

Transcriptomic and methylomic features in asthmatic and nonasthmatic twins

To the editor,

Asthma is the most prevalent chronic lung inflammatory disorder characterized by reversible airflow obstruction, affecting 358 million people worldwide,¹ with aggravating factors like obesity, attention-deficit/hyperactivity disorder, socioeconomic status like poor healthcare affordability and facility, smoking and alcohol intake.² RNA sequencing (RNA-seq) in atopic asthma,³ childhood asthma⁴ and adult-onset severe asthma⁵ has improved our understanding of cellular and molecular pathways involved. Genome-wide DNA methylation studies investigating 5-methylcytosine in CpG sites have linked methylation in lung cells with asthma endotypes and genetic risk.⁶ Studies focusing on integration of genomics and interactomes have also been performed in asthma.⁷ However, this multifaceted asthma phenotype together with genetic heterogeneity and environmental influences makes it challenging to fully understand the features that trigger and influence asthma development and progression. Towards addressing these challenges, we investigated transcriptomic and methylomic data in a twin cohort of asthmatic and nonasthmatic individuals.

In this exploratory study, the individuals were identified as asthmatics (GINA score ≥ 1) according to the Global Strategy for Asthma Management and Prevention guidelines (<http://ginasthma.org>). Participants with any viral or bacterial infections or immune disorder were excluded. All the participants with asthma were on a low dose of inhaled corticosteroid and rescue inhaler (albuterol) only. We collected PBMC from 16 female monozygotic twin pairs (5 concordant asthmatic pairs; 8 concordant nonasthmatic pairs and 3 pairs discordant for asthma) among which 13 individuals were asthmatic (GINA score 1 $n = 4$; 2 $n = 4$; 3 $n = 5$, Table 1). The transcriptomic RNA-seq was done on all 16 twin pairs, and the methylomic whole-genome bisulfite sequencing was done on 8 twin pairs (3 concordant asthmatic pairs; 5 concordant nonasthmatic pairs) wherein 6 individuals were asthmatics (GINA score 1 $n = 1$; 2 $n = 1$; 3 $n = 4$). Methods are detailed in Data S1.

Unsupervised hierarchical clustering of transcriptomic profiles shows 9/16 twin pairs clustering together, while only 2/8 twin pairs cluster together in methylomic profiles. 2/3 twin pairs discordant for asthma clustered together in transcriptomic profiles reflects similarity based on twinning (Figure 1A). The demographics age, BMI and weight are detailed in Table 1.

The transcriptomic analysis of total 1665 annotated genes (File S1) yielded 117 differentially expressed (46 up-regulated; 71 down-regulated, Table S1) genes in comparison with asthmatics over

nonasthmatics (Figure 1B). The expression profiles of the top 10 differentially expressed genes are represented in Figure S2. The KEGG pathway enrichment analysis of these differentially expressed genes yielded pathways such as ribosome, glycosaminoglycan biosynthesis—chondroitin sulphate and SNARE interactions in vesicular transport to be significantly enriched (adj P -value $< .05$, Figure S1) which is linked with eosinophilic dysregulation in asthma.⁸ Among 14562 CpGs identified by methylomics, 96 were found to be differentially methylated between asthmatics and nonasthmatics (Figure 1C, Table S2). Among the dysregulated features in our cohort, the genes *CTNNA3*, *SEMA3D*, *CDHR3*, *TGFB1*, *IRF3*, *CALM3*, *FCRL1*, *IFT22* and *IL4RA*, and the CpGs on *MUC20*, *SLC12A7*, *GSE1*, *DRAIC* and *SC5D* have been reported in asthma.^{9,10} The identification of functionally uncharacterized *ENSG00000231345* and *ENSG00000226143* genes in our cohort requires further investigations.

The integration of transcriptomic and methylomic significant features from the 8 twin pairs with both RNA-seq and methyl-seq assays was done using regularized canonical correlation analysis (rCCA) to identify interactions between features. Among the total 13873 interactions, strong correlations (absolute $r > .7$; positive = 813 and negative = 2528, Figure S3) were noted. For visuality, only the top 100 interactions (positive = 63, negative = 37) are shown in Figure 1D. The expression of *CALM3* was found to be correlated with the CpG methylation status of *MUC20* ($r = -.7421$) and *GSE1* ($r = -.7240$), which were all previously reported to be asthma related.¹⁰ Similarly, the expression of *IFT22* gene was found to be strongly correlated with the CpG methylation status of *DRAIC* ($r = .8286$) in our cohort. The CpG methylation status of *ATXN2* ($r = -.7188$) and *MAD1L1* ($r = -.7163$) was strongly correlated with *CALM3* and *MYL6* expression ($r = -.7124$). The previously unreported *ZCCHC24* ($r = .8554$) and *EVX1-AS* ($r = .8424$) correlation with the CpG methylation status of *MUC20* requires further investigations in the context of asthma. The top 50 strongest correlated features from rCCA from our cohort are reported in Table S3.

The dysregulated features presented here reflect the differences between nonasthmatics and a range of asthmatics (mild to moderate) which is different from most prior studies that have focused on severe or acute asthma (GINA score 3+). In this study, the PBMC analysed represents an overall disease rather than a cell type-specific phenotype, which needs to be considered while interpreting dissimilarities from previous studies. Although monozygotic twins essentially share their genetic makeup and, if living together, also share much of their early life environmental

	Asthma Median [1st-3rd quartile] (N = 13)	Nonasthma Median [1st-3rd quartile] (N = 19)	P-value ^a
Age (y)	52 [46-62]	48 [31-58]	.691
Weight (kg) (N = 13)	75.4 [66.70-93.40]	62.6 [54.35-62.95]	.746
BMI (kg/m ²) (N = 13)	28.49 [22.25-32.31]	24.72 [21.48-26.37]	.962
GINA score	1 = 4; 2 = 4, 3 = 5	0 = 19	.019

^aBy generalized linear model with t test.

TABLE 1 Demographic information of participants

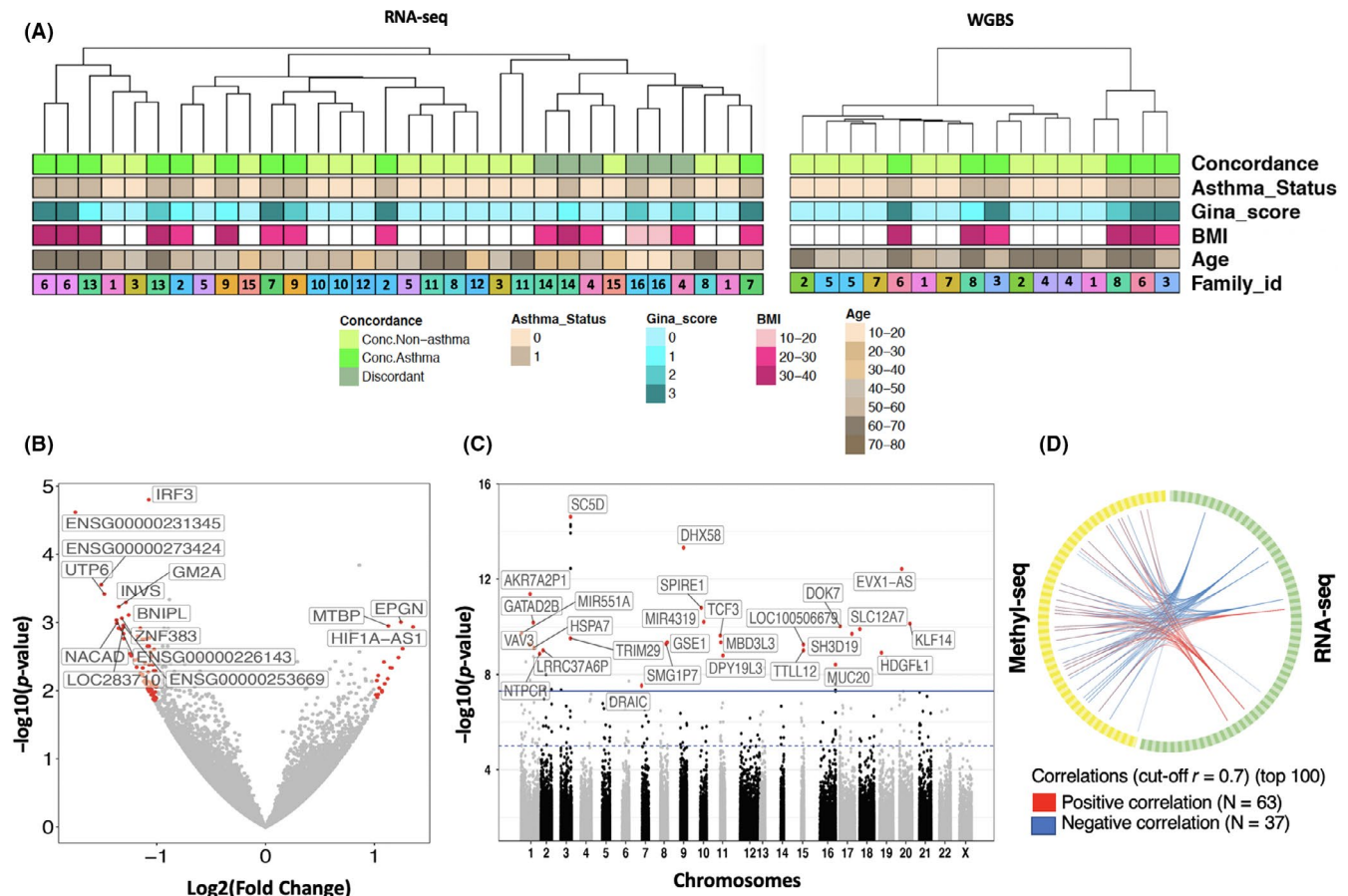


FIGURE 1 A, Unsupervised hierarchical clustering of RNA-seq and methyl-seq samples with complete linkage. B, Volcano plot of differentially expressed genes (grey = nonsignificant genes, blue = significantly genes with absolute $\log_2$ (fold change) ≥ 1 and FDR-adjusted P -value $< .01$, top 30 gene annotations are visualized). C, Manhattan plot of CpGs across the chromosomes. D, Circos plot showing the relation among all the features with strong Spearman's correlations (absolute $r > .7$; top 100 only) represented as links (red = positive; blue = negative). (Extended figure legends can be found in online Data S2)

exposure, we did not record information about twins living together. Twin pairs discordant for asthma would be ideal for exploration of genetic vs environmental influence on asthma. But, with only 3 twin pairs discordant for asthma (mild to moderate; GINA scores of 1 $n = 1$; and 2 $n = 2$) in our cohort, the sample size is too small to draw statistical conclusions. The inclusion of concordant twins as specific population but without sufficient number of discordant twin pairs was not relevant for the current study. Hence, increasing the number of twin pairs discordant for asthma

is crucial for such future investigations. The analysis stratified by different GINA scores was not possible due to small sample size; hence, only asthmatic (GINA score ≥ 1) and nonasthmatics were compared. Larger studies including male samples and sufficient number of participants with different GINA scores are needed to study specific asthma phenotypes which would greatly benefit the asthma-specific signature identification and evaluation across asthma severities. Furthermore, the biological validation of identified differential features would strengthen our findings.

In conclusion, this exploratory study investigating transcriptomic and methylomic profiles in a twin cohort has enabled the identification of features of asthma vs nonasthmatics. While our specific cohort adds insight to our understanding of the phenotypes of mild-to-moderate asthma, the recognition of previously known *CALM3*, *DRAIC*, *IFT22*, *MUC20* and *SC5D* asthma features in our cohort supports our findings. Additionally, the multi-omic integration approach enables correlation-based identification of candidate features and their interactions not yet known in asthma disease.

ACKNOWLEDGMENTS

The authors thank the participants and the SRI and Washington State Twin Registries.

CONFLICT OF INTEREST

ESB and FB are employees of Janssen Research and Development LLC, PA, USA. RNB is an employee of Genentech, CA, USA. CM is an employee of BioProcessing Solutions, NJ, USA. KCN has received funding from, is currently funded by, or cofounded the following: the NIH, Food Allergy Research & Education, End Allergies Together, Before Brands, Alladapt Immunotherapeutics, Adare Pharmaceuticals, AstraZeneca, Novartis, Genentech, Astellas, DBV Technologies, ForTra, Aimmune Therapeutics, Regeneron, Nestle, Sanofi, ImmuneWorks, Cour, and the Environmental Protection Agency. All other authors declare that they have no relevant conflicts of interest.

FUNDING INFORMATION

This research is funded by NIH NHLBI R01HL118612, NIEHS R21ES022153-02 and part by Janssen Biotech, Inc, and Sean N. Parker Center for Allergy and Asthma Research.

ETHICAL APPROVAL

All the biological samples were collected and studied under Stanford approved protocol (IRB 8629) in full compliance with CFR and GCP guidelines.

Gopal Krishna R. Dhondalay¹ 

Bryan Bunning¹

Rebecca N. Bauer²

Elliot S. Barnathan³

Christopher Maniscalco⁴

Frédéric Baribaud³

Kari C. Nadeau¹ 

Sandra Andorf¹ 

¹Sean N. Parker Center for Allergy and Asthma Research, Stanford University School of Medicine, Stanford, CA, USA

²Genentech, South San Francisco, CA, USA

³Janssen Research & Development LLC, Spring House, PA, USA

⁴BioProcessing Solutions, Piscataway, NJ, USA

Correspondence

Sandra Andorf, Sean N. Parker Center for Allergy and Asthma Research, Stanford University School of Medicine, Stanford, CA, USA.

Email: andorf@stanford.edu

ORCID

Gopal Krishna R. Dhondalay  <https://orcid.org/0000-0002-1020-0175>

Kari C. Nadeau  <https://orcid.org/0000-0002-2146-2955>

Sandra Andorf  <https://orcid.org/0000-0002-3093-2568>

REFERENCES

1. Burnett R, Chen H, Szyszkowicz M, et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc Natl Acad Sci USA*. 2018;115(38):201803222.
2. Papi A, Brightling C, Pedersen SE, Reddel HK. Asthma. *Lancet*. 2018;391(10122):783-800.
3. Yick CY, Zwiderman AH, Kunst PW, et al. Gene expression profiling of laser microdissected airway smooth muscle tissue in asthma and atopy. *Allergy Eur J Allergy Clin Immunol*. 2014;69(9):1233-1240.
4. Yeh YL, Su MW, Chiang BL, Yang YH, Tsai CH, Lee YL. Genetic profiles of transcriptomic clusters of childhood asthma determine specific severe subtype. *Clin Exp Allergy*. 2018;48(9):1164-1172.
5. Hekking PWP, Loza MJ, Pavlidis S, et al. Pathway discovery using transcriptomic profiles in adult-onset severe asthma. *J Allergy Clin Immunol*. 2018;141(4):1280-1290.
6. Nicodemus-Johnson J, Myers RA, Sakabe NJ, et al. DNA methylation in lung cells is associated with asthma endotypes and genetic risk. *JCI Insight*. 2016;1(20):e90151.
7. Hwang S, Son SW, Kim SC, Kim YJ, Jeong H, Lee D. A protein interaction network associated with asthma. *J Theor Biol*. 2008;252(4):722-731.
8. Adamko D, Odemuyiwa SO, Moqbel R. The eosinophil as a therapeutic target in asthma: beginning of the end, or end of the beginning? *Curr Opin Pharmacol*. 2003;3(3):227-232.
9. Park H, Tantisira KG. Genetic signatures of asthma exacerbation. *Allergy Asthma Immunol Res*. 2017;9(3):191-199.
10. Forno E, Wang T, Qi C, et al. DNA methylation in nasal epithelium, atopy, and atopic asthma in children: a genome-wide study. *Lancet Respir Med*. 2019;7(4):336-346.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

SUPPLEMENTARY INFORMATION

Data S1: Methods

The peripheral blood mononuclear cells were isolated from individuals and the total RNA was extracted using RNeasy kit (Qiagen Inc., Valencia, CA) according to manufacturer's protocol. RNA quality and quantity were analyzed with 2100 Bioanalyzer using Agilent RNA 600 Nano kit (Agilent Technologies Inc., Palo Alto, CA).

For the RNA-seq, the libraries were sequenced on Illumina NextSeq500 instrument using 2x100bp paired-end reads, following the manufacturer's protocols. An in-house modified ENCODE pipeline (www.encodeproject.org) with GENCODE v24 primary assembly annotation (<https://www.encodeproject.org/references/ENCSR884DHJ/>) was used for mapping the reads against the hg19 reference genome using the TopHat2 aligner¹. For downstream analysis only annotated genes from Ensemble were used, but in-silico predicted genes, pseudo-genes and non-coding RNAs were excluded from analysis. Additionally, the genes whose expression was not measured in at least 80% of the samples in our cohort were removed. After these steps, a total of 1665 genes were obtained. The Bioconductor (<https://www.bioconductor.org>) package *DESeq2*² was used for differential expression analysis of log₂ transformed data between asthmatics and non-asthmatics by adjusting for age, BMI, weight and twin pair. Any gene passing the threshold of absolute log₂(fold change) ≥ 1 and Benjamini and Hochberg FDR-adjusted *P*-value < 0.01 was considered to be significantly differentially expressed. For pathway enrichment analysis, the *gage* and *kegg* packages from Bioconductor are used, wherein pathways with an adjusted *P*-value of < 0.1 were considered significant.

The whole-genome bisulfite sequencing was done on directional bisulfite-converted libraries using Illumina NextSeq500 instrument with 2x75bp paired-end reads, following the manufacturer's protocols. The reads were analyzed by mapping them against the hg19 reference genome using the Bowtie2 aligner³ and Bismark⁴ for extraction of methylation sites. The *methyKit* R package⁵ was used for finding differentially methylated (absolute log₂(fold change) ≥ 1 and Bonferroni-corrected *P*-value < 0.01) sites by adjusting for age, BMI, weight and twin pair.

The multi-omic integration of features from transcriptomic and methylomic data for asthma status was done using the *mixOmics* R package⁶. Briefly, the median expression values of

these features (differentially expressed genes and CpGs) from asthmatic and non-asthmatic individuals was used for inter-omic correlation using regularized canonical correlation analysis (rCCA)⁷ using sparse partial least square discriminant analysis with centroid method and a cut-off of absolute Spearman's correlation $r > 0.7$. The bi-directional and self-correlations were removed, and the results were visualized as a circos plot.

References:

1. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 2013;14(4):R36. doi:10.1186/gb-2013-14-4-r36
2. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):1-21. doi:10.1186/s13059-014-0550-8
3. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357-359. doi:10.1038/nmeth.1923
4. Krueger F, Andrews SR. Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics.* 2011;27(11):1571-1572. doi:10.1093/bioinformatics/btr167
5. Akalin A, Kormaksson M, Li S, et al. MethylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 2012;13(10):R87. doi:10.1186/gb-2012-13-10-R87
6. Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLOS Comput Biol.* 2017;13(11):e1005752. doi:10.1371/journal.pcbi.1005752
7. Golugula A, Lee G, Master SR, Feldman MD, Tomaszewski JE, Madabhushi A. Supervised regularized canonical correlation analysis: Integrating histologic and proteomic data for predicting biochemical failures. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS.* 2011:6434-6437. doi:10.1109/IEMBS.2011.6091588

Data S2: Extended figure legends:

- (A) Unsupervised hierarchical clustering with complete linkage for RNA-seq and methyl-seq samples. Samples are annotated for asthma status, concordance for asthma, GINA score, BMI, weight in kg, age in years, and twinning. White represents unavailable data. The individuals of a twin pair are coded with the same number.
- (B) Volcano plot of differentially expressed genes represented as dots. Grey dots represent the non-significant genes. Red dots are the differentially expressed genes above the threshold of absolute $\log_2(\text{Fold Change}) \geq 1$ and FDR-adjusted P -value < 0.01 . The top 15 annotated significantly differentially expressed genes are represented for visualization.
- (C) Manhattan plot of CpG sites across the chromosomes is shown. The Bonferroni-corrected optimal P -value threshold for genome-wide significance of $< 1e-05$ is represented as blue dashed line and significance threshold P -value $< 5e-08$ is represented as blue solid line. The CpG sites with highest significance per gene are highlighted in red and annotated with the gene name. For the visualization only top 25 genes with highest adjusted P -values are represented. CpG on *MUC20* and *DRAIC* genes are represented for visualization.
- (D) Integration of transcriptomic and methylomic data. The circos plot of features from transcriptomic data (green) and methylomic data (yellow) across asthmatic and non-asthmatics samples. The top 100 strongest correlations between the features are represented as links (red = positive; blue = negative). A larger version with all strong correlation (absolute $r > 0.7$) can be found in supplementary Figure S3.

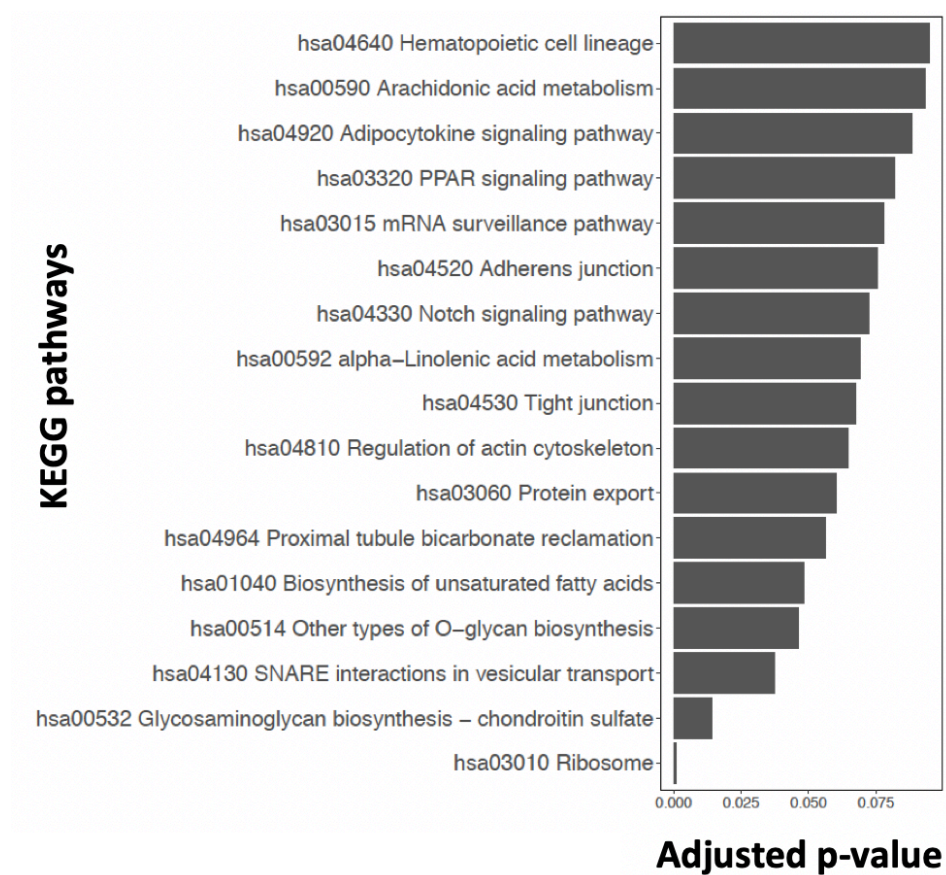


Figure S1: Pathway enrichment analysis

Pathway enrichment analysis of KEGG pathways of 117 differentially expressed genes across asthmatics and non-asthmatics.

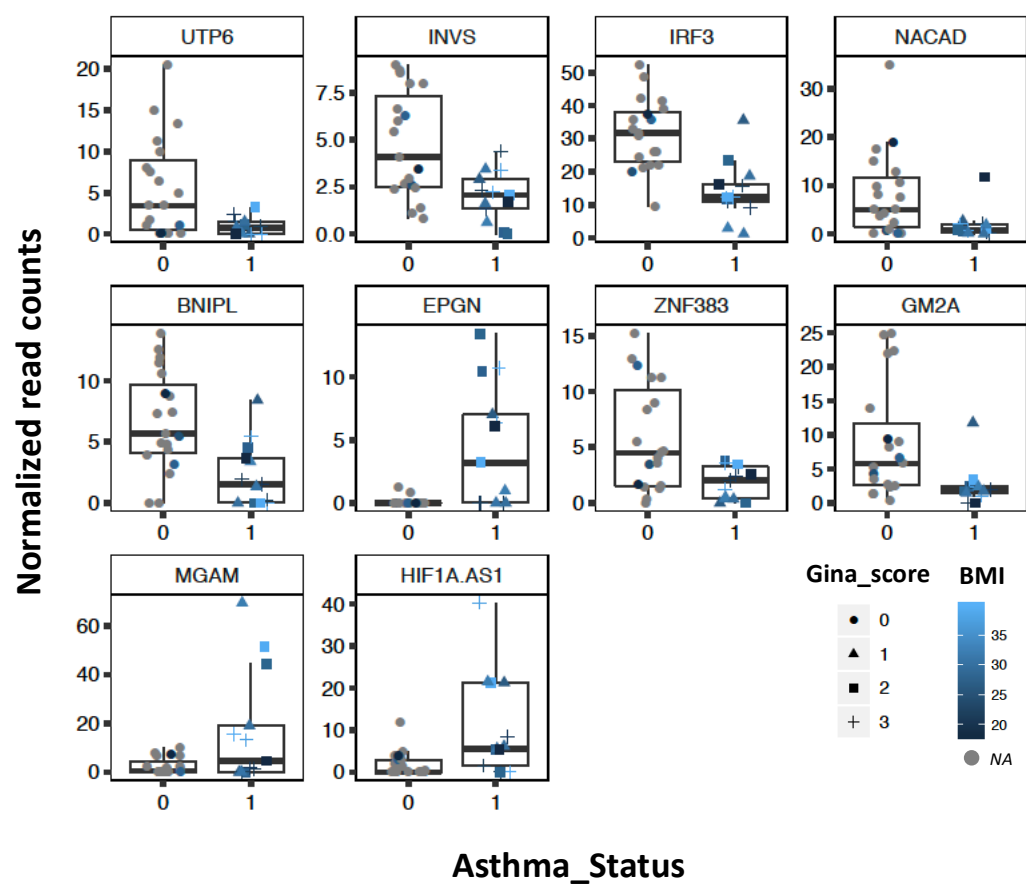


Figure S2: Expression profile of top10 differentially expressed genes

The transcriptomic expression profile of the top10 annotated differentially expressed genes. X-axis is the asthma status of individuals (0 = non-asthma; 1 = asthma) and y-axis are the quantile normalized read counts.

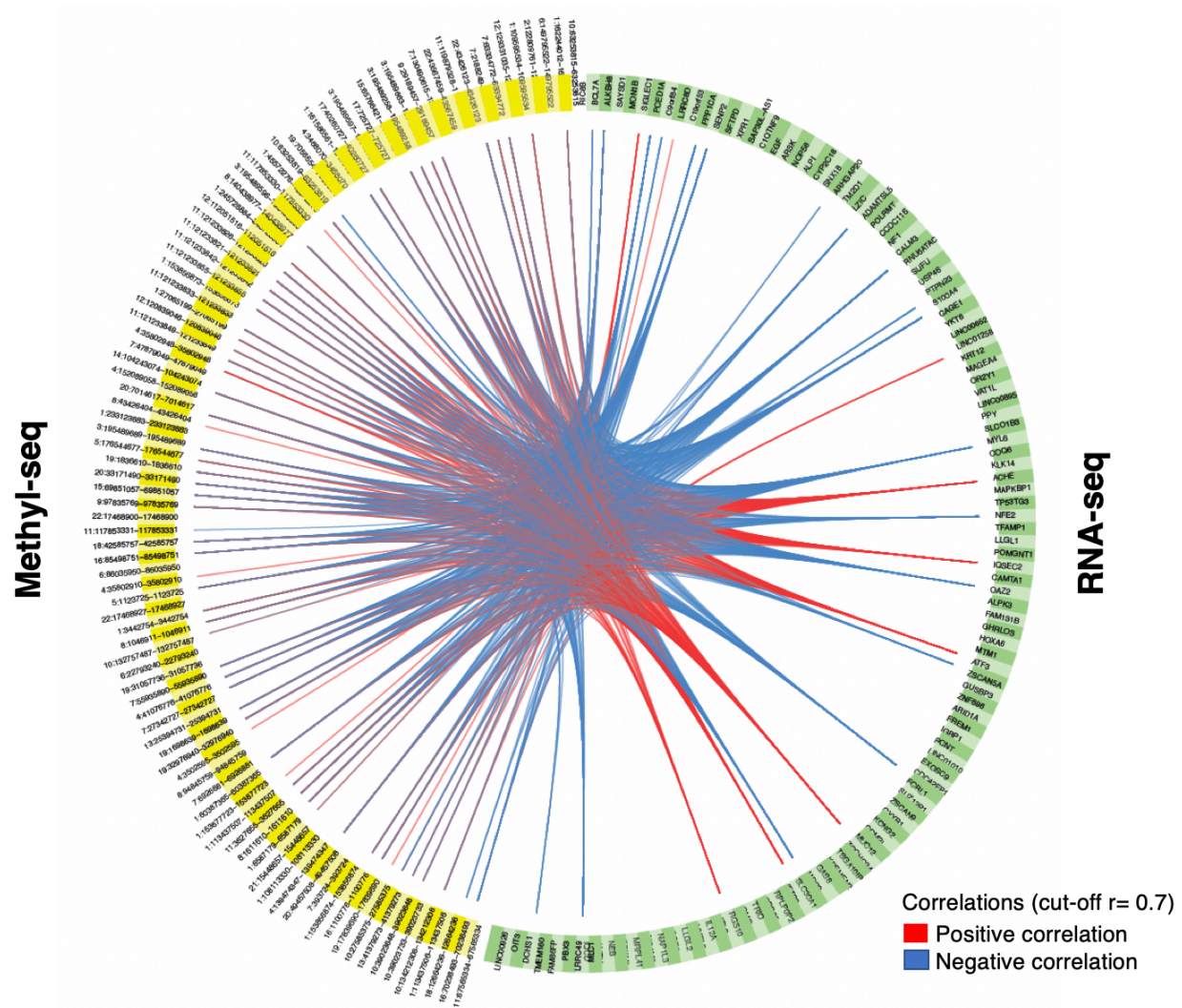


Figure S3: Regularized canonical correlation analysis

Integration of transcriptomic and methylomic data. The circos plot of features from transcriptomic data (green) and methylomic data (yellow) across asthmatic and non-asthmatics samples. Strong correlations (absolute correlation > 0.7) between the features are represented as links (red = positive; blue = negative).

Table S1: Differentially expressed genes between asthma and non-asthma (N=117)

	Gene symbol	Ensembl_ID	log2(Fold Change)	P-value	p-adj [§]
1	IRF3	ENSG00000126456	-1.07413	0.00002	0.00000
2	ENSG00000231345	ENSG00000231345	-1.74923	0.00002	0.00000
3	B3GNTL1	ENSG00000175711	0.86204	0.00014	0.00000
4	ENSG00000273424	ENSG00000273424	-1.51128	0.00028	0.00000
5	UTP6	ENSG00000108651	-1.48289	0.00038	0.00000
6	LOC101927596	ENSG00000231621	-0.88554	0.00046	0.00000
7	INVS	ENSG00000119509	-1.28553	0.00051	0.00000
8	GM2A	ENSG00000196743	-1.34924	0.00058	0.00000
9	PPL	ENSG00000118898	-0.84605	0.00067	0.00000
10	BNIP1	ENSG00000163141	-1.25864	0.00077	0.00000
11	HIST2H2BC	ENSG00000261716	-0.92181	0.00081	0.00000
12	ZNF383	ENSG00000188283	-1.32322	0.00086	0.00000
13	ENSG00000226143	ENSG00000226143	-1.37093	0.00092	0.00000
14	EPGN	ENSG00000182585	1.24411	0.00098	0.00000
15	NACAD	ENSG00000136274	-1.36707	0.00103	0.00000
16	ENSG00000253669	ENSG00000253669	-1.30805	0.00109	0.00000
17	MTBP	ENSG00000172167	1.12803	0.00112	0.00000
18	HIF1A-AS1	ENSG00000258777	1.35553	0.00117	0.00001
19	LOC283710	ENSG00000259772	-1.30524	0.00117	0.00001
20	VEGFB	ENSG00000173511	-1.35228	0.00120	0.00001
21	ENSG00000131484	ENSG00000131484	-1.15191	0.00121	0.00001
22	PABPC4-AS1	ENSG00000228060	-0.93324	0.00121	0.00001
23	ENSG00000181227	ENSG00000181227	-0.89553	0.00122	0.00001
24	RHBDF1	ENSG00000007384	-1.32902	0.00124	0.00001
25	NQO2	ENSG00000124588	0.78081	0.00133	0.00001
26	DMAC2	ENSG00000105341	-0.91176	0.00138	0.00001
27	AMT	ENSG00000145020	-0.95680	0.00143	0.00001
28	ZNF436-AS1	ENSG00000249087	-1.30685	0.00146	0.00001
29	FRAT1	ENSG00000165879	-1.17492	0.00158	0.00001
30	ARPC5L	ENSG00000136950	-1.06647	0.00167	0.00001
31	TMEM176B	ENSG00000106565	-1.30416	0.00171	0.00001
32	CSRP2	ENSG00000175183	-1.09747	0.00173	0.00001
33	ENSG00000244479	ENSG00000244479	-1.13299	0.00174	0.00001
34	SNX32	ENSG00000172803	-1.10946	0.00175	0.00002
35	CYP2A6	ENSG00000255974	-1.07139	0.00222	0.00002

36	MUS81	ENSG00000172732	-0.90060	0.00222	0.00002
37	CRIP1	ENSG00000213145	-0.94614	0.00222	0.00002
38	IL17RC	ENSG00000163702	-1.08613	0.00223	0.00003
39	MGAM	ENSG00000257335	1.25926	0.00241	0.00003
40	TRMT1L	ENSG00000121486	-1.02440	0.00249	0.00003
41	ENSG00000261222	ENSG00000261222	0.86267	0.00268	0.00003
42	ENSG00000270605	ENSG00000270605	-0.94101	0.00269	0.00003
43	ENSG00000266978	ENSG00000266978	-1.05847	0.00269	0.00003
44	ENSG00000272993	ENSG00000272993	-0.92580	0.00278	0.00003
45	ENSG00000228862	ENSG00000228862	-1.24602	0.00285	0.00003
46	DPY19L3	ENSG00000178904	-1.23584	0.00308	0.00003
47	OR2A1	ENSG00000221970	-0.92307	0.00317	0.00003
48	GPC1	ENSG00000063660	-1.04437	0.00317	0.00003
49	CA11	ENSG00000063180	1.22177	0.00325	0.00003
50	RBMS1	ENSG00000153250	-0.60369	0.00335	0.00003
51	ENSG00000259321	ENSG00000259321	-1.18016	0.00346	0.00004
52	VPS33A	ENSG00000139719	-1.20354	0.00353	0.00004
53	OLFM2	ENSG00000105088	-0.96595	0.00355	0.00004
54	MIR302CHG	ENSG00000249532	0.64152	0.00371	0.00004
55	DERL3	ENSG00000099958	-0.67977	0.00373	0.00004
56	ENSG00000279541	ENSG00000279541	0.29230	0.00376	0.00005
57	PWP1	ENSG00000136045	1.08746	0.00382	0.00005
58	CRYBA1	ENSG00000108255	-1.11042	0.00392	0.00005
59	SDHAP2	ENSG00000242086	-0.97387	0.00392	0.00005
60	ZNF79	ENSG00000196152	-0.87257	0.00398	0.00005
61	ENSG00000276718	ENSG00000276718	-0.98203	0.00400	0.00005
62	ENSG00000253908	ENSG00000253908	-0.74311	0.00403	0.00006
63	TMBIM1	ENSG00000135926	0.77725	0.00404	0.00006
64	HIST1H4C	ENSG00000197061	-1.06892	0.00416	0.00007
65	HDAC4	ENSG00000068024	-1.03690	0.00418	0.00007
66	RAB26	ENSG00000167964	-0.93789	0.00434	0.00007
67	UBE2S	ENSG00000108106	-0.91923	0.00448	0.00007
68	WASHC2C	ENSG00000172661	-0.94851	0.00449	0.00007
69	ZNF487	ENSG00000243660	-1.18579	0.00452	0.00007
70	SH2B1	ENSG00000178188	1.14687	0.00457	0.00008
71	PHLDB3	ENSG00000176531	1.15967	0.00463	0.00008
72	POT1	ENSG00000128513	-1.12947	0.00464	0.00008
73	ENSG00000228366	ENSG00000228366	0.59097	0.00480	0.00008
74	LINC00324	ENSG00000178977	0.91038	0.00484	0.00008

75	DIPK1B	ENSG00000165716	-1.01375	0.00505	0.00008
76	ENSG00000226756	ENSG00000226756	-1.02745	0.00507	0.00008
77	LHFPL2	ENSG00000145685	-0.60300	0.00519	0.00008
78	ENSG00000225630	ENSG00000225630	-0.83934	0.00522	0.00008
79	MATN1	ENSG00000162510	0.70115	0.00528	0.00008
80	ENSG00000250711	ENSG00000250711	-1.08024	0.00540	0.00008
81	KLHL17	ENSG00000187961	-1.15215	0.00554	0.00008
82	ENSG00000224272	ENSG00000224272	-1.14169	0.00575	0.00009
83	TMEM41A	ENSG00000163900	-1.13178	0.00577	0.00009
84	PNPLA6	ENSG00000032444	0.78508	0.00578	0.00009
85	RASD1	ENSG00000108551	-0.79343	0.00579	0.00009
86	ENSG00000237593	ENSG00000237593	-0.94890	0.00591	0.00009
87	TPTEP1	ENSG00000100181	1.03558	0.00598	0.00009
88	ENSG00000204113	ENSG00000204113	0.71719	0.00598	0.00010
89	ENSG00000263179	ENSG00000263179	-1.13657	0.00601	0.00011
90	SMPD1	ENSG00000166311	-0.87411	0.00613	0.00011
91	ENSG00000272037	ENSG00000272037	-0.65159	0.00616	0.00011
92	CTNNBIP1	ENSG00000178585	1.12358	0.00665	0.00011
93	KLF7	ENSG00000118263	0.91855	0.00671	0.00012
94	ING1	ENSG00000153487	-0.73436	0.00680	0.00012
95	ENSG00000225822	ENSG00000225822	-1.10388	0.00685	0.00012
96	CTNS	ENSG00000040531	-0.97575	0.00685	0.00012
97	KIAA0100	ENSG00000007202	-0.82564	0.00690	0.00012
98	PQBP1	ENSG00000102103	-0.95404	0.00691	0.00012
99	GBP3	ENSG00000117226	1.02195	0.00710	0.00012
100	RSRP1	ENSG00000117616	-0.61419	0.00713	0.00012
101	EIF1B	ENSG00000114784	0.92160	0.00714	0.00012
102	GTF2H2C	ENSG00000183474	-1.11418	0.00721	0.00012
103	ENSG00000233293	ENSG00000233293	-1.12101	0.00728	0.00013
104	C12orf76	ENSG00000174456	-0.38519	0.00728	0.00013
105	ZMIZ2	ENSG00000122515	0.97576	0.00731	0.00014
106	ENSG00000232134	ENSG00000232134	-1.07363	0.00736	0.00014
107	TMEM121B	ENSG00000183307	-1.03507	0.00737	0.00015
108	ENSG00000186481	ENSG00000186481	-1.10132	0.00742	0.00016
109	ENSG00000262721	ENSG00000262721	-1.11096	0.00760	0.00016
110	PPP1R35	ENSG00000160813	-0.94891	0.00764	0.00017
111	PIM1	ENSG00000137193	0.81520	0.00771	0.00019
112	ATP6AP2	ENSG00000182220	0.66736	0.00778	0.00019
113	ANKRD52	ENSG00000139645	1.09607	0.00793	0.00020

114	PTP4A1	ENSG00000112245	0.76672	0.00796	0.00021
115	ENSG00000278974	ENSG00000278974	0.40141	0.00801	0.00022
116	MAP6D1	ENSG00000180834	-1.05325	0.00802	0.00023
117	HIST1H3E	ENSG00000274750	-1.09723	0.00816	0.00023

\$ Benjamini and Hochberg FDR-adjusted P -value.

Table S2: Differentially methylated CpG sites (N=96)

	Gene symbol	Gene ID	Chr:Start-End	Δ Meth*	P-value	p-adj ^s
1	SC5D	NM_001024956	11:121233833-121233833	21.3373	0.0000000	0.0000000
2	SC5D	NM_001024956	11:121233842-121233842	22.1358	0.0000000	0.0000000
3	SC5D	NM_001024956	11:121233849-121233849	21.4598	0.0000000	0.0000000
4	SC5D	NM_001024956	11:121233826-121233826	22.7206	0.0000000	0.0000000
5	DHX58	NM_024119	17:40260727-40260727	24.8737	0.0000000	0.0000000
6	SC5D	NM_001024956	11:121233821-121233821	21.2410	0.0000000	0.0000001
7	SC5D	NM_001024956	11:121233855-121233855	21.4274	0.0000000	0.0000001
8	EVX1-AS	NR_120507	7:27342727-27342727	29.1334	0.0000000	0.0000001
9	AKR7A2P1	NR_002796	1:113437507-113437507	-25.5199	0.0000000	0.0000005
10	SPIRE1	NM_001128627	18:12664236-12664236	-17.3251	0.0000000	0.0000018
11	GATAD2B	NM_020699	1:153877723-153877723	-24.3729	0.0000000	0.0000062
12	MIR4319	NR_036203	18:42585757-42585757	22.1174	0.0000000	0.0000062
13	KLF14	NM_138693	7:130490615-130490615	21.0924	0.0000000	0.0000065
14	DOK7	NM_001256896	4:3502595-3502595	-23.9296	0.0000000	0.0000076
15	SLC12A7	NM_006598	5:1123725-1123725	30.7202	0.0000000	0.0000094
16	MIR551A	NR_030277	1:3442754-3442754	19.6824	0.0000000	0.0000132
17	SH3D19	NM_001128924	4:152089058-152089058	-23.2492	0.0000000	0.0000132
18	TCF3	NM_003200	19:1698639-1698639	22.0571	0.0000000	0.0000145
19	TRIM29	NM_012101	11:119879328-119879328	-24.7971	0.0000000	0.0000181
20	GSE1	NM_001134473	16:85498751-85498751	23.3372	0.0000000	0.0000245
21	MBD3L3	NM_001164425	19:7056554-7056554	18.7477	0.0000000	0.0000245
22	VAV3	NM_001079874	1:108113330-108113330	-13.6523	0.0000000	0.0000263
23	SMG1P7	NR_033959	16:70238493-70238493	19.6890	0.0000000	0.0000263
24	LOC100506679	NR_125362	22:43426123-43426123	-24.4916	0.0000000	0.0000263
25	HSPA7	NR_024151	1:161586561-161586561	-19.9657	0.0000000	0.0000342
26	LRR37A6P	NR_003525	10:27585375-27585375	-22.8981	0.0000000	0.0000407
27	TTLL12	NM_015140	22:43567459-43567459	-32.4169	0.0000000	0.0000407
28	HDGFL1	NM_138574	6:22793240-22793240	-21.5604	0.0000000	0.0000500
29	NTPCR	NM_032324	1:233123883-233123883	-23.7219	0.0000000	0.0000536
30	DPY19L3	NM_207325	19:32976940-32976940	25.2282	0.0000000	0.0000609
31	GATAD2B	NM_020699	1:153856873-153856873	-21.2980	0.0000000	0.0000659
32	NT5E	NM_001204813	6:86035950-86035950	-25.5273	0.0000000	0.0000696
33	CYP2J2	NM_000775	1:60387365-60387365	18.9775	0.0000000	0.0001038
34	MUC20	NM_152673	3:195489597-195489597	18.7982	0.0000000	0.0001298
35	REXO1	NM_020695	19:1836610-1836610	23.0034	0.0000000	0.0001438
36	LOC286083	NR_111948	8:1046911-1046911	-24.3442	0.0000000	0.0001523

37	BCAS4	NM_017843	20:49457508-49457508	21.4213	0.0000000	0.0001655
38	TMEM26-AS1	NR_120645	10:63253815-63253815	-25.6507	0.0000000	0.0003064
39	AKR7A2P1	NR_002796	1:113437506-113437506	-20.4248	0.0000000	0.0005005
40	ZFYVE21	NM_001198953	14:104243074-104243074	14.9458	0.0000000	0.0005502
41	DOK7	NM_001301071	4:3468070-3468070	-20.4910	0.0000000	0.0006111
42	MUC20	NM_152673	3:195489663-195489663	19.5940	0.0000000	0.0006908
43	ACTR3BP5	NR_045000	10:39023733-39023733	-19.9897	0.0000000	0.0010588
44	MIR378C	NR_036180	10:132757487-132757487	-14.3324	0.0000000	0.0010928
45	LOC101927708	NR_126335	11:3627655-3627655	18.3001	0.0000000	0.0011122
46	MUC20	NM_152673	3:195489258-195489258	18.7529	0.0000000	0.0011122
47	MUC20	NM_152673	3:195489596-195489596	21.3900	0.0000001	0.0012287
48	C7orf69	NM_001302627	7:47879049-47879049	-17.2898	0.0000001	0.0012287
49	DLGAP2-AS1	NR_103863	8:1611610-1611610	18.6627	0.0000001	0.0012940
50	GLT1D1	NM_144669	12:129331035-129331035	-11.6981	0.0000001	0.0015437
51	KIF26B	NM_018012	1:245728884-245728884	-10.7233	0.0000001	0.0016197
52	MAD1L1	NM_003550	7:2188249-2188249	17.2742	0.0000001	0.0016819
53	TMEM67	NM_153704	8:94845759-94845759	-16.1740	0.0000001	0.0017817
54	ACTR3BP5	NR_045000	10:39023648-39023648	-18.3770	0.0000001	0.0022150
55	MSI1	NM_002442	12:120839046-120839046	-14.6053	0.0000001	0.0022229
56	CCZ1B	NM_198097	7:6926881-6926881	-23.1377	0.0000001	0.0022229
57	LINC01005	NR_039987	7:63334772-63334772	19.6503	0.0000001	0.0022777
58	SEPT14	NM_207366	7:55935890-55935890	24.2724	0.0000001	0.0023412
59	DPP8	NM_197960	15:65766421-65766421	-16.9904	0.0000001	0.0023552
60	ATXN2	NM_002973	12:112051516-112051516	12.5304	0.0000001	0.0025050
61	MIR4654	NR_039798	1:162244012-162244012	-18.4723	0.0000001	0.0027538
62	ZNF536	NM_014717	19:31057736-31057736	-20.7924	0.0000002	0.0028693
63	RNF17	NM_001184993	13:25394731-25394731	-21.9197	0.0000002	0.0029896
64	GAB4	NM_001037814	22:17468927-17468927	-17.6435	0.0000002	0.0029896
65	PWWP2B	NM_138499	10:134212308-134212308	-25.9947	0.0000002	0.0030840
66	PLEKHG5	NM_198681	1:6587179-6587179	19.6891	0.0000002	0.0031842
67	LOC442497	NR_033960	7:393724-393724	-17.2625	0.0000002	0.0032546
68	LOC439933	NR_122079	4:35802948-35802948	-15.8217	0.0000002	0.0032756
69	FAM86C2P	NR_024249	11:67565334-67565334	20.8940	0.0000002	0.0034976
70	ANKRD20A11P	NR_027270	21:15448657-15448657	-15.4665	0.0000002	0.0034987
71	MAP1S	NM_018174	19:17839690-17839690	20.6471	0.0000003	0.0039708
72	APBB2	NM_173075	4:41076776-41076776	18.0482	0.0000003	0.0039708
73	MIR621	NR_030352	13:41379273-41379273	-11.6571	0.0000003	0.0041028
74	LINC00499	NR_051987	4:139474347-139474347	-11.5002	0.0000003	0.0041028
75	ZC3H12D	NM_207360	6:149795522-149795522	-18.7129	0.0000003	0.0042044

76	MIR6081	NR_106729	9:97835769-97835769	21.5662	0.0000003	0.0044739
77	UROD	NR_036510	1:45572276-45572276	-20.1559	0.0000003	0.0046112
78	SSTR5	NM_001172560	16:1100776-1100776	18.1755	0.0000003	0.0048589
79	IL10RA	NM_001558	11:117853331-117853331	-28.9848	0.0000003	0.0049133
80	POTEA	NM_001002920	8:43426404-43426404	-23.9464	0.0000004	0.0049602
81	LINGO2	NM_001258282	9:29189457-29189457	-10.7338	0.0000004	0.0050253
82	MAP1LC3A	NM_032514	20:33171490-33171490	11.7989	0.0000004	0.0050606
83	WDR47	NM_001142551	1:109595534-109595534	-20.4999	0.0000004	0.0053646
84	MUC20	NM_152673	3:195489689-195489689	25.4622	0.0000004	0.0053646
85	GLOD4	NM_016080	17:725727-725727	-20.4847	0.0000004	0.0057307
86	KCNK9	NM_001282534	8:140438977-140438977	-14.1567	0.0000004	0.0057307
87	ARID1A	NM_006015	1:27065199-27065199	24.1724	0.0000004	0.0057407
88	LOC439933	NR_122079	4:35802910-35802910	-16.5366	0.0000004	0.0057458
89	LINC01428	NR_110609	20:7014617-7014617	12.8859	0.0000005	0.0061995
90	DRAIC	NR_026979	15:69851057-69851057	19.3187	0.0000005	0.0068208
91	NSD1	NM_172349	5:176544677-176544677	14.6312	0.0000006	0.0071344
92	IL10RA	NM_001558	11:117853330-117853330	-24.1293	0.0000006	0.0075539
93	TMEM26-AS1	NR_120645	10:63253819-63253819	-22.1134	0.0000006	0.0078622
94	GAB4	NM_001037814	22:17468900-17468900	-16.9066	0.0000007	0.0080576
95	TSN	NR_048556	2:122809761-122809761	-11.9142	0.0000007	0.0080801
96	GATAD2B	NM_020699	1:153856874-153856874	-24.1445	0.0000007	0.0083038

* ΔMeth = Mean difference between asthma Vs nonasthma samples.

\$ Bonferroni-corrected *P*-value.

Table S3: Canonical correlations among 1665 RNA-seq features and 14562 Methyl-seq features. Top 50 strongest absolute Spearman's correlations (r) are presented.

	Feature 1			r^s	Feature 2		
	ID	Symbol	Omic		ID	Symbol	Omic
1	3:195489258-195489258	MUC20	Methyl-seq	-0.7421	CALM3	CALM3	RNA-seq
2	16:85498751-85498751	GSE1	Methyl-seq	-0.7240	CALM3	CALM3	RNA-seq
3	CALM3	CALM3	RNA-seq	-0.7188	12:112051516-112051516	ATXN2	Methyl-seq
4	7:2188249-2188249	MAD1L1	Methyl-seq	-0.7188	CALM3	CALM3	RNA-seq
5	CALM3	CALM3	RNA-seq	-0.7129	7:27342727-27342727	EVX1-AS	Methyl-seq
6	CALM3	CALM3	RNA-seq	-0.7128	3:195489663-195489663	MUC20	Methyl-seq
7	MYL6	MYL6	RNA-seq	-0.7124	3:195489258-195489258	MUC20	Methyl-seq
8	CALM3	CALM3	RNA-seq	-0.7080	14:104243074-104243074	ZFYVE21	Methyl-seq
9	19:1836610-1836610	REXO1	Methyl-seq	-0.7036	CALM3	CALM3	RNA-seq
10	12:112051516-112051516	ATXN2	Methyl-seq	0.7883	20:33171490-33171490	MAP1LC3A	Methyl-seq
11	12:129331035-129331035	GLT1D1	Methyl-seq	0.7884	3:195489258-195489258	MUC20	Methyl-seq
12	16:85498751-85498751	GSE1	Methyl-seq	0.7899	17:40260727-40260727	DHX58	Methyl-seq
13	ATF3	ATF3	RNA-seq	0.7900	12:112051516-112051516	ATXN2	Methyl-seq
14	15:69851057-69851057	DRAIC	Methyl-seq	0.7902	7:2188249-2188249	MAD1L1	Methyl-seq
15	12:112051516-112051516	ATXN2	Methyl-seq	0.7902	15:69851057-69851057	DRAIC	Methyl-seq
16	18:42585757-42585757	MIR4319	Methyl-seq	0.7918	3:195489258-195489258	MUC20	Methyl-seq
17	14:104243074-104243074	ZFYVE21	Methyl-seq	0.7932	19:1836610-1836610	REXO1	Methyl-seq
18	ZCCHC24	ZCCHC24	RNA-seq	0.7940	16:85498751-85498751	GSE1	Methyl-seq
19	15:69851057-69851057	DRAIC	Methyl-seq	0.7958	16:85498751-85498751	GSE1	Methyl-seq
20	3:195489258-195489258	MUC20	Methyl-seq	0.7981	5:1123725-1123725	SLC12A7	Methyl-seq
21	19:1836610-1836610	REXO1	Methyl-seq	0.7985	3:195489663-195489663	MUC20	Methyl-seq
22	19:1836610-1836610	REXO1	Methyl-seq	0.7987	7:27342727-27342727	EVX1-AS	Methyl-seq
23	ATF3	ATF3	RNA-seq	0.8035	16:85498751-85498751	GSE1	Methyl-seq

24	14:104243074-104243074	ZFYVE21	Methyl-seq	0.8037	7:27342727-27342727	EVX1-AS	Methyl-seq
25	19:1836610-1836610	REXO1	Methyl-seq	0.8053	7:2188249-2188249	MAD1L1	Methyl-seq
26	12:112051516-112051516	ATXN2	Methyl-seq	0.8053	19:1836610-1836610	REXO1	Methyl-seq
27	IQSEC2	IQSEC2	RNA-seq	0.8053	16:85498751-85498751	GSE1	Methyl-seq
28	IFT22	IFT22	RNA-seq	0.8091	20:7014617-7014617	LINC01428	Methyl-seq
29	17:40260727-40260727	DHX58	Methyl-seq	0.8096	3:195489258-195489258	MUC20	Methyl-seq
30	14:104243074-104243074	ZFYVE21	Methyl-seq	0.8104	7:2188249-2188249	MAD1L1	Methyl-seq
31	12:112051516-112051516	ATXN2	Methyl-seq	0.8104	14:104243074-104243074	ZFYVE21	Methyl-seq
32	16:85498751-85498751	GSE1	Methyl-seq	0.8111	19:1836610-1836610	REXO1	Methyl-seq
33	20:33171490-33171490	MAP1LC3A	Methyl-seq	0.8138	3:195489258-195489258	MUC20	Methyl-seq
34	15:69851057-69851057	DRAIC	Methyl-seq	0.8157	3:195489258-195489258	MUC20	Methyl-seq
35	3:195489663-195489663	MUC20	Methyl-seq	0.8158	7:2188249-2188249	MAD1L1	Methyl-seq
36	12:112051516-112051516	ATXN2	Methyl-seq	0.8158	3:195489663-195489663	MUC20	Methyl-seq
37	7:2188249-2188249	MAD1L1	Methyl-seq	0.8160	7:27342727-27342727	EVX1-AS	Methyl-seq
38	IQSEC2	IQSEC2	RNA-seq	0.8160	3:195489258-195489258	MUC20	Methyl-seq
39	14:104243074-104243074	ZFYVE21	Methyl-seq	0.8162	16:85498751-85498751	GSE1	Methyl-seq
40	16:85498751-85498751	GSE1	Methyl-seq	0.8217	3:195489663-195489663	MUC20	Methyl-seq
41	16:85498751-85498751	GSE1	Methyl-seq	0.8218	7:27342727-27342727	EVX1-AS	Methyl-seq
42	IFT22	IFT22	RNA-seq	0.8228	20:33171490-33171490	MAP1LC3A	Methyl-seq
43	IFT22	IFT22	RNA-seq	0.8286	15:69851057-69851057	DRAIC	Methyl-seq
44	12:112051516-112051516	ATXN2	Methyl-seq	0.8287	16:85498751-85498751	GSE1	Methyl-seq
45	19:1836610-1836610	REXO1	Methyl-seq	0.8313	3:195489258-195489258	MUC20	Methyl-seq
46	14:104243074-104243074	ZFYVE21	Methyl-seq	0.8366	3:195489258-195489258	MUC20	Methyl-seq
47	3:195489258-195489258	MUC20	Methyl-seq	0.8424	7:27342727-27342727	EVX1-AS	Methyl-seq
48	3:195489258-195489258	MUC20	Methyl-seq	0.8494	7:2188249-2188249	MAD1L1	Methyl-seq
49	IFT22	IFT22	RNA-seq	0.8494	17:40260727-40260727	DHX58	Methyl-seq
50	ZCCHC24	ZCCHC24	RNA-seq	0.8554	3:195489258-195489258	MUC20	Methyl-seq

\$ Spearman's correlation