UM - Flu Challenge Data Analysis
Initial findings

Dec 2, 2008

A Hero, Y, Huang
Dept of EECS and Bioinformatics Program
University of Michigan

Summary

1.Normalized and performed analysis of mRNA only (IA., AA, CC, and Metabolites not available).

2. FDR gene screening – flu has much higher number ofsx/asx discriminating biomarkers than RSV or HRV studies.

4. Prediction - milestones attained using 5 top ranked FDR screened genes and naïve quadratic classifier..

3. Mixed component analysis (MCA) – early and late acting biomarkers identified

20% FDR screening: Early d.e.

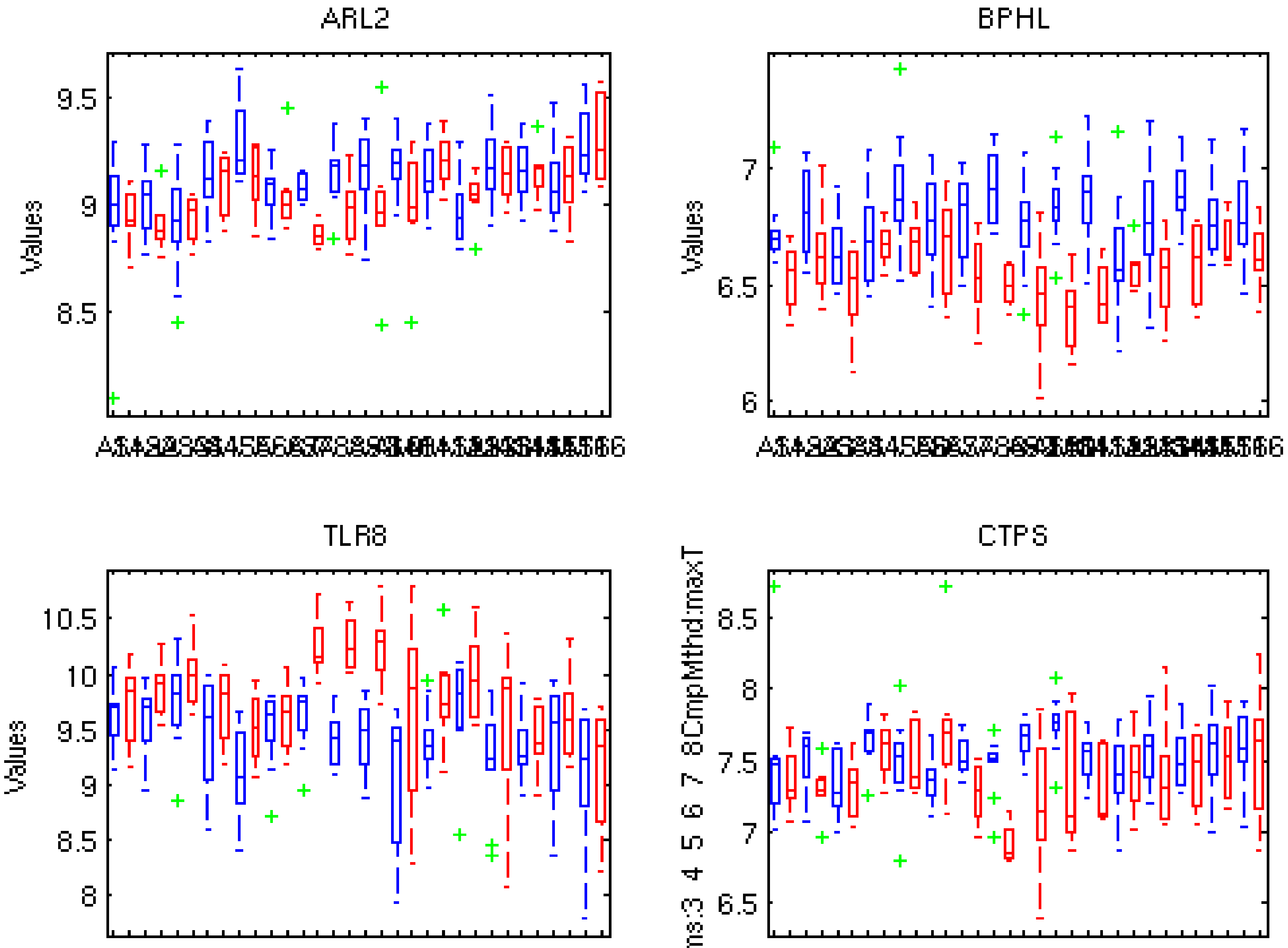
	Persistent dominance	Instantaneous
HRV	372/778	1/1'
RSV	18/236	0/16
INF	1590/1693	206/535

KEY: X/Y

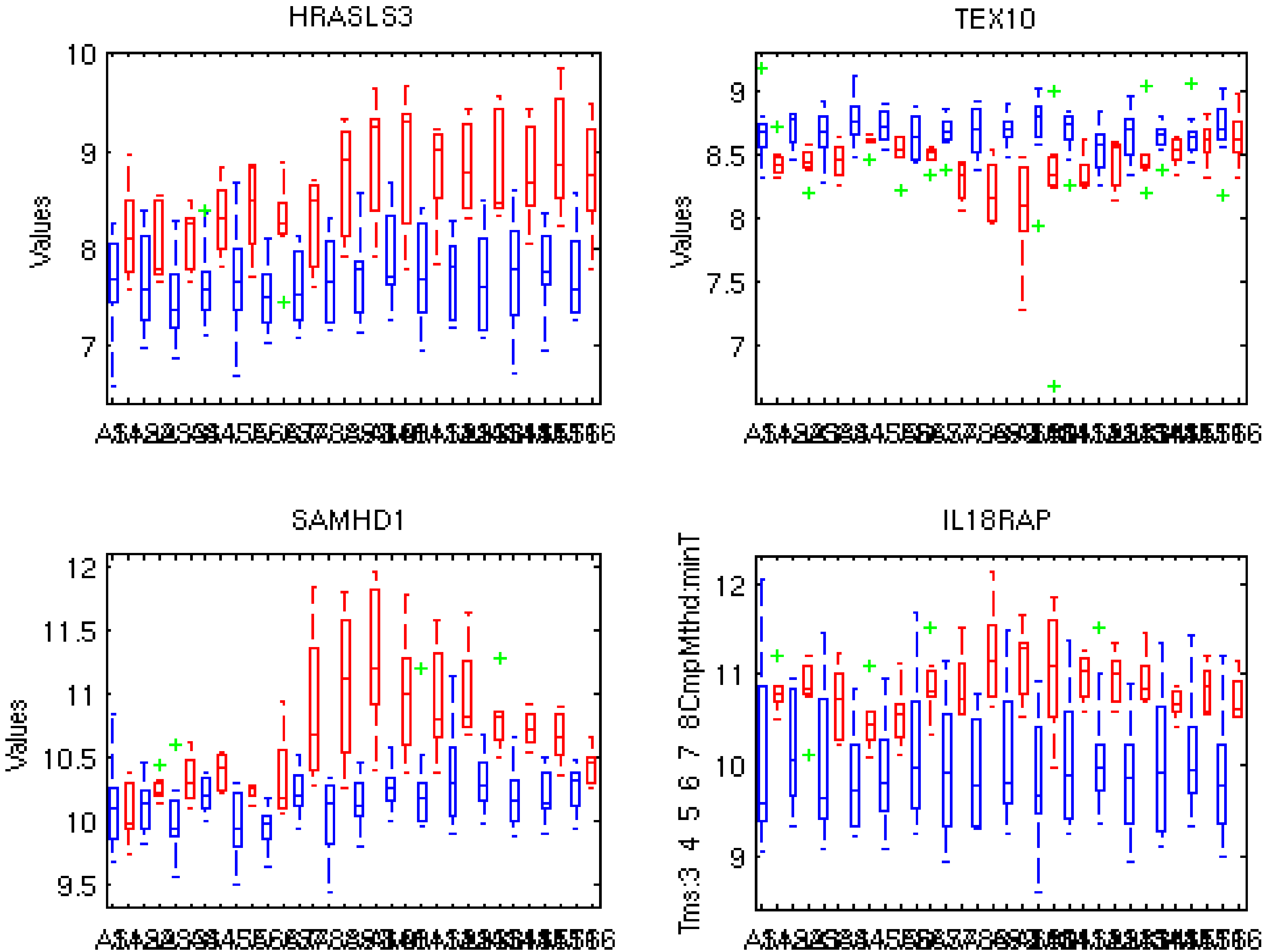
X d.e. between .1T and 0.8T

Y d.e. between .1T and T

Flu: 4 early d.e: instantaneous



Flu: 4 early d.e: persistent



Prediction milestones: 5 genes

FDR early screened: HOXA6', 'EXTL2'

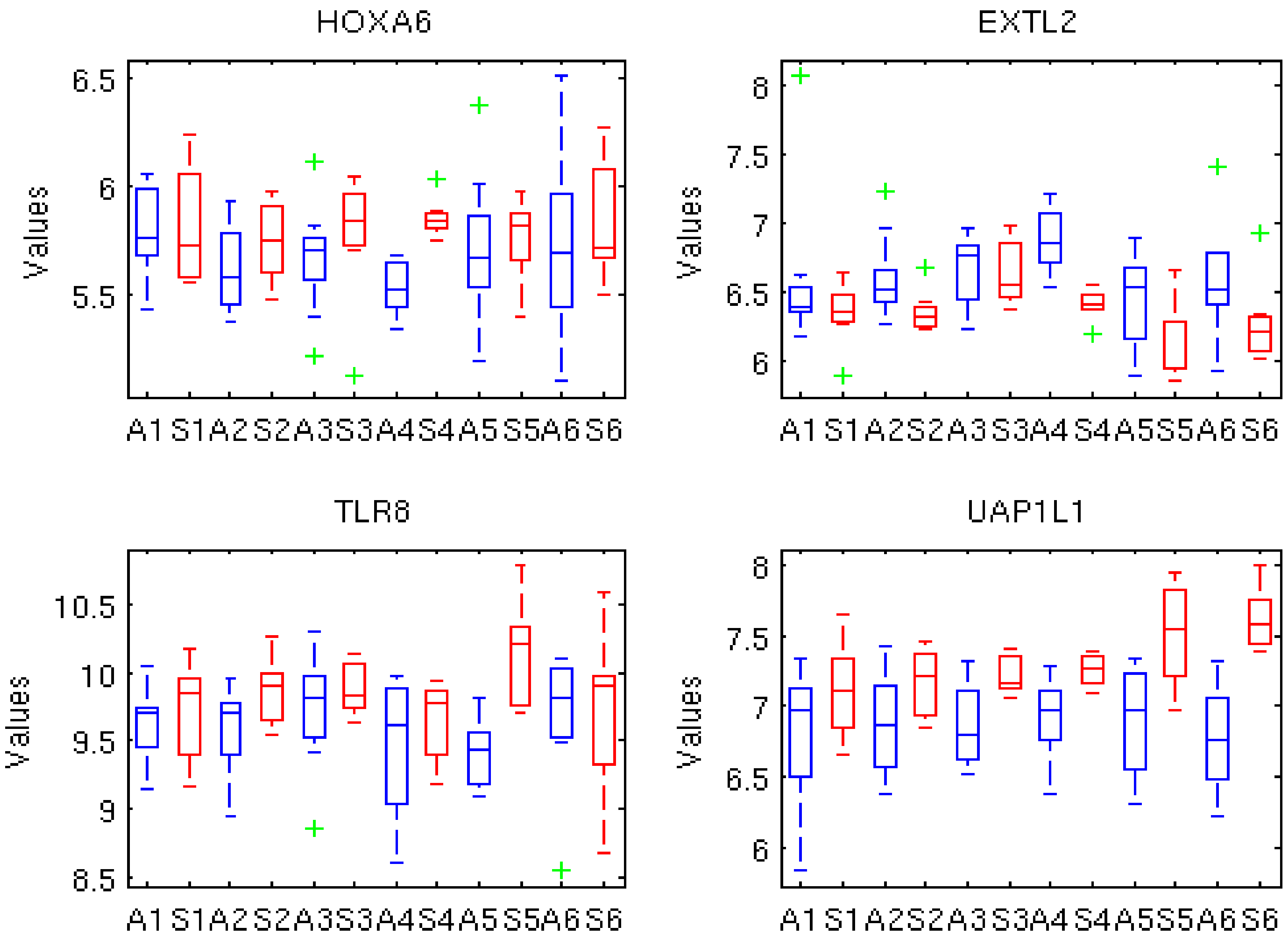
FDR late screened: 'TLR8', 'UAP1L1', 'HRASLS3'

Naïve (diagonal) quadratic classifier

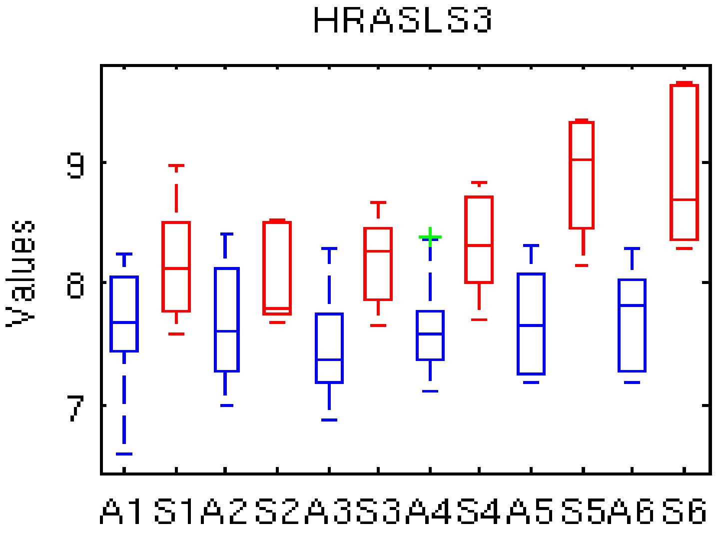
achieves milestones

	1-alpha	1-beta
0.1T	1.00	0.72
0.2T	1.00	0.97
0.8T	0.98	1.00

Predictor gene plots (blue are asx, red are sx)



Predictor Gene Plots (ctd)



Mixed Component Analysis (MCA)

MCA is a recent method [1] orginally developed for hyperspectral imaging in geo remote sensing (under AFOSR-MURI)

[1] N. Dobigeon, J.-Y. Tourneret, S. Massaoui, M. Coulon and A.O. Hero, 'Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery,' IEEE Trans. on Signal Pocessing, submitted Sept 2008.

Properties

- decomposes the overall expression into different possibly overlapping components
- extracts component loading proportionality factors for each chip
- like Bayes factor regression methods (BFRM), MCA operates without subject or sx/asx labels.

MCA vs BFRM

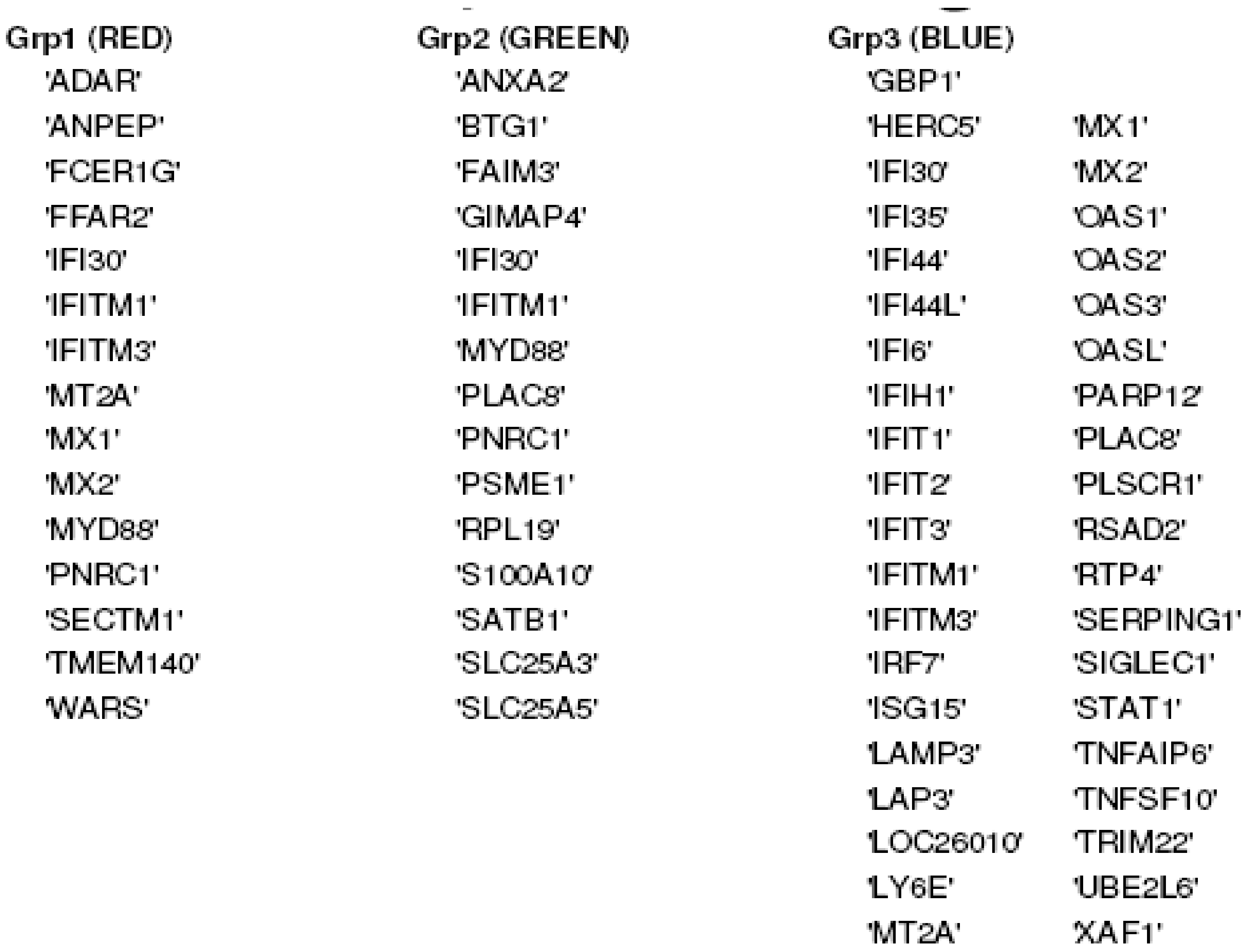
MCA is similar to BFRM

- MCA and BFRM both use similar Bayesian model
- $Y=MA+N$
- Y: Ngenes by Nchips observed matrix
- M: Ngenes by Ncomp factor components (unknown)
- A: Ncomp by Nchip loadings (unknown)

MCA is different than BFRM

- A lies on positive simplex ($A>0$ and colsum(A) =1)
- Dimensionality reduction (PCA) is used to estimate A
- MCA leverages on hyperspectral imaging technology (fast scalable algorithms)

Endmembers (gene factors)



Grp1 (RED)

'ADAR'
'ANPEP'
'FCER1G'
'FFAR2'
'IFI30'
'IFITM1'
'IFITM3'
'MT2A'
'MX1'
'MX2'
'MYD88'
'PNRC1'
'SECTM1'
'TMEM140'
'WARS'

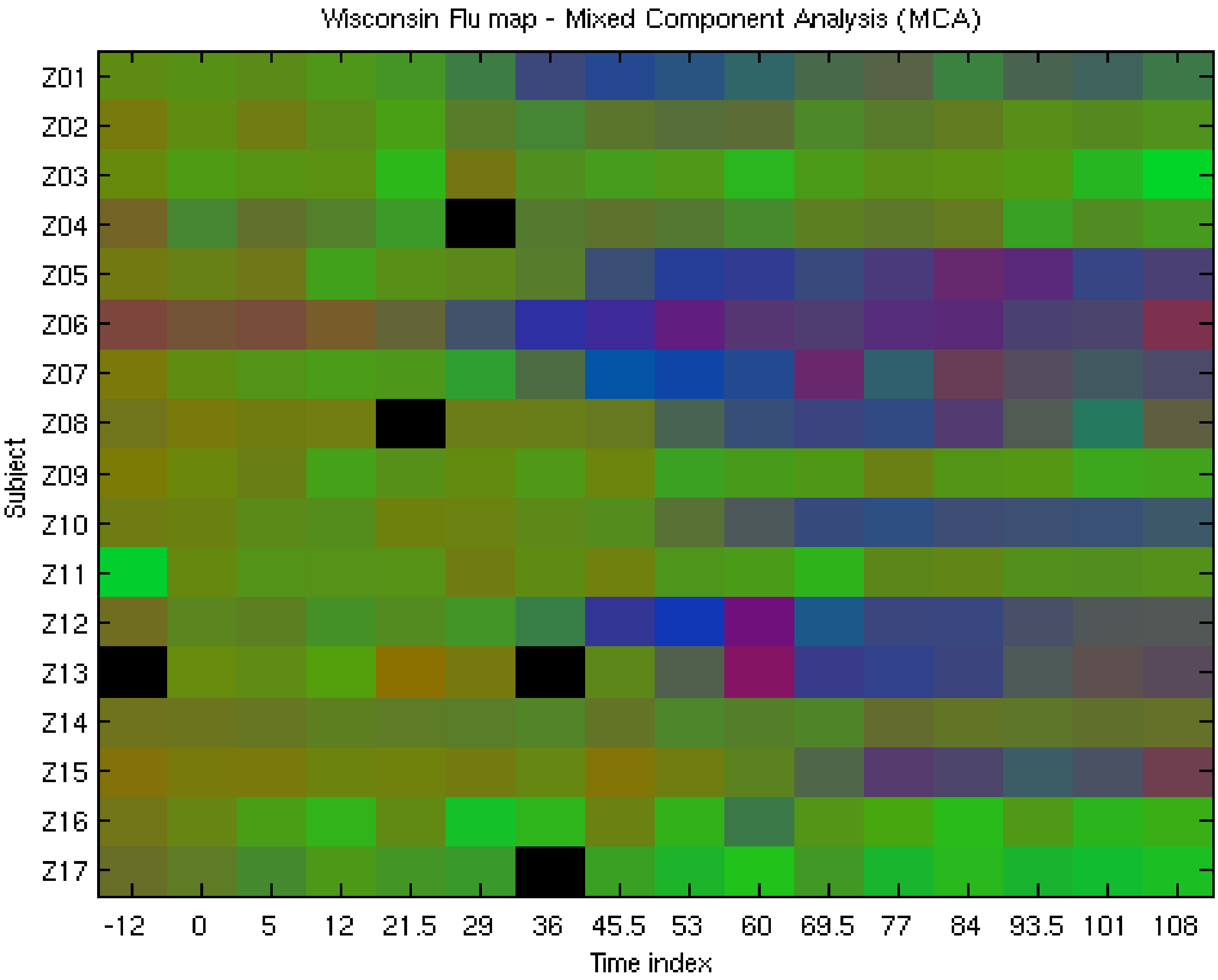
Grp2 (GREEN)

'ANXA2'
'BTG1'
'FAIM3'
'GIMAP4'
'IFI30'
'IFITM1'
'MYD88'
'PLAC8'
'PNRC1'
'PSME1'
'RPL19'
'S100A10'
'SATB1'
'SLC25A3'
'SLC25A5'

Grp3 (BLUE)

'GBP1'
'HERC5'
'IFI30'
'IFI35'
'IFI44'
'IFI44L'
'IFI6'
'IFIH1'
'IFIT1'
'IFIT2'
'IFIT3'
'IFITM1'
'IFITM3'
'IRF7'
'ISG15'
'LAMP3'
'LAP3'
'LOC26010'
'LY6E'
'MT2A'
'MX1'
'MX2'
'OAS1'
'OAS2'
'OAS3'
'OASL'
'PARP12'
'PLAC8'
'PLSCR1'
'RSAD2'
'RTP4'
'SERPING1'
'SIGLEC1'
'STAT1'
'TNFAIP6'
'TNFSF10'
'TRIM22'
'UBE2L6'
'XAF1'

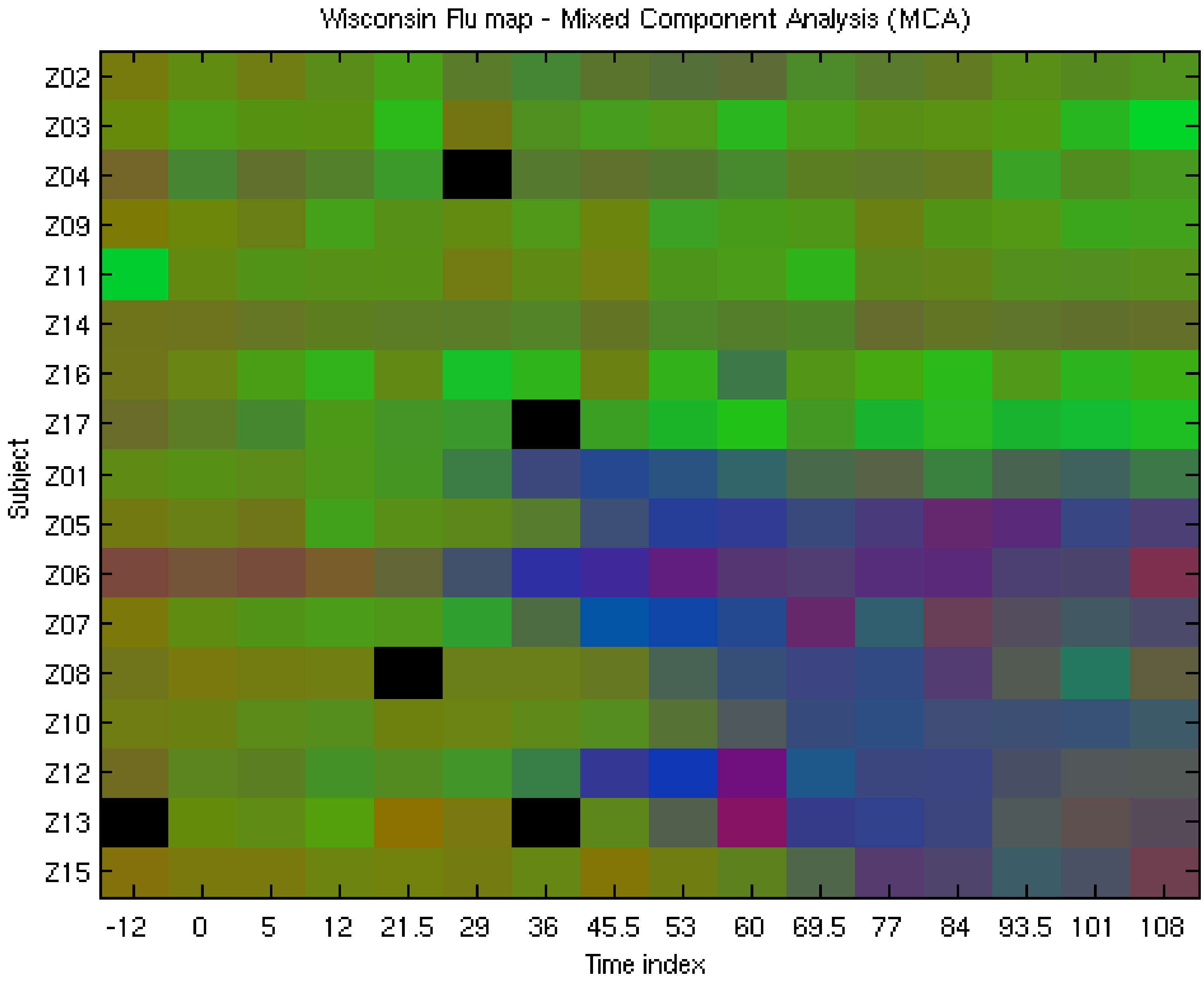
RGB Heatmap (subj by time)



INF Symptom Scores

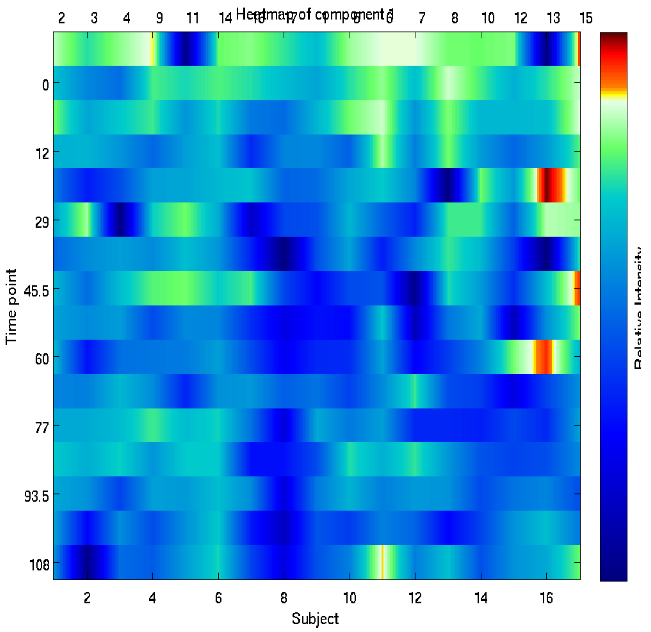
flu001	0	1	2	4	5	9	6	5	4	2	1	0
flu002	2	0	0	0	0	0	0	0	0	0	0	0
flu003	0	0	0	0	0	0	0	0	0	0	0	0
flu004	0	0	0	0	0	0	0	0	0	0	0	0
flu005	0	3	1	4	6	9	11	11	12	9	9	1
flu006	0	0	1	6	7	6	7	7	5	5	4	2
flu007	0	0	0	6	11	12	6	4	2	3	3	3
flu008	0	0	0	0	2	6	10	8	10	10	6	5
flu009	0	0	0	0	1	1	1	1	1	1	1	0
flu010	0	0	0	0	0	3	4	3	4	5	1	1
flu011	0	0	0	0	2	2	1	0	0	0	0	0
flu012	0	0	0	0	0	2	2	4	3	3	2	2
flu013	0	0	0	0	0	1	1	2	2	2	1	1
flu014	0	0	0	0	0	0	0	0	1	0	0	0
flu015	0	1	0	0	0	2	2	0	1	2	1	0
flu016	0	0	0	0	0	1	0	0	0	0	0	0
flu017	0	0	0	0	0	0	0	0	0	0	0	0

Heatmap – grouped by asx/sx labels

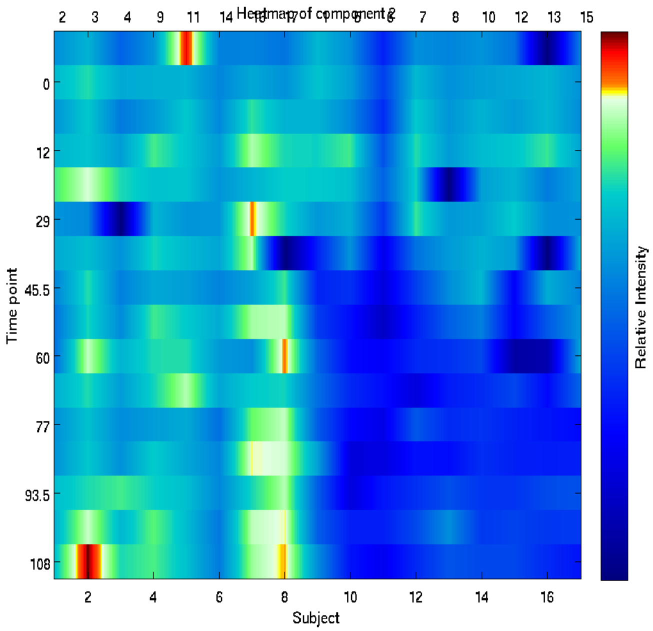


Chip loading factors (time by subj)

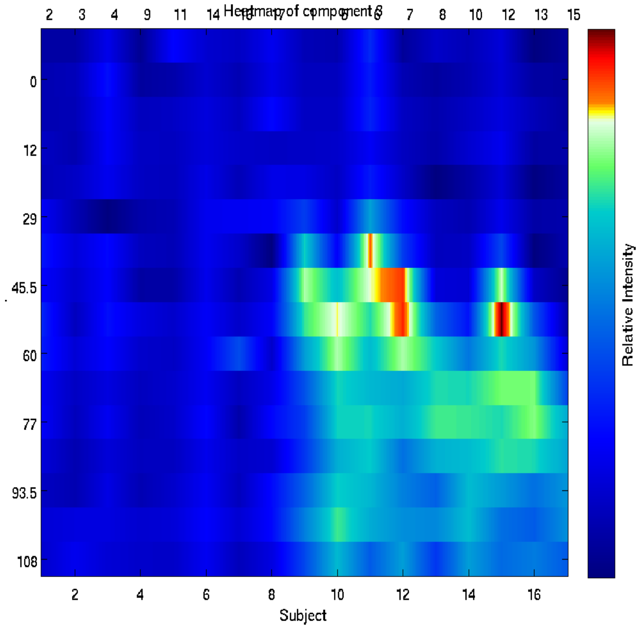
Grp1



Grp2



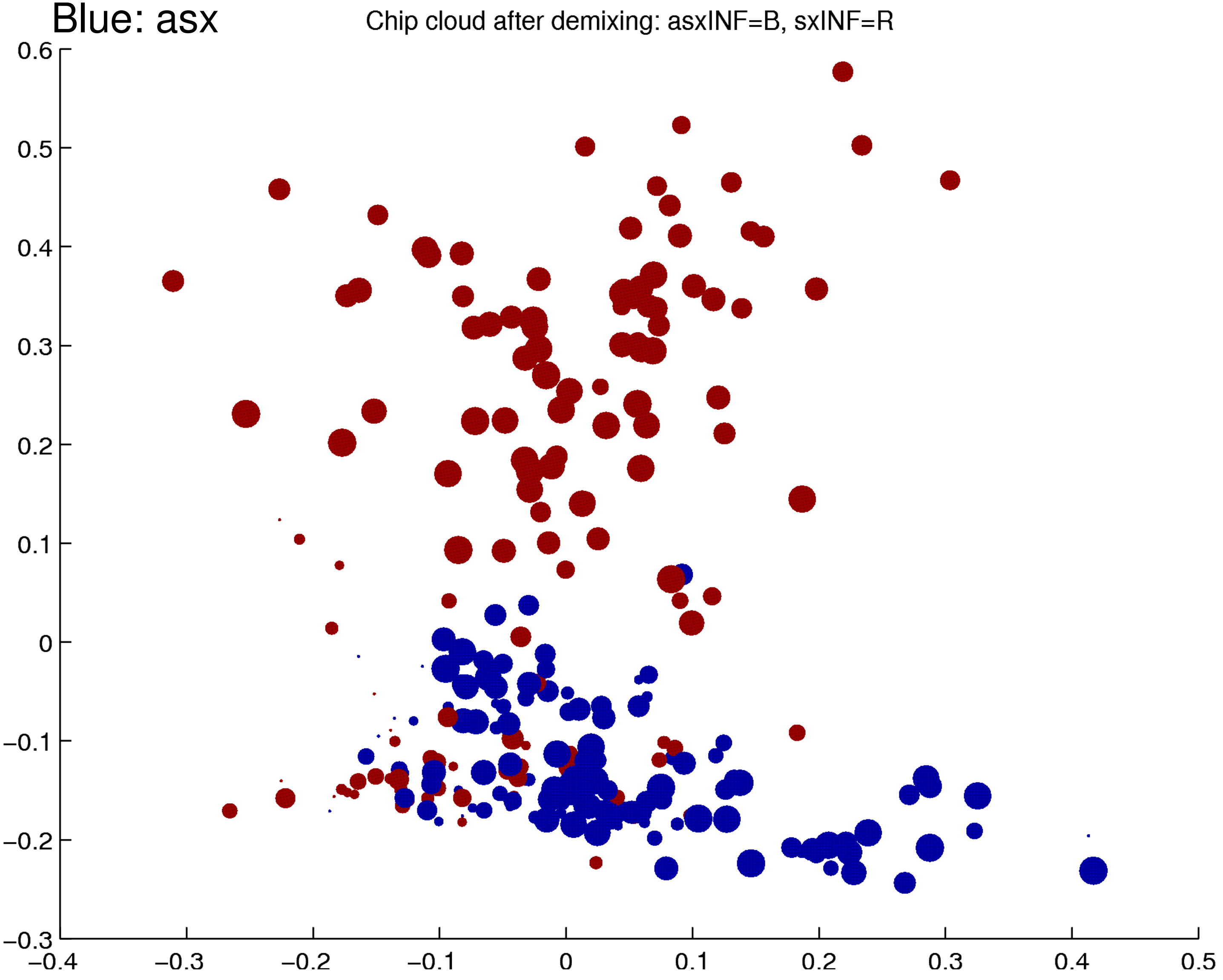
Grp3



MCA INF chip scatter plot

Red: sx

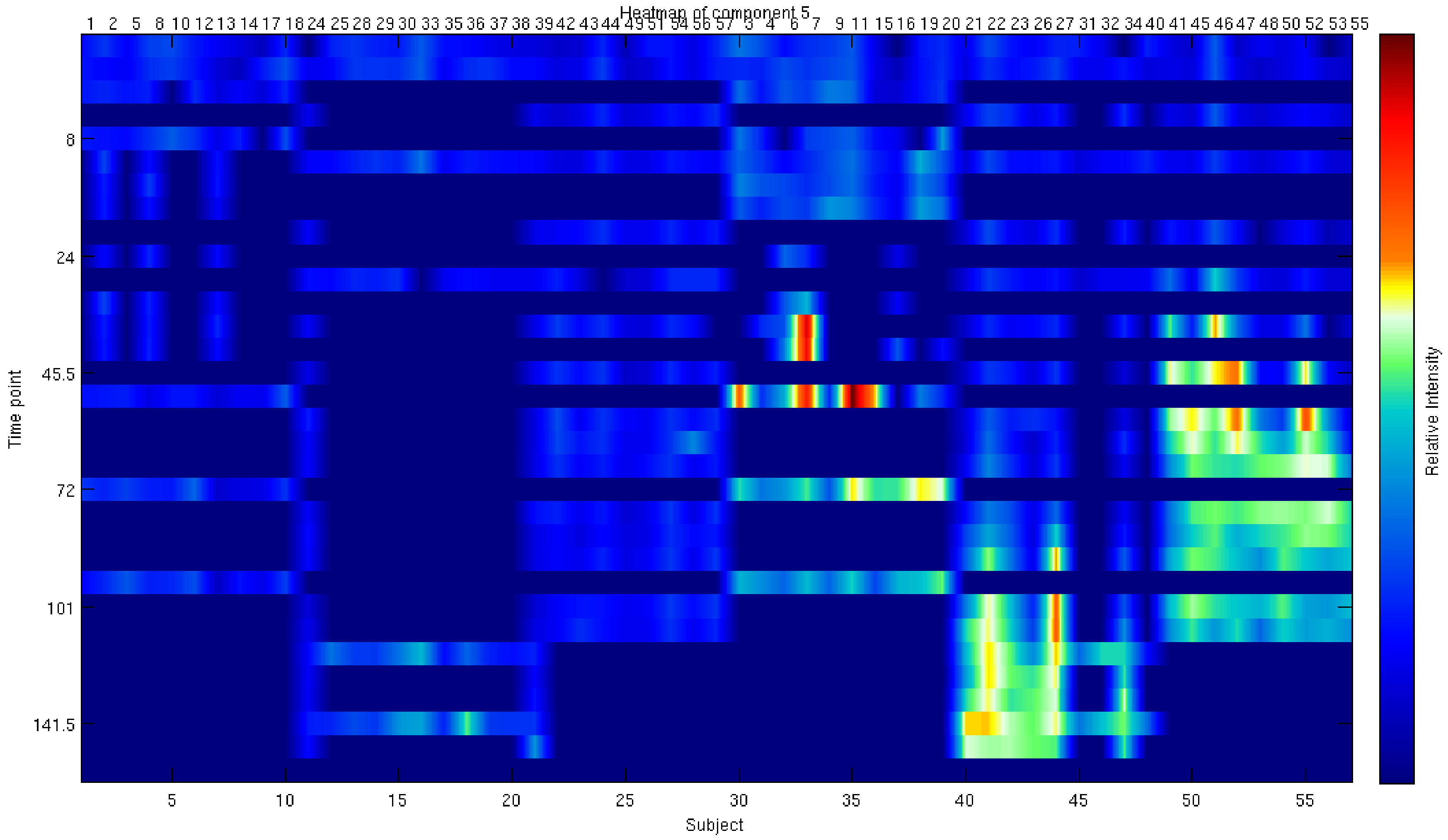
Blue: asx



Grp1	Grp2	Grp3
'ADAR'	'ANXA2'	'GBP1'
'ANPEP'	'BTG1'	'HERC5'
'FCER1G'	'FAIM3'	'IFI30'
'FFAR2'	'GIMAP4'	'IFI35'
'IFI30'	'IFI30'	'IFI44'
'IFITM1'	'IFITM1'	'IFI44L'
'IFITM3'	'MYD88'	'IFI6'
'MT2A'	'PLAC8'	'IFIH1'
'MX1'	'PNRC1'	'IFIT1'
'MX2'	'PSME1'	'IFIT2'
'MYD88'	'RPL19'	'IFIT3'
'PNRC1'	'S100A10'	'IFITM1'
'SECTM1'	'SATB1'	'IFITM3'
'TMEM140'	'SLC25A3'	'IRF7'
'WARS'	'SLC25A5'	'ISG15'
		'LAMP3'
		'LAP3'
		'LOC26010'
		'LY6E'
		'MT2A'
		'MX1'
		'MX2'
		'OAS1'
		'OAS2'
		'OAS3'
		'OASL'
		'PARP12'
		'PLAC8'
		'PLSCR1'
		'RSAD2'

Joint HRV/RSV/INF MCA

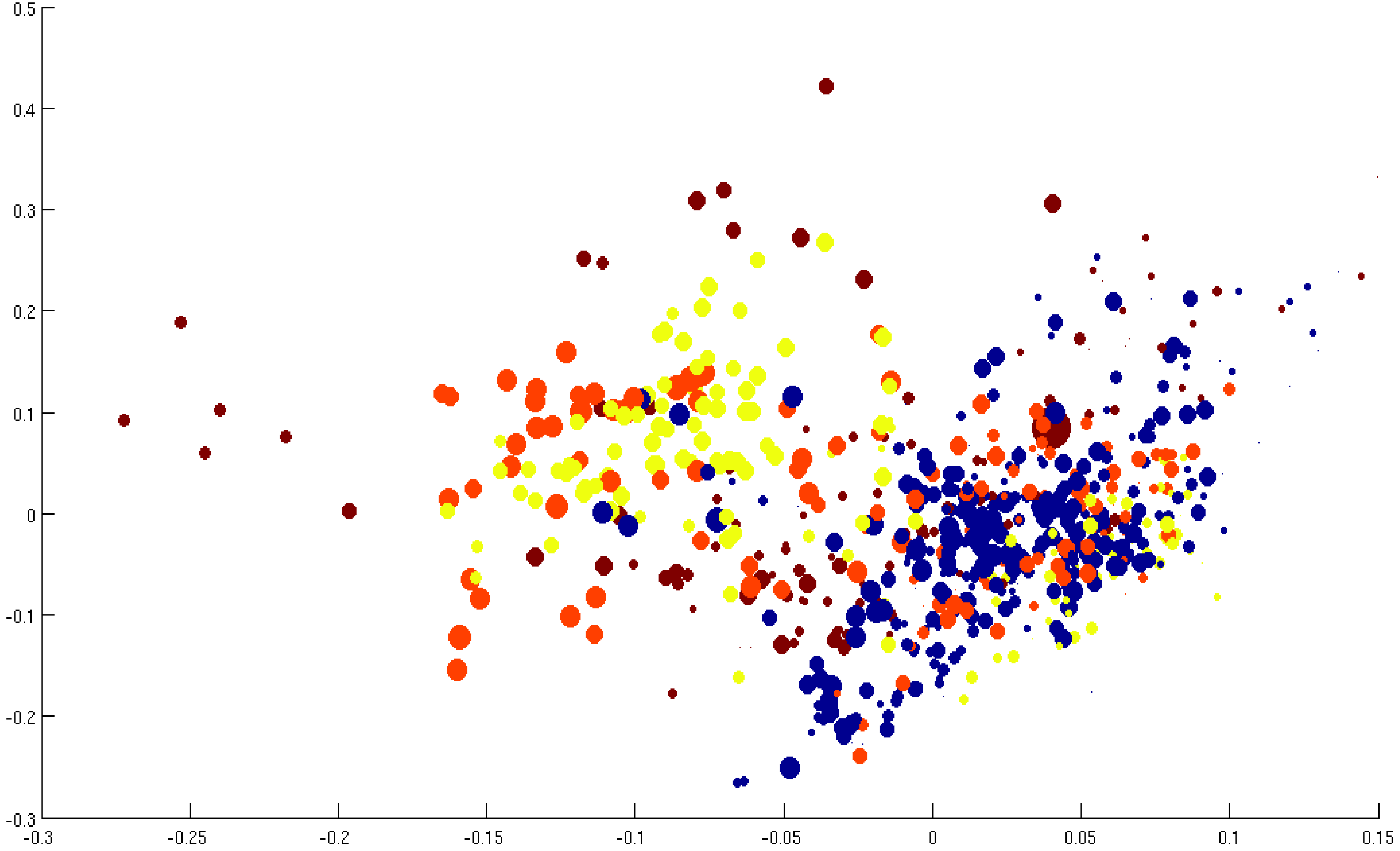
Factor 5



Joint HRV/RSV/INF MCA

5 factor chip scatter

Chip cloud after demixing: asx=B, sxHRV=red, sxRSV=orange, sxINF=yellow



All virus comparison: sx/asx d.e. early genes

Define “Early genes” as persistenty d.e. over 0.1T, 0.2T, 0.8T at FDR of 20%

Number of common HRV, RSV, INF at FDR of 20%: 0

HRV	RSV	INF	HI	HR	RI	HRI
44	5	471	1	0	0	0

Common HRV/INF gene: 'DTX4'

All virus comparison: sx/asx d.e. late genes

Define "Late genes" as persistenty d.e. over 0.8T to T
at FDR of 20%

Persistently d.e. genes in HRV, RSV, INF at FDR of 20%: 514

HRV	RSV	INF	HI	HR	RI	HRI
1770	1258	4629	1323	565	904	514

Persistently d.e. genes in HRV, RSV, INF at FDR of 1%: 24

HRV	RSV	INF	HI	HR	RI	HRI
169	111	2013	142	24	90	24

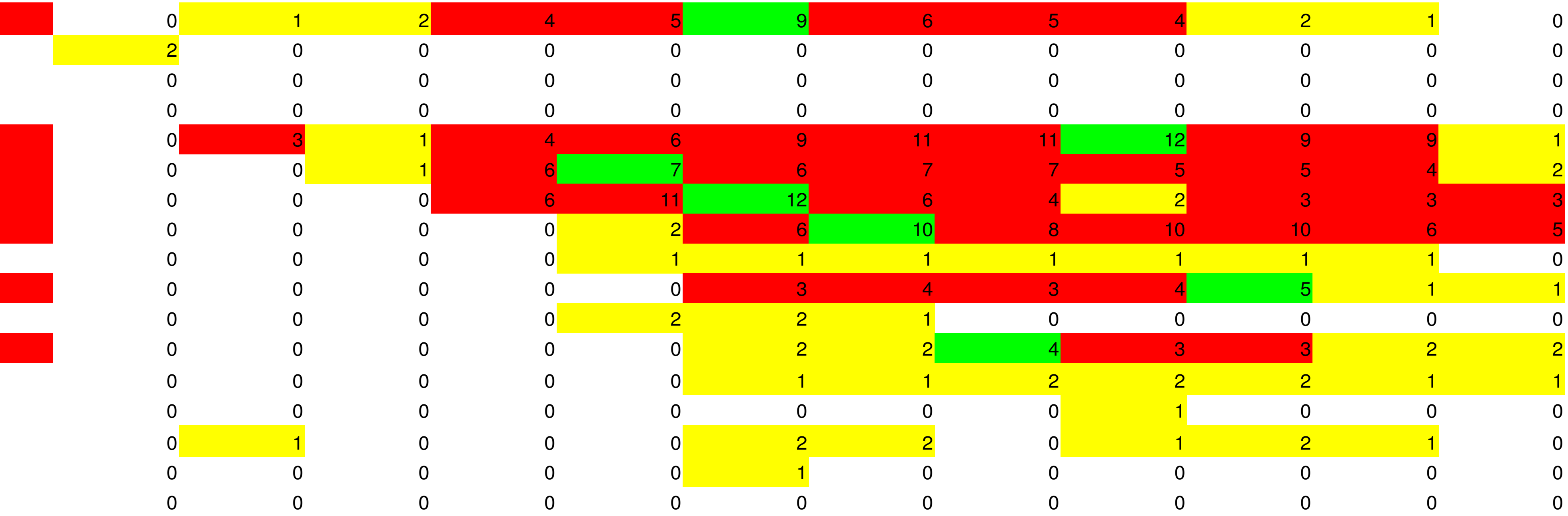
Highest 24 all-virus late genes

'C11orf75'
'C1QA'
'CENTA2'
'EPB41L3'
'FER1L3'
'FLVCR2'
'GBP1'
'IFI16'
'IFI30'
'IFI35'
'IFIT3'
'IRF7'
'MAFB'
'NPC2'
'PARP12'
'PLSCR1'
'PSMB9'
'SAMHD1'
'SCO2'
'SERPING1'
'SP110'
'TBC1D8'
'UBE2L6'
'WARS'

RSV symptom scores

blood draw time	ONCE			0/5/12hrs	21.5 hrs	29 hrs/36h	45.5hrs	53h 60h	69.5 h	77h 84h	93 h	101h 108h	117.5 h	125h 132 h	141.5 h		165.5h		
Subject	minus1	minus1bar	minus1bpr	0am	1am	1pm	2am	2pm	3am	3pm	4am	4pm	5am	5pm	6am	6pm	7am	7pm	CLASS
RSV001	0	0	0	0	0	1	0	0	0	0	2	3	5	6	8	11	10	10	YES
RSV002	0	0	0	0	1	2	2	1	2	3	12	13	18	16	20	20	13	13	YES
RSV003	0	0	0	0	1	0	0	0	0	1	2	2	2	2	6	4	4	4	YES
RSV004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NO
RSV005	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	No
RSV006	0	0	0	0	0	0	0	0	0	0	1	2	6	6	8	7	9	9	YES
RSV007	0	0	0	0	0	0	0	2	6	10	20	19	14	12	15	12	12	12	YES
RSV008	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	Yes
RSV009	0	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	No
RSV010	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	NO
RSV011	0	0	0	0	0	0	0	0	0	0	1	1	2	3	5	5	2	2	YES
RSV012	0	0	0	0	0	0	0	0	0	0	2	1	4	2	4	3	4	4	YES
RSV013	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	0	0	NO
RSV014	0	0	0	0	0	0	0	0	0	0	2	2	6	5	4	2	5	5	YES
RSV015	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	NO
RSV016	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	NO
RSV017	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NO
RSV018	0	0	0	0	0	1	0	1	1	1	2	1	0	0	0	0	1	1	NO
RSV019	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NO
RSV020	0	0	0	0	2	0	0	1	1	0	0	1	2	1	1	6	6	6	YES

INF Symptom Scores



0	1	2	4	5	9	6	5	4	2	1	0
2	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	3	1	4	6	9	11	11	12	9	9	1
0	0	1	6	7	6	7	7	5	5	4	2
0	0	0	6	11	12	6	4	2	3	3	3
0	0	0	0	2	6	10	8	10	10	6	5
0	0	0	0	1	1	1	1	1	1	1	0
0	0	0	0	0	3	4	3	4	5	1	1
0	0	0	0	2	2	1	0	0	0	0	0
0	0	0	0	0	2	2	4	3	3	2	2
0	0	0	0	0	1	1	2	2	2	1	1
0	0	0	0	0	0	0	0	1	0	0	0
0	1	0	0	0	2	2	0	1	2	1	0
0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0