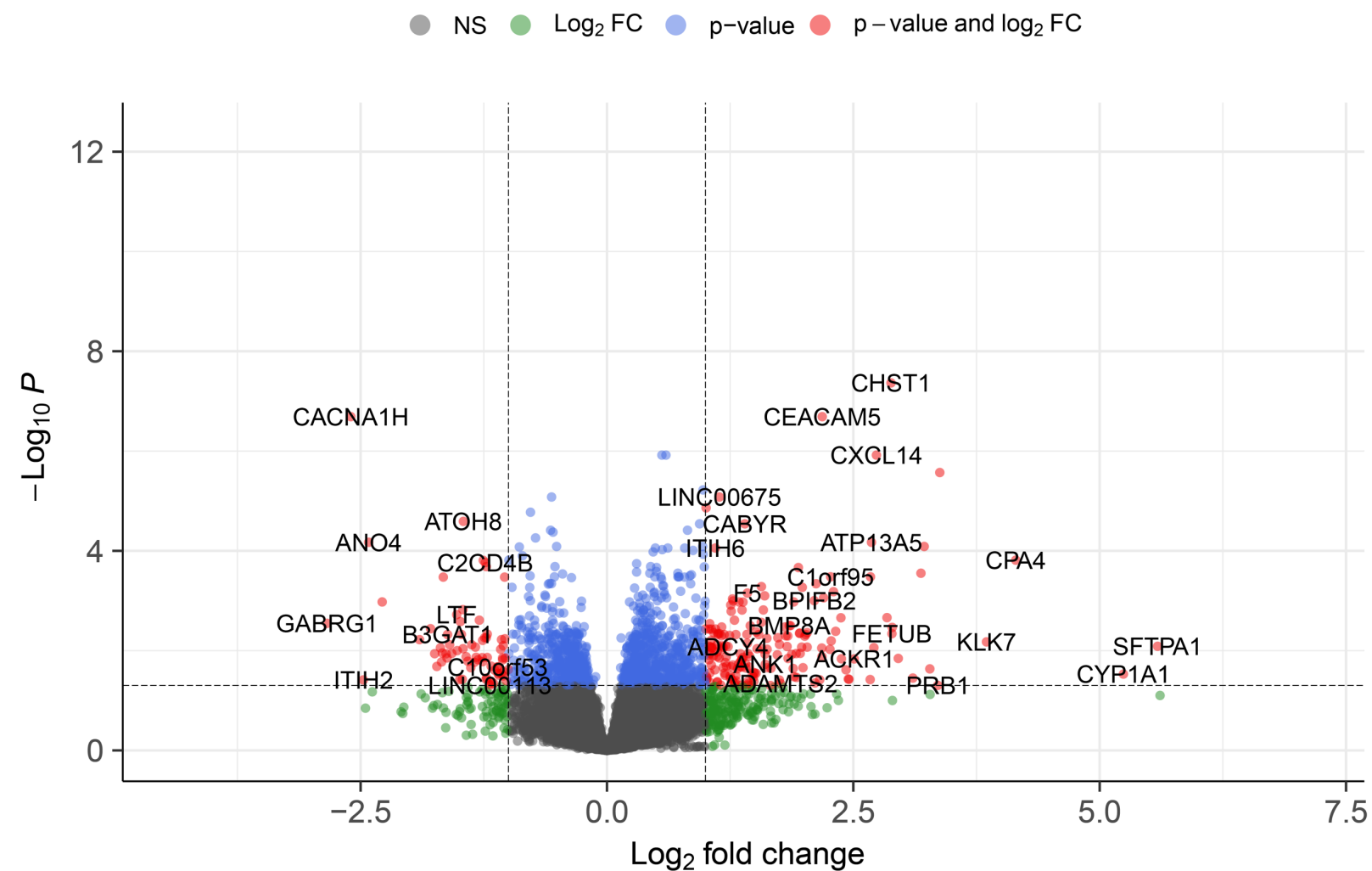


Supplemental Table 1: SARP Participant Demographics (n=152)

		Healthy Controls (n=41, 27%)	Mild to Moderate Asthma (n=48, 32%)	Severe Asthma (n=63, 41%)	p-value
Age (years)		45.8 (29.1 – 52.8)	35.7 (27.4 – 47.1)	50.1 (35.9. 53.6)	0.01
% Female		24(59%)	33 (65%)	38 (63%)	0.83
Race	Native American	0 (0%)	1 (2%)	0 (0%)	0.19
	Asian	4 (10%)	1 (2%)	1 (2%)	
	African American	7 (17%)	13 (25%)	20 (33%)	
	White	29 (71%)	36 (71%)	37 (62%)	
	Hispanic	1 (2%)	0 (0%)	1 (2%)	
	Other	0 (0%)	0 (0%)	1 (2%)	
BMI		26.9 (23.3 – 30.0)	27.7 (22.8 – 33.6)	31.0 (26.1 – 37.8)	0.01
FEV1 % predicted		99.4% (90.6 – 109.9%)	88.1% (76.2 – 102.0%)	64.0% (56.0 – 81.0%)	<0.01
Blood Absolute Eosinophil Count (cells/mL)		130 (80 - 180)	176 (87 - 303)	175 (133 – 306)	0.04
FeNO (ppb)		15 (10 - 18)	31 (17 – 47)	20 (12 – 45)	<0.01

Continuous Variables are shown as median (IQR). Categorical Variables as n (%). P-values for categorical variables are Kruskal-Wallis; p-values for categorical variables are Fisher's Exact



Supplemental Figure 1. Volcano plot showcasing differential expression results comparing MMA and SA using bulk RNA-seq data. Genes which are labeled or are represented by a red point have an absolute Log₂ fold change greater than 1 and a p-adjusted value of < .05. Log₂ fold change greater than 1 indicates greater expression in severe asthma patients. Plot generated using EnhancedVolcano package for R.

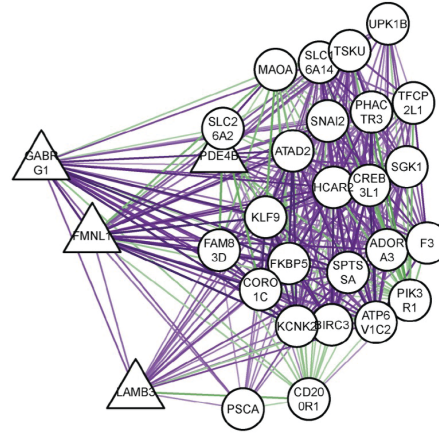
A**Severe Asthma vs Healthy Control**

Genes associated with Significant Latent Factor

LYPD2	PLD1M4	
SOC53	CIT	
DUSP5	CTTNBP2	
IRF4	CABYR	
HS6ST2	SCAMP5	
MAFF	CKAP2	
STK17B	ABCG2	FKBP5
ATP10A	GRIN3B	TFCP2L1
VEGFA	APBB1	PHACTR3
TNFRSF1B	NGEF	FAM83D
SH3TC2	PSENEN	SNAI2
LYNX1	CFI	KLF9
HPGD	LGALS4	HCAR2
PLCB4	PDE10A	SLC16A14
CXCL6	ATP6V1C2	ATAD2
CGREF1	RPS26	GABRG1
DUSP2	NEURL1B	PDE4B
RUNX3	SLC26A4	SPTSSA
GOS2	VWA3B	ADORA3
GPRC5A	FOXN4	PIK3R1
KCNK2	KCNK2	FMNL1
ATP6V1C2	KCNK6	MAOA
TSKU	UPK1B	PSCA
CORO1C	CORO1C	LAMB3
DPYSL3	TSKU	BIRC3
FA2H	FGFR1	CD200R1
MATN2	FA2H	KCNK2
CREB3L1	MATN2	UPK1B
S100A14	PLCB4	ATP6V1C2
KCNE3	F3	CORO1C
KRT24	CD200R1	TSKU
PLCH2	SGK1	FGFR1
BEST4	SLC26A2	F3
PDIAS	S100A14	SGK1
PDE10A	CREB3L1	SLC26A2
CEP152	BEST4	CREB3L1
GRIK2		

6 42 54

Significant Latent Factor

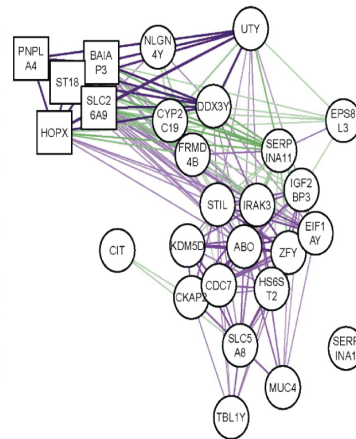
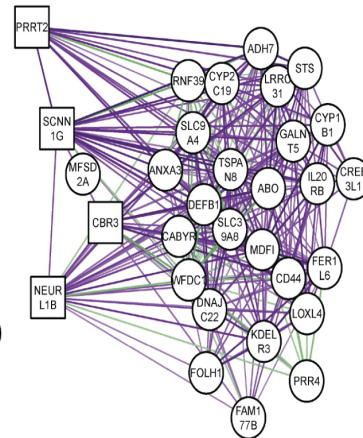
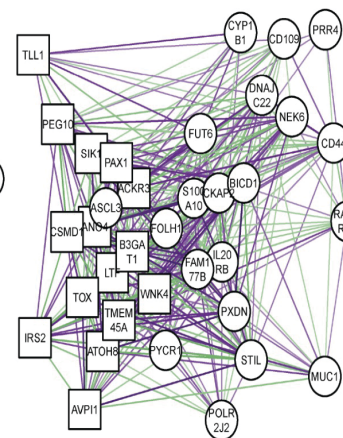
Z54**B****Severe Asthma vs Mild Moderate Asthma**

Genes associated with Significant Latent Factor

DDX3Y				
RPS4Y1				
UTY				
KDM5D	ADH7	DACH1	ITF	TIMP1
ZFY	LOXL4	C2orf72	ATOH8	ITIH6
E1F1AY	RNF39	PRKAR2B	SIK1	NGEF
NLGN4Y	NEURL1B	SLC16A12	IRS2	SPATA32
TBL1Y	GALNT5	F2R	ANO4	SMPDL3A
TMSB4Y	LRRC31	CD44	AVP11	CHST1
ZNF718	SLC39A8	TOX	ACSM1	SMIM22
CEL	KDELR3	GSTA3	S100A10	DDX60L
ZNF208	PRRT2	CD44	IL20RB	CHPT1
BAIAP3	DEFB1	SLC13A3	GIPC2	AVPR1A
ST18	SLC9A4	FA2H	PAX1	AHNAK2
FRMD4B	CBR3	FADS1	B3GAT1	TLR3
PNPLA4	FOLH1	ADIRF	WNK4	PTGES
RBPD1	CABYR	ERAP2	PEG10	IFIT5
HOPX	SCN1G	WFDCL1	REEP1	SERPINB4
SERPINA1	STS	CATSPERB	CSMD1	NT5E
SLC26A9	MDF1	CCDC81	TOX	CYP2F1
CKAP2	CREB3L1	GALNT5	TMEM45A	KLHL14
MUC1	FER1L6	AKAP12	RASGRP3	VNN1
CDC7	TSPAN8	FUT9	CD44	CD44
EPS8L3	CD44	DNAJC22	DNAJC22	DNAJC22
CHST1	DNAJC22	CABYR	MUC1	CABYR
PXDN	MUC1	MUC1	PXDN	CKAP2
CYP2C19	FAM177B	CHST1	FAM177B	MUC1
GAD1	CYP2C19	PXDN	CYP2C19	CDC7
ABO	GAD1	FAM177B	PRR4	FAM177B
ANXA3	ABO	KDELR3	BICD1	CYP2C19
SERPINA11	ANXA3	CYP1B1	STIL	GAD1
SLC5A8	ANXA3	ABO	ASCL3	KDELR3
HS6ST2	PRR4	ASCL3	CD109	ABO
CLCA4	MFSD2A	CYP1B1	IGF2BP3	CYP1B1
MUC4	SLC5A8	PRR4	CD109	SLC5A8
CIT	WFDCL1	STIL	NEK6	MFSD2A
IRAK3	STIL	ASCL3	FUT6	ADORA3
ASCL3	ASCL3	IL20RB	POLR2J2	MUC4
IGF2BP3	IL20RB		FOLH1	CIT

3 8 30 56 63

Significant Latent Factor

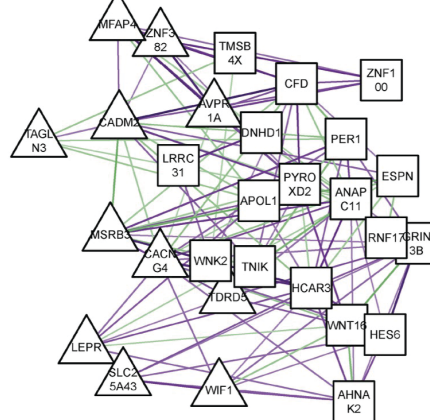
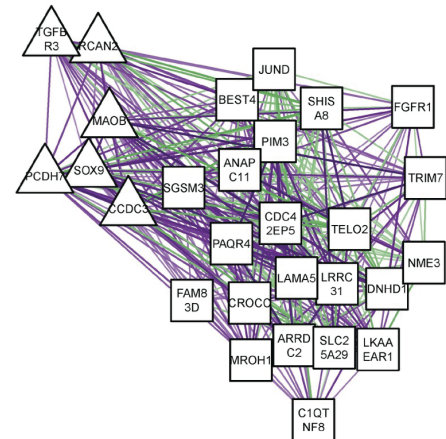
Z3**Z8****Z56****C****Mild-Moderate Asthma vs Healthy Control**

Genes associated with Significant Latent Factor

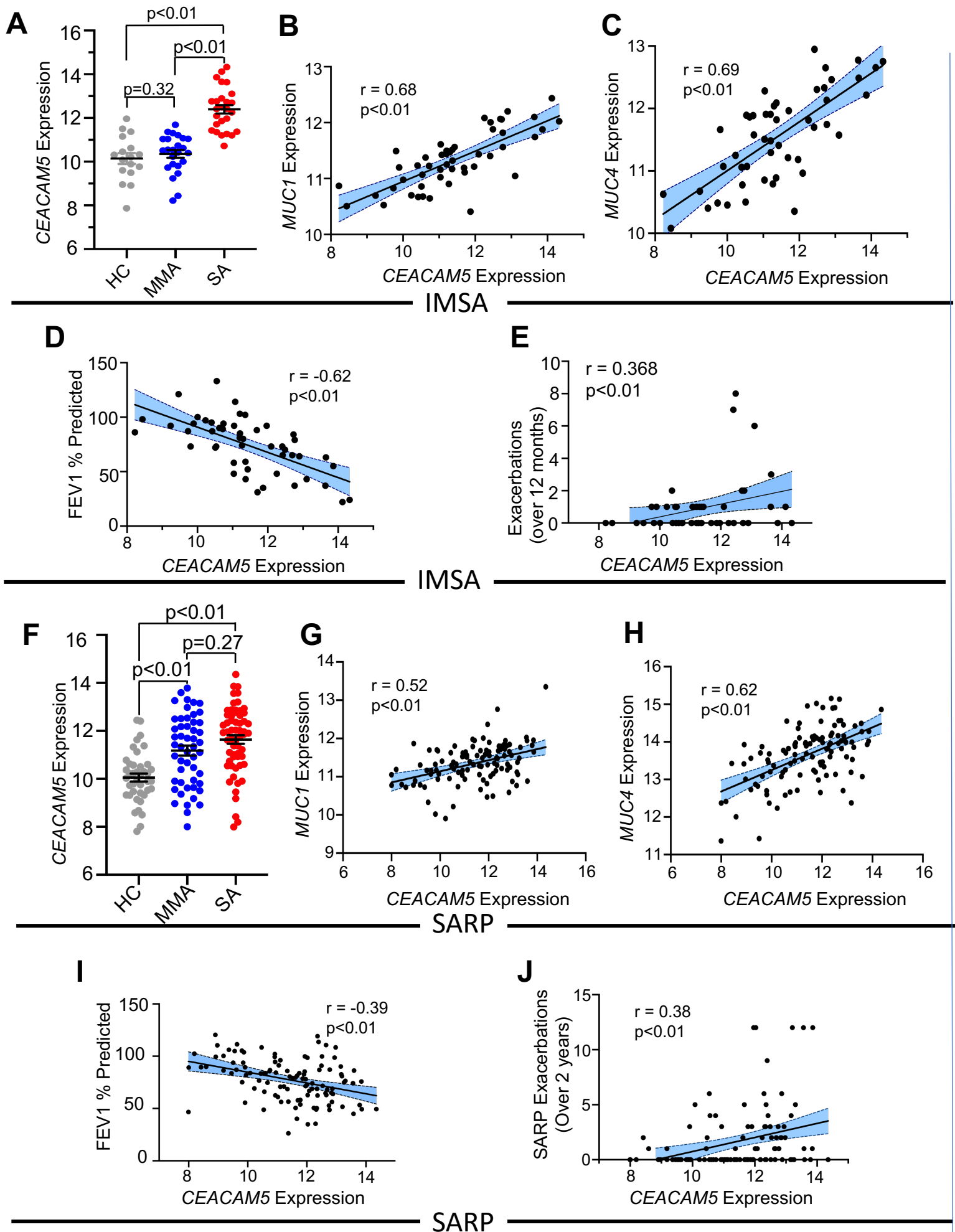
CORO2B	TAGLN3	
NR2F1	LEPR	
OLFML2A	CADM2	
SLC6A14	MSRB3	
MDGA1	WIF1	
SPOCK3	ZNF382	
PYROXD2	PML	
PCDH20	HILA-DQB1	
DTNA	ZNF100	
C8orf48	AVPR1A	
MMP1	CCNO	
C5AR1	PDZD4	
GDA	SPARCL1	
ARRDC3	SLC25A43	
TUBA1A	TDRD5	
ATP6V0C	COL14A1	
MMRN1	CACNG4	
ZNF273	MFAP4	
NELL2	WNK2	
WNT2	CFD	
PIK3R1	AHNAK2	
ARRDC2	DNHD1	
CROCC	SLC25A29	
SLC25A29	APOL1	
APOL1	ANAPC11	
NME3	C1QTNF8	
HCAR3	TRIM7	
ALOX15B	GRIN3B	
ERRF1	PER1	
PNPLA7	ESPN	
PLEKHG4B	LRRRC31	
AKR1C1	NBL1	
ZNF846	PAQR4	
GPR34	PYROXD2	
LDLR	HCAR3	
SLC31A2	ALOX15B	
SMC1A	TSPAN8	
	METRN	
	GPR34	

15 26 33 67

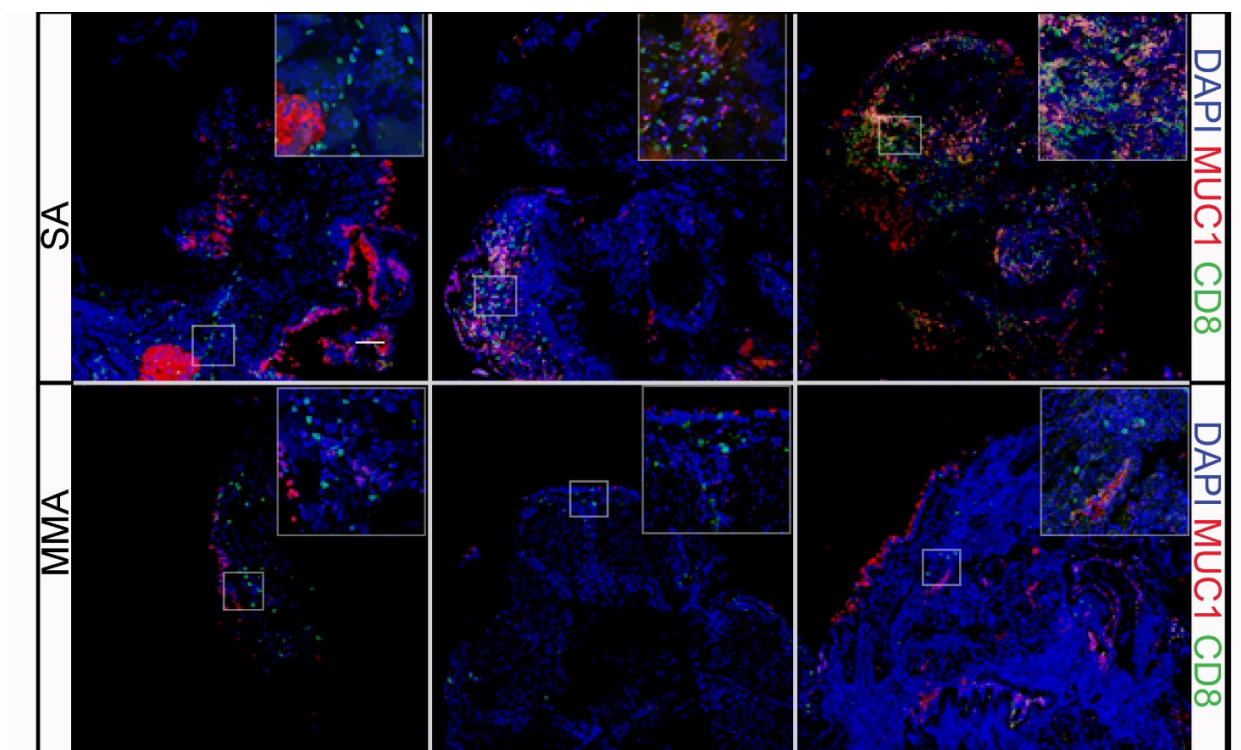
Significant Latent Factor

Z26**Z67**

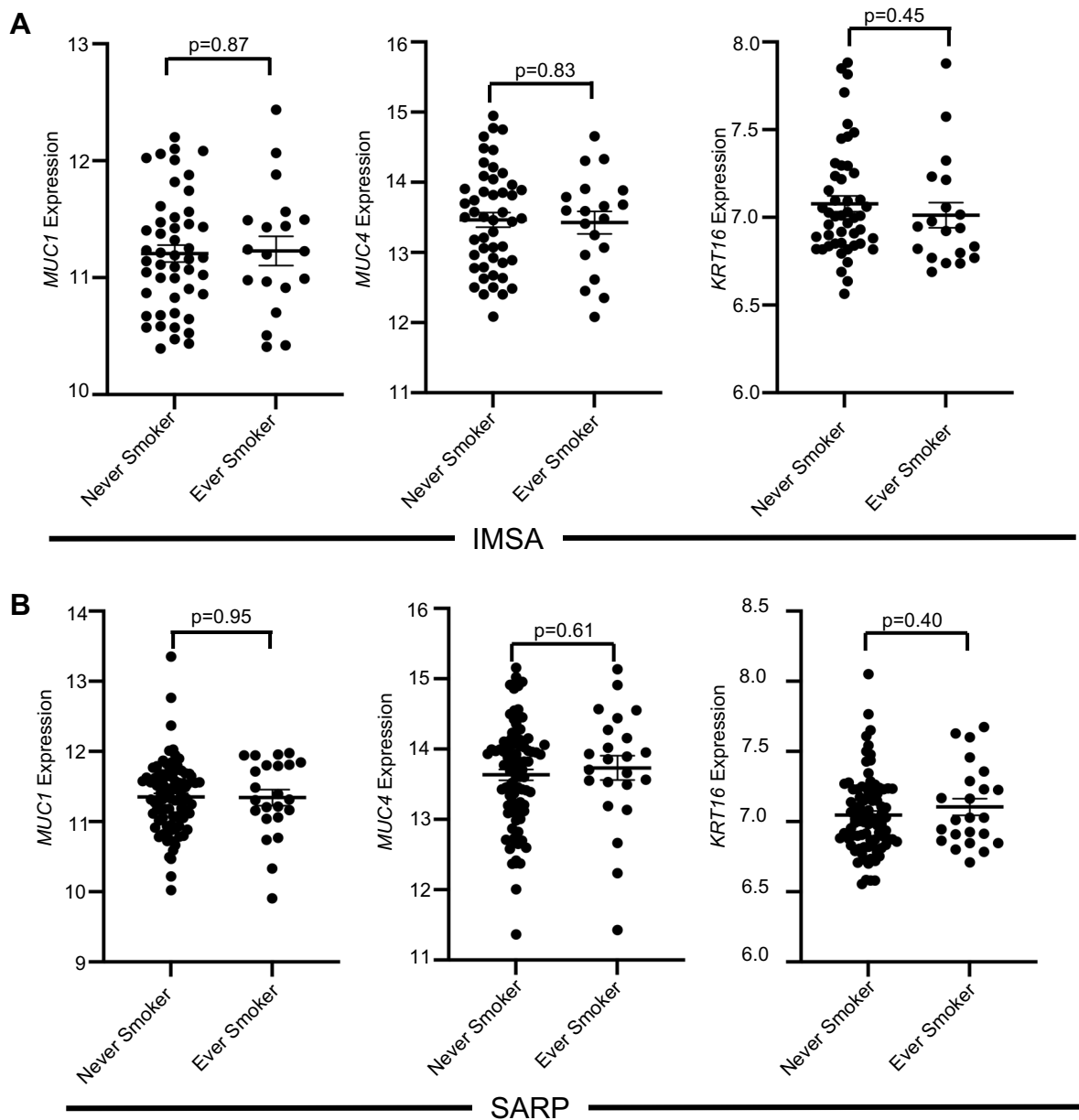
Supplemental Figure 2. Latent factor gene lists for each significant standalone latent factor as discovered by SLIDE and used in the analyses. Red features correspond to fold change in the direction of the higher severity for a given pairwise comparison, and vice versa for blue. To the right are correlation networks for latent factors not included in Figure 3 (labeled as Zx). Triangle nodes represent a logfc indicating greater expression in HC patients, whereas squares indicate greater expression in MMA and circles indicate the same for SA. Purple edges are positive correlation, and green edges are negative correlation. Edges are only present if they have a strength of $R > |.4|$. Edge thickness indicates strength of correlation. Nodes only present if they passed a loading and AUC threshold.



Supplemental Figure 3. Expression of *CEACAM5* gene and relationship with *MUC1* and *MUC4* gene expression and clinical measures in the two asthma cohorts. Shown are (A) *CEACAM5* RNA expression in bulk RNA-seq data of airway brushings by asthma severity and plotted against (B) *MUC1* and (C) *MUC4* expression, and asthma clinical correlations (D) FEV1% predicted (E) number of exacerbations in the 12-month period of cohort enrollment. (F-I) Similar data plotted using RNA-seq data of airway brushings in SARP. (J) Number of cumulative prospective exacerbations occurring from baseline visit through year 3 in SARP was plotted against *CEACAM5* expression. Data plotted as Mean + SEM. Spearman's nonparametric correlations, linear regression (solid line) with 95% CI (shaded blue) shown on each scatter plot. Kruskal-Wallis with Dunn's post-hoc testing, Spearman's nonparametric correlation.



Supplemental Figure 4. Representative immunofluorescence images of MUC1 and CD8 proteins (inset). Images were derived from human endobronchial biopsies obtained from SA and MMA patients. Scale 200 μ m.



Supplemental Figure 5. Prior smoking history does not affect MUC1, MUC4 or KRT16 expression. Expression of MUC1, MUC4 and KRT16 compared in the IMSA and SARP cohorts between participants who reported never smoking (Never Smoker) and participants who had previously smoked but quit (Ever Smoker). Student's t-test with Welch's correction.