# Maternal smoking during pregnancy and DNA methylation in offspring

**Stephanie J. London, M.D., Dr.P.H.**

**Epidemiology Branch**
**National Institute of Environmental Health Sciences**

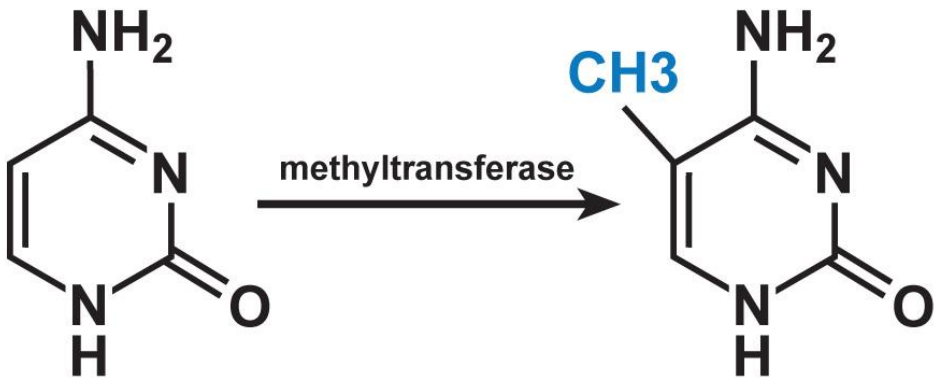
# Epigenetics

- Modifications to DNA that do not involve changes to the DNA sequence
  - Can affect gene expression
  - Can be inherited
- Various types of epigenetic modifications
  - Methylation – best studied in humans

# DNA Methylation

- Addition of methyl group to 5 position of cytosine

- C-phosphate-G sites (CpGs)



# Epigenetics and the *In Utero* Environment

- Much of the epigenome established in early development
- *In utero* exposures can impact disease susceptibility
- DNA methylation one proposed mechanism
  - Mouse models, esp. Agouti
    - Bisphenol A
    - Dietary methyl donors
  - Few human data

# Maternal Smoking During Pregnancy respiratory outcomes in children

- US Surgeon General's Report 2014 sufficient evidence of causality for:
  - ↓ Lung function
  - Early lower respiratory illnesses (parental smoking)
  - Childhood Asthma (parental smoking)

# Mechanism for effects of maternal smoking in pregnancy on offspring

- Not completely understood
- Epigenetic modifications may play a role

# Our Study

- Association of maternal smoking during pregnancy and DNA methylation in newborn cord blood DNA using the Illumina Methyl450K beadchip
- First study of an *in utero* exposure using this platform
- Joubert et al. *Environmental Health Perspectives* 2012

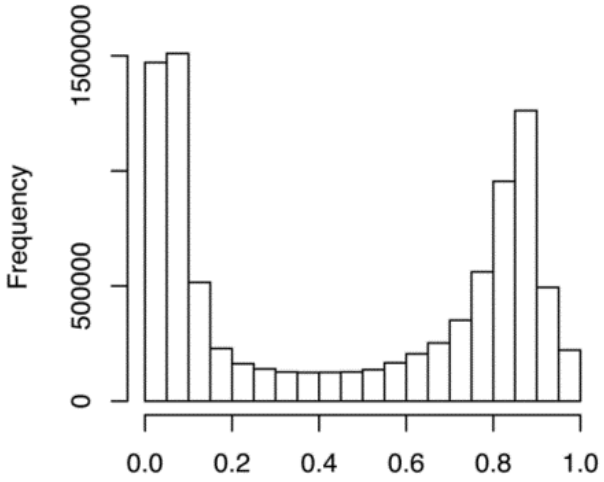
# Illumina Human Methylation450 (Illumina 450K)

Design guided by expert panel (2010) to provide reasonable genome wide coverage:

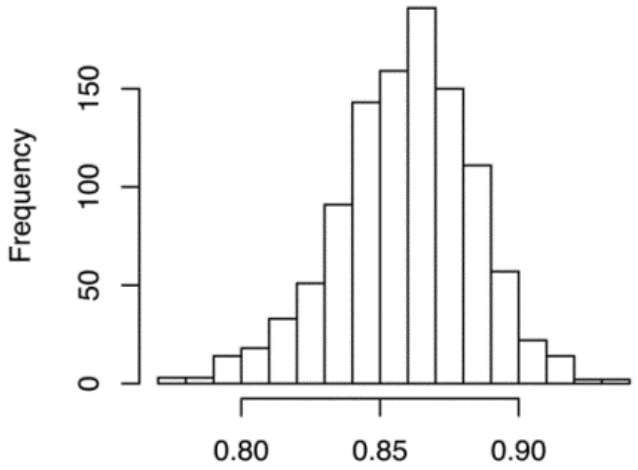
- 99% of Refseq genes
- 96% of CpG Islands
- CpG sites outside of CpG islands
- Other sites

# DNA Methylation Measurements

- 485,577 CpGs
- Beta (0 to 1)




Beta, all CpGs



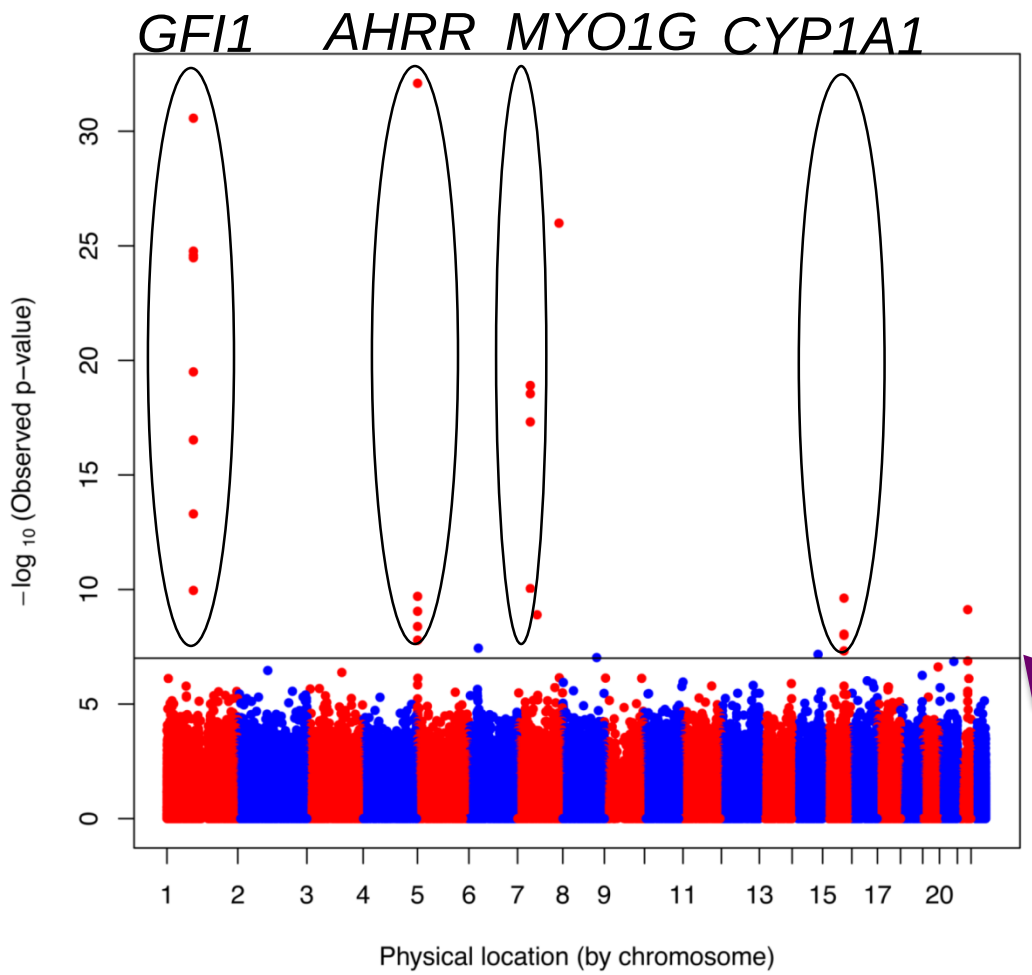

Beta, one CpG

# Norwegian Mother and Child Study (MoBa)

- MoBa – pregnancy cohort, N≈107,000
- Subset of 1,062 newborns with cotinine (biomarker of smoking) measurements in maternal plasma from week 18 of pregnancy
  - 13% consistent with active smoking
    - Generally light smokers

# Maternal smoking in pregnancy and newborn DNA methylation: MoBa

26 CpGs at 10 loci  $P < 1*10^{-7}$




**6 other loci:**

*RUNX1*

CNTNAP2

*HLA-DPB2*

*EXT1*

*TTC7B*

*ENSG00000225718*

Bonferroni correction 480K tests

Joubert et al., *EHP* 2012

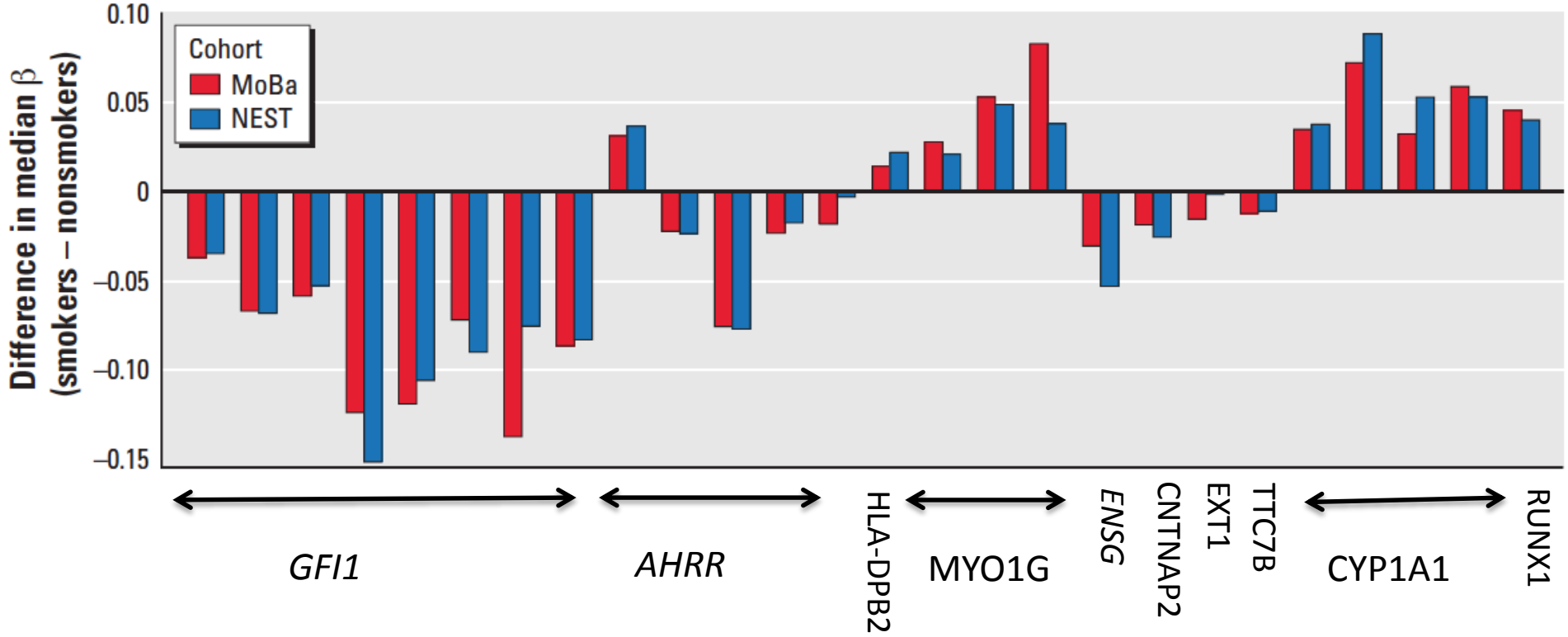
# Replication Study

- Duke Newborn Epigenetics Study (NEST)
  - Newborn cord blood DNA sampled on maternal smoking:
    - 18 w/ mothers who smoked during pregnancy
    - 18 w/ non-smoking mothers

# Striking Degree of Replication

- 3 genes (5 CpGs) replicated at strict Bonferroni look-up significance (26 tests)
    - ***AHRR*** and ***CYP1A1***: Aryl hydrocarbon receptor (AhR) signaling pathway
    - ***GFI1***: Diverse developmental processes
- Replication P values systematically smaller than expected by chance: Kolmogorov $P<0.00011$

# Differential Methylation by Maternal Smoking at all 26 CpGs: Similar Direction and Magnitude in Discovery and Replication Populations



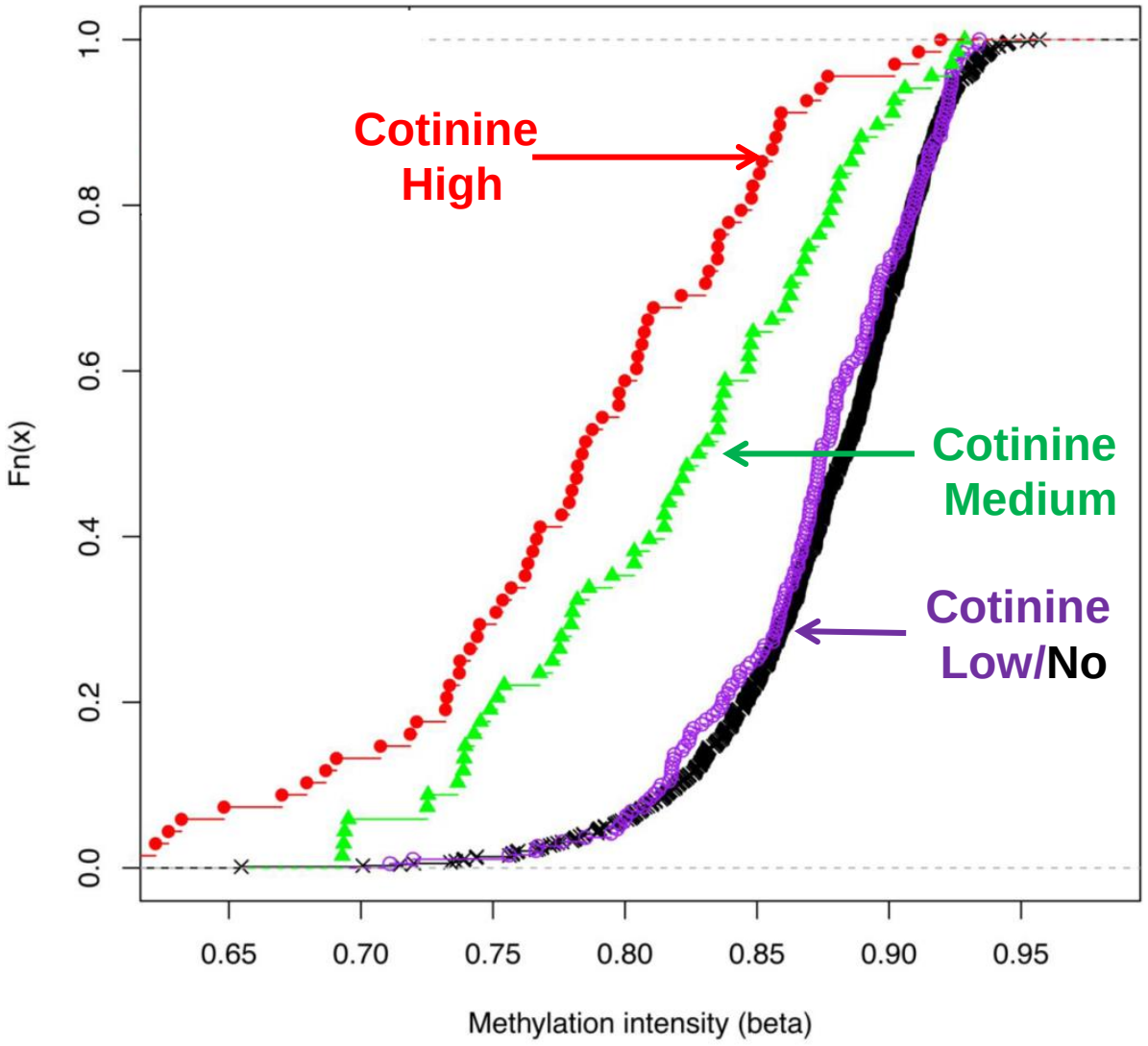

Effect size correlation – 0.97

Joubert et al., *EHP* 2012

# Subsequent consistent replication

- Nearly all of our findings, including those just below Bonferroni threshold, have confirmed in subsequent studies of newborns

# Smoking has dose-response effect on methylation Example: *AHRR* cg05575921




Jonkheree-Terpstra trend test
$p < 2.2 \times 10^{-16}$

Joubert et al., *EHP* 2012

# Same top CpG (*AHRR* cg05575921) associated with adult smoking

- 6+ studies with Illumina450K
- ↓ methylation in smokers
- Shenker et al. 2012
  - Discovery in whole blood
  - Lung tissue: *AHRR* ↓ methylation and ↑ expression in smokers vs. nonsmokers

# Conclusions

- Identified, and replicated, DNA methylation differences at birth from mothers' smoking in pregnancy
  - *AHRR* and *CYP1A1* in AhR pathway
    - Opposing effects of smoking on methylation in these two genes consistent with opposing effects in the AhR pathway
  - But also novel genes for smoking effects
    - *GFI1, CNTNAP2, MYO1G*
- Many of the same signals also seen for personal smoking in adults
- May support epigenetic mechanisms for effects of maternal smoking in pregnancy

# Follow-up questions

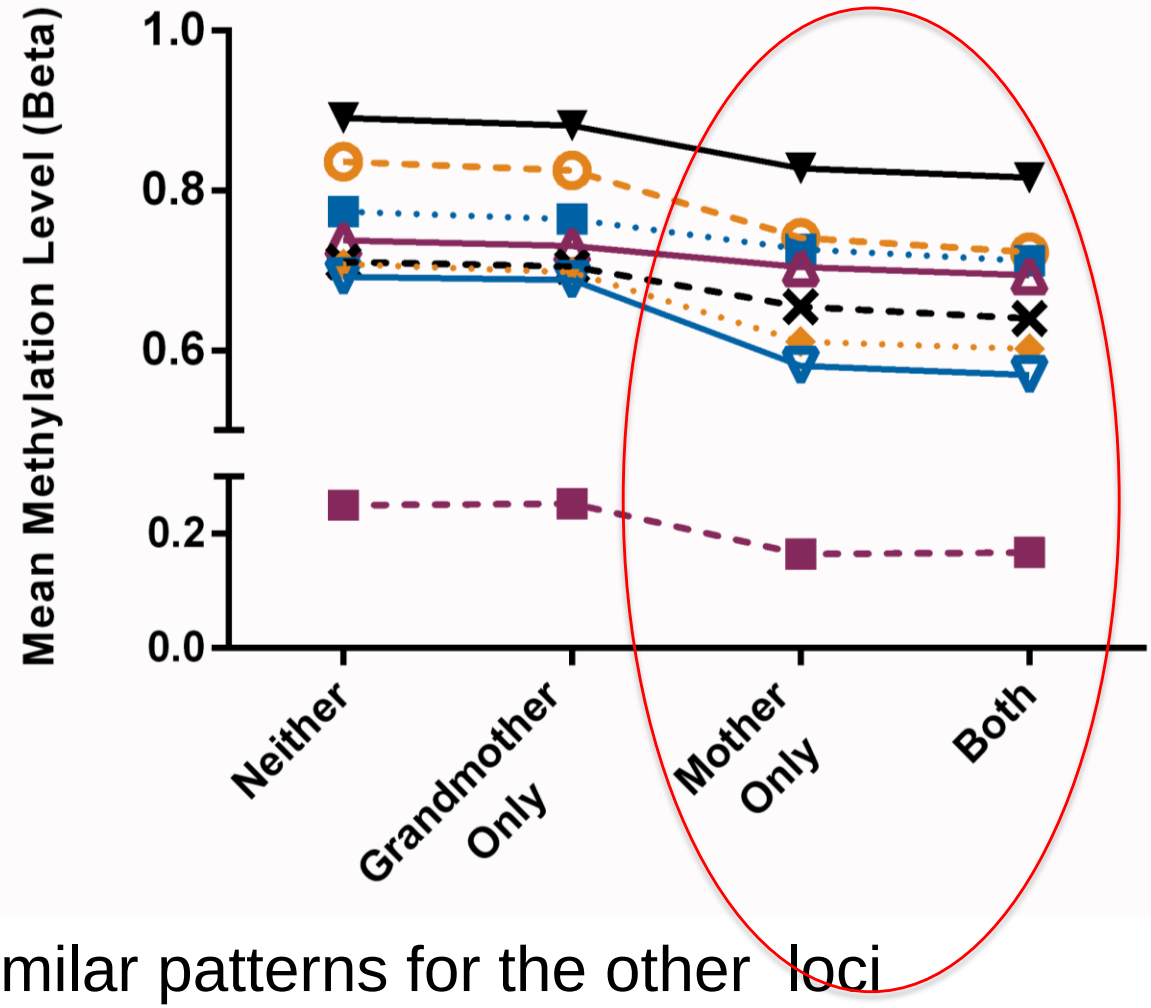
1. Do these findings reflect *in utero* exposure effects or epigenetic inheritance of smoking related methylation marks?
2. If *in utero* effects, is sustained exposure during pregnancy required?

# *In utero* exposure vs. epigenetic inheritance

- For the 26 Bonferroni significant CpGs, we evaluated methylation at birth in relation to smoking by:
  - Maternal grandmother
  - Father
  - Mother ending before pregnancy
  - Mother ending in early pregnancy (confirmed by cotinine measured ~18 weeks)
  - Mother sustained smoking in pregnancy
- Joubert et al. *Cancer Epidemiology Biomarkers and Prevention* (CEBP) 2014

# Conclusion:
# Little evidence for epigenetic inheritance

# Maternal grandmother smoking does not contribute




Ex: 7 CpGs in GFI1

Similar patterns for the other loci

Joubert *et al. CEBP* 2014

# Other evidence against epigenetic inheritance in these data

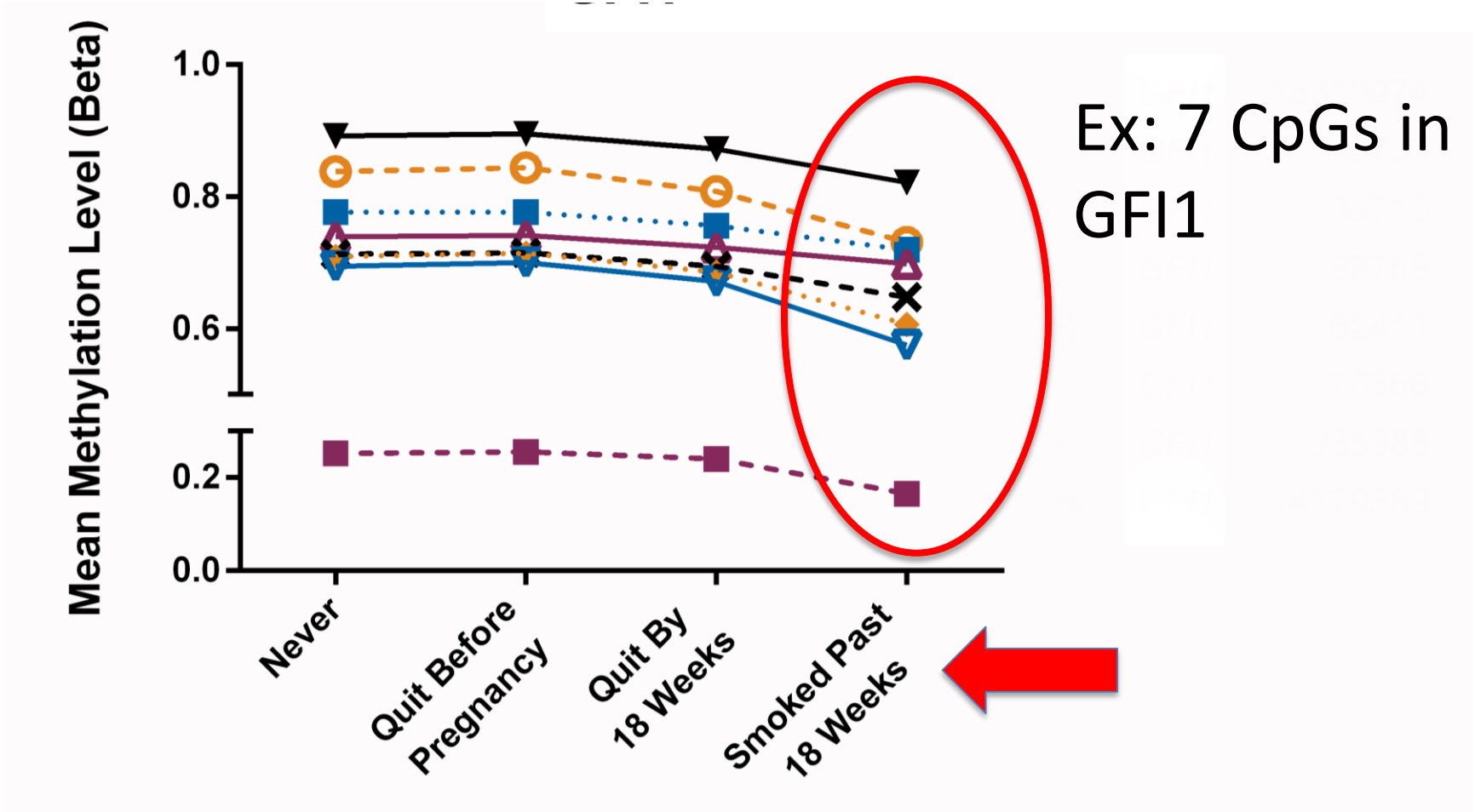
- No effect on DNA methylation in offspring for:
  - Father's smoking
  - Smoking by the mother that ended before pregnancy

Joubert *et al. CEBP* 2014

# Timing of exposure

- Methylation differences much stronger for sustained smoking during pregnancy than for smoking ending by week 18 (relative to no smoking during pregnancy)

# Strongest signals for sustained smoking




Similar patterns for the other loci

Joubert *et al. CEBP* 2014

# Conclusions

- Methylation differences by maternal smoking appear to reflect *in utero* exposure
  - Sustained exposure across pregnancy required
- No evidence for transgenerational epigenetic inheritance
  - Of great interest, but little evidence in mammals to date

# Next steps

Follow-up in consortia

# Consortia in genome wide studies

- In genome wide association studies (GWAS), combining studies using meta-analysis in consortia has led to the discovery of many novel loci that would have remained unknown if studies had published separately

# Consortia in genome wide methylation

- Same approach should be powerful for genome wide methylation studies
  - Although there are some technical issues for methylation that do not plague GWAS genotyping such as:
    - Batch effects
    - Normalization
    - Cell type correction
  - These can be addressed

# Issues for meta-analysis of methylation data: Batch

- GWAS genotyping calls much more robust than quantifying methylation
- Virtually no batch effects for genotyping calls
  - Robust to different methods of DNA extraction, age of extracted DNA
  - Cases and controls do not need to be randomized across plates
- Batch effects an issue with methylation

# Issues for meta-analysis of methylation data: Normalization

- Uniform technology across SNPs on a GWAS genotyping platform
- On the Illumina 450K, probe chemistry changed midway.
  - 1/3 of probes are Type 1, 2/3 are Type 2
  - Type 1 and Type 2 probes also differ in their genomic locations
  - So methylation distributions differ by Type 1/2

# Normalization

- Although type 1/type 2 should NOT influence estimation for individual CpGs, normalization now required
  - Many different normalization schemes published
  - People tended to be believe that the one they use was best – strong opinions with few data
  - Different schemes could introduce heterogeneity
- We (Mike Wu et al, Epigenetics 2014) found that normalization versus not had no influence on replicates nor on our smoking findings
  - But we now normalize for practical reasons

# Issues in Meta-Analysis of Methylation Data: Cell type

- Methylation values varies by cell type – genotypes do not
- Most studies used whole blood – most do not have complete blood counts (CBC)
- Even CBC do not break down lymphocytes into the major types
- Adjustment methods are imperfect
  - Available reference panel - 6 adult Swedish men
    - Reinius et al. 2012

# Issues in meta-analysis of methylation data: meaning of λ

- In GWAS elevated λ has a fairly clear interpretation
  - Uncontrolled population stratification
  - Or some error in the analysis, need to filter very rare variants, etc
- In methylation data interpretation of λ is unclear
  - Often > 1.2.
  - Often 1.0 when the QQ plot looks terrible

# PACE Consortium:

## **P**regnancy **A**nd **C**hild **E**pigenetics

- Spring 2013 – invited all studies with Illumina 450K data available on newborns to participate in consortium analysis of maternal smoking and methylation in offspring. 12 attended.
- Consortium has expanded since then

# Preliminary results: Maternal smoking in pregnancy and DNA methylation in newborns

PACE Consortium

# Harmonized Smoking Variables

- Mostly questionnaire-based smoking information from the mother
  - Cotinine used in one study
- Sustained smoking in pregnancy (primary variable)
  - Defined as woman smoking most of pregnancy
  - Methylation differences reflected sustained smoking by the mother in MoBa1 analysis (Joubert et al., *CEBP* 2014)
- Any smoking during pregnancy
  - Includes cohorts with too few sustained smokers or timing information too limited to identify sustained smokers

# PACE smoking meta-analysis participating cohorts

- ALSPAC
- BAMSE
- CHAMACOS
- CHS
- EARLI
- GALA
- GECKO
- GENERATION R
- IOW
- MEDALL-EDEN
- MEDALL-INMA
- MEDALL-PIAMA
- MOBA1 – NIEHS-MoBa
- MOBA2 – NIEHS-MoBa
- MOBA3 – IARC-MoBa
- Norway Clefts
- NEST
- SEED
- Project VIVA

# PACE: smoking sample size

- Total **Newborns: 6,615**
- **Sustained** maternal smoking in pregnancy: **915**
- **Any** maternal smoking in pregnancy: **1,561**

- Data on **3,226 older children** for look-up of top findings
  - Do the signals seen in newborns persist into later childhood?

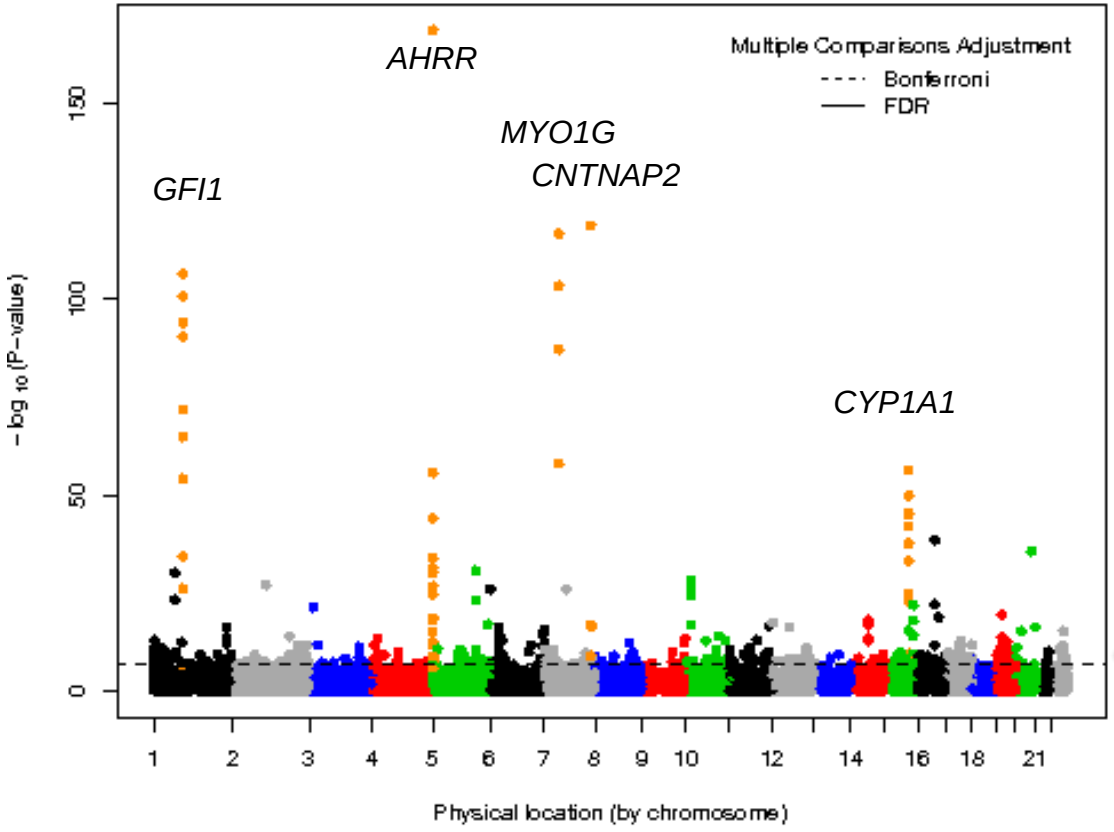
# Models

- Robust linear regression
  - Cpg ~ smoking + batch variable + maternal age + parity + SES variable + ancestry (as needed)
- Cohort specific normalization
  - Results very similar using raw betas

# Models

- Robust linear regression
  - Cpg ~ smoking + batch variable + maternal age + parity + SES variable + ancestry (as needed)
- Common model – no transformation or normalization of the methylation beta values
  - Batch correction by including term in model
- Favorite model - whichever normalization and batch correction cohorts prefer

# Methylation at birth and sustained smoking in pregnancy - raw betas - not normalized



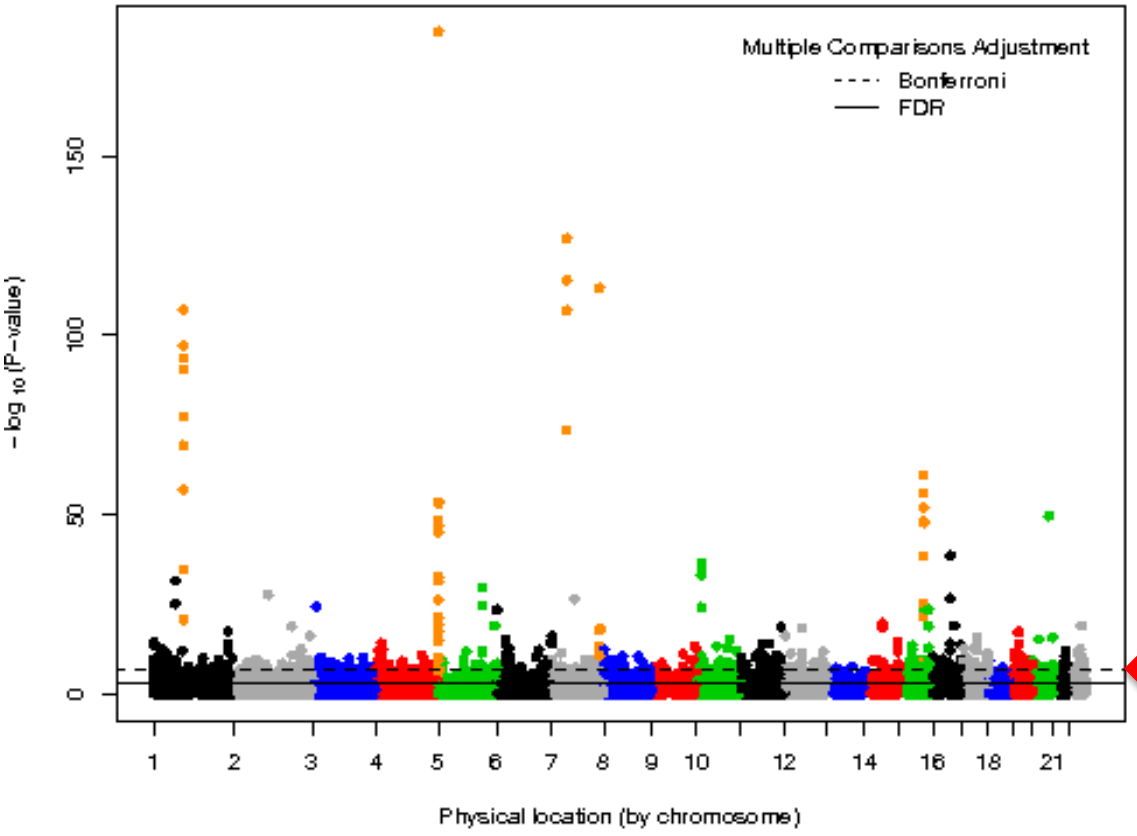

**445** Bonferroni significant CpGs*
(many more FDR significant CpGs)
Lambda = 2.1
Top hit: *AHRR* cg05575921

Joubert et al. 2012 loci in gold

Bonferroni significance
$P < 1*10^{-7}$

* Compared to **26** Bonferroni-significant CpGs from Joubert et al., *EHP* 2012

# Methylation at birth and sustained smoking in pregnancy: Normalized betas – various normalization schemes



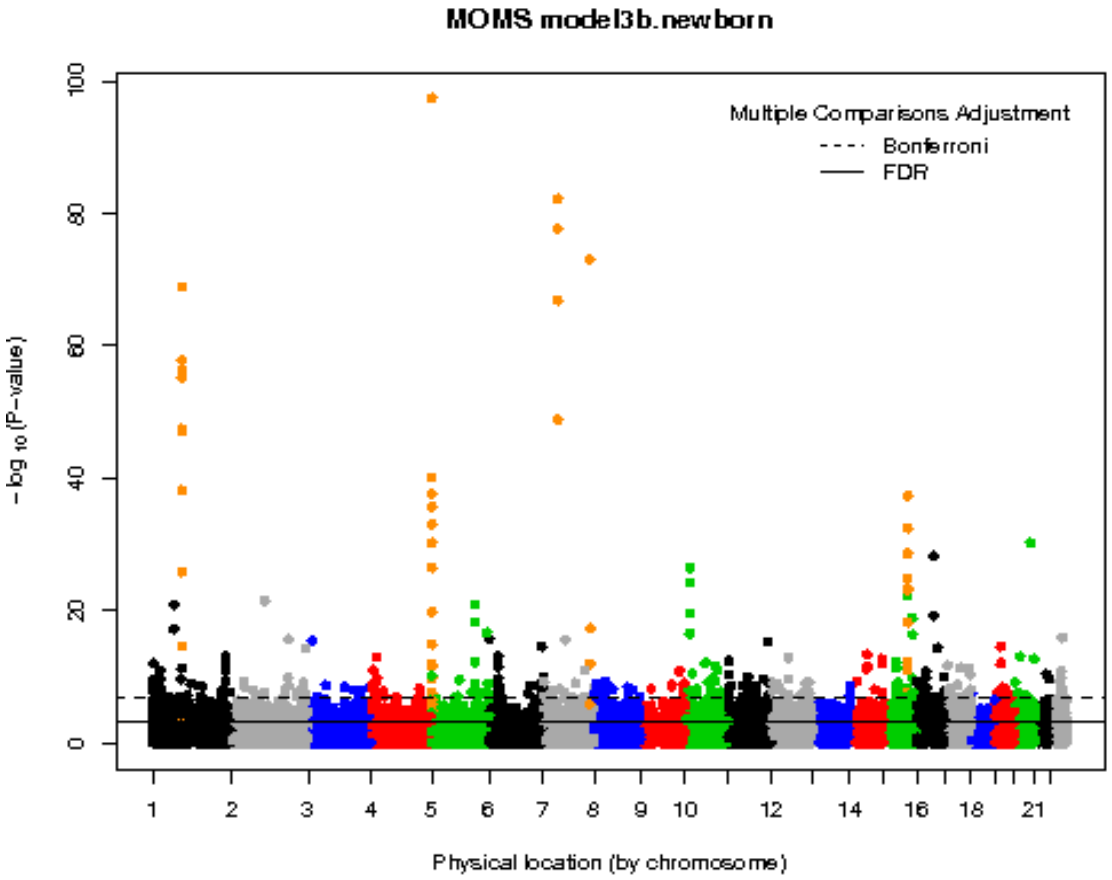

Similar results
Slightly better lambda value

538 Bonferroni significant CpGs
Lambda = 1.6
Top hit: *AHRR* cg05575921

Joubert et al. 2012 loci in gold

Bonferroni significance
$P < 1*10^{-7}$

# Methylation at birth and ANY smoking in pregnancy Normalized




As expected given Joubert et al. *CEBP* 2014:

Fewer findings for ANY smoking than for SUSTAINED smoking

339 Bonferoni significant CpGs
Lambda = 1.4
Top hit: *AHRR* cg05575921

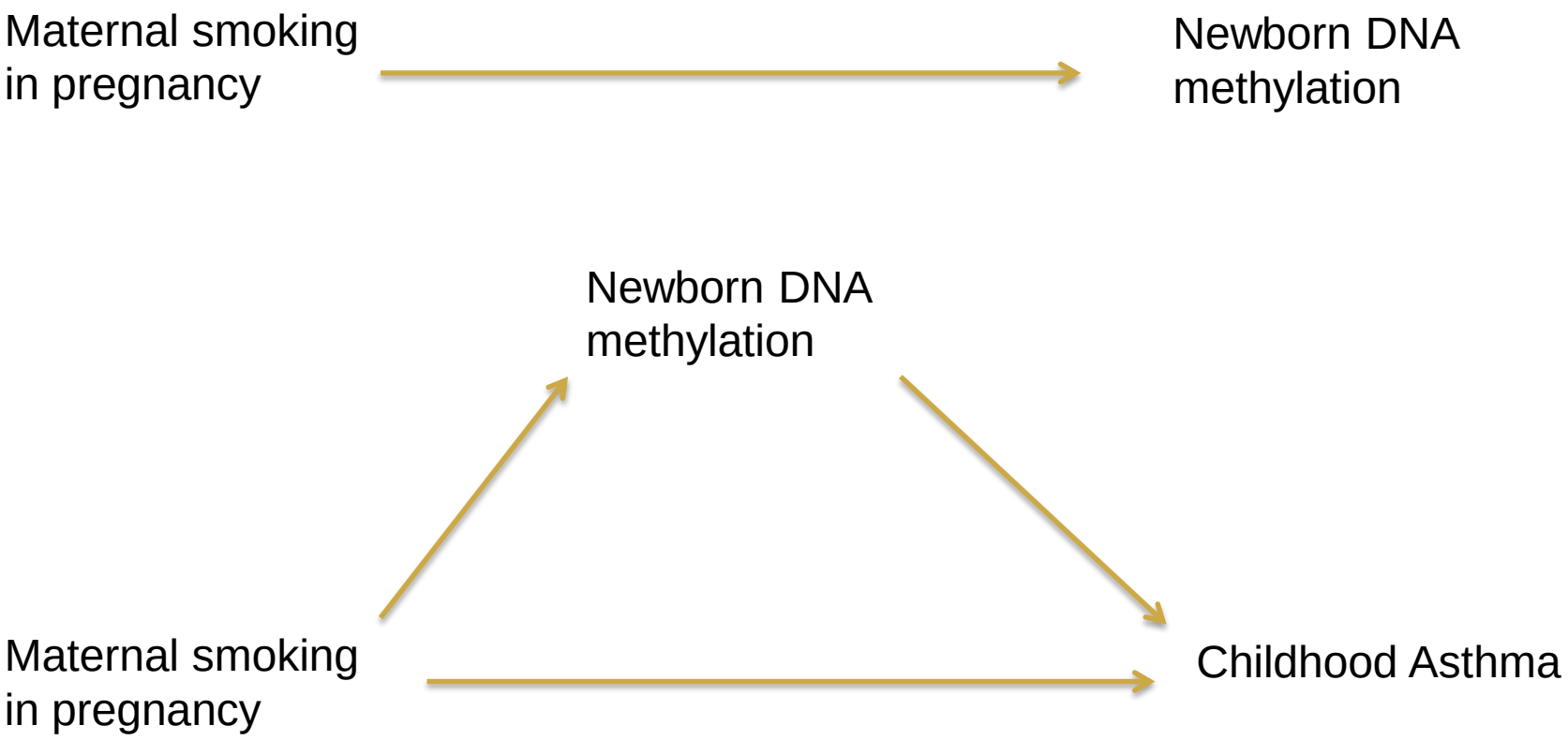
Joubert et al. 2012 loci in gold

# Smoking meta-analysis

- Findings robust to cell type correction
- Manuscript will be submitted soon
  - Larger number of statistically significant findings than we can easily summarize or integrate
    - Pathway and network analysis not so useful for summarizing
  - Correlated with gene expression

# Mediation in methylation studies

# Does methylation mediate asthma risk from maternal smoking?



# Maternal Smoking and Methylation: Mediator or Biomarker?

- Are the differentially methylated loci on the causal pathway between maternal smoking and related health outcomes?
- OR are these signal just robust biomarkers incorporating dose and duration of exposure?
- Unknown -
  - BUT they will likely look like mediators in analyses with self-reported smoking and ANY outcome
  - Measurement error methods need to be applied and interpretations moderated

# Asthma and methylation signals for *in utero* methylation

- The significant hits for in utero smoking in newborns are enriched for inflammatory pathways
- Asthma genetic signals are not over-represented

# Does methylation mediate genetic signals?

- Some reports that methylation mediates genetic associations with disease
  - Ex: Rheumatoid Arthritis – Liu *et al. Nat Biotechnol* 2013
- Problem in inference – batch effects
  - Essentially NO batch confounding for GWAS genotype
  - Batch confounding with case control status an issue with methylation platforms
  - Can make methylation appear to mediate genotype associations at a given locus because retains correlation with outcome whereas genotype does not

# Methylation and asthma in the PACE Consortium

- Meta-analysis of asthma and wheezing phenotypes at preschool and school age underway.
- Large numbers of cases needed – as with GWAS

# MoBa Methylation Project Collaborators

- NIEHS
  - Bonnie Joubert
  - Shyamal Peddada
  - Douglas Bell
  - Xuting Wang
  - Lea Cupul-Uicab
- Norwegian Institute of Public Health
  - Siri Håberg
  - Wenche Nystad
  - Per Magnus Ueland
  - Stein Emil Vollset
  - Roy Nilsen
- University of North Carolina at Chapel Hill
  - Michael Wu
- Duke University
  - Cathrine Hoyo
  - Susan Murphy
  - Zhiqang Huang
- Bevital A/S Laboratory
  - Øivind Midttun

# PACE Consortium –smoking participants

Carrie Breton

Olena Gruzieva

Goran Pershagen

Caroline Relton

Rebecca Richmond

Nina Holland

Paul Yousefi

Kelly Beluski

Dani Fallin

Andrea Baccarelli

Rosalyn Wright

Alan Just

Cathrine Hoyo

Susan Murphy

Liesbeth Duijts

Janine Felix

Vincent Jaddoe

Sam Oh

Leanne Kupers

Harold Sneider

Wilfried Karmaus

Hongmei Zhang

Bonnie Joubert

Sarah Reese

Siri Haberg

Sarah Reese

Dirkje Postma

Cheng Xu

Christina Markunas

Allen Wilcox

Wenche Nystad

Jack Taylor

Chris Ladd-Acosta

Shyamal Peddada

Stephanie London

And others

Thanks!

london2@niehs.nih.gov