천식연구회 - COPD연구회 공동 심포지움

Omics study

강원대 김우진

- GWAS
  - Discovery - Nat Genet
  - Causal variant - Blue
  - Interaction - Blue
- DNA Methylation
  - Causal inference testing - JACI
  - Longitudinal changes - Nat Med
  - EWAS - Lancet RM, Epigenomics
- Gene expression
  - Emphysema - IJCOPD
  - Single cell RNA-seq - Blue
- Multi-omics
  - ERJ

Cf) proteome, metabolome, microbiome

GWAS on COPD, asthma, lung functions

Multiethnic meta-analysis identifies ancestry-specific and cross-ancestry loci for pulmonary function

Nat Commun 2018;9:2976

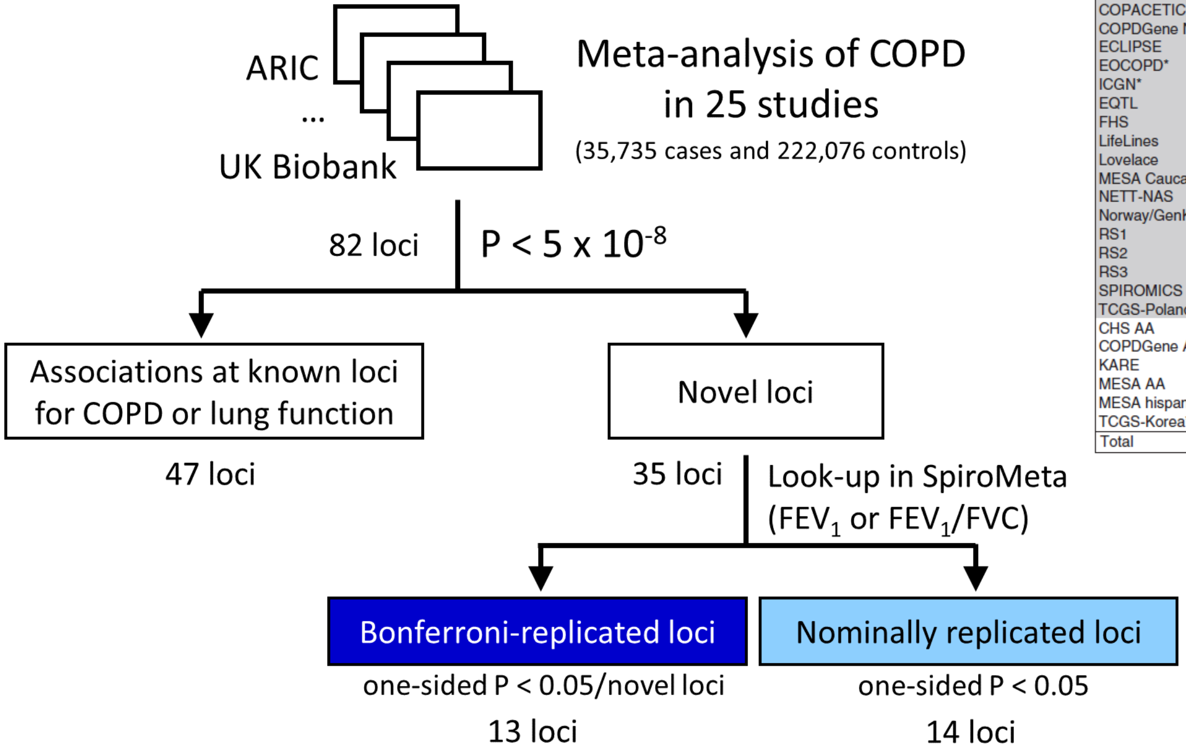
A Genome-Wide Association Study in Hispanics/Latinos Identifies Novel Signals for Lung Function. The Hispanic Community Health Study/Study of Latinos

AJRCCM 2018;198:208-19

An admixture mapping meta-analysis implicates genetic variation at 18q21 with asthma susceptibility in Latinos

JACI 2018 Epub

COPD GWAS



ARIC

…

UK Biobank

Meta-analysis of COPD in 25 studies

(35,735 cases and 222,076 controls)

82 loci

$P < 5 \times 10^{-8}$

Associations at known loci for COPD or lung function

47 loci

Novel loci

35 loci

Look-up in SpiroMeta (FEV_1 or FEV_1/FVC)

Bonferroni-replicated loci

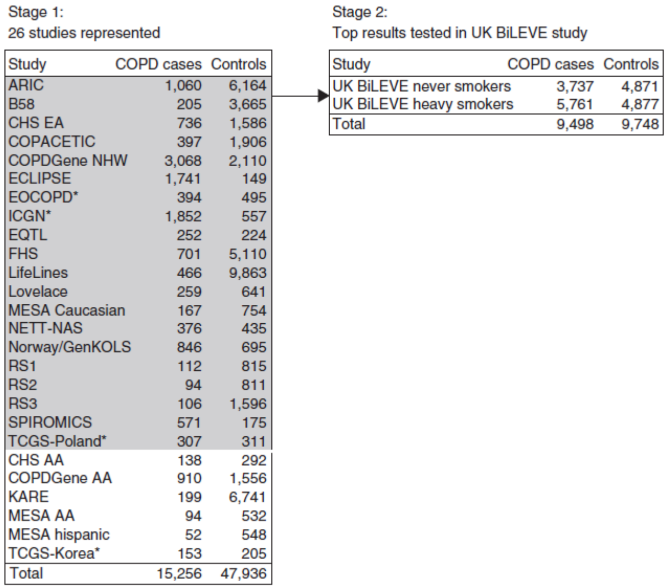
one-sided P < 0.05/novel loci

13 loci

Nominally replicated loci

one-sided P < 0.05

14 loci



Stage 1:
26 studies represented

Study	COPD cases	Controls
ARIC	1,060	6,164
B58	205	3,665
CHS EA	736	1,586
COPACETIC	397	1,906
COPDGene NHW	3,068	2,110
ECLIPSE	1,741	149
EOCOPD*	394	495
ICGN*	1,852	557
EQTL	252	224
FHS	701	5,110
LifeLines	466	9,863
Lovelace	259	641
MESA Caucasian	167	754
NETT-NAS	376	435
Norway/GenKOLS	846	695
RS1	112	815
RS2	94	811
RS3	106	1,596
SPIROMICS	571	175
TCGS-Poland*	307	311
CHS AA	138	292
COPDGene AA	910	1,556
KARE	199	6,741
MESA AA	94	532
MESA hispanic	52	548
TCGS-Korea*	153	205
Total	15,256	47,936

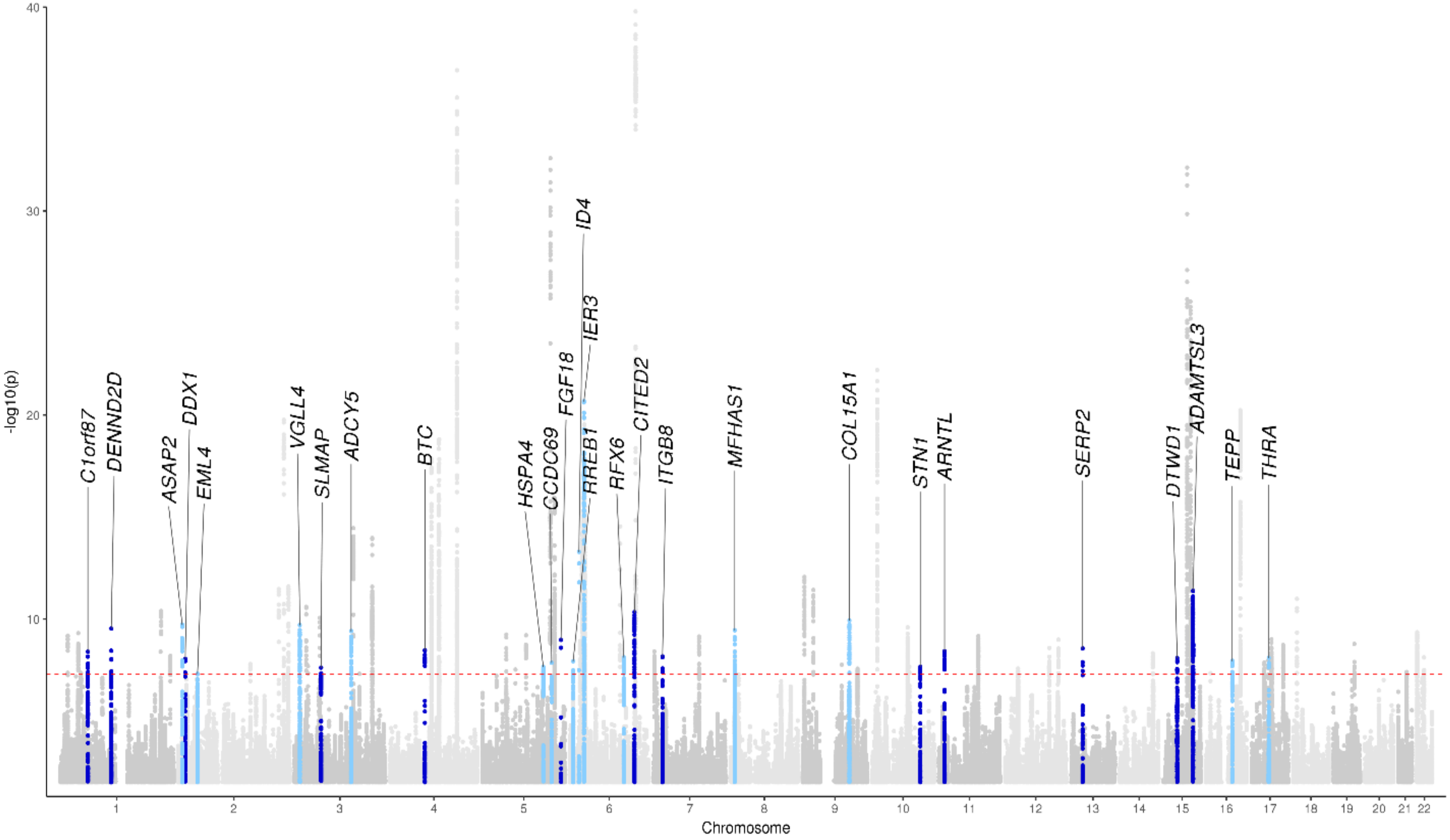
Stage 2:
Top results tested in UK BiLEVE study

Study	COPD cases	Controls
UK BiLEVE never smokers	3,737	4,871
UK BiLEVE heavy smokers	5,761	4,877
Total	9,498	9,748

Nat Genet 2017

Nat Genet, in press

COPD GWAS replicated by lung function in SpiroMeta



Nat Genet, in press

Identification of target genes

a.

82 genome-wide
significant loci

⇩

- Gene expression
- Methylation
- Coding associations
- Chromatin interaction
- DNAse hypersensitivity
- Similarity in gene sets

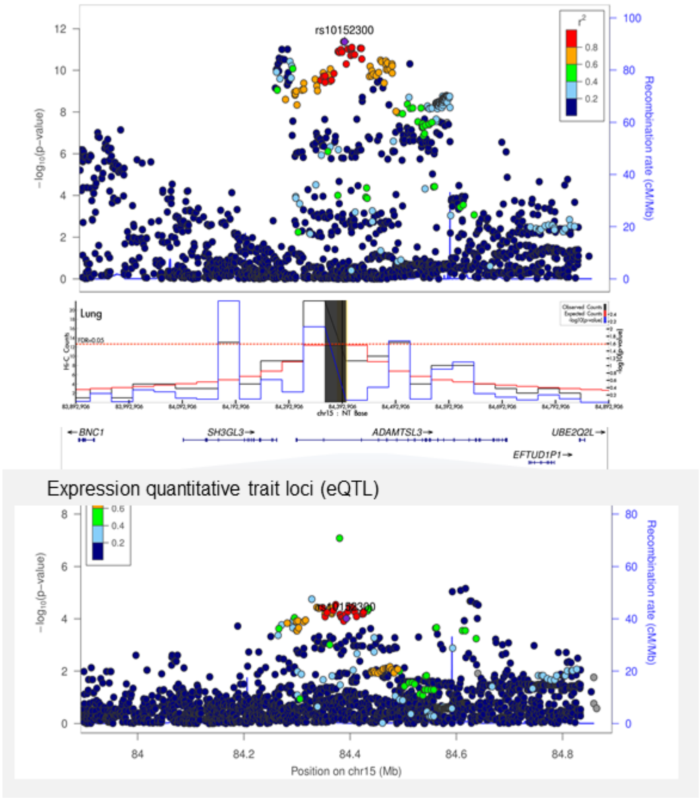
⇩

106 protein-coding genes supported by gene expression

50 protein-coding genes supported by ≥2 other datasets

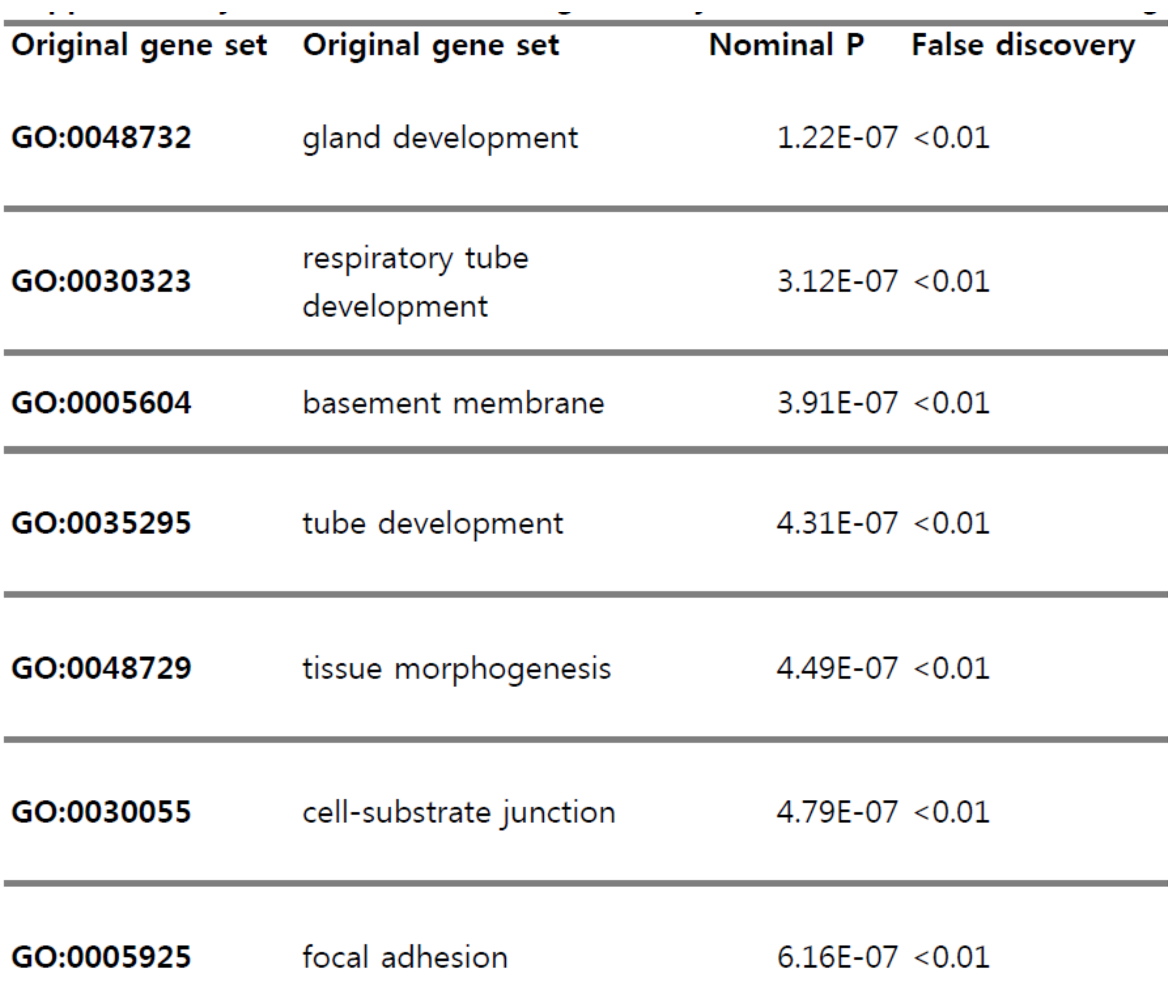
Locus	Gene	GREx	mQTL	Cod.	Hi-C	DHS	GSet
1p13	*DRAM2*	●					
1p13	*OVGP1*		●	●	●	●	
6q24	*HECA*				●	●	
15q25	*ADAMTSL3*	●	●		●	●	●
15q25	*BNC1*	●				●	●
15q25	*HOMER2*					●	●

b.



Nat Genet, in press

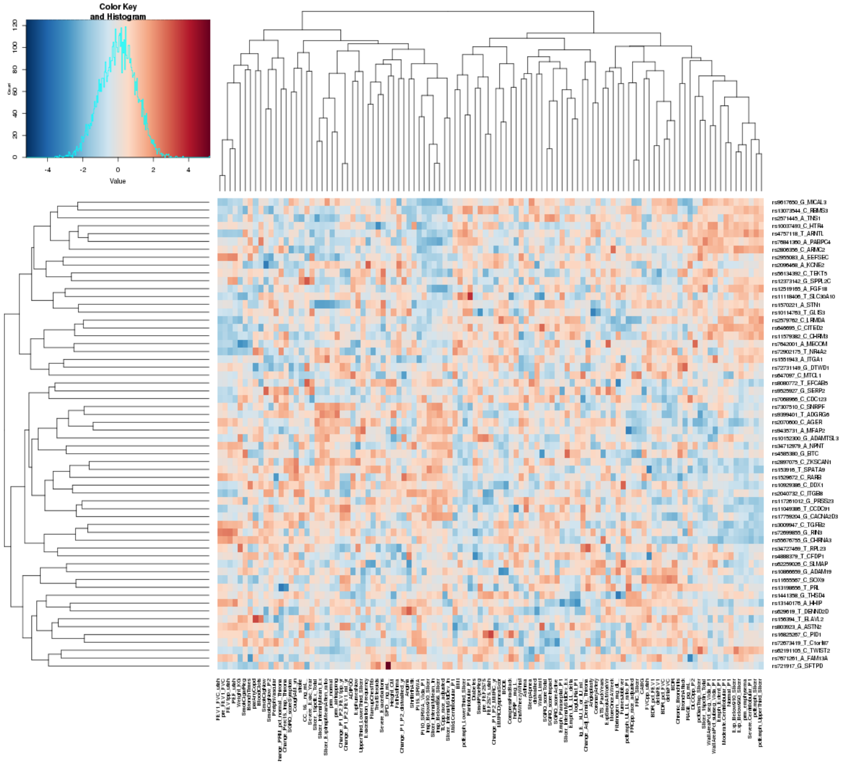
Associated pathways

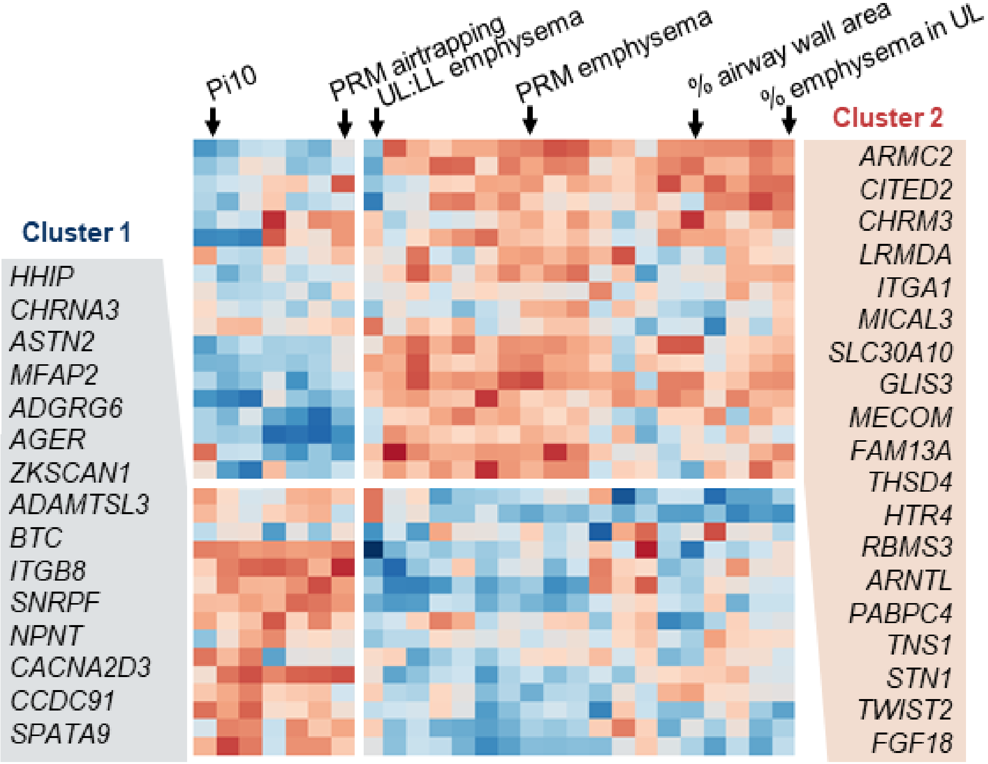


Original gene set	Original gene set	Nominal P	False discovery
GO:0048732	gland development	1.22E-07	<0.01
GO:0030323	respiratory tube development	3.12E-07	<0.01
GO:0005604	basement membrane	3.91E-07	<0.01
GO:0035295	tube development	4.31E-07	<0.01
GO:0048729	tissue morphogenesis	4.49E-07	<0.01
GO:0030055	cell-substrate junction	4.79E-07	<0.01
GO:0005925	focal adhesion	6.16E-07	<0.01

Gene sets significantly enriched at FDR < 0.05 using DEPICT

COPD-related phenotypes

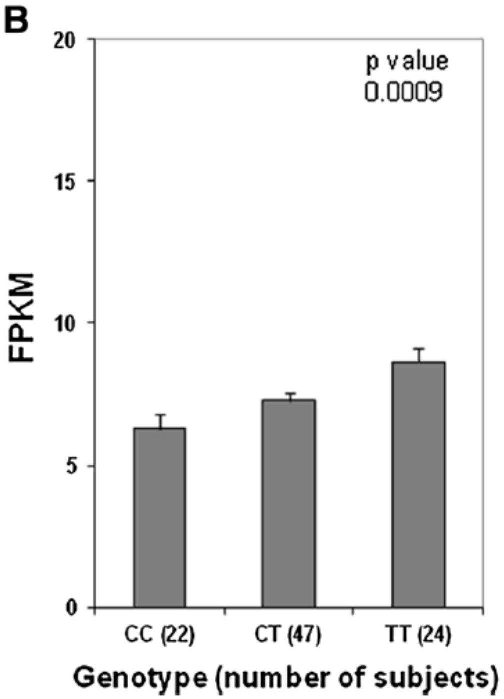




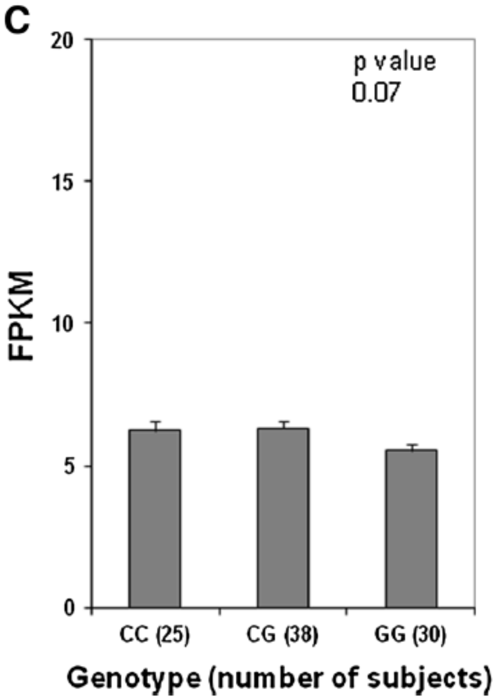
Nat Genet, in press

FAM13A association and expression in Korean

SNP	Gene	Tertile 1
rs2609264	*FAM13A*	1.48 (1.18–1.85), 0.0006
rs2609261	*FAM13A*	1.39 (1.11–1.73), 0.004
rs2609260	*FAM13A*	1.38 (1.10–1.73), 0.0046
rs10231916	*ACN9*	1.25 (1.00–1.56), 0.051
rs10229181	*ACN9*	1.29 (1.03–1.61), 0.028
rs8031759	*MCTP2*	1.15 (0.79–1.65), 0.47
rs886282	*AKAP1*	1.04 (0.76–1.41), 0.82



FAM13A

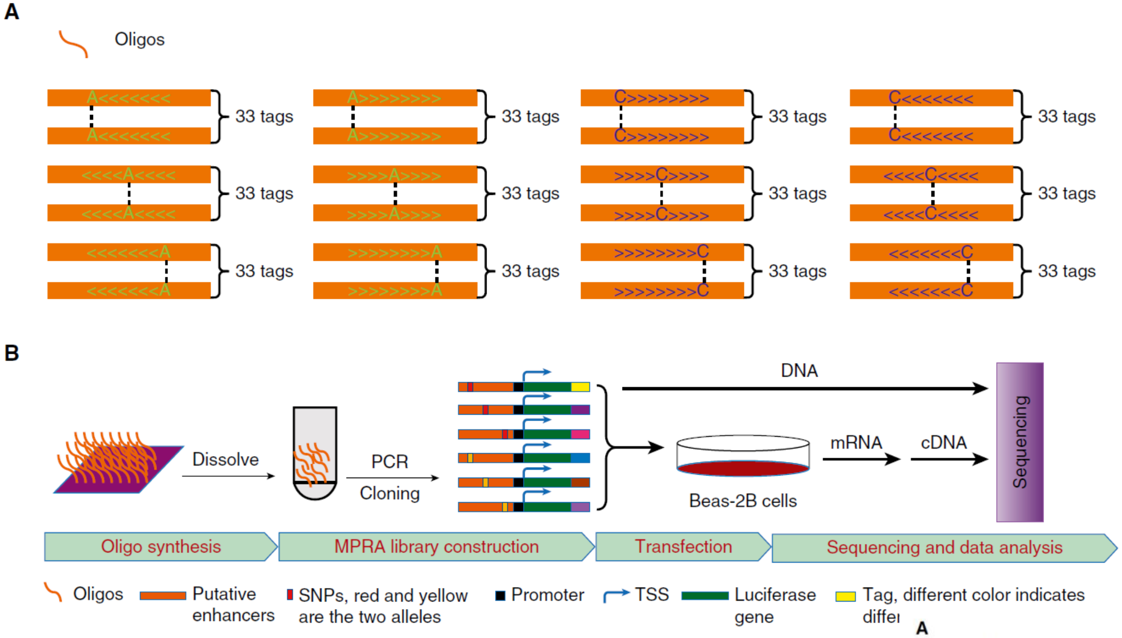


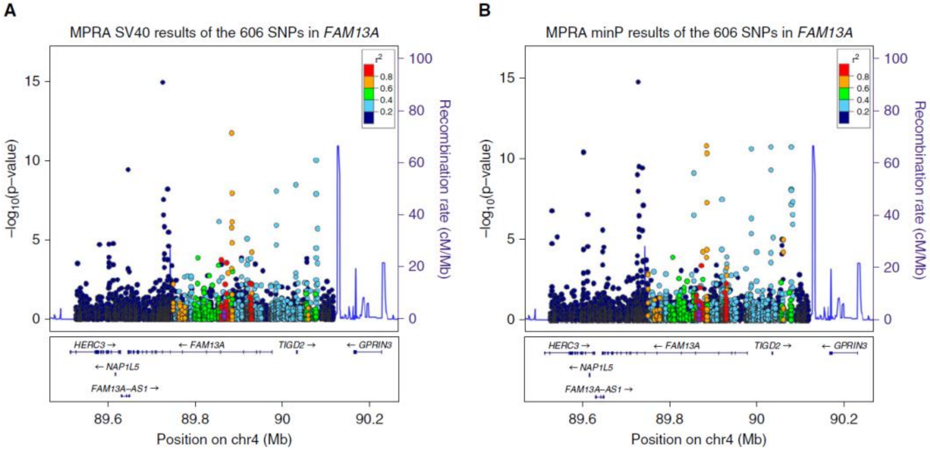
ACN9

Lung (2014) 192:473–480

Identification of Functional Variants in the *FAM13A* Chronic Obstructive Pulmonary Disease Genome-Wide Association Study Locus by Massively Parallel Reporter Assays

Peter J. Castaldi[1,2*], Feng Guo[1*], Dandi Qiao[1*], Fei Du[1], Zun Zar Chi Naing[1], Yan Li[1], Betty Pham[1], Tarjei S. Mikkelsen[3‡], Michael H. Cho[1,2], Edwin K. Silverman[1,2], Xiaobo Zhou[1,2]

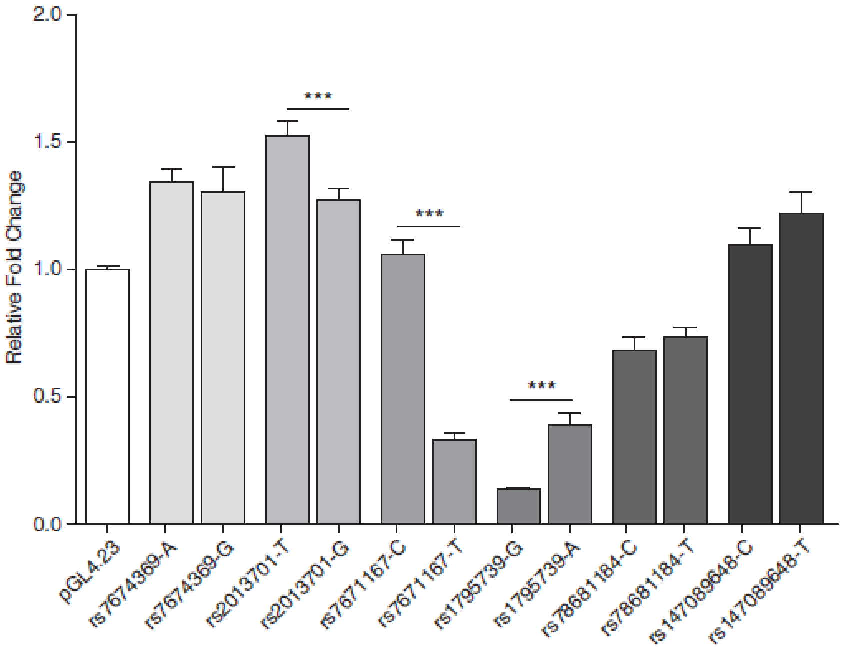


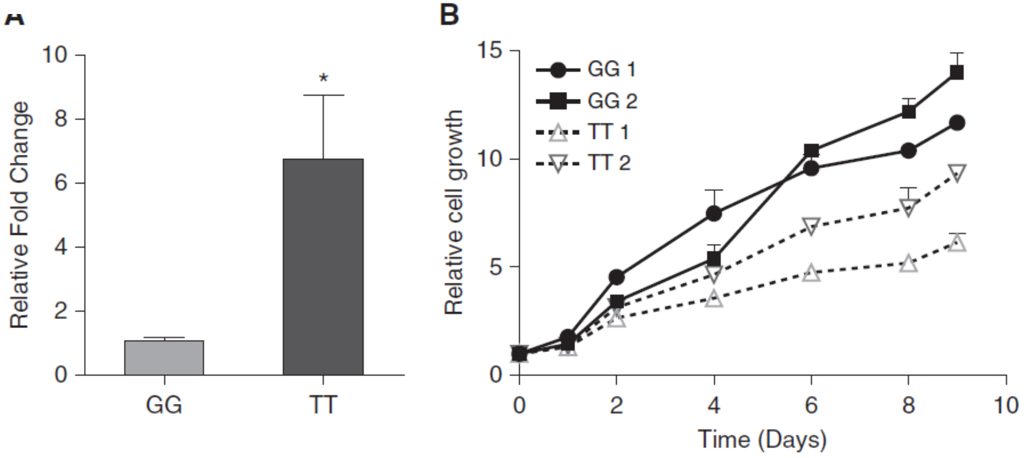


45 variants

Am J Respir Crit Care Med Vol 199, Iss 1, pp 52–61, Jan 1, 2019

Enhancer activity in 16HBE cells

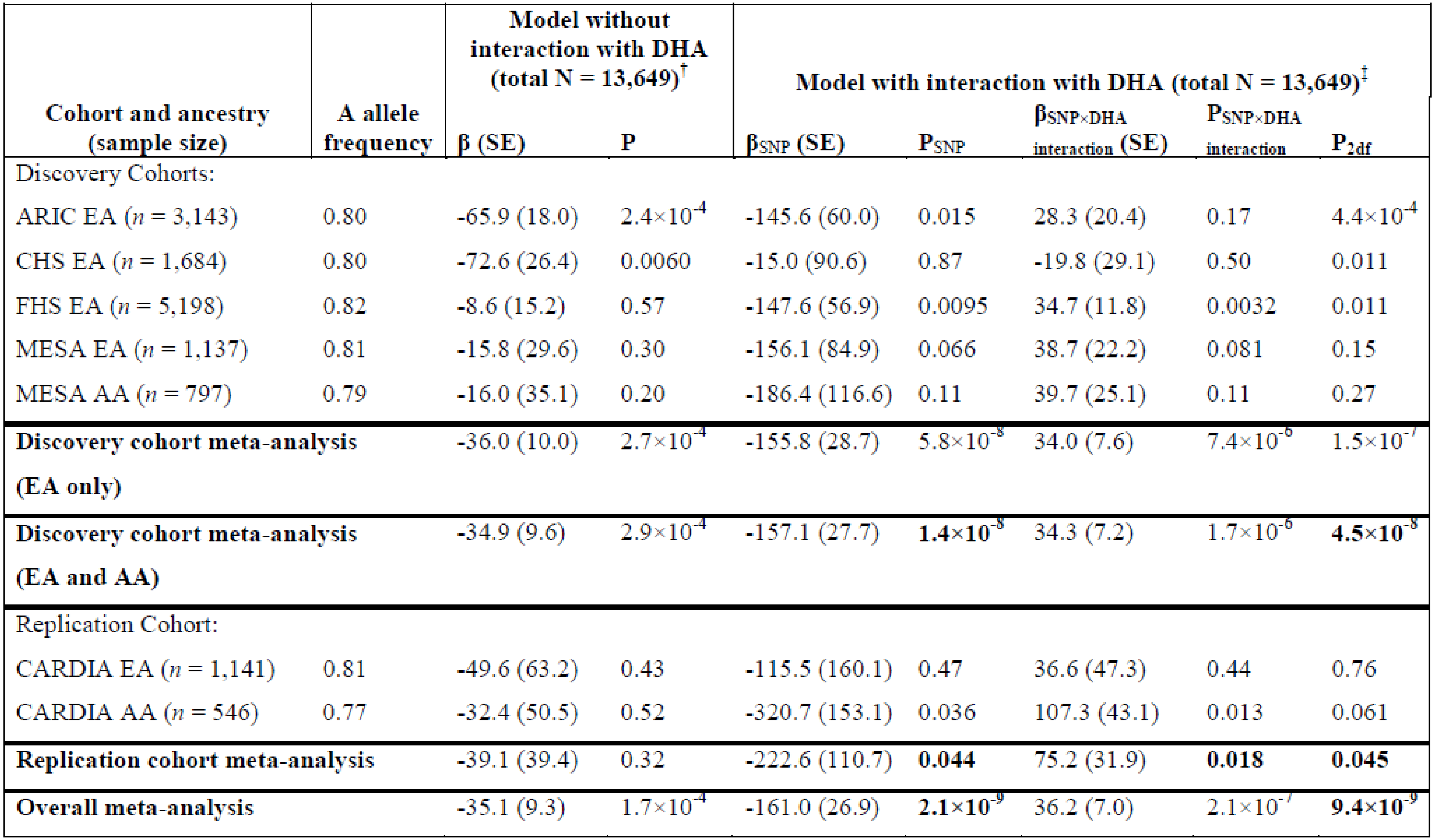




Am J Respir Crit Care Med Vol 199, Iss 1, pp 52–61, Jan 1, 2019

Interactions between genetics variants and n-3 PUFAs

DPP10



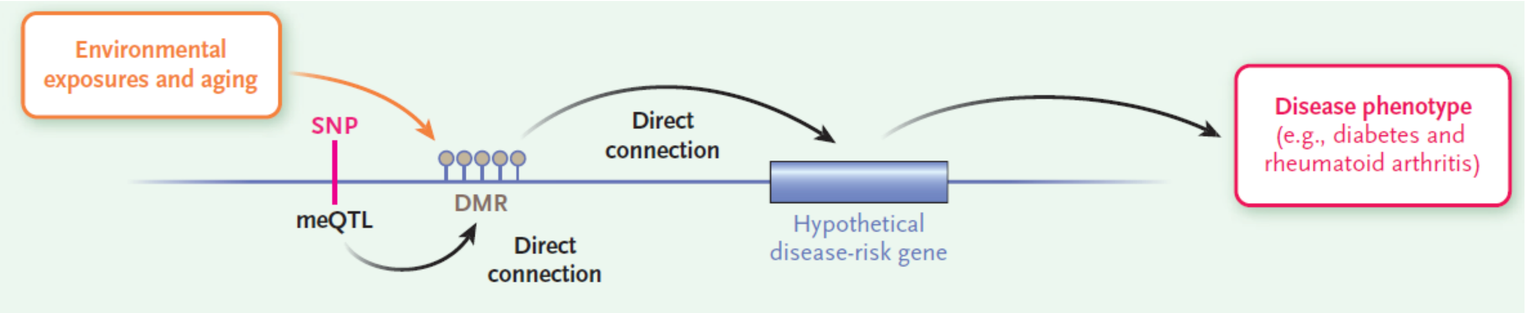
Cohort and ancestry (sample size)	A allele frequency	Model without interaction with DHA (total N = 13,649)†		Model with interaction with DHA (total N = 13,649)‡				
		β (SE)	P	β_{SNP} (SE)	P_{SNP}	$\beta_{SNP\times DHA\ interaction}$ (SE)	$P_{SNP\times DHA\ interaction}$	P_{2df}
Discovery Cohorts:								
ARIC EA (*n* = 3,143)	0.80	-65.9 (18.0)	2.4×10^{-4}	-145.6 (60.0)	0.015	28.3 (20.4)	0.17	4.4×10^{-4}
CHS EA (*n* = 1,684)	0.80	-72.6 (26.4)	0.0060	-15.0 (90.6)	0.87	-19.8 (29.1)	0.50	0.011
FHS EA (*n* = 5,198)	0.82	-8.6 (15.2)	0.57	-147.6 (56.9)	0.0095	34.7 (11.8)	0.0032	0.011
MESA EA (*n* = 1,137)	0.81	-15.8 (29.6)	0.30	-156.1 (84.9)	0.066	38.7 (22.2)	0.081	0.15
MESA AA (*n* = 797)	0.79	-16.0 (35.1)	0.20	-186.4 (116.6)	0.11	39.7 (25.1)	0.11	0.27
Discovery cohort meta-analysis (EA only)		-36.0 (10.0)	2.7×10^{-4}	-155.8 (28.7)	5.8×10^{-8}	34.0 (7.6)	7.4×10^{-6}	1.5×10^{-7}
Discovery cohort meta-analysis (EA and AA)		-34.9 (9.6)	2.9×10^{-4}	-157.1 (27.7)	$\mathbf{1.4\times10^{-8}}$	34.3 (7.2)	1.7×10^{-6}	$\mathbf{4.5\times10^{-8}}$
Replication Cohort:								
CARDIA EA (*n* = 1,141)	0.81	-49.6 (63.2)	0.43	-115.5 (160.1)	0.47	36.6 (47.3)	0.44	0.76
CARDIA AA (*n* = 546)	0.77	-32.4 (50.5)	0.52	-320.7 (153.1)	0.036	107.3 (43.1)	0.013	0.061
Replication cohort meta-analysis		-39.1 (39.4)	0.32	-222.6 (110.7)	**0.044**	75.2 (31.9)	**0.018**	**0.045**
Overall meta-analysis		-35.1 (9.3)	1.7×10^{-4}	-161.0 (26.9)	$\mathbf{2.1\times10^{-9}}$	36.2 (7.0)	2.1×10^{-7}	$\mathbf{9.4\times10^{-9}}$

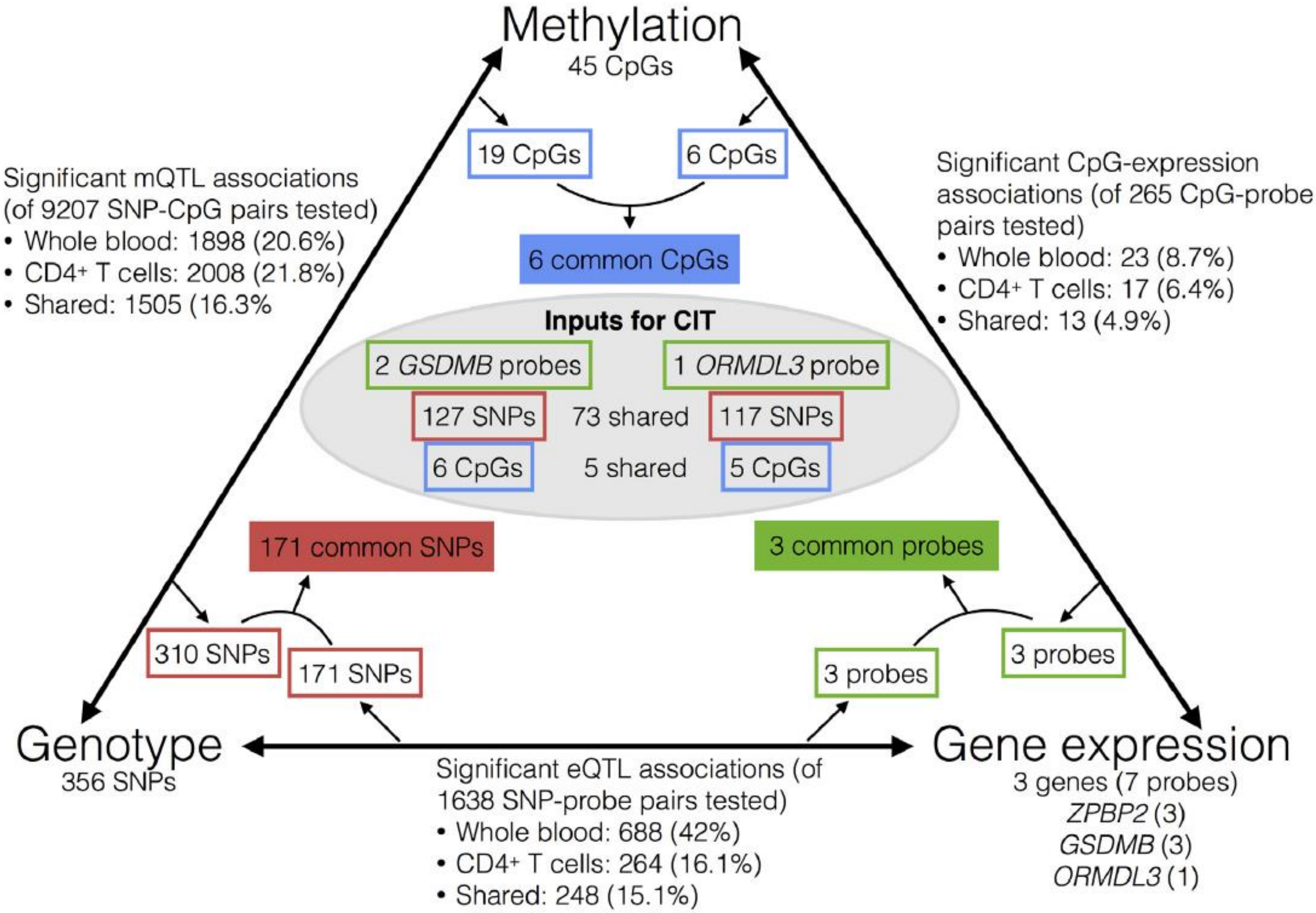
AJRCCM, in press

GWAS summary

- Discovery
  - <u>Sample size</u>
- On-going studies
  - Causal variant/Functional studies
  - Gene-Environment Interaction
  - Subtyping
  - WGS

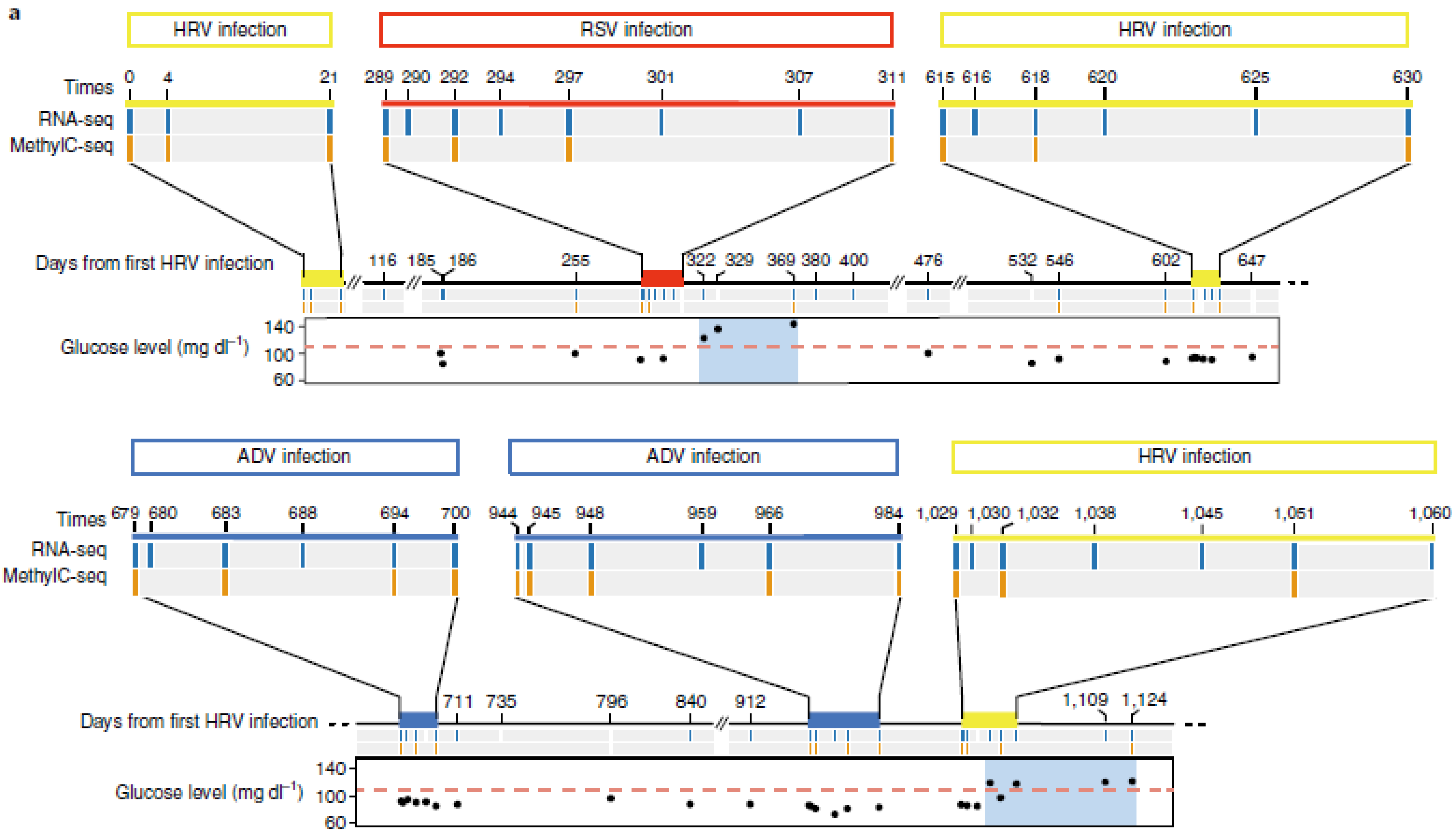
- Epigenome (DNA methylation)
- Transcriptome
  - Influenced by
    - Aging
    - Environment
    - Cell-type




Methylation
45 CpGs
19 CpGs
6 CpGs
6 common CpGs
Significant mQTL associations (of 9207 SNP-CpG pairs tested)
• Whole blood: 1898 (20.6%)
• CD4+ T cells: 2008 (21.8%)
• Shared: 1505 (16.3%
Significant CpG-expression associations (of 265 CpG-probe pairs tested)
• Whole blood: 23 (8.7%)
• CD4+ T cells: 17 (6.4%)
• Shared: 13 (4.9%)
Inputs for CIT
2 GSDMB probes
1 ORMDL3 probe
127 SNPs
73 shared
117 SNPs
6 CpGs
5 shared
5 CpGs
171 common SNPs
3 common probes
310 SNPs
171 SNPs
3 probes
3 probes
Genotype
356 SNPs
Gene expression
3 genes (7 probes)
ZPBP2 (3)
GSDMB (3)
ORMDL3 (1)
Significant eQTL associations (of 1638 SNP-probe pairs tested)
• Whole blood: 688 (42%)
• CD4+ T cells: 264 (16.1%)
• Shared: 248 (15.1%)

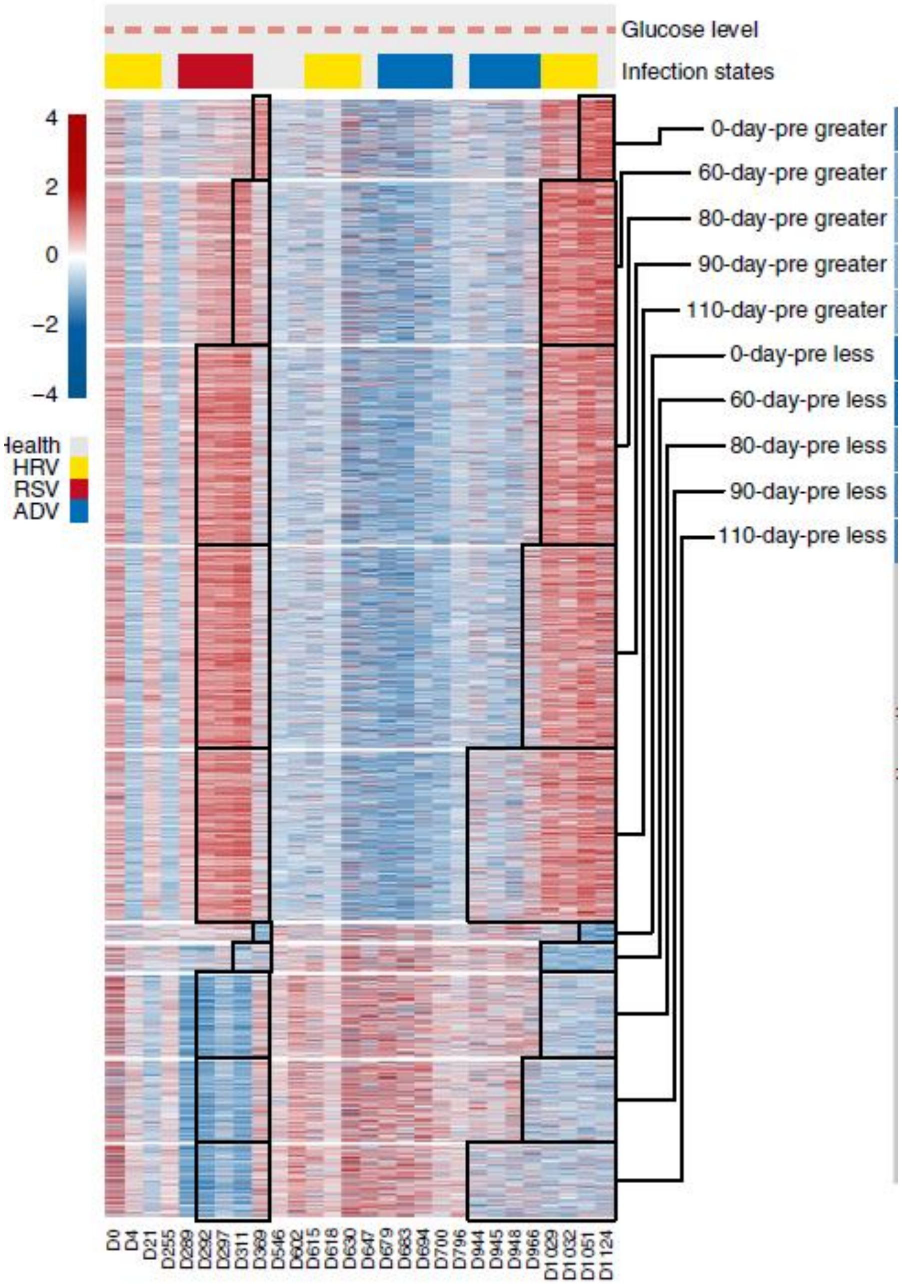
Longitudinal personal DNA methylome dynamics in a human with a chronic condition

Rui Chen[1,3,4], Lin Xia[2,4], Kailing Tu[2,4], Meixue Duan[2], Kimberly Kukurba[1], Jennifer Li-Pook-Than[1], Dan Xie[1,2]* and Michael Snyder[1]*

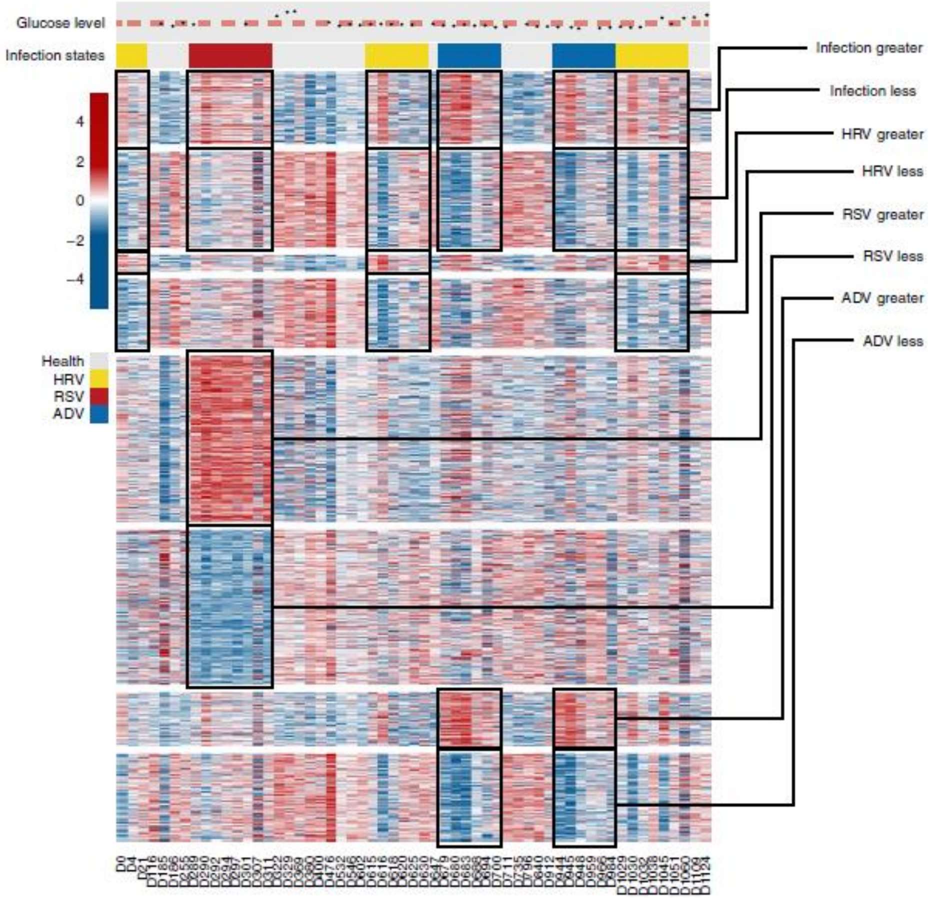


28 methylome 57 transcriptome

NATURE MEDICINE | VOL 24 | DECEMBER 2018 | 1930–1939



Methylomic changes



Gene expression changes

DNA methylation in childhood asthma: an epigenome-wide meta-analysis

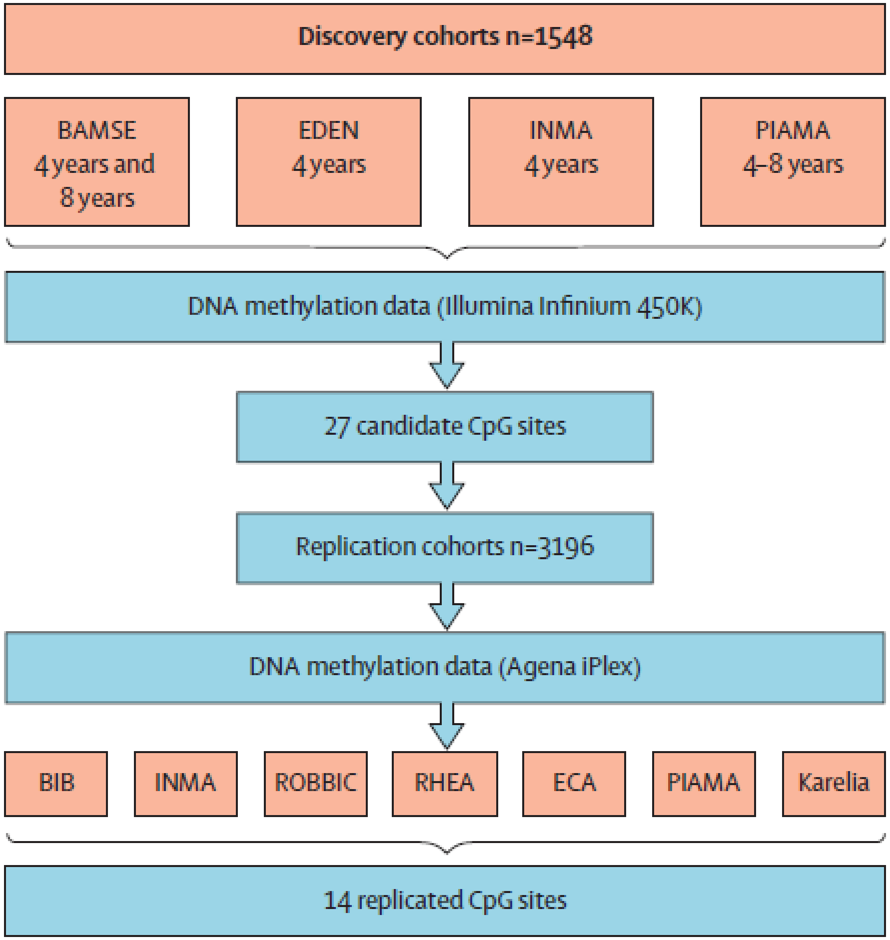
Cheng-Jian Xu, Cilla Söderhäll, Mariona Bustamante, Nour Baïz, Olena Gruzieva, Ulrike Gehring, Dan Mason, Leda Chatzi, Mikel Basterrechea, Sabrina Llop, Maties Torrent, Francesco Forastiere, Maria Pia Fantini, Karin C Lødrup Carlsen, Tari Haahtela, Andréanne Morin, Marjan Kerkhof, Simon Kebede Merid, Bianca van Rijkom, Soesma A Jankipersadsing, Marc Jan Bonder, Stephane Ballereau, Cornelis J Vermeulen, Raul Aguirre-Gamboa, Johan C de Jongste, Henriette A Smit, Ashish Kumar, Göran Pershagen, Stefano Guerra, Judith Garcia-Aymerich, Dario Greco, Lovisa Reinius,

rri Alenius, Nanna Fyhrquist, Nathanaël Lemonnier, Johann Pellet,

en, Yang Li, Cisca Wijmenga, Mihai G Netea, Miriam F Moffatt,

atherine Laprise, Kai-Håkon Carlsen, Davide Gori, Daniela Porta,

Brunekreef, Juha Kere, Martijn C Nawijn, Isabella Annesi-Maesano,



Discovery cohorts n=1548

BAMSE 4 years and 8 years

EDEN 4 years

INMA 4 years

PIAMA 4–8 years

DNA methylation data (Illumina Infinium 450K)

27 candidate CpG sites

Replication cohorts n=3196

DNA methylation data (Agena iPlex)

BIB

INMA

ROBBIC

RHEA

ECA

PIAMA

Karelia

14 replicated CpG sites

Cord blood methylation predicting early childhood asthma (EDEN [n=146], INMA [n=496,] BIB [n=296], ROBBIC [n=378])

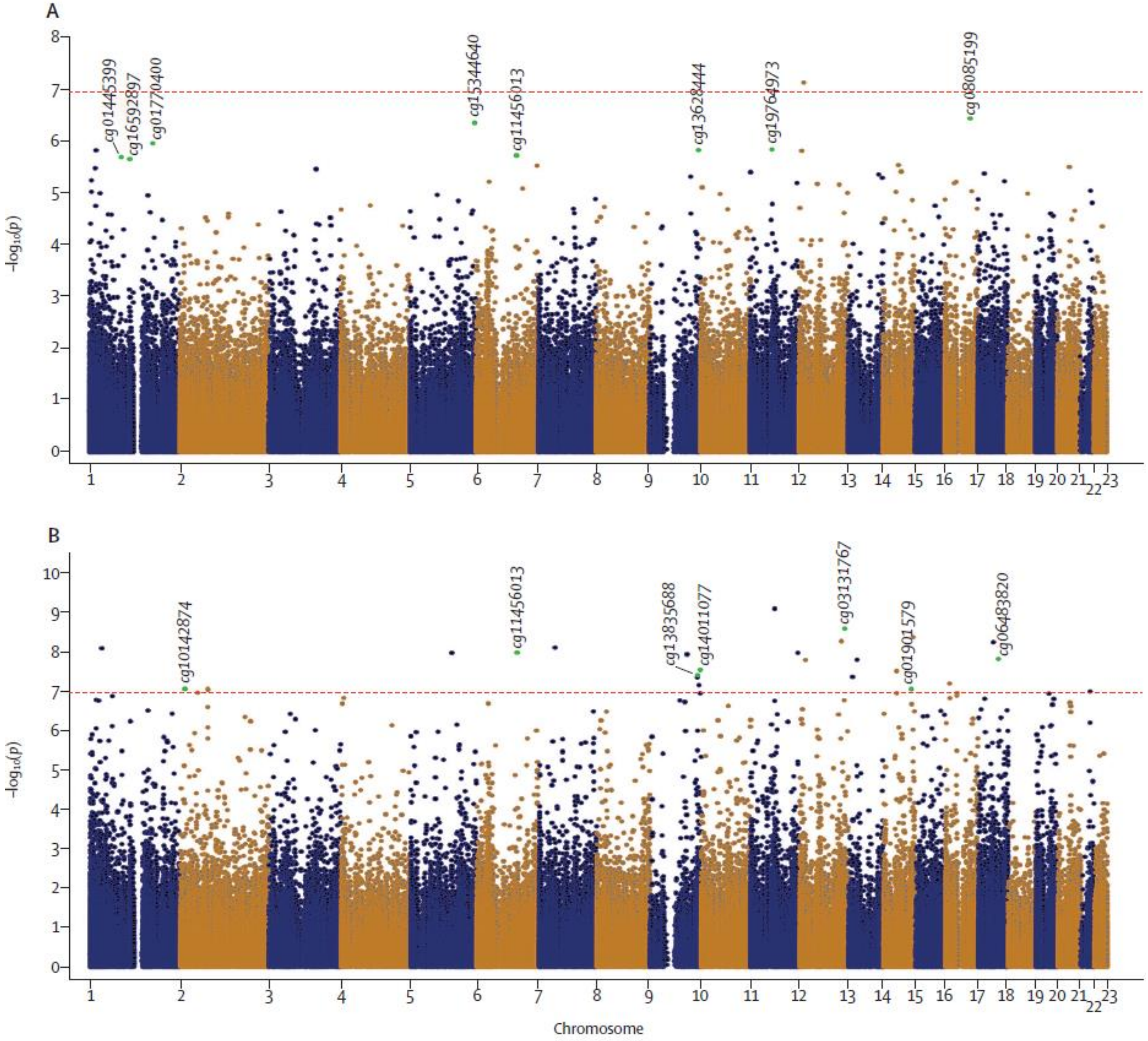
Extra replication (nasel epithelial [n=455], eosinophils [n=24], whole blood with full age range [n=167, 1–79 years])

Functional genomics and gene expression (BAMSE [n=260], INMA [n=119], BIOS [n=2367])

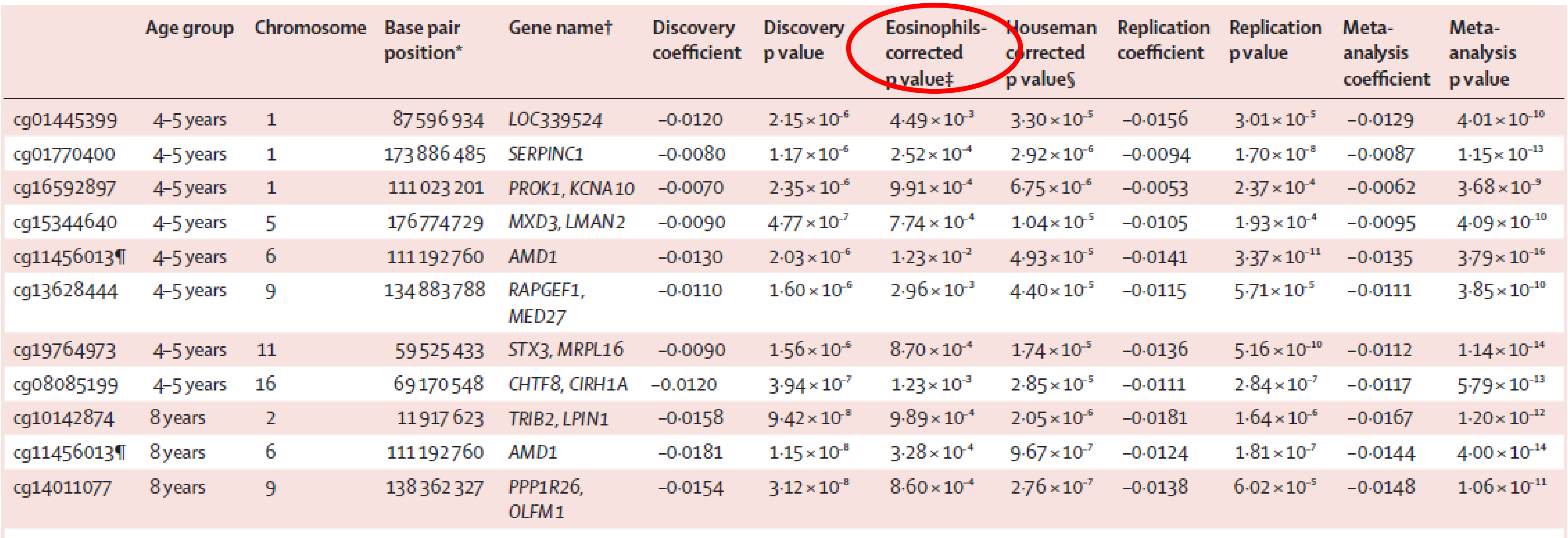
Gene set and functional enrichment analyses

Genetic association and causality (blood, n=496)

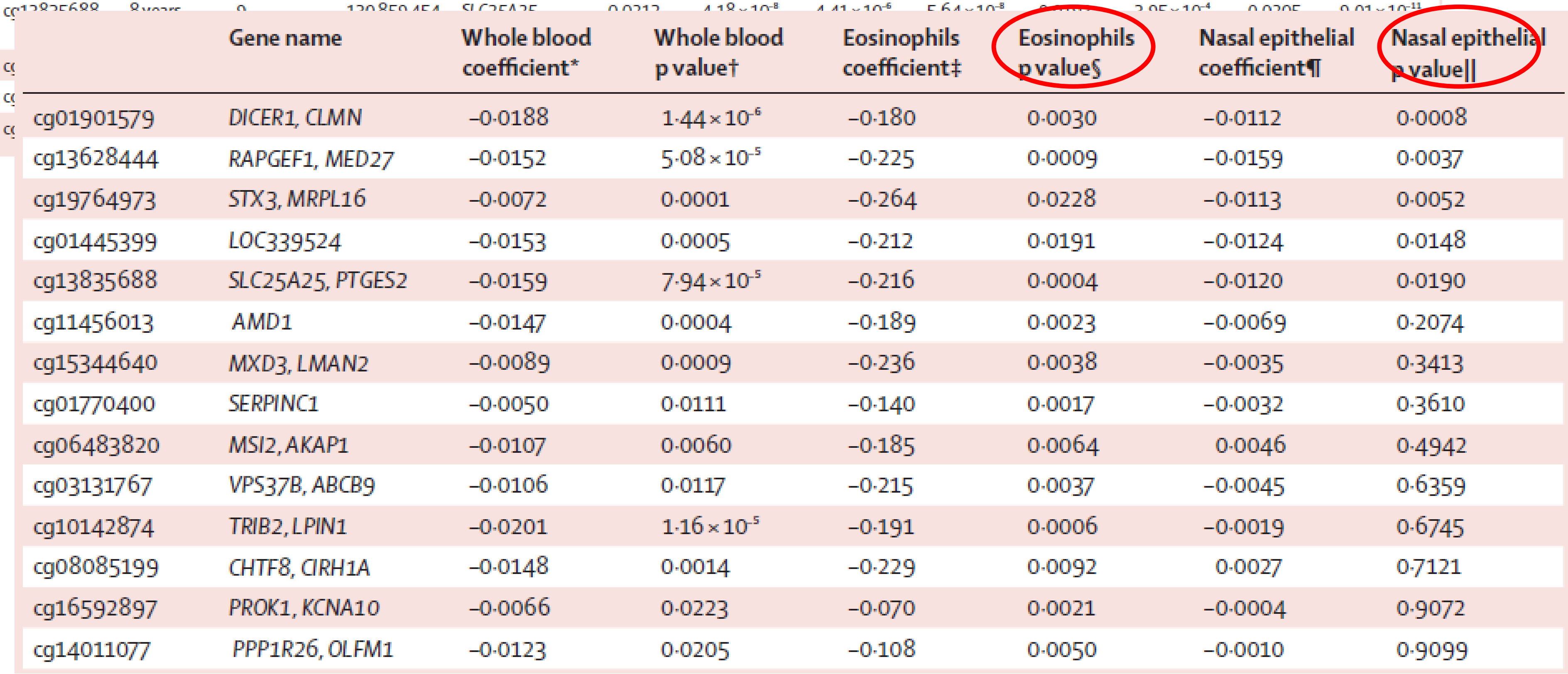
Lancet Respir Med 2018; 6: 379–88



Lancet Respir Med 2018;
6: 379–88

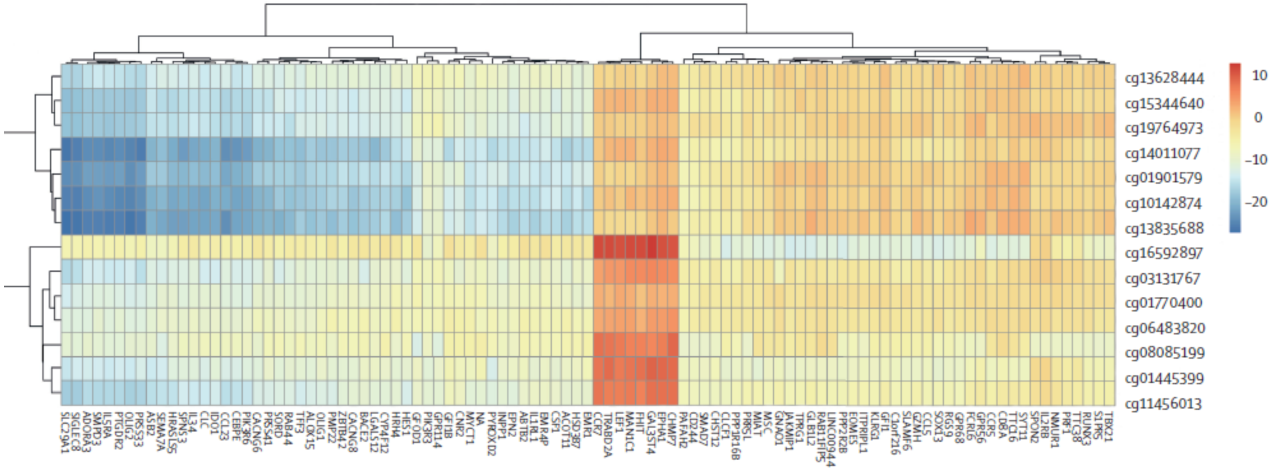


	Age group	Chromosome	Base pair position*	Gene name†	Discovery coefficient	Discovery p value	Eosinophils-corrected p value‡	Houseman corrected p value§	Replication coefficient	Replication p value	Meta-analysis coefficient	Meta-analysis p value
cg01445399	4–5 years	1	87 596 934	*LOC339524*	–0·0120	$2·15×10^{-6}$	$4·49×10^{-3}$	$3·30×10^{-5}$	–0·0156	$3·01×10^{-5}$	–0·0129	$4·01×10^{-10}$
cg01770400	4–5 years	1	173 886 485	*SERPINC1*	–0·0080	$1·17×10^{-6}$	$2·52×10^{-4}$	$2·92×10^{-6}$	–0·0094	$1·70×10^{-8}$	–0·0087	$1·15×10^{-13}$
cg16592897	4–5 years	1	111 023 201	*PROK1, KCNA10*	–0·0070	$2·35×10^{-6}$	$9·91×10^{-4}$	$6·75×10^{-6}$	–0·0053	$2·37×10^{-4}$	–0·0062	$3·68×10^{-9}$
cg15344640	4–5 years	5	176 774 729	*MXD3, LMAN2*	–0·0090	$4·77×10^{-7}$	$7·74×10^{-4}$	$1·04×10^{-5}$	–0·0105	$1·93×10^{-4}$	–0·0095	$4·09×10^{-10}$
cg11456013¶	4–5 years	6	111 192 760	*AMD1*	–0·0130	$2·03×10^{-6}$	$1·23×10^{-2}$	$4·93×10^{-5}$	–0·0141	$3·37×10^{-11}$	–0·0135	$3·79×10^{-16}$
cg13628444	4–5 years	9	134 883 788	*RAPGEF1, MED27*	–0·0110	$1·60×10^{-6}$	$2·96×10^{-3}$	$4·40×10^{-5}$	–0·0115	$5·71×10^{-5}$	–0·0111	$3·85×10^{-10}$
cg19764973	4–5 years	11	59 525 433	*STX3, MRPL16*	–0·0090	$1·56×10^{-6}$	$8·70×10^{-4}$	$1·74×10^{-5}$	–0·0136	$5·16×10^{-10}$	–0·0112	$1·14×10^{-14}$
cg08085199	4–5 years	16	69 170 548	*CHTF8, CIRH1A*	–0.0120	$3·94×10^{-7}$	$1·23×10^{-3}$	$2·85×10^{-5}$	–0·0111	$2·84×10^{-7}$	–0·0117	$5·79×10^{-13}$
cg10142874	8 years	2	11 917 623	*TRIB2, LPIN1*	–0·0158	$9·42×10^{-8}$	$9·89×10^{-4}$	$2·05×10^{-6}$	–0·0181	$1·64×10^{-6}$	–0·0167	$1·20×10^{-12}$
cg11456013¶	8 years	6	111 192 760	*AMD1*	–0·0181	$1·15×10^{-8}$	$3·28×10^{-4}$	$9·67×10^{-7}$	–0·0124	$1·81×10^{-7}$	–0·0144	$4·00×10^{-14}$
cg14011077	8 years	9	138 362 327	*PPP1R26, OLFM1*	–0·0154	$3·12×10^{-8}$	$8·60×10^{-4}$	$2·76×10^{-7}$	–0·0138	$6·02×10^{-5}$	–0·0148	$1·06×10^{-11}$

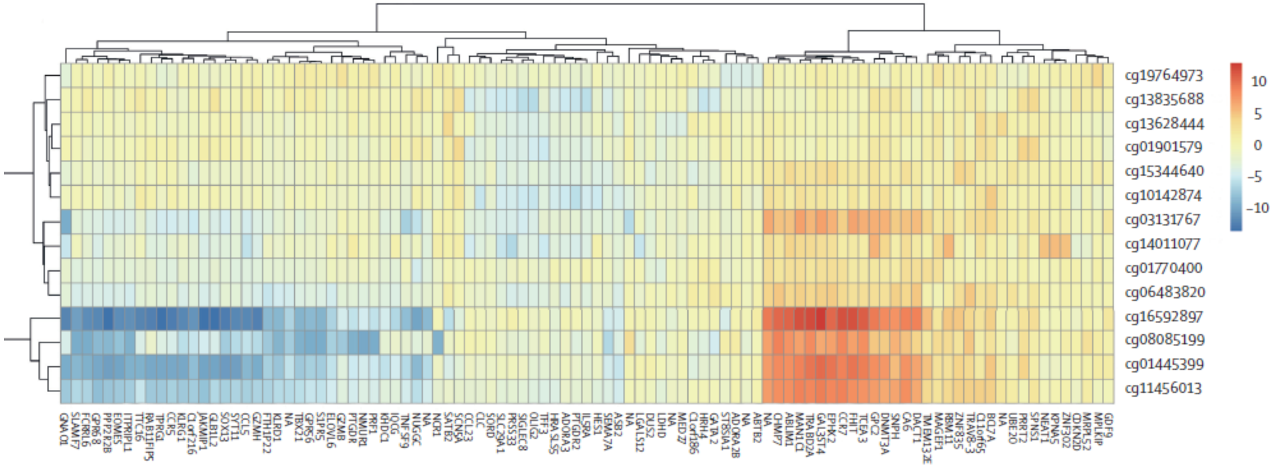


	Gene name	Whole blood coefficient*	Whole blood p value†	Eosinophils coefficient‡	Eosinophils p value§	Nasal epithelial coefficient¶	Nasal epithelial p value‖
cg01901579	*DICER1, CLMN*	–0·0188	$1·44×10^{-6}$	–0·180	0·0030	–0·0112	0·0008
cg13628444	*RAPGEF1, MED27*	–0·0152	$5·08×10^{-5}$	–0·225	0·0009	–0·0159	0·0037
cg19764973	*STX3, MRPL16*	–0·0072	0·0001	–0·264	0·0228	–0·0113	0·0052
cg01445399	*LOC339524*	–0·0153	0·0005	–0·212	0·0191	–0·0124	0·0148
cg13835688	*SLC25A25, PTGES2*	–0·0159	$7·94×10^{-5}$	–0·216	0·0004	–0·0120	0·0190
cg11456013	*AMD1*	–0·0147	0·0004	–0·189	0·0023	–0·0069	0·2074
cg15344640	*MXD3, LMAN2*	–0·0089	0·0009	–0·236	0·0038	–0·0035	0·3413
cg01770400	*SERPINC1*	–0·0050	0·0111	–0·140	0·0017	–0·0032	0·3610
cg06483820	*MSI2, AKAP1*	–0·0107	0·0060	–0·185	0·0064	0·0046	0·4942
cg03131767	*VPS37B, ABCB9*	–0·0106	0·0117	–0·215	0·0037	–0·0045	0·6359
cg10142874	*TRIB2, LPIN1*	–0·0201	$1·16×10^{-5}$	–0·191	0·0006	–0·0019	0·6745
cg08085199	*CHTF8, CIRH1A*	–0·0148	0·0014	–0·229	0·0092	0·0027	0·7121
cg16592897	*PROK1, KCNA10*	–0·0066	0·0223	–0·070	0·0021	–0·0004	0·9072
cg14011077	*PPP1R26, OLFM1*	–0·0123	0·0205	–0·108	0·0050	–0·0010	0·9099

A Heatmap of *Z* score of overlapped genes (without Eos correction)



B Heatmap of *Z* score of overlapped genes (with Eos correction)



Eosinophil gene signature
Reduced activity of naïve T cell
Increased activity of effector and memory CD8 T cell and NK cell

Lancet Respir Med 2018; 6: 379–88

Epigenome-wide association study of chronic obstructive pulmonary disease and lung function in Koreans

Mi Kyeong Lee[1,2], Yoonki Hong[2], Sun-Young Kim[3], Woo Jin Kim[‡,2] & Stephanie J London[*,‡,1]

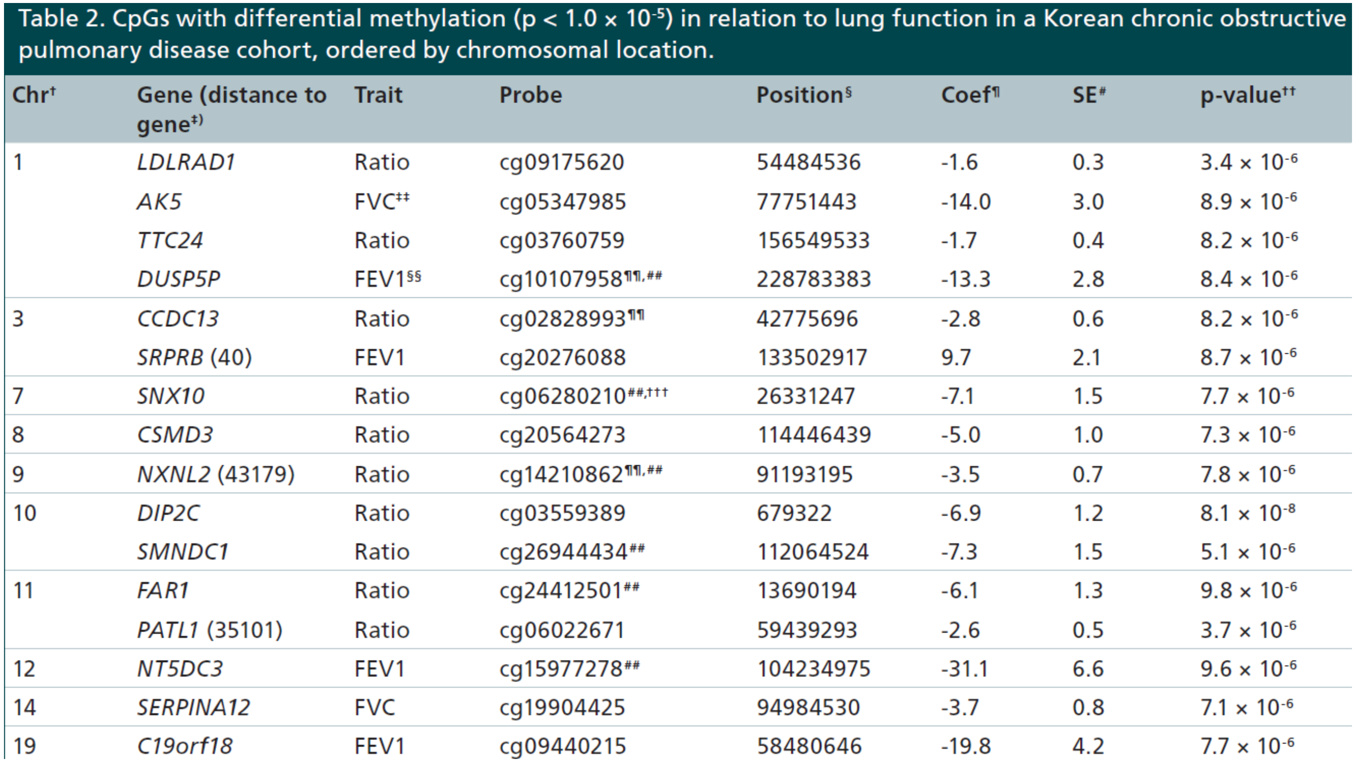
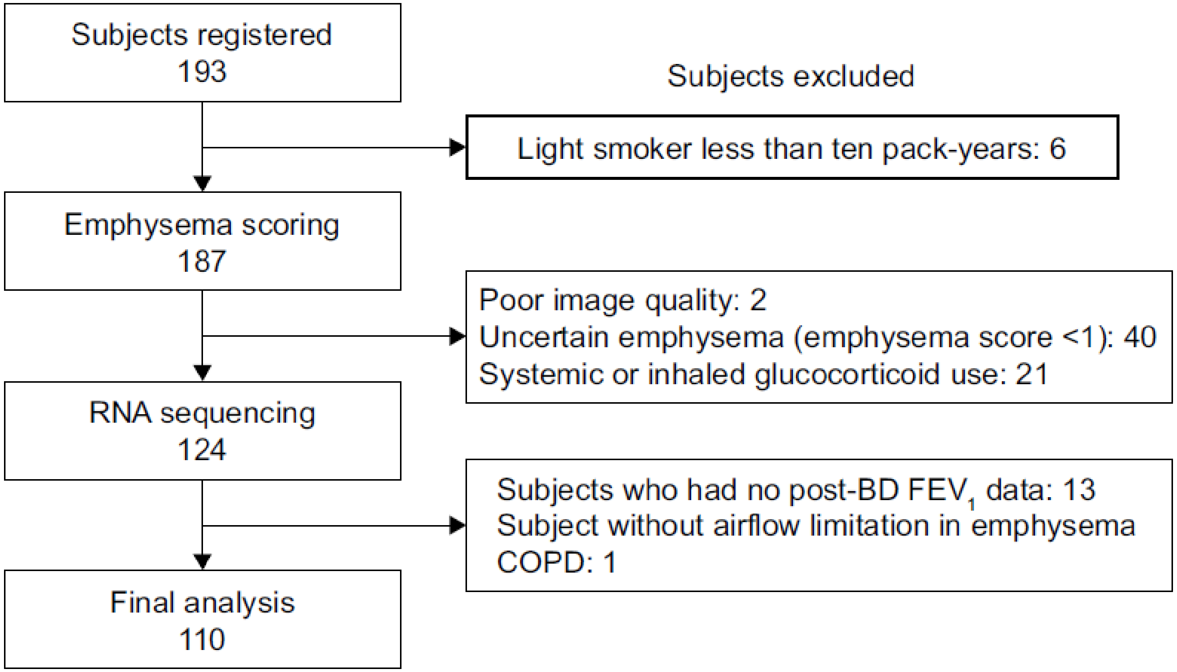


Table 2. CpGs with differential methylation ($p < 1.0 \times 10^{-5}$) in relation to lung function in a Korean chronic obstructive pulmonary disease cohort, ordered by chromosomal location.

Chr[†]	Gene (distance to gene[‡])	Trait	Probe	Position[§]	Coef[¶]	SE[#]	p-value[††]
1	*LDLRAD1*	Ratio	cg09175620	54484536	-1.6	0.3	3.4×10^{-6}
	AK5	FVC[‡‡]	cg05347985	77751443	-14.0	3.0	8.9×10^{-6}
	TTC24	Ratio	cg03760759	156549533	-1.7	0.4	8.2×10^{-6}
	DUSP5P	FEV1[§§]	cg10107958[¶¶,##]	228783383	-13.3	2.8	8.4×10^{-6}
3	*CCDC13*	Ratio	cg02828993[¶¶]	42775696	-2.8	0.6	8.2×10^{-6}
	SRPRB (40)	FEV1	cg20276088	133502917	9.7	2.1	8.7×10^{-6}
7	*SNX10*	Ratio	cg06280210[##,†††]	26331247	-7.1	1.5	7.7×10^{-6}
8	*CSMD3*	Ratio	cg20564273	114446439	-5.0	1.0	7.3×10^{-6}
9	*NXNL2* (43179)	Ratio	cg14210862[¶¶,##]	91193195	-3.5	0.7	7.8×10^{-6}
10	*DIP2C*	Ratio	cg03559389	679322	-6.9	1.2	8.1×10^{-8}
	SMNDC1	Ratio	cg26944434[##]	112064524	-7.3	1.5	5.1×10^{-6}
11	*FAR1*	Ratio	cg24412501[##]	13690194	-6.1	1.3	9.8×10^{-6}
	PATL1 (35101)	Ratio	cg06022671	59439293	-2.6	0.5	3.7×10^{-6}
12	*NT5DC3*	FEV1	cg15977278[##]	104234975	-31.1	6.6	9.6×10^{-6}
14	*SERPINA12*	FVC	cg19904425	94984530	-3.7	0.8	7.1×10^{-6}
19	*C19orf18*	FEV1	cg09440215	58480646	-19.8	4.2	7.7×10^{-6}

Epigenomics (2017) 9(7), 971–984

Gene expression profile of human lung in a relatively early stage of COPD with emphysema



Subjects registered
193

Subjects excluded

Light smoker less than ten pack-years: 6

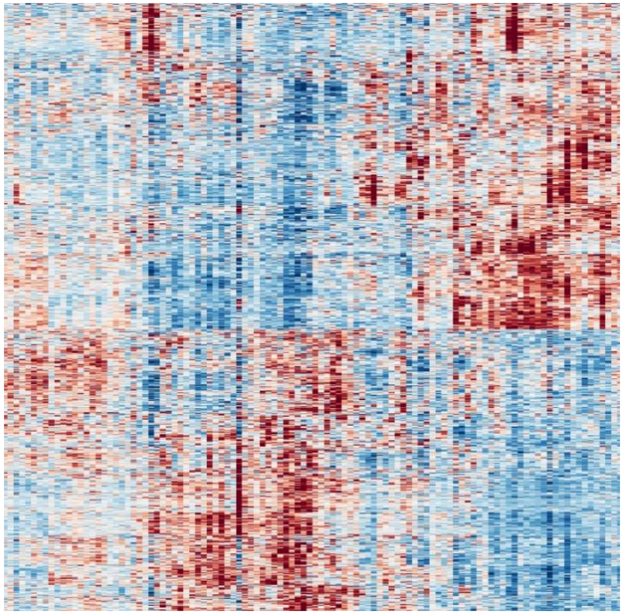
Emphysema scoring
187

Poor image quality: 2
Uncertain emphysema (emphysema score <1): 40
Systemic or inhaled glucocorticoid use: 21

RNA sequencing
124

Subjects who had no post-BD FEV_1 data: 13
Subject without airflow limitation in emphysema COPD: 1

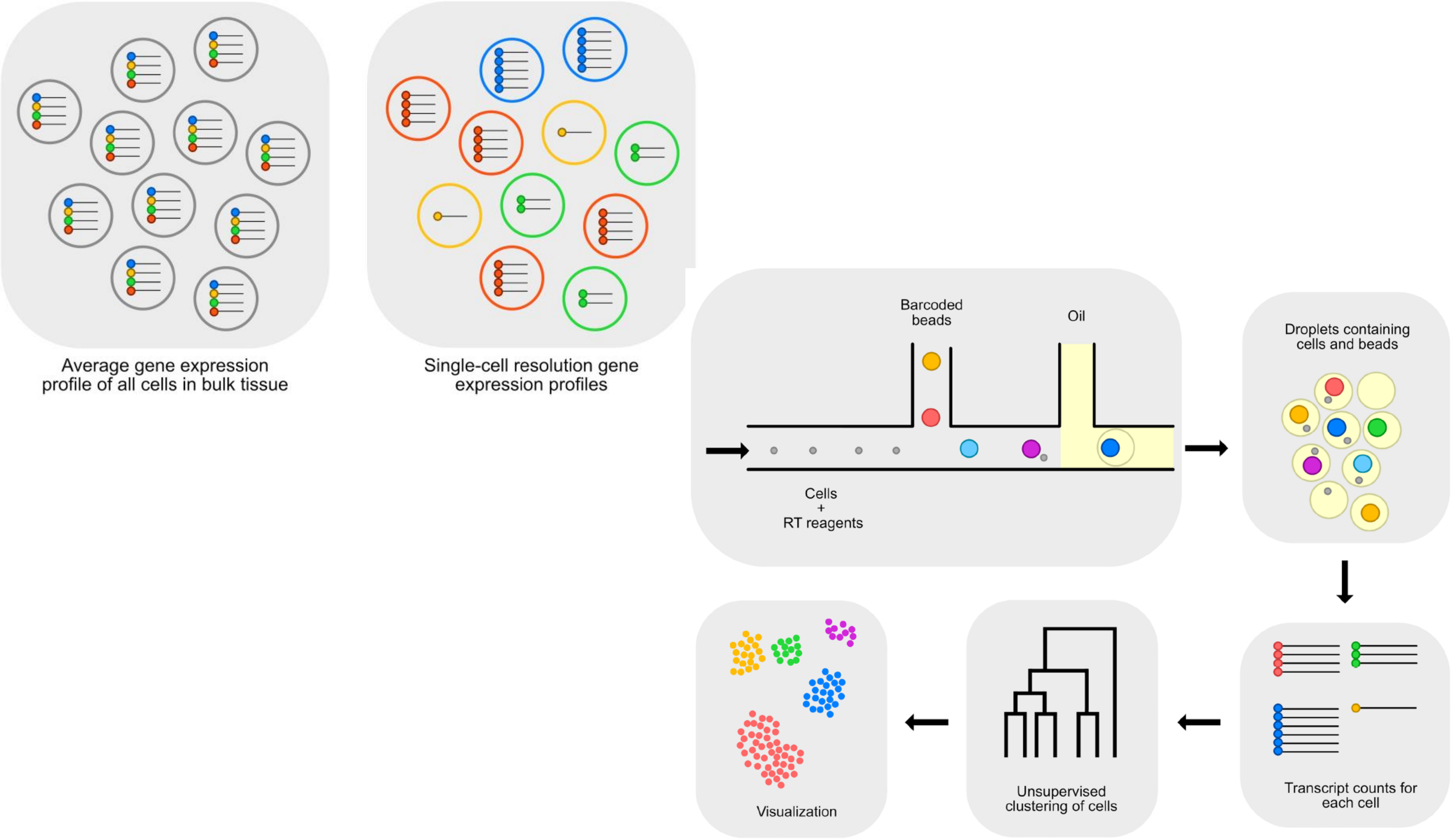
Final analysis
110



Terms	Count of genes involved	Fold enrichment	FDR
Representative DAVID results for pathways with decreased expression in COPD with emphysema subjects compared with healthy smokers			
GO:0005739–mitochondrion	71	1.96	1.03×10^{-4}
GO:0070062–extracellular exosome	120	1.57	4.14×10^{-4}

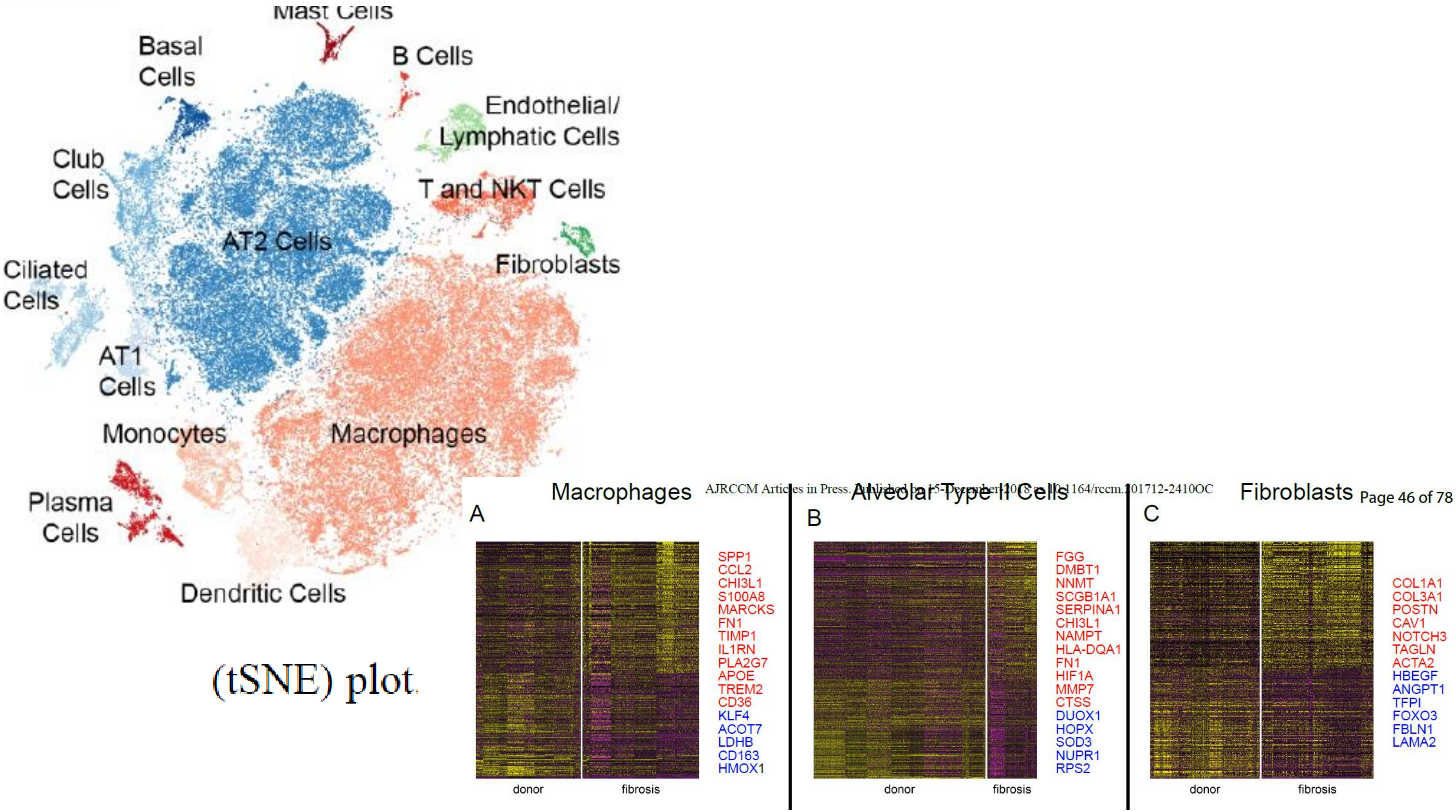
International Journal of COPD 2018:13 2643–2655

Transcriptomics and single-cell RNA-sequencing



Respirology (2019) **24**, 29–36

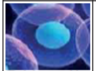
Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis

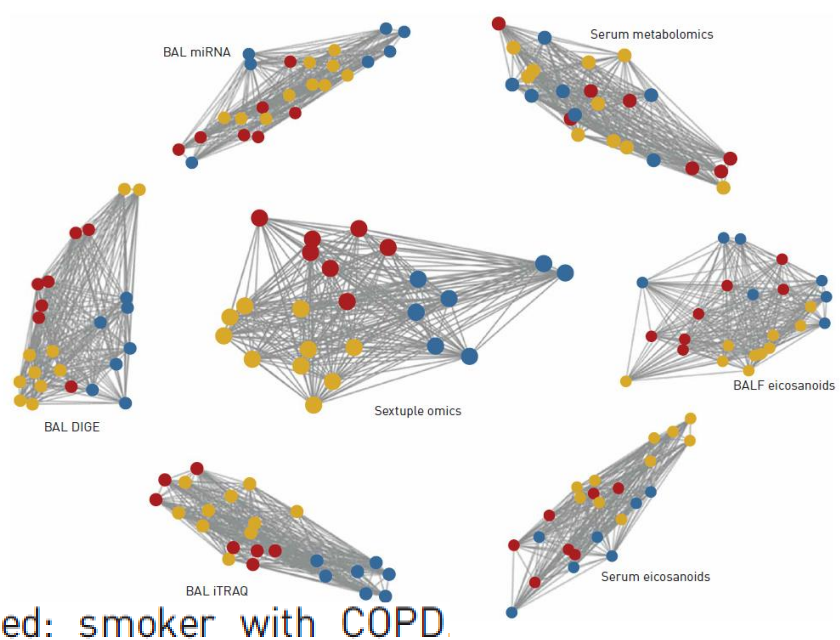


(tSNE) plot.

AJRCCM, in press

Integration of multi-omics datasets enables molecular classification of COPD

	mRNA	miRNA	Proteins		Metabolites	
	Micro-array	Micro-array	DIGE	iTRAQ/TMT	Metabolome	Eicosanoids
 Serum					*	*
 BALF						*
 BAL cells	*	*	*	*		
 BEC				*		
 Exosome		*				



yellow: smoker with normal spirometry; red: smoker with COPD

Eur Respir J 2018; 51: 1701930

Summary

- EWAS
  - Methylation change due to environmental exposure
  - Disease associated CpG loci
- <u>Single cell multi-omics</u>
  - To understand biology of airway diseases