus infection life cycle
- X Role of host factors identified by RNAi in this dataset
- Mapping RNA-seq reads onto WNV genome
 - good quantitation of viral load?
- Wei Zheng's prior analysis
 - differential expression (DE) at gene level => low power
 - discussion of pathway analysis methods

HSF and HRF in this dataset



In all samples, HSF increase and HRF decrease after infection.

HSF and HRF in this dataset



Flavivirus genome structure: no polyA

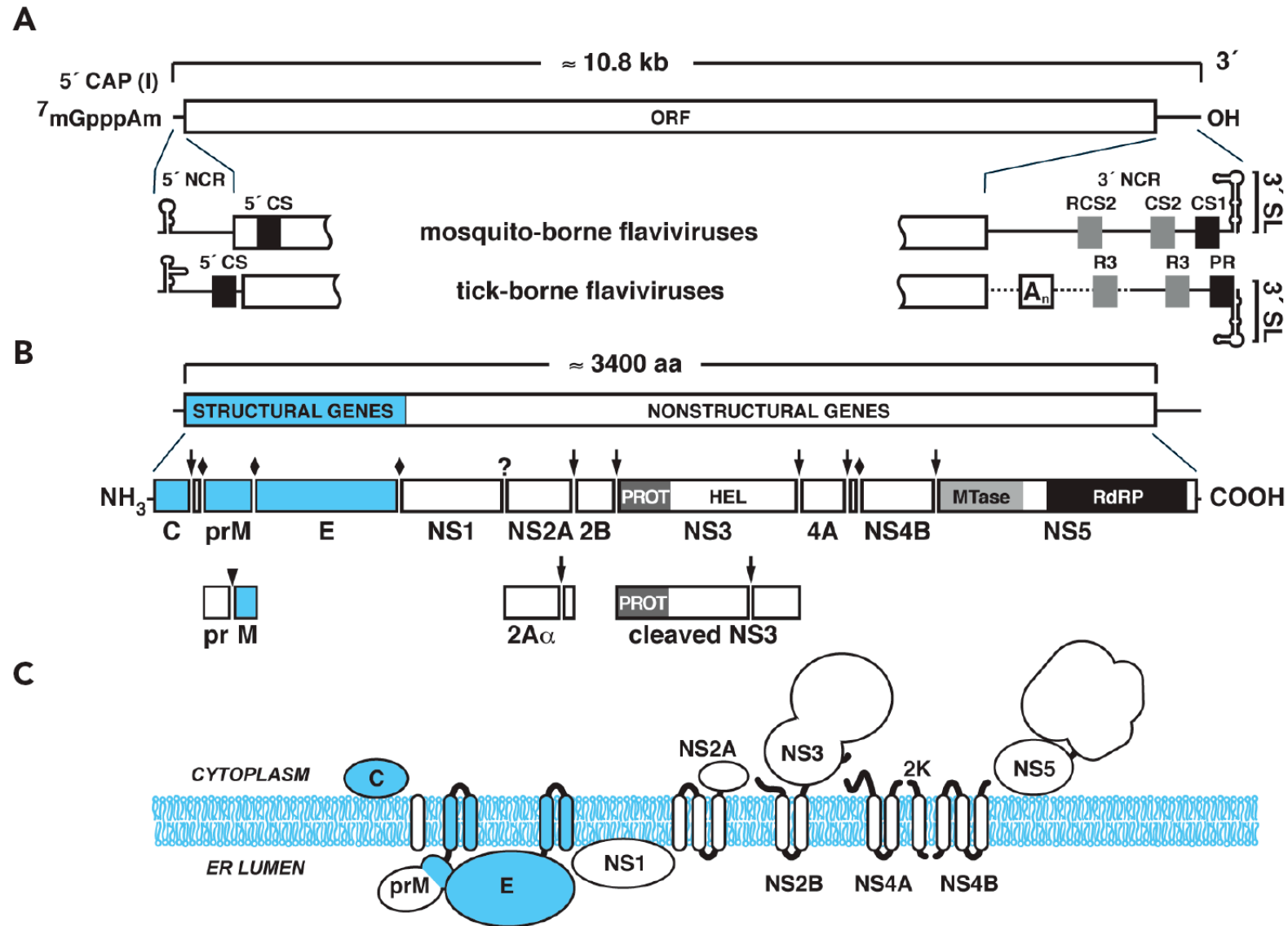
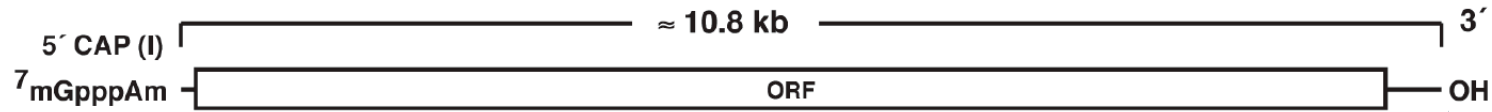


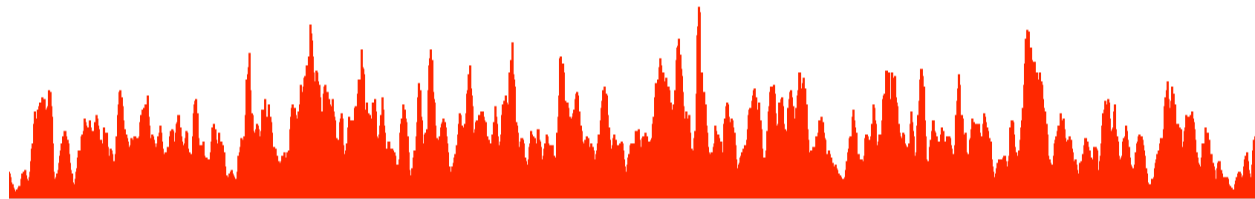
Figure 33.5. Fields Virology 5th ed. (2007)

RNAseq reads mapped to WNV genome (Sample with max RPKM)

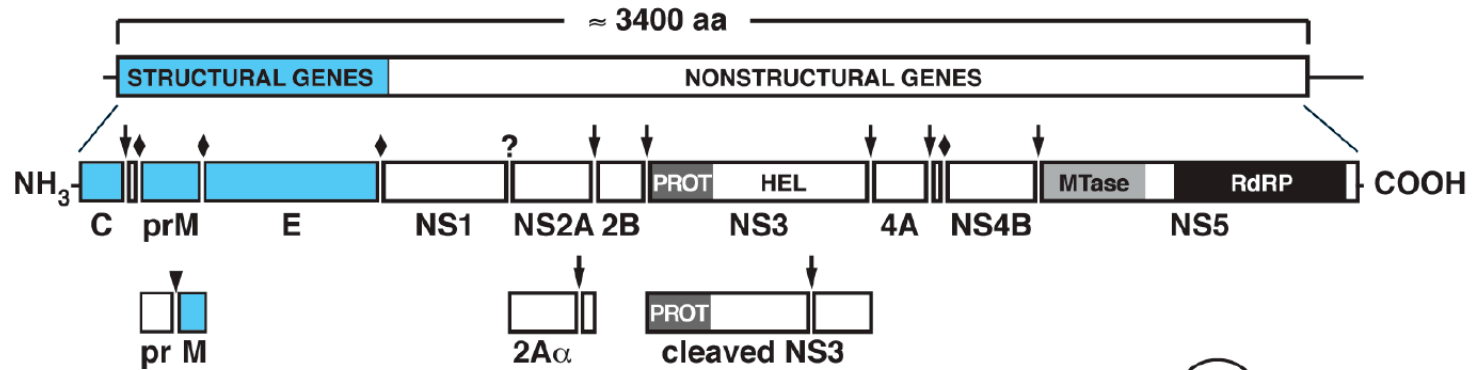
A



Sample AC
WNV-infected
RPKM 1430.4



B



C

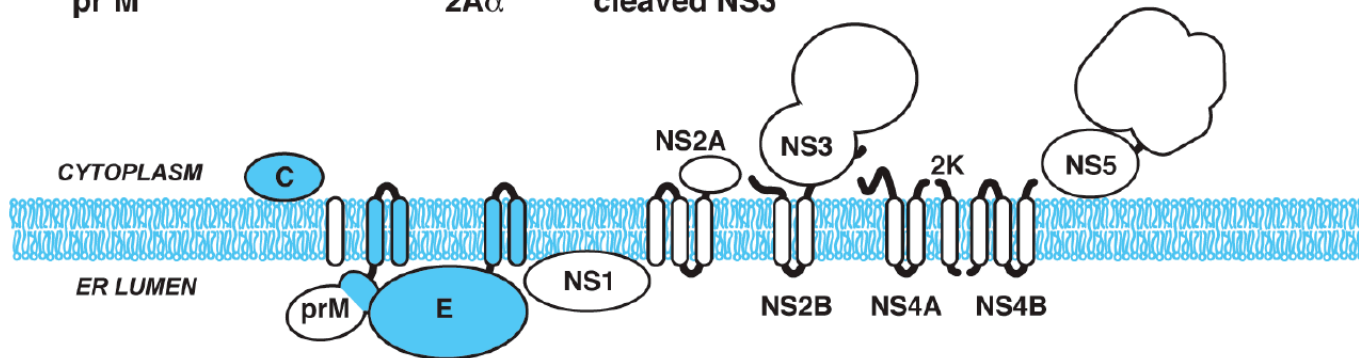
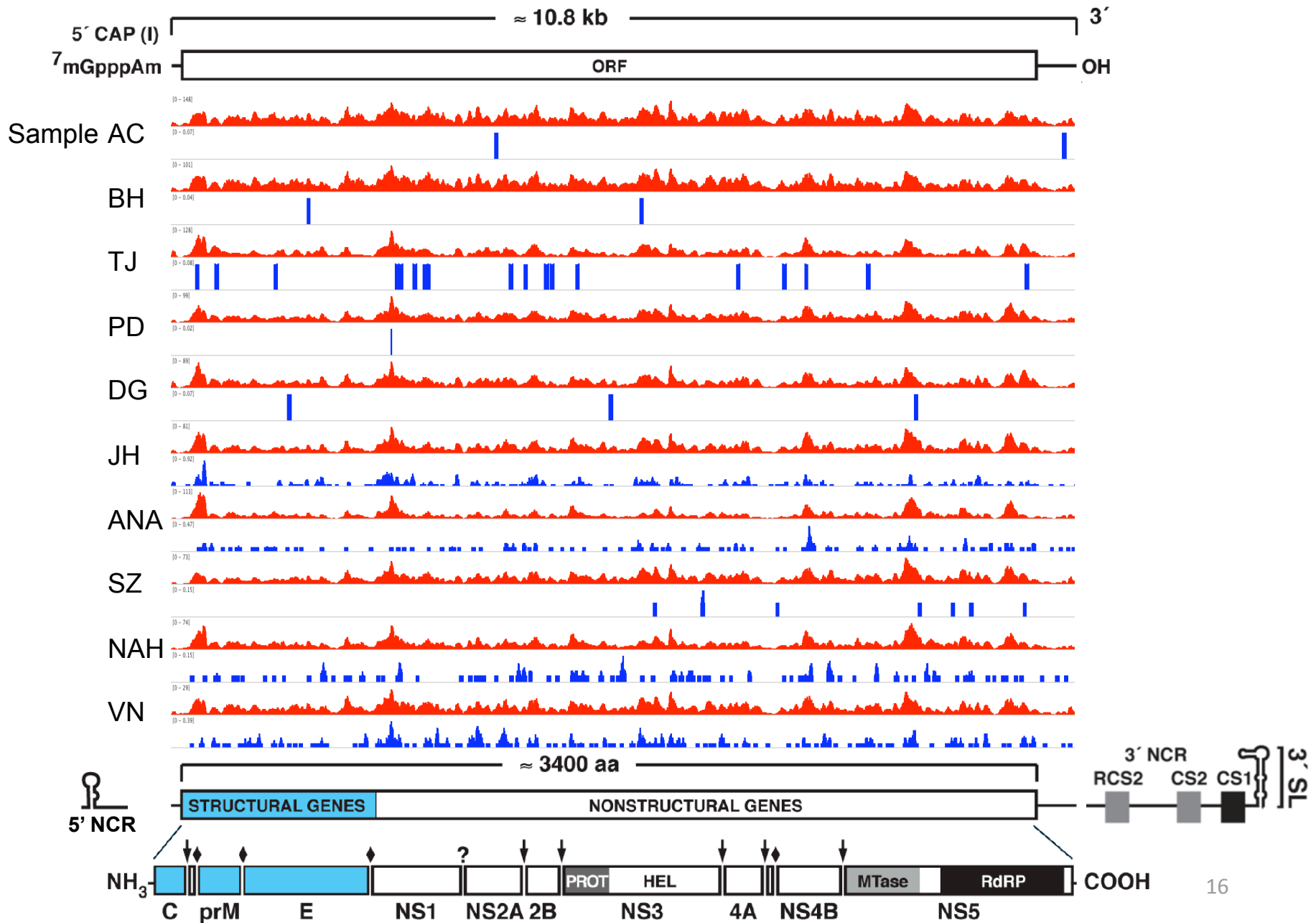


Figure 33.5. Fields Virology 5th ed. (2007)

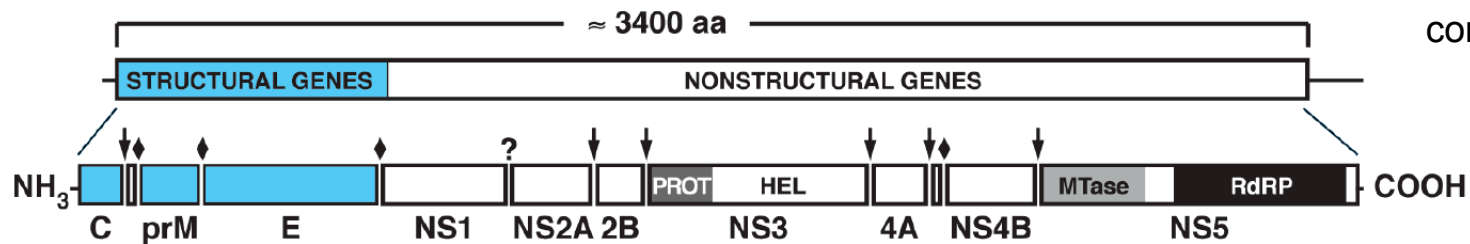
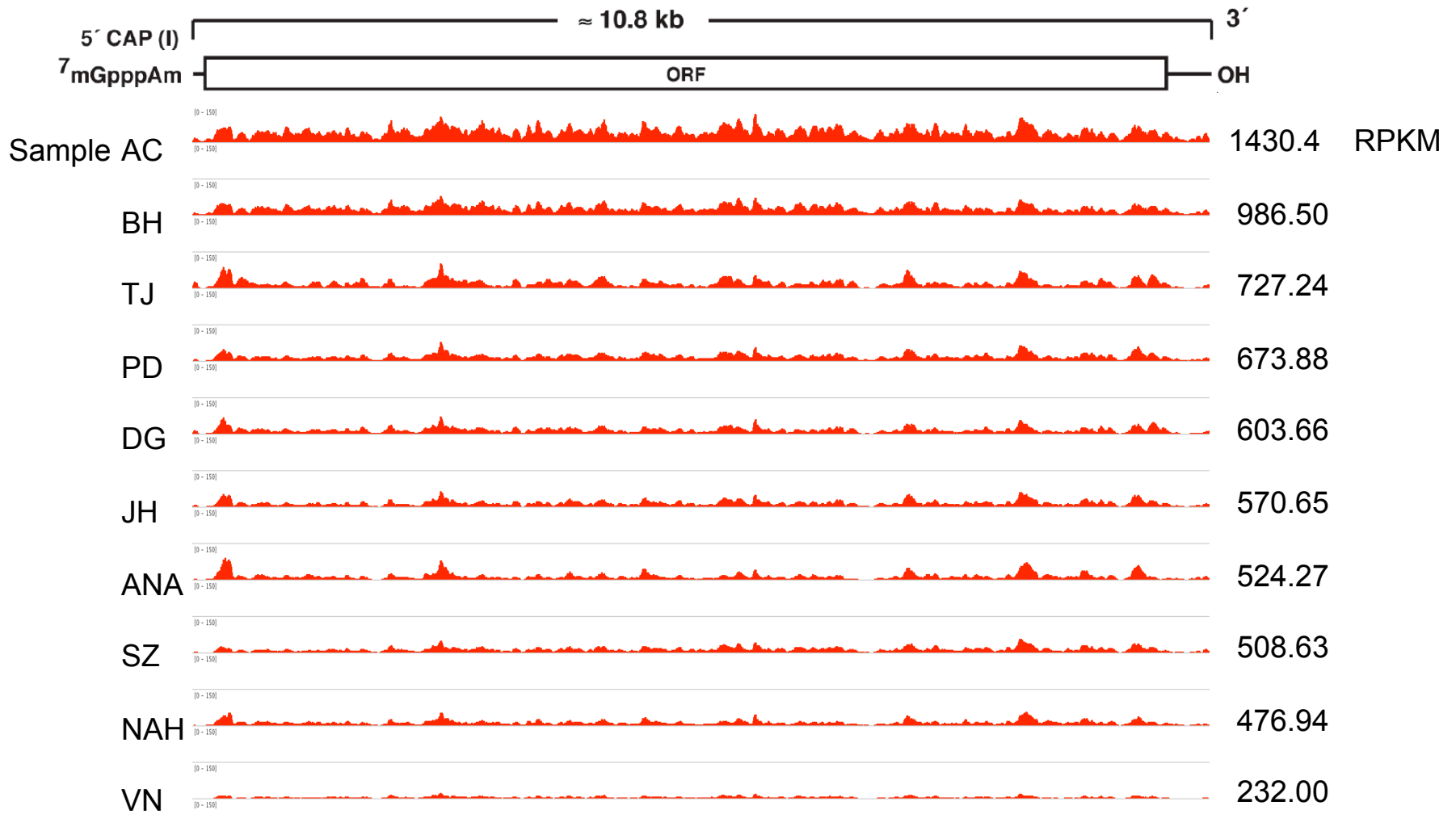
RNAseq reads mapped to WNV genome (autoscaled)

infected
control



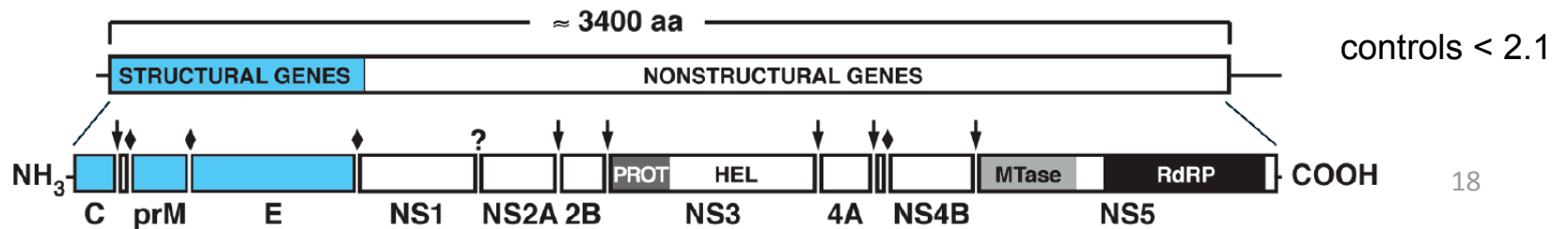
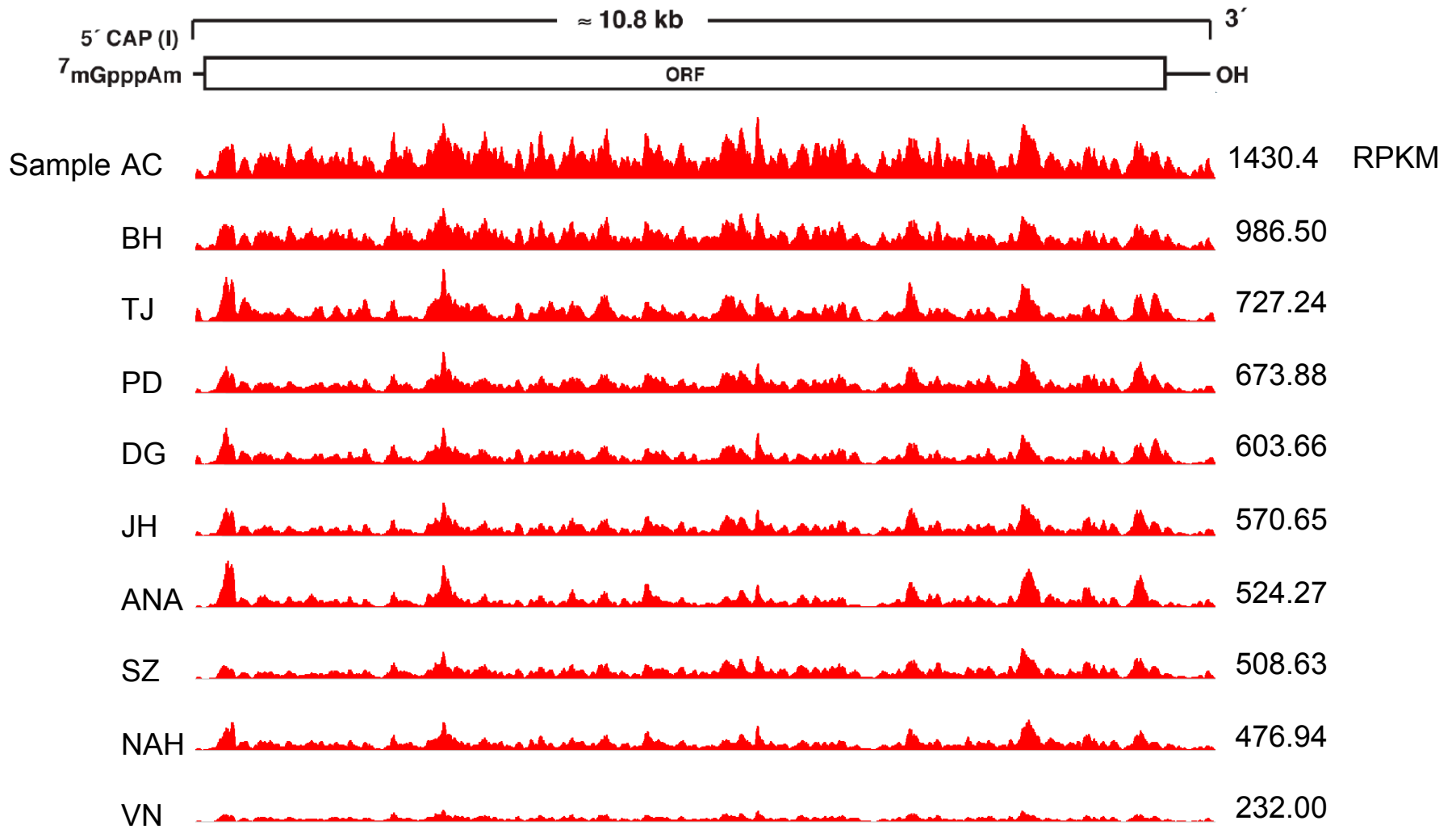
RNAseq reads mapped to WNV genome (normalized)

infected
control



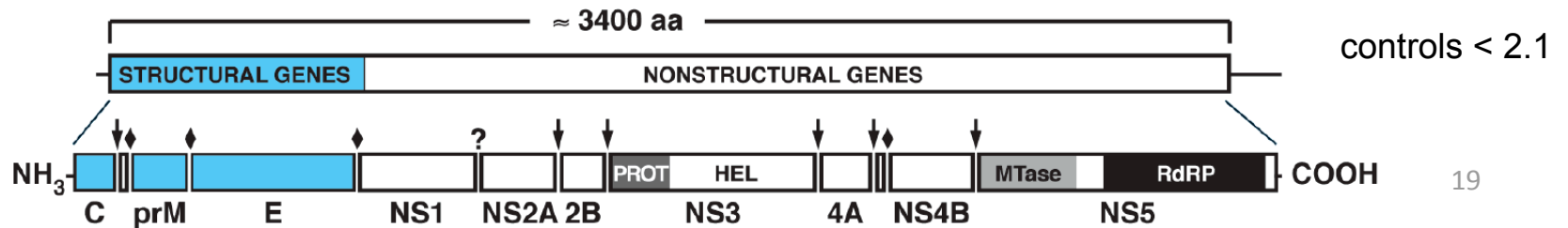
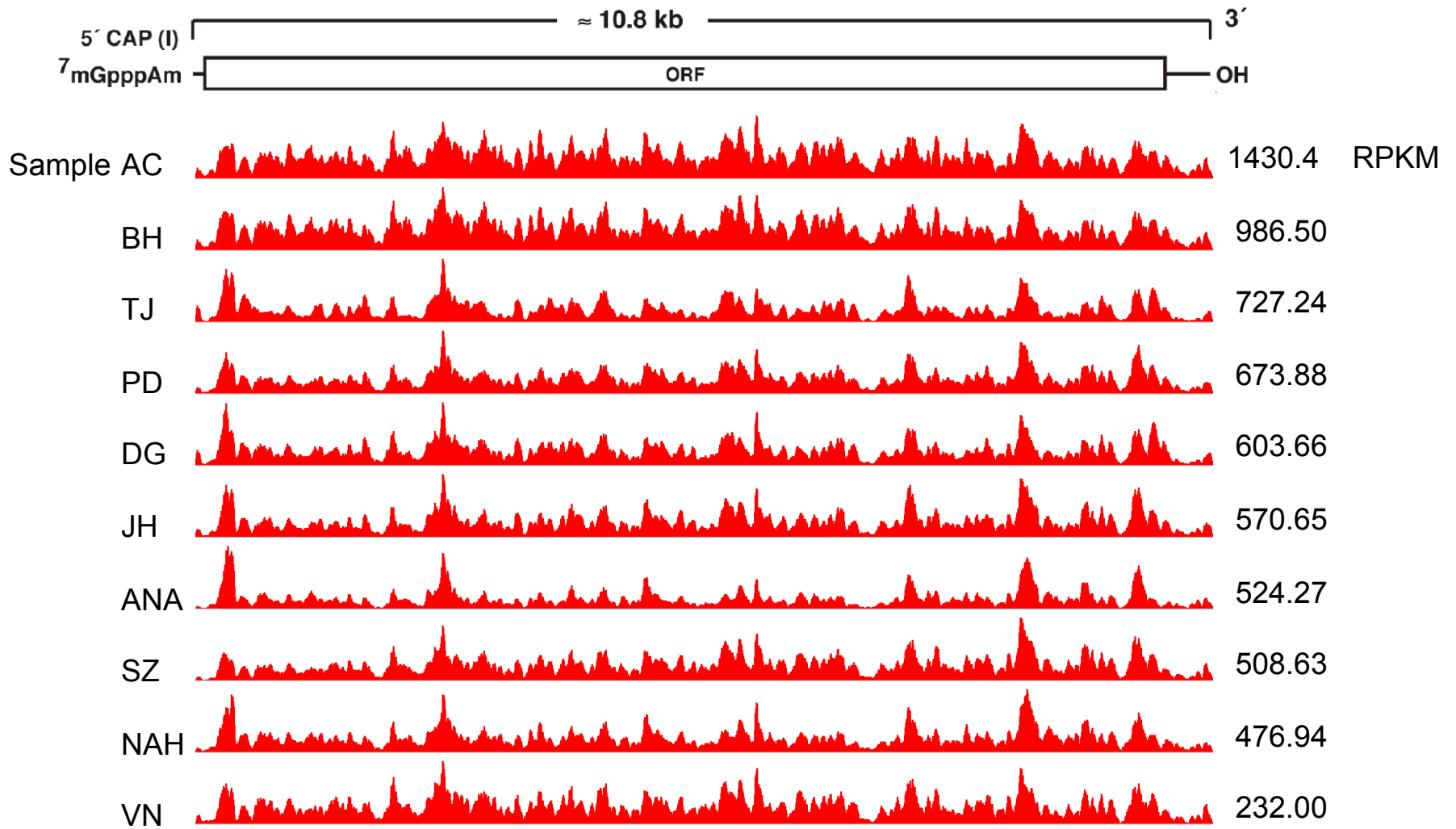
RNAseq reads mapped to WNV genome (normalized)

infected
control

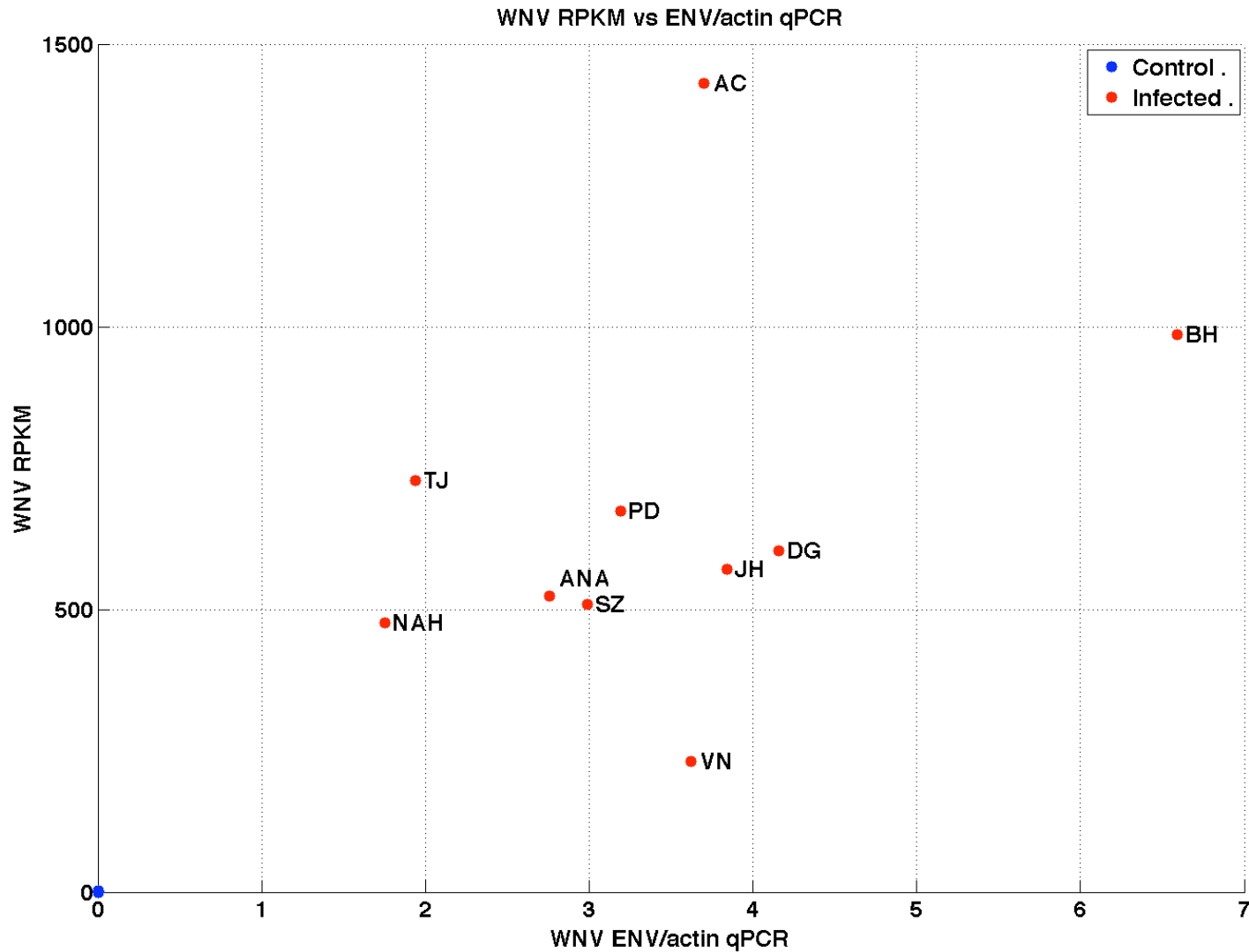


RNAseq reads mapped to WNV genome (autoscaled)

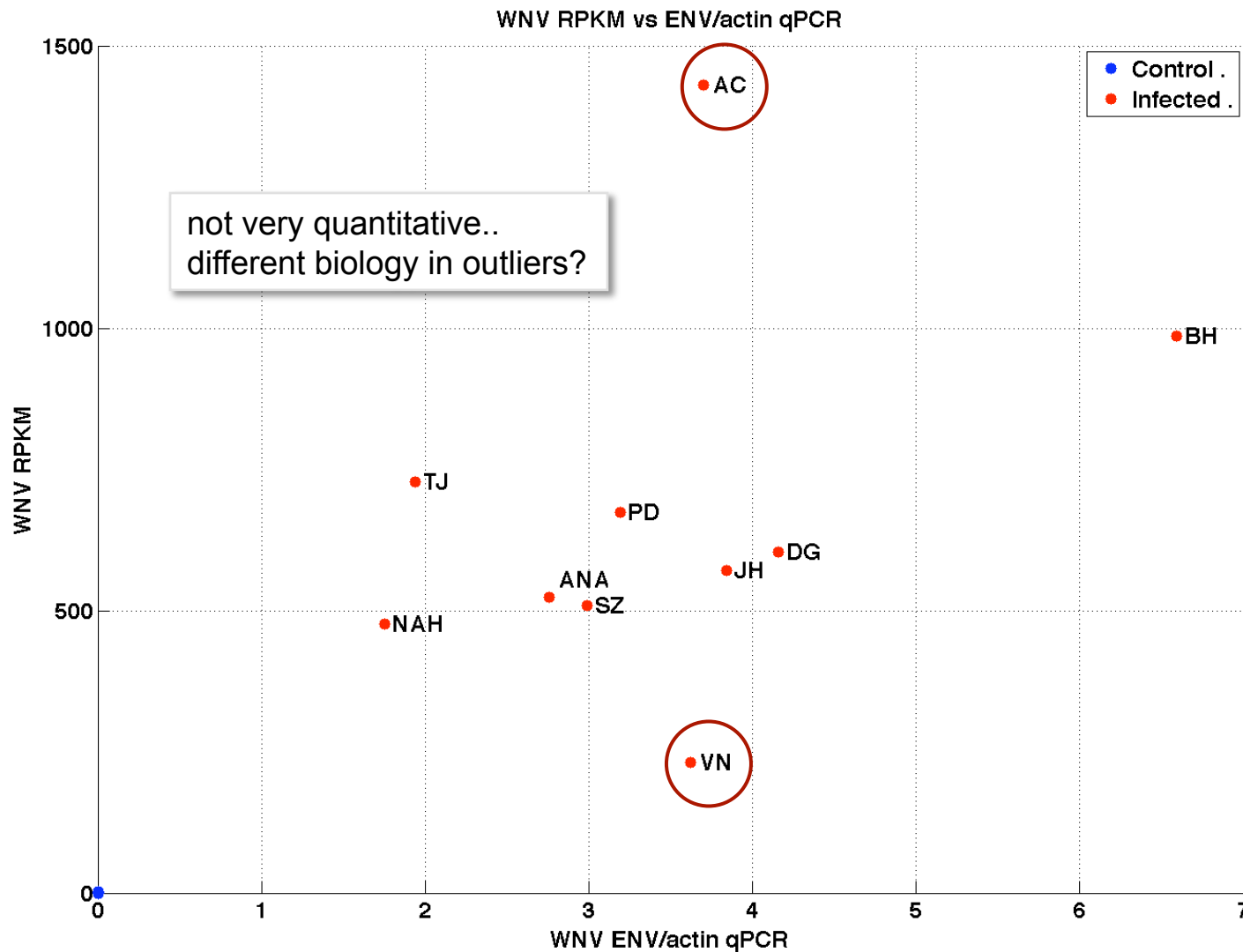
infected
control



Correlation of RNAseq reads mapped to WNV Genome vs. qPCR of WNV envelope gene



Correlation of RNAseq reads mapped to WNV Genome vs. qPCR of WNV envelope gene

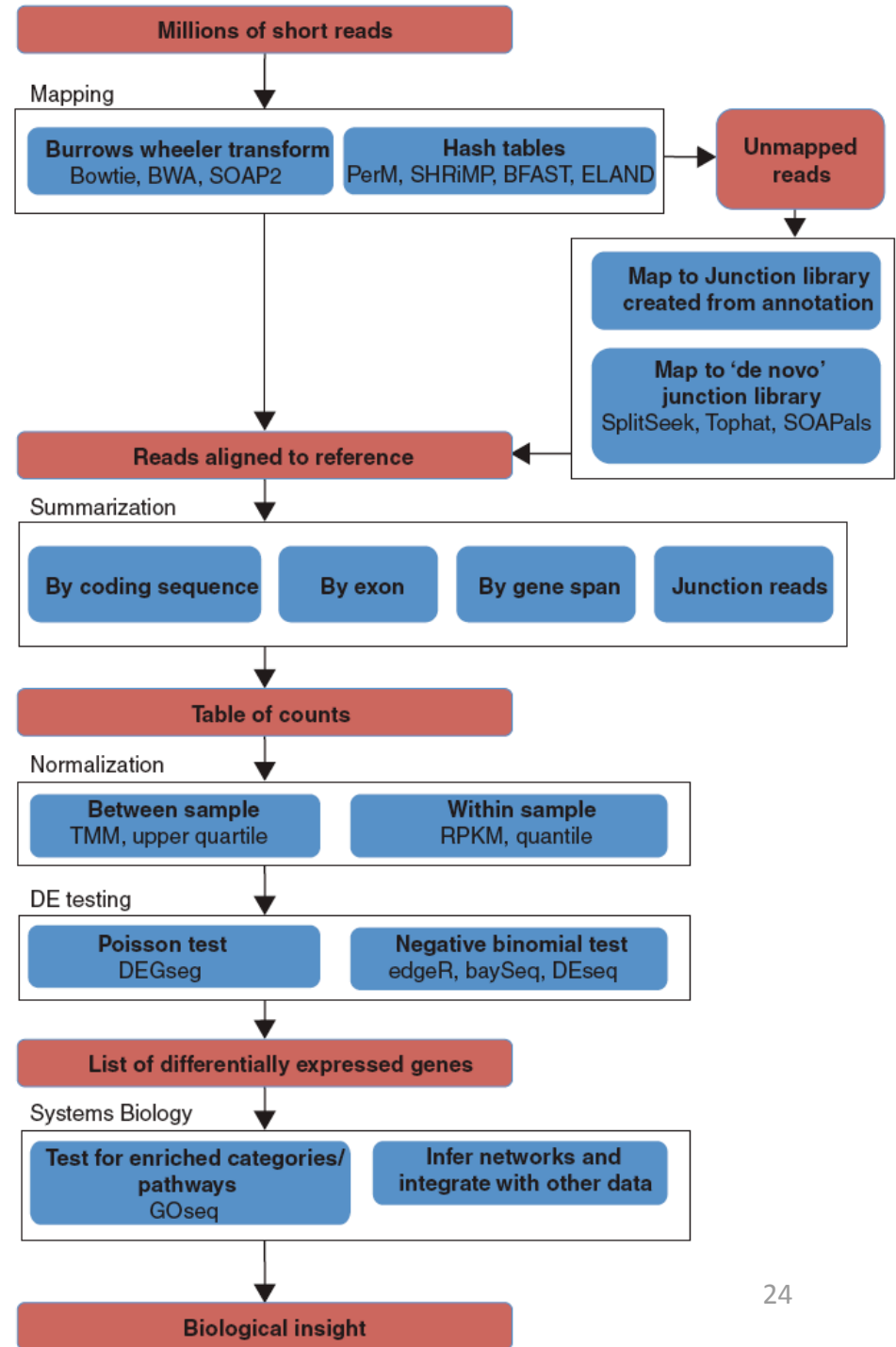


Step	My Analysis	Wei Zheng's Analysis
Read Mapping	Bowtie on hg19+WNV - no novel splice junctions + mapping to WNV	TopHat on hg19 + novel splice junctions - no mapping to WNV
RPKM / FPKM calculation	RSEQtools mrQuantifier GENCODE v3c annotation	Cufflinks Ensembl annotation
Cross-sample Normalization	---	TMM Trimmed Mean of M
Differential Expression (DE)	---	edgeR: negative binomial model pairwise case/control comparisons With FDR≤0.2, 82 DE transcripts in common across 10 samples, 60 with GO annotation
GO enrichment	---	EASE on top 500 transcripts ranked by sum of significance rank in 10 paired tests

Next Steps

- With permissive FDR of 0.2, only ~80 genes in common across all 10 samples
 - => need to pool genes into gene sets / pathways to increase statistical power
 - Chao suggested Gene Set Enrichment Analysis (GSEA), but it is designed for microarray datasets
 - => I need to research pathway-level DE tools for RNAseq data
- In parallel, I need to understand WNV biology better, so that I can select a meaningful set of pathways to test.
- Ruth's response to Wei's analysis was that most of the GO terms, etc, reflect known WNV biology
 - => need to use known WNV pathways as a mask, remove known biology, and look for novel genes / pathways of interest
- So two topics from here on
 - Differential Expression analysis
 - WNV biology

Analyzing differential expression of RNAseq data



End of WNV RNA-seq slides

- Additional topics if there is time
 - Epigenomics roadmap
 - Connecting cross-species cell lineages via ontologies

Human Epigenome Atlas

Release 2,
embargoed
until Aug 2011

		Bisulfite-Seq	MeDip Seq	MRE-Seq	RRBS	DNase Hypersensitivity	mRNA-seq	smRNA-Seq	Exon Arrays	ChIP-Seq Input	Histone H3K27me3	Histone H3K36me3	Histone H3K4me1	Histone H3K4me3	Histone H3K9ac	Histone H3K9me3	Histone H2AK5ac	Histone H2BK5ac	Histone H2BK120ac	Histone H2BK12ac	Histone H2BK15ac	Histone H2BK20ac	Histone H3K14ac	Histone H3K18ac	Histone H3K23ac	Histone H3K27me2	Histone H3K27ac	Histone H3K4ac	Histone H3K4me2	Histone H3K56ac	Histone H3K79me1	Histone H3K79me2	Histone H4K20me1	Histone H4K5ac	Histone H4K8ac	Histone H4K91ac	
Adipose Derived Mesenchymal Stem Cells	<input type="checkbox"/>								1	1	1	1	1	1																							
Adipose Nuclei	<input type="checkbox"/>								1		1	1	1	1																							
Adult Kidney	<input type="checkbox"/>			1					1		1		1		1																						
Adult Liver	<input type="checkbox"/>								1	2	2	2	3	2	2																						
Bone Marrow Derived Mesenchymal Stem Cells	<input type="checkbox"/>									1	1	1	1	1	1																						
CD15 Primary Cells	<input type="checkbox"/>									1			1																								
CD19 Primary Cells	<input type="checkbox"/>			1					1	2	2		2		2																						
CD3 Primary Cells	<input type="checkbox"/>								1	2	2	2	2	2	2																						
CD34 Cultured Cells	<input type="checkbox"/>								1	1	1	1	1	1	1																						
CD34 Primary Cells	<input type="checkbox"/>			2					1	2	2	2	2	2	2																						
CD4 Memory Primary Cells	<input type="checkbox"/>								1																												
Colonic Mucosa	<input type="checkbox"/>			1					1	1	1	1	1	1	1																						
Duodenum Mucosa	<input type="checkbox"/>			1					1	1	1	1	1	1	1																						
ES-13 Cell Line	<input type="checkbox"/>			2					2	2	2	1	2	2	2																						
ES-WA7 Cell Line	<input type="checkbox"/>			2						1	1	1	1	1	1																						
Fetal Brain	<input type="checkbox"/>									1	1			1																							
Fetal Heart	<input type="checkbox"/>								1	1	1																										
Fetal Kidney	<input type="checkbox"/>									1	1	1	1	1	1	1																					
Fetal Lung	<input type="checkbox"/>									1	1	1	2	1	1																						
H1 Cell Line	<input type="checkbox"/>			4					2	2	2	2	3	2	1																						
HUES1 Cell Line	<input type="checkbox"/>			1																																	
HUES28 Cell Line	<input type="checkbox"/>			1																																	
HUES3 Cell Line	<input type="checkbox"/>			1																																	
HUES48 Cell Line	<input type="checkbox"/>			1																																	
HUES49 Cell Line	<input type="checkbox"/>			1																																	
HUES53 Cell Line	<input type="checkbox"/>			1																																	
HUES6 Cell Line	<input type="checkbox"/>			1																																	
HUES62 Cell Line	<input type="checkbox"/>			1																																	
HUES63 Cell Line	<input type="checkbox"/>			1																																	
HUES64 Cell Line	<input type="checkbox"/>			1																																	
HUES65 Cell Line	<input type="checkbox"/>			1																																	
HUES66 Cell Line	<input type="checkbox"/>			1																																	
HUES8 Cell Line	<input type="checkbox"/>			1																																	
HUES9 Cell Line	<input type="checkbox"/>			1																																	

Human Epigenome Atlas

Release 2,
embargoed
until Aug 2011

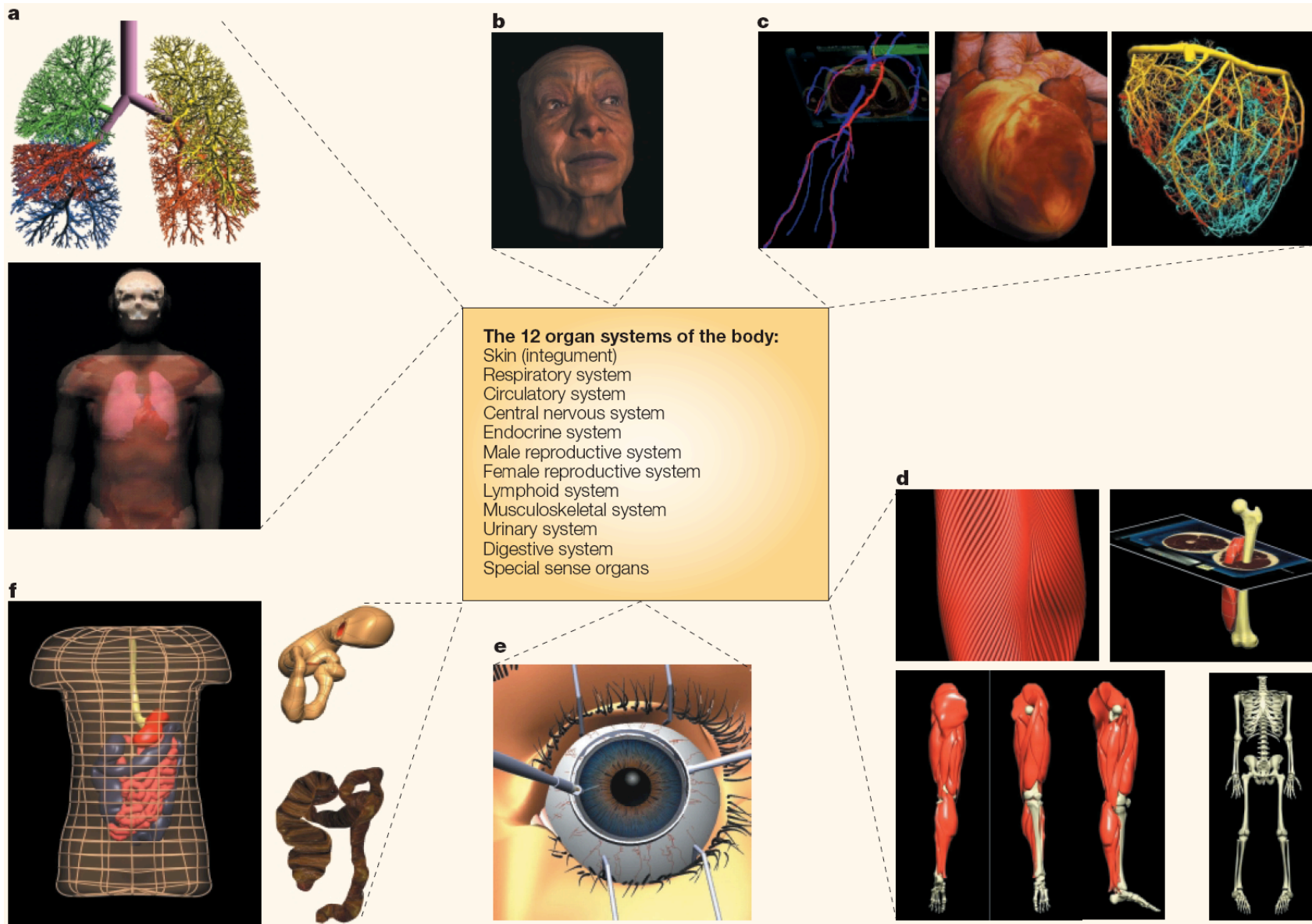
		Bisulfite-Seq	MeDip Seq	MRE-Seq	RRBS	DNase Hypersensitivity	mRNA-seq	smRNA-Seq	Exon Arrays	ChIP-Seq Input	Histone H3K27me3	Histone H3K36me3	Histone H3K4me1	Histone H3K4me3	Histone H3K9ac	Histone H2AK9me3	Histone H2AK5ac	Histone H2BK5ac	Histone H2BK120ac	Histone H2BK12ac	Histone H2BK15ac	Histone H2BK20ac	Histone H3K14ac	Histone H3K18ac	Histone H3K23ac	Histone H3K27me2	Histone H3K27ac	Histone H3K4ac	Histone H3K4me2	Histone H3K56ac	Histone H3K79me1	Histone H3K79me2	Histone H4K20me1	Histone H4K5ac	Histone H4K8ac	Histone H4K91ac		
Mobilized CD34 Primary Cells	<input type="checkbox"/>								4	3	3	2	6		4																							
Mobilized CD4 Primary Cells	<input type="checkbox"/>								1																													
Muscle Satellite Cultured Cells	<input type="checkbox"/>								3	3	3	3	3	3														3										
Pancreatic Islets	<input type="checkbox"/>			1					1	1	1	1	1	1	1																							
Rectal Mucosa	<input type="checkbox"/>			1					2	2	2	2	2	1	2																							
Rectal Smooth Muscle	<input type="checkbox"/>			1					1	1	1		1	1	1																							
Skeletal Muscle	<input type="checkbox"/>			2					2	3	3	3	3	3	3																							
Stomach Smooth Muscle	<input type="checkbox"/>			1					1	1	1	1	1		1																							
Th17 Primary Cells	<input type="checkbox"/>								1							1																						
Treg Primary Cells	<input type="checkbox"/>								1	1	1		1		1																							
iPS-11a Cell Line	<input type="checkbox"/>			1					1				1																									
iPS-11b Cell Line	<input type="checkbox"/>			1																																		
iPS-15b Cell Line	<input type="checkbox"/>			2					2	1	1		1	1	1																							
iPS-17a Cell Line	<input type="checkbox"/>			1																																		
iPS-17b Cell Line	<input type="checkbox"/>			1																																		
iPS-18a Cell Line	<input type="checkbox"/>			1																																		
iPS-18b Cell Line	<input type="checkbox"/>			1					1																													
iPS-18c Cell Line	<input type="checkbox"/>			2					1	1	1		2	1	1																							
iPS-20b Cell Line	<input type="checkbox"/>			3					2	2	1		1	1	1																							
H1 Cell Line	<input type="checkbox"/>	1							7	3	3	3	3	2	3	2	2	1	2	2	2		2		2	1	1	2	1	3	2	1	1			1		
H1-BMP4 Cell Line	<input type="checkbox"/>	1																																				
H9 Cell Line	<input type="checkbox"/>										1		1	1												1		1										
IMR90 Cell Line	<input type="checkbox"/>	1							8	2	2	3	2	2	3	2		2	3	3	2	2	2	2		3	2	2	2	4	2	2	2	4	2			
Breast Luminal Epithelial Cells	<input type="checkbox"/>	4	5			1																																
Breast Myoepithelial Cells	<input type="checkbox"/>	4	3			1			2	2	2	2	1	1	2																							
Breast Stem Cells	<input type="checkbox"/>	4	4																																			
Breast vHMEC	<input type="checkbox"/>	1	1			1	1		2			2	1		1																							
CD4 Memory Primary Cells	<input type="checkbox"/>	2	2																																			
CD4 Naive Primary Cells	<input type="checkbox"/>	2	3								1			1																								
CD8 Naive Primary Cells	<input type="checkbox"/>	2	3						1	1		1	2	1	1																							
H1 Cell Line	<input type="checkbox"/>	2	3			1	1		2	1	2	1	2	1	2																							
Peripheral Blood Mononuclear Primary Cells	<input type="checkbox"/>	2	3						3	1	1	1		2	2																							

Human Epigenome Atlas

Release 2,
embargoed
until Aug 2011

		Bisulfite-Seq	MeDip Seq	MRE-Seq	RRBS	DNase Hypersensitivity	mRNA-seq	smRNA-Seq	Exon Arrays	ChIP-Seq Input	Histone H3K27me3	Histone H3K36me3	Histone H3K4me1	Histone H3K4me3	Histone H3K9ac	Histone H3K9me3	Histone H2AK5ac	Histone H2BK5ac	Histone H2BK120ac	Histone H2BK12ac	Histone H2BK15ac	Histone H2BK20ac	Histone H3K14ac	Histone H3K18ac	Histone H3K23ac	Histone H3K23me2	Histone H3K27ac	Histone H3K4ac	Histone H3K4me2	Histone H3K56ac	Histone H3K79me1	Histone H3K79me2	Histone H4K20me1	Histone H4K5ac	Histone H4K8ac	Histone H4K91ac	
WU	Fetal Adrenal Gland	<input type="checkbox"/>				1																															
	Fetal Brain	<input type="checkbox"/>				7																															
	Fetal Heart	<input type="checkbox"/>				2																															
	Fetal Kidney	<input type="checkbox"/>				2																															
	Fetal Lung	<input type="checkbox"/>				10																															
	IMR90 Cell Line	<input type="checkbox"/>				4																															
	Mobilized CD34 Primary Cells	<input type="checkbox"/>				14																															

Connecting Ontologies: Physiome Project



Connection to Pseudogenes

- None

Three (of five) star presentation

- You must use Helvetica font
- Slides must be numbered
- Presentation must be uploaded to wiki

- Presentation must be about pseudogenes
- Presenter must bring cup cakes

Postdoc Project History

Past

- Caulobacter RodZ (Emonet / Jacobs-Wagner)
- CheV review (Emonet)
- Sci. Signaling paper on network dynamics
- NRG review on non-coding annotation
- modENCODE manuscript wrangling

Future

- unfinished protein sectors work
- WNV RNA-seq / DE for brainseq?
- epigenomics
- 1st author research