Programming of the respiratory epithelium *in utero* - insight from the amniotic epithelial methylome

Patricia Agudelo-Romero[1,2,3*†], Thomas Iosifidis[1,4,5,*], James Lim[1], Nina Kresoje[1], David G. Hancock[1,6,7], Guicheng Zhang[1,4], Abhinav Sharma[8], Talya Conradie[1], Yuliya V. Karpievitch[1,4,9], Desiree T. Silva[7,10,11,12], Anthony Bosco[13,14,] Susan L. Prescott[7,10,15,16], Peter N. LeSouëf [1,17], Elizabeth Kicic-Starcevich[1], Anthony Kicic[1,5,18,19], David J. Martino[1,9,#], Stephen M. Stick[6,#] .

[*]co-first authors, [#]co-senior authors, [†]corresponding author

Affiliations:

1: Wal-yan Respiratory Research Centre, The Kids Research Institute Australia, Perth, 6009, Western Australia, Australia.

2: School of Molecular Sciences, The University of Western Australia, Perth, Western Australia 6009, Australia.

3: European Virus Bioinformatics Centre, Friedrich-Schiller-Universitat Jena, 07737 Jena, Germany.

4: School of Population Health, Curtin University, Perth, 6102, Western Australia, Australia.

5: Centre for Cell Therapy and Regenerative Medicine, School of Medicine and Pharmacology, The University of Western Australia, Perth, 6009, Western Australia, Australia.

6: Department of Respiratory and Sleep Medicine, Perth Children's Hospital Perth, Perth, 6009, Western Australia, Australia.

7: Medical School, The University of Western Australia, Perth, 6009, Western Australia, Australia.

8: DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, SAMRC Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, 7700, South Africa.

9: School of Biomedical Sciences, The University of Western Australia, Perth, 6009, Western Australia, Australia.

10: The Kids Research Institute Australia, Perth, 6009, Western Australia, Australia.

11: School of Medical and Health Sciences, Edith Cowan University, Perth, 6027, Western Australia, Australia.

12: Department of Paediatrics and Neonatology, Joondalup Health Campus, Perth, 6027, Western Australia, Australia.

13: Asthma and Airway Disease Research Center University of Arizona, Tucson, 85721, Arizona, United States.

14: Department of Immunobiology, The University of Arizona College of Medicine, Tucson, 85721, Arizona, United States.

15: Nova Institute for Health, Baltimore, 21231, Maryland, United States.

16: Department of Family and Community Medicine, University of Maryland, Baltimore, 21231, Maryland, United States.

17: School of Paediatrics and Child Health, University of Western Australia, Perth, 6027, Western Australia, Australia.

18: Occupation, Environment and Safety, School of Population Health, Curtin University, Perth, 6102, Western Australia, Australia.

19: Department of Respiratory and Sleep Medicine, Perth Children's Hospital, Perth, 6009, Western Australia, Australia.

***Co-first authors**

[#]**Co-senior authors**

[†]**Correspondence:**

To whom correspondence should be addressed: Dr Patricia Agudelo-Romero; Address: The Kids Research Institute Australia, Perth Children's Hospital, 15 Hospital Avenue, Nedlands, Western Australia, 6009; Email: patricia.agudeloromero@thekids.org.au

Sources of support

This work was supported by grants from the National Health and Medical Research Council of Australia (APP1157548), Department of Health (Western Australia)-Future Health Research and Innovation Fund (2020, 2021, 2022). S.M.S. is supported by an NHMRC Investigator Grant (NHMRC2007725). P.A-R. received funding from the Google Cloud Education Program, a Telethon Kids Institute Theme Collaboration Award grant (PR030564), the Branchi Family Foundation, and a Future Health Research and Innovation (FHRI) Fellowship by the Department of Health (IF2024-25/1), Government of Western Australia. T.I. is supported through the Channel 7 Telethon Trust, Stan Perron Charitable Foundation People Fellowship and previously supported by the Future Health Research Innovation Fund (FHRIF 2020-2023) and Imogen Miranda Suleski Fellowship. A.K is a Rothwell Family Fellow and D.G.H. is a Stan Perron/Perth Children's Hospital Foundation (PCHF) Fellow. D.M. is supported by FHRIF. A.B. is supported by the NIH (R21 AI176305-01A1, R01AI099108-11A1). The ORIGINS birth cohort has received core funding support from the Paul Ramsay Foundation and the Commonwealth Government of Australia through the Channel 7 Telethon Trust. Substantial in-kind support has been provided by The Kids Research Institute Australia and Joondalup Health Campus.

Running title

Amnion as a surrogate for the airway epithelium

Subject Category

1.12 Clinical asthma
1.17 Epidemiology (Pediatric): Risk Factors < PEDIATRICS
3.40 Airway epithelium < CELL AND MOLECULAR BIOLOGY
3.05 Epigenetics < CELL AND MOLECULAR BIOLOGY
3.02 Bioinformatics/Biological Computing < CELL AND MOLECULAR BIOLOGY

Keywords: Epithelium; airway; amnion; in-utero; exposures; epigenetics; asthma; smoking; DNA methylation; EWAS.

Conflict of Interest declaration: A.B. is a co-founder, equity holder, and director of the startup company Respiradigm Pty Ltd that is related to this work. A.B. is the founder of the startup company INSiGENe Pty Ltd that is related to this work. All other the authors declare that they have no affiliations with or involvement in any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript.

Abstract

Background: Dysregulation of the airway epithelium contributes to recurrent wheezing and asthma and may have developmental origins. Here, we investigated the relationship between the placental amniotic and nasal epithelial methylation landscapes to determine whether amniotic epithelium provides insight into fetal programming of respiratory tissue.

Methods: We conducted high-throughput target-capture DNA methylation sequencing of 84 matched pairs of placental amniotic and neonatal nasal brushings samples within the Airway Epithelium Respiratory Illnesses and Allergy (AERIAL) cohort. Comparative analysis of tissue-specific methylation profiles, and conservation of methylation changes associated with gestational exposures (maternal smoking and maternal asthma), was explored.

Results: Between amniotic and nasal tissues, we identified 4,897 differentially methylated regions (FDR $\leq$ 0.05 and $\log_2$FC $\geq$ |0.2|) that were generally hypermethylated in the nasal epithelium. Despite these extensive tissue-specific differences, filtering for loci with non-significant differential methylation (FDR $\geq$ 0.1) revealed 1,493,976 CpG loci (~20% of the measured methylome) with highly concordant methylation ratios levels between tissues (Pearson's R $\geq$ 0.8). These loci included genes crucial to epithelial and lung development. Within these conserved regions, associations with maternal asthma and prenatal smoking were consistently represented in both tissues.

Conclusions: The conserved methylome signatures support the use of amniotic tissue as a valuable tissue for investigating the developmental programming of airway vulnerability, potentially leading to early risk stratification and targeted interventions for childhood asthma.

114 **Introduction**

115 Asthma is a chronic inflammatory airway disease, among the most frequent causes of
116 hospital admissions in children and represents a significant global disease burden (1–3). A
117 substantial body of evidence suggests prenatal exposures play a critical role in shaping newborn
118 lung function (4), and contribute to the airway epithelial cell dysfunction observed in children
119 with asthma (5). Intrauterine exposure to air pollution and tobacco smoke are well-established risk
120 factors for childhood asthma that impair fetal lung development and increase airway
121 hyperresponsiveness in offspring (6). Similarly, poorly controlled maternal asthma in pregnancy
122 has been associated with childhood wheeze and increased risk for asthma in offspring through
123 alterations in fetal development (7) Animal models also suggest intrauterine chronic inflammatory
124 conditions might directly influence fetal-placental development (8) and aberrant prenatal
125 programing of the airway epithelium tissues (9). Understanding these developmental programming
126 mechanisms is a research priority needed to facilitate early detection and targeted interventions.

127 Epigenetic factors such as DNA methylation play a crucial role in fetal lung development,
128 and changes in DNA methylation have been observed in the airway epithelium in children with
129 asthma (10, 11). Environmental exposures such as cigarette smoke and pollution induce reversible
130 epigenetic changes in airway epithelial cells via oxidative stress (12, 13), disrupting developmental
131 gene expression patterns, and contributing to asthma development (14, 15). Methylation may
132 therefore be an important epigenetic mechanism through which to understand the complex
133 interplay between the prenatal environment and fetal epithelial programming in the development
134 of childhood asthma.

135 A key challenge to studying the prenatal origins of epithelial dysfunction relates to the
136 challenges of sampling the newborn airway epithelial cells, which are difficult to obtain before the
137 onset of key postnatal events such as colonization of the mucosal surfaces, and exposure to
138 airborne antigens. To address this, we explored the utility of fetal amniotic epithelial tissue as a
139 surrogate for airway epithelial cells (16). The placental amnion is exposed to the maternal
140 environment and may harbor epigenetic signatures (17, 18) that reflect processes occurring in the
141 developing fetal respiratory tract. Amniotic epithelium cells possess stem cell like characteristics
142 and are also attractive targets for cellular therapy (19). In this observational study, we explored the
143 landscape of DNA methylation in amniotic epithelium and matched nasal epithelial samples from
144 a cohort of healthy term newborns from the Airway Epithelium Respiratory Illnesses and Allergy
145 (AERIAL) study (20). We conducted a comparative analysis of genome-wide DNA methylation
146 profiles, and examined the relative conservation of DNA methylation patterning at specific genes
147 crucial for lung and epithelial development. We also explored whether methylation changes
148 associations with key maternal exposures that influence asthma development were conserved in
149 both tissues. Our results inform ongoing developmental programming studies on epithelium
150 vulnerability.

151 **Methods**

152 **Study Design**

153 AERIAL (20) is a prospective birth cohort nested within The ORIGINS birth cohort (21,
154 22). Matched pairs of amniotic membrane biopsies, nasal brushings and maternal urine were
155 obtained from 84 newborns and their mothers. Maternal asthma history and tobacco smoke
156 exposure during pregnancy was ascertained from at least 2 consecutive questionnaires
157 administered at 20, 28 and 36 weeks of pregnancy. This study was conducted in accordance with
158 the Declaration of Helsinki and was approved by the Ramsey Health Care HREC WA-SA (#1908).
159 All parents, guardians, or next of kin provided written informed consent to participate in this study
160 and collect maternal and newborn samples for the downstream analyses presented in this article.

161 **Sample Collection and Processing**

162 Placentas were processed within 48 hours post-birth on average (20.4 ± 13.1 hours [SD
163 (Standard Deviation)]), with the chorion membrane manually separated (23) and the amnion
164 membrane sampled. Matched nasal epithelial samples from newborns were collected within six
165 weeks from birth (15.9 ± 7.9 days [SD]) post-birth as detailed in the study protocol (20). All
166 samples were cryopreserved at -80°C until DNA extraction was performed.

167 **Nucleic Acid Isolation**

168 Genomic DNA was extracted from both nasal brushings and amniotic membrane samples
169 using the chemagic 360 automated nucleic acid isolation system (Revvity, Baesweiler, Germany)
170 and the chemagic™ DNA Blood 400 Kit H96 (Revvity, part# CMG-1091), following the
171 manufacturer's instructions, and then stored at -80 °C until analysis. DNA quantity was assessed

172 using the Qubit HS dsDNA Assay (Q32854, Thermo Fisher Scientific) on a Qubit fluorometer
173 (Thermo Fisher Scientific, Waltham, MA) (Figure E1 in the online supplement).

174 **DNA methylation measures**

175 Libraries were constructed using capture DNA methylation sequencing with enzymatic
176 conversion (EM-seq) and target enrichment employing the TWIST Human Methylome Panel.
177 Libraries were prepared from 200 ng of genomic DNA using the NEB Next Enzymatic Methyl-
178 seq Library Preparation Protocol (Twist Bioscience and New England Biolabs, CA, USA)
179 following the manufacturer's instructions. Targeted capture was performed using the Twist
180 Targeted Methylation Sequencing protocol. Pre-capture libraries were pooled in 8-plex format and
181 hybridized at 60°C for 16 hours. Subsequently, the hybridized pools were washed, and PCR
182 amplified for six cycles according to the manufacturer's protocol. Finally, capture libraries were
183 sequenced at the Genomics WA facility on an Illumina NovaSeq 6000 (Illumina, CA, USA) using
184 a pair-end configuration with 150 base pair (bp) length reads (Figure E1 in the online supplement).

185 **Bioinformatic Analysis**

186 Bioinformatics analyses utilized computing and data resource provided by the Australian
187 BioCommons Leadership Share (ABLeS) program (24) and Pawsey Supercomputing Research
188 Centre (25, 26). Figure E1 summarizes the main steps carried out during the bioinformatics
189 analyses. Briefly, raw methylation FASTQ files were processed using the nf-core/methylseq v2.3.0
190 pipeline. Methylseq pipeline was executed under Nextflow v23.04.2 (Di Tommaso et al., 2017)
191 and the Human Genome Reference Consortium Human Build 38 (GRCh38) using the BWA-
192 meth/MethylDackel workflow. Subsequent metrics and filtering were performed using target-
193 methylseq-qc pipeline v2.1.0 (28), which included targeted methylation coverage assessment via

194 *Picard-Profiler* mode and generation of targeted BED files based on BED intervals from the
195 TWIST Human Methylome Panel using *bed-filter* mode. The Kids Research Institute Australia
196 cluster and Google Cloud infrastructure were utilized for further analyses. Relevant scripts for data
197 pre-processing are available on GitHub (https://github.com/wal-
198 yan/AERIAL/tree/main/RespEpithelium_Methylome).

199 **Statistical Analyses**

200 Statistical analyses were conducted in R language v4.3.2 and RStudio v2023.03.0+386
201 (29). Figure E1 in the online supplement illustrates the flow of the analyses performed. Where
202 specific packages are not mentioned, analyses were conducted using base R functions. During data
203 quality control (QC), 2.58% of CpG sites (n=204,963) with zero coverage in any sample were
204 removed. Additionally, 9.70% of CpG loci (n=770,067) with very low coverage (<5) or extremely
205 high coverage (>500) were excluded. The remaining 87.72% of CpG sites (n=6,961,516) that
206 passed QC were used for subsequent downstream analysis. Mitochondrial and non-standard
207 chromosomes were removed from the dataset. Methylation ratios were derived from sequencing
208 counts and expressed as beta (β) as follows:

209 $$\frac{methylated\ alleles}{(unmethylated + methylated) * 100}$$

210 With $\log_2$ transformation to *M*-values for statistical analysis. Sample quality control was performed
211 by sex inference by extracting methylation calls on sex chromosomes and comparing to self-
212 reported sex to identify potential sample mix-ups. Dimensional reduction was conducted using
213 Principal Component Analysis (PCA) on M values (Figure E1, Step1, in the online supplement).
214 Differential methylation was evaluated both within regions (differentially methylated regions,

215 DMRs) and at individual CpG sites (differentially methylated positions, DMPs) using linear
216 regression. DMRs were detected using the DMRcate package v2.16.1 (30) and the model built
217 using tissue type as the primary variable and adjusting only for gender, as DMRcate method
218 corrects per sample using modelMatrixMeth function from edgeR package (31), therefore
219 participant ID were not included to avoid overfitting. To define DMRs, a minimum of 50 CpG
220 sites was used, applying a stringent threshold for genome-wide significance (Fold Discovery Rate
221 (FDR) ≤ 0.01 and mean difference $\geq |1|$) (Figure E1, Step2, in the online supplement). For CpG-
222 level analysis, the limma package v3.58.1 (32) was used to construct a linear model with tissue
223 type as the primary variable including gender as fixed effect. We also adjusted for the non-
224 independence of samples from the same individual by applying the *duplicateCorrelation()* method
225 from the limma, using individual ID as a blocking variable. The estimated correlation was included
226 during model fitting using *lmfit()* to improve the accuracy of differential analysis. Tissue-specific
227 DMPs were identified using a threshold of FDR ≤ 0.05 and $\log_2$ Fold Change ($\log_2$FC) $> |0.2|$
228 (Figure E1, Step3, in the online supplement). To identify conserved CpGs, we first retained loci
229 with no significant differences between tissues (FDR ≥ 0.1). Among these, we then selected CpGs
230 with high concordance in methylation levels between tissues (Pearson's correlation, $R \geq 0.8$). Note
231 that the FDR threshold refers to the differential methylation analysis between tissues, not the
232 correlation with exposures. Gestational exposures (maternal smoking and maternal asthma) were
233 then evaluated separately within conserved regions using the same limma-based linear modelling
234 approach describe above, adjusting for sex and accounting for repeated measures by participant
235 ID. Significant associations were defined as FDR ≤ 0.05 and $\log_2$FC $\geq |0.2|$ (Figure E1, Step 4, in
236 the online supplement).

237 DMPs were annotated to genes using the annotatr v1.28.0 (33), or internal functions for DMR
238 calling using DMRcate v2.16.1 (30). Gene Ontology (GO) pathway enrichment analysis was
239 conducted using the rGREAT v2.4.0 (34), applying the 'Two nearest genes' rule where gene
240 regulatory domains were defined by extending in both directions to the nearest transcriptional start
241 site, but no more than 20kb in either direction. Statistically enriched pathway terms were
242 summarized using the Simplify Enrichment package v1.12.0 (35). Supervised Partial-Least Square
243 (PLS) Discriminant Analysis (DA) was implemented with mixOmics v6.26.0 (36) to assess
244 associations with maternal asthma history and *in utero s*moking exposure. Visualization of BAM
245 files was performed using the Integrative Genome Viewer (IGV) software v2.8.9 (37)

246 **Urinary Cotinine Levels**

247 Maternal smoking status was determined via positive response to questionnaire and
248 validated with urinary cotinine levels (38) . Urinary cotinine levels were measured using the
249 Salimetrics® High Sensitivity Salivary Cotinine Enzyme Immunoassay Kit (Cat No. 1-2002-5,
250 SALIMETRICS, State College, Pennsylvania 16803, USA), following the manufacturer's
251 protocol. Urine samples collected at 20 and 36 weeks of pregnancy as part of the ORIGINS
252 biobank protocol (21, 22) were preserved at -80°C until analyzed. A standard curve was then
253 established using a 4-parameter non-linear regression curve fit, and cotinine concentrations
254 calculated using GraphPad Prism software v9.3.1. The assay lower limit of cotinine detection was
255 0.15 ng/mL. Positive urinary cotinine tests were >10 ng/mL with those below 10 ng/ml as non-
256 smokers.

257 **Data Deposition**

Due to ethical and privacy restrictions, raw sequencing datasets cannot be made available. However, de-identified methylation beta values sufficient to reproduce the findings are available upon request from the authors, while associated scripts are publicly accessible on GitHub (https://github.com/wal-yan/AERIAL/tree/main/RespEpithelium_Methylome).

Results

Demographics and Sequencing quality

The characteristics of the study population are presented in Table 1. While the newborns' ethnic backgrounds were diverse, the cohort was predominantly Caucasian, with 51.2% of mothers and 48.8% of fathers identifying as such. Newborns had a mean gestational age of 39.12 ± 1.28 weeks [SD], with three preterm births recorded in the cohort, occurring between 35 and 36 weeks of gestation. Sequencing quality statistics were similar across tissues (Table 2), with low duplication rates and 93% of bases covered at a depth of at least 30X. The average number of paired-end sequences generated was approximately 183 ± 28.90 [SD] million for amnion samples and 168.67 ± 33.50 [SD] million for nasal samples. The methylation status was assessed across 7,936,546 million individual CpG dinucleotides, with 6,961,516 remaining after quality control filtering. Fold-80 base penalty metrics were between 1.2-1.6 across all samples indicating good uniformity of coverage for low-diversity libraries.

Differential Methylation Landscape Between Amniotic and Nasal Tissues

We began by examining the general features of epithelium methylation landscapes between amniotic and nasal tissues, which revealed similar enrichment of methylated CpG sites around transcriptional start sites (Figure 1A), and the distribution of methylated and unmethylated CpG islands was similar across tissues (Figure 1B). These patterns were consistent across gene promoters and other regulatory features such as gene enhancers (Figure 1A and Figure E2A in the online supplement). Despite the overall similarities in CpG enrichment, dimensional reduction analysis of all 6,961,516 CpG sites across the genome revealed tissue type as the primary source

284 of variation in DNA methylation profiles (PERMANOVA: pseudo-F=91.52, R^2 0.36, p-
285 value=0.001) suggesting substantial tissue-specific differences (Figure 1C).

286 To determine the extent of tissue-specific differences in DNA methylation patterns, we
287 conducted differential analysis comparing amniotic and nasal epithelium at individual CpG sites
288 and identifying the differentially methylated regions (DMR). This analysis revealed 716,075
289 individual CpGs that contributed to 4,897 genome-wide significant DMRs (Table E1 in the online
290 supplement). Notably, in the nasal epithelium most regions exhibited higher methylation, at 3,847
291 regions (78.50%), while 1,050 regions (21.50%) showed a loss in methylation, compared to the
292 amniotic epithelium. The volcano plot highlights the most hypomethylated region in amnion
293 relative to nasal tissue as an intergenic region on chromosome 15 (chr15:70473876-70475721),
294 while the most hypermethylated region was within the *NAXD* gene on chromosome 13
295 (chr13:110626965-110629183) (Figure 2A). Genomic track visualization revealed broad and
296 consistent methylation differences across these loci (Figure 2B). The intergenic region on
297 chromosome 15 exhibited higher methylation levels in amnion, whereas *NAXD* on chromosome
298 13 showed increased methylation in nasal tissue (Figure 2B). To gain broader insights into the
299 biological processes associated with these methylation differences, we annotated DMRs to the
300 nearest genes and performed Gene Ontology (GO) enrichment analysis. This revealed strong
301 enrichment of pathways related to development and differentiation, with additional ontologies
302 related to stimuli and metabolic processes (Figure 2C and Figure E2B in the online supplement).

303 **Conservation of Methylation Across A Subset of Lung Specific Genes**

304 We identified a conserved methylation signature (Figure 3A and Table E2 in the online
305 supplement), comprising 1,493,976 CpG loci, representing approximately 21% of the total CpG

306 landscape (n=6,961,516). This conserved pattern was distributed across all genomic regulatory
307 features, with the highest representation in introns (20.09%), CpG islands (14.10%), and promoters
308 (12.10%) (Figure E2A in the online supplement). To explore whether there are conserved
309 methylation patterns in genes biologically relevant to epithelial and lung development, we queried
310 the Genotype-Tissue Expression (GTEx) consortium database (https://gtexportal.org/home/) to
311 extract a set of genes uniquely expressed in lung tissue compared to all other anatomical sites (total
312 459 genes). Among these gene sets, 159 genes (~34.6%) exhibited highly concordant methylation
313 patterns (Pearson's R ≥ 0.8) between the two tissues and general conservation of the methylation
314 landscape (Figure 3B). For instance, Surfactant Protein A1 (*SFTPA1)* exhibited intermediate levels
315 of DNA methylation in both amnion and nasal tissues (Figure E2B in the online supplement),
316 while the epithelial specific gene, Secretoglobin Family 3A Member 1 (*SCGB3A1),* known to
317 regulate cell proliferation was predominantly hypomethylated in both tissues (Figure E2B in the
318 online supplement), indicating conservation in subsets of biologically relevant genes. We
319 conducted a pathway enrichment analysis of the conserved signature reveling enrichment of terms
320 related to anatomical and morphogenic processes, cellular organization, metabolism and cell
321 cycling, as well as the regulation of signal transduction and stimulus response (Figure 3D and
322 Figure E2C in the online supplement).

323 **Amnion tissue preserves methylation signatures associated with exposures during pregnancy**

324 Within the conserved portion of the methylome (Table E2 in the online supplement), we
325 investigated statistical associations with two well-characterized exposures known to increase the
326 risk of asthma development in offspring: maternal asthma history and *in utero s*moking exposure
327 (39). In this cohort 9 out of 84 mothers reported a history of asthma. Binary logistic regression
328 analysis, adjusted for multiple comparisons (FDR ≤ 0.05), identified a small but significant

329 associations at 31 CpG sites (25 hypermethylated and 6 hypomethylated) with maternal asthma
330 (Table 3 and Table E3 in the online supplement). Principal component analysis (PCA) of the
331 conserved CpG sites significantly associates with maternal asthma history showed that this
332 exposure was the major driver of variation in DNA methylation at these CpGs in both tissues
333 (PERMANOVA: pseudo-F=29.09, R^2 0.15, p-value=0.001) (Figure 4A). Enrichment analysis
334 linked these DMPs to pathways involved in cell communication, signal transduction, and response
335 to stimuli. We applied Partial Least Squares - Discriminant Analysis (PLS-DA), which identified
336 the 15 most relevant features contributing to group separation (Figure 4B). Notably, methylation
337 levels of 13 CpGs decreased in the maternal asthma group, while two CpGs increased (Figure 4C).
338 We visualized the top three significant CpG sites: chr16:86536881 (*MTHFSD*), chr2:10529804,
339 and chr3:59456728 (*CFAP20DC-DT*), exhibiting the largest effect sizes which demonstrated
340 consistency across both amnion and nasal tissues in relation to maternal asthma history (Figure
341 4C).

342 Next, we used the urinary cotinine measurements at 20 and 36 weeks of gestation as a
343 marker of intrauterine smoke exposure. In our cohort, 4 of 88 mothers had cotidine levels
344 consistent with tobacco smoking. Binary logistic regression analysis identified 164 CpG
345 significant associations (126 hypermethylated and 38 hypomethylated) with maternal smoking
346 exposure (Table 4 and Table E4 in the online supplement). Principal component analysis (PCA)
347 of the conserved CpG sites significantly associated with smoking exposure identified it as primary
348 source of the variation in DNA methylation in both tissues (PERMANOVA: pseudo-F=27.23, R^2
349 0.15, p-value=0.001) (Figure 5A), with significant enrichment of pathways involved in regulation
350 of localization, cell activation, proliferation, and lung development. By applying PLS-DA the 15
351 most relevant features contributing to group separation (Figure 5B) comprised 14 CpG sites with

352 decreased levels in the exposed group, while one CpG increased. Visualization of the top three
353 significant CpG sites: chr12:116575870 (*MAP1LC3B2*), chr14:74362734, and chr17:36093007
354 (*CCL3*), also demonstrated the consistency of these changes across both amnion and nasal tissues
355 between individuals with and without prenatal smoking exposure (Figure 5C).

356

357 **Discussion**

358 The goal of this study was to explore the utility of placental amnion tissue as an easily
359 accessible fetal-derived tissue for studying fetal programming of epithelial function in the context
360 of asthma development. To achieve this, we investigated the relationship between the placental
361 amniotic and nasal epithelial methylation landscapes in *ex vivo* samples from newborns. The key
362 observations from this study are as follows.

363 Substantial tissue-specific methylation differences existed between the two tissues,
364 characterized by a predominantly hypomethylated profile (reduced methylation) in amniotic
365 epithelium samples. Given that DNA methylation levels typically increase in more terminally
366 differentiated cell types, reflecting a more restrictive chromatin structure, this hypomethylation in
367 amniotic tissue aligns with its more pluripotent state and stem-cell like characteristics reported in
368 regenerative medicine studies (19). Despite these considerable differences, general features of the
369 methylation landscape were similar, including enrichment of methylated CpGs around
370 transcriptional start sites and low methylation across CpG Islands. consistent with patterns reported
371 across multiple human somatic tissues (40). At the same time, histone modifications have been
372 reported to contribute to epithelial plasticity, which is involved in both development and disease
373 progression (41, 42). This divergence in epigenetic control may facilitate tissue-specific
374 adaptations and responses to environmental changes, underscoring the dynamic nature of the
375 epigenome.

376 Around 20% of the regions captured by the EM-seq target capture assay exhibited very
377 high concordance of methylation levels, which we deemed the 'conserved methylome' between
378 amniotic and nasal tissues (Figure 3A, Table E2 in the online supplement). This shared methylation
379 signature may reflect common developmental and regulatory pathways shaped by the intrauterine

380 environment. Amniotic fluid contains bioactive molecules that play essential roles in fetal
381 development, particularly influencing critical cellular processes in lung epithelial cells (43).
382 Additionally, the amniotic membrane contributes to the production of surfactant proteins and lipids
383 vital for fetal lung maturation (44). Notably, within this conserved fraction, genes that play a
384 critical role in lung development were represented, supporting the utility for amniotic tissue as a
385 window for developmental programming of the respiratory epithelium (Figure 3B). However, we
386 acknowledge that the full spectrum of biologically relevant genes with conserved methylation
387 patters remains to be determined, a limitation of the present analysis. Future studies using more
388 comprehensive and unbiased methods, such as whole-genome bisulfite sequencing, could help
389 uncover additional loci of interest.

390 Moving beyond this comparative analysis, we examined how maternal exposures known
391 to influence the development of asthma in early life are reflected in the conserved methylation
392 patterns of both tissues (Figure 5 and Figure 6). Strengthening support for the utility of amniotic
393 tissue to reveal intrauterine influences on respiratory development, we report that statistical
394 associations were equivalent across tissues within this conserved fraction of the methylome (Table
395 E3 in the online supplement). While previous studies have reported epigenetic alterations in the
396 offspring linked to maternal asthma and smoking during pregnancy (45, 46), our findings indicate
397 that these signals are already embedded in the shared, conserved methylation signature between
398 amnion and nasal epithelium. This suggests that *in utero* programming effects persist within the
399 post-natal nasal epithelium and may be tractable longitudinally, as compelling targets for early
400 asthma risk stratification. It is important to acknowledge the observational nature of this study
401 precludes establishing causality between methylation patterns in the nasal and amnion tissue, *in*
402 *utero* exposures and epithelial dysfunction. Furthermore, we acknowledge that our approach pre-

403 selects for the conserved methylome, and therefore the extent to which *in utero* exposures
404 differentially influence these tissues is unclear, and therefore we do not provide an estimate for the
405 overall level of concordance for exposure related effects. The latter reflects another limitation of
406 this study due to the limited sample size, and the incomplete sampling of the full methylome when
407 using target capture assays.

408 In conclusion, our study advances the field by providing cogent evidence for using
409 amniotic tissue to provide insight regarding developmental programming of the respiratory
410 epithelium. As well as identifying potential mechanisms for intrinsic vulnerability, a similar
411 strategy could be used to assess prenatal interventions to enhance respiratory epithelial innate
412 immune function. These approaches circumvent the challenges of obtaining airway epithelial cells
413 from newborns, providing opportunities for early risk stratification that may eventually contribute
414 to interventions for childhood wheezing and asthma.
415

Acknowledgments

We would like to thank the contributions of the AERIAL families, together with the cohort study investigator Jose A. Caparros-Martin, Lidija Turkovic, and our consumers Rael Rivers, Kate McGee and Olivia Gleeson, and the Western Australian Epithelial Research Program Consumer Reference Group for their contributions towards project inception/design and its adaptation through COVID-19 and other related amendments.

We also extend our gratitude to the hardworking and dedicated AERIAL research team for recruitment, liaising, and sample collection and processing over the duration of this study, with special mention to Minda Amin, Bailee Renouf, Courtney Kidd and Ashleigh Heng-Chin for supporting the bioinformaticians in the analysis.

We are grateful to all the ORIGINS families who support the project. We would also like to acknowledge and thank the following teams and individuals who have made ORIGINS possible: ORIGINS project team; Joondalup Health Campus (JHC); members of ORIGINS Community Reference and Participant Reference Groups; Research Interest Groups and the ORIGINS Scientific Committee; The Kids Research Institute Australia; City of Wanneroo; City of Joondalup; and Professor Fiona Stanley.

This study is a sub-project of ORIGINS. This unique long-term study, a collaboration between The Kids Research Institute Australia and Joondalup Health Campus, is one of the most comprehensive studies of pregnant women and their families in Australia to date, recruiting 10,000 families over a decade from the Joondalup and Wanneroo communities of Western Australia.

437 **References**

438 1. Inoue Y, Shimojo N. Epidemiology of virus-induced wheezing/asthma in children. *Front*
439 *Microbiol* 2013;4.
440 2. Alvarez-Alvarez I, Niu H, Guillen-Grima F, Aguinaga-Ontoso I. Meta-analysis of
441 prevalence of wheezing and recurrent wheezing in infants. *Allergol Immunopathol (Madr)*
442 2018;46:210–217.
443 3. Masoli M, Fabian D, Holt S, Beasley R. The global burden of asthma: Executive summary
444 of the GINA Dissemination Committee Report. *Allergy: European Journal of Allergy and*
445 *Clinical Immunology* 2004;59:469–478.
446 4. Stick SM, Burton PR, Gurrin L, Sly PD, LeSouëf PN. Effects of maternal smoking during
447 pregnancy and a family history of asthma on respiratory function in newborn infants. *Lancet*
448 1996;348:1060–1064.
449 5. Iosifidis T, Sutanto EN, Buckley AG, Coleman L, Gill EE, Lee AH, *et al.* Aberrant cell
450 migration contributes to defective airway epithelial repair in childhood wheeze. *JCI Insight*
451 2020;5.
452 6. Wang B, Chen H, Chan YL, Wang G, Oliver BG. Why Do Intrauterine Exposure to Air
453 Pollution and Cigarette Smoke Increase the Risk of Asthma? *Front Cell Dev Biol* 2020;8:.
454 7. Clifton VL. Managing asthma in pregnancy: effects on future child health. *Lancet Respir*
455 *Med* 2019;7:485–486.
456 8. Clifton VL, Moss TJM, Wooldridge AL, Gatford KL, Liravi B, Kim D, *et al.* Development
457 of an experimental model of maternal allergic asthma during pregnancy. *Journal of*
458 *Physiology* 2016;594:1311–1325.
459 9. Zazara DE, Wegmann M, Giannou AD, Hierweger AM, Alawi M, Thiele K, *et al.* A
460 prenatally disrupted airway epithelium orchestrates the fetal origin of asthma in mice.
461 *Journal of Allergy and Clinical Immunology* 2020;145:1641–1654.
462 10. Stefanowicz D, Hackett TL, Garmaroudi FS, Günther OP, Neumann S, Sutanto EN, *et al.*
463 DNA Methylation Profiles of Airway Epithelial Cells and PBMCs from Healthy, Atopic
464 and Asthmatic Children. *PLoS One* 2012;7.
465 11. Reese SE, Xu CJ, den Dekker HT, Lee MK, Sikdar S, Ruiz-Arenas C, *et al.* Epigenome-
466 wide meta-analysis of DNA methylation and childhood asthma. *Journal of Allergy and*
467 *Clinical Immunology* 2019;143:2062–2074.
468 12. Knopik VS, MaCcani MA, Francazio S, McGeary JE. The epigenetics of maternal cigarette
469 smoking during pregnancy and effects on child development. *Dev Psychopathol*
470 2012;24:1377–1390.
471 13. Nonkovic N, Marceau K, McGeary JE, Ramos AM, Palmer RHC, Heath AC, *et al.* Maternal
472 smoking during pregnancy is associated with DNA methylation in early adolescence: A
473 sibling comparison design. *Dev Psychol* 2024; doi:10.1037/DEV0001747.
474 14. Cherneva R V., Kostadinov D. Epigenetic Targets for Therapeutic Approaches in COPD
475 and Asthma. Nutrigenomics – Possible or Illusive. *Folia Med (Plovdiv)* 2019;61:358–369.
476 15. Ordovas-Montanes J, Dwyer DF, Nyquist SK, Buchheit KM, Vukovic M, Deb C, *et al.*
477 Allergic inflammatory memory in human respiratory epithelial progenitor cells. *Nature*
478 2018;560:649–654.
479 16. Suzuki M, Maekawa R, Patterson NE, Reynolds DM, Calder BR, Reznik SE, *et al.* Amnion
480 as a surrogate tissue reporter of the effects of maternal preeclampsia on the fetus. *Clin*
481 *Epigenetics* 2016;8:.

482 17. Ching T, Song MA, Tiirikainen M, Molnar J, Berry M, Towner D, *et al.* Genome-wide
483 hypermethylation coupled with promoter hypomethylation in the chorioamniotic
484 membranes of early onset pre-eclampsia. *Mol Hum Reprod* 2014;20:885–904.
485 18. Kim J, Pitlick MM, Christine PJ, Schaefer AR, Saleme C, Comas B, *et al.* Genome-wide
486 analysis of DNA methylation in human amnion. *The Scientific World Journal* 2013;2013:.
487 19. Miki T. Stem cell characteristics and the therapeutic potential of amniotic epithelial cells.
488 *American Journal of Reproductive Immunology* 2018;80:e13003.
489 20. Kicic-Starcevich E, Hancock DG, Iosifidis T, Agudelo-Romero P, Caparros-Martin JA,
490 Karpievitch Y V., *et al.* Airway epithelium respiratory illnesses and allergy (AERIAL) birth
491 cohort: study protocol. *Frontiers in Allergy* 2024;5.
492 21. Silva DT, Hagemann E, Davis JA, Gibson LY, Srinivasjois R, Palmer DJ, *et al.* Introducing
493 the ORIGINS project: a community-based interventional birth cohort. *Rev Environ Health*
494 2020;35:281–293.
495 22. D'Vaz N, Kidd C, Miller S, Amin M, Davis JA, Talati Z, *et al.* The ORIGINS Project
496 Biobank: A Collaborative Bio Resource for Investigating the Developmental Origins of
497 Health and Disease. *Int J Environ Res Public Health* 2023;20.
498 23. Magatti M, Caruso M, De Munari S, Vertua E, De D, Manuelpillai U, *et al.* Human
499 Amniotic Membrane-Derived Mesenchymal and Epithelial Cells Exert Different Effects on
500 Monocyte-Derived Dendritic Cell Differentiation and Function. *Cell Transplant*
501 2015;24:1733–1752.
502 24. Gustafsson OJR, Al Bkhetan Z, Francis R, Manos S. *Enabling national step changes in*
503 *bioinformatics through ABLeS, the Australian BioCommons Leadership Share*. *Zenodo*
504 Melbourne, Australia: 2023. doi:https://doi.org/10.5281/zenodo.10139651.
505 25. Pawsey Supercomputing Research Centre. *Nimbus Research Cloud*. Perth, Western
506 Australia: 2023. at <https://doi.org/10.48569/v0j3-qd51>.
507 26. Pawsey Supercomputing Research Centre. *Acacia Object Storage*. Perth, Western
508 Australia: 2023. at <https://doi.org/10.48569/nfe9-a426>.
509 27. DI Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow
510 enables reproducible computational workflows. *Nat Biotechnol* 2017;35:316–319.
511 28. Sharma A, Conradie T, Martino D, Stick S, Agudelo-Romero P. target-methylseq-qc: a
512 lightweight pipeline for collecting metrics from targeted sequence mapping files. *J Open*
513 *Source Softw* 2025;10:7608.
514 29. R Core Team. R: A language and environment for statistical computing, R Foundation for
515 Statistical. *R Foundation for Statistical Computing* 2021;at <https://www.r-project.org/>.
516 30. Peters TJ, Buckley MJ, Chen Y, Smyth GK, Goodnow CC, Clark SJ. Calling differentially
517 methylated regions from whole genome bisulphite sequencing with DMRcate. *Nucleic*
518 *Acids Res* 2021;49.
519 31. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential
520 expression analysis of digital gene expression data. *Bioinformatics* 2009;26:139–140.
521 32. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, *et al.* Limma powers differential
522 expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*
523 2015;43:e47.
524 33. Cavalcante RG, Sartor MA. annotatr: genomic regions in context. *Bioinformatics*
525 2017;33:2381.
526 34. Gu Z, Hübschmann D. rGREAT: an R/bioconductor package for functional enrichment on
527 genomic regions. *Bioinformatics* 2023;39.

35. Gu Z, Hübschmann D. simplifyEnrichment: A Bioconductor Package for Clustering and Visualizing Functional Enrichment Results. *Genomics Proteomics Bioinformatics* 2023;21:190.
36. Rohart F, Gautier B, Singh A, Lê Cao KA. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol* 2017;13.
37. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, *et al.* Integrative genomics viewer. *Nat Biotechnol* 2011;29:24–26.
38. Noakes PS, Thomas R, Lane C, Mori TA, Barden AE, Devadason SG, *et al.* Association of maternal smoking with increased infant oxidative stress at 3 months of age. *Thorax* 2007;62:714–717.
39. Bobolea I, Arismendi E, Valero A, Agustí A. Early Life Origins of Asthma: A Review of Potential Effectors. *J Investig Allergol Clin Immunol* 2019;29:168–179.
40. Lokk K, Modhukur V, Rajashekar B, Märtens K, Mägi R, Kolde R, *et al.* DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biol* 2014;15.
41. Kiesslich T, Pichler M, Neureiter D. Epigenetic control of epithelial-mesenchymal-transition in human cancer. *Mol Clin Oncol* 2013;1:3–11.
42. Yuan S, Natesan R, Sanchez-Rivera FJ, Li J, Bhanu N V., Yamazoe T, *et al.* Global regulation of the histone mark H3K36ME2 underlies epithelial plasticity and metastatic progression. *Cancer Discov* 2020;10:854–871.
43. Li B, Chen X, Ma W, Li N, Chen Q, Liao J, *et al.* An endogenous amniotic fluid-derived 10-amino acid peptide improves lung development and hyperoxia injury. *Transl Pediatr* 2023;12:46–55.
44. Lemke A, Castillo-Sánchez JC, Prodinger F, Ceranic A, Hennerbichler-Lugscheider S, Pérez-Gil J, *et al.* Human amniotic membrane as newly identified source of amniotic fluid pulmonary surfactant. *Sci Rep* 2017;7.
45. Magnaye KM, Clay SM, Nicodemus-Johnson J, Naughton KA, Huffman J, Altman MC, *et al.* DNA methylation signatures in airway cells from adult children of asthmatic mothers reflect subtypes of severe asthma. *Proc Natl Acad Sci U S A* 2022;119.
46. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, *et al.* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am J Hum Genet* 2016;98:680–696.

565 **Table 1. Demographic characteristics of the Airway Epithelium Respiratory Illnesses and**

566 **Allergy (AERIAL) study.**

Variables	Categories
Gender	**Female** 50 % (n=42) **Male** 50 % (n=42)
Ethnicity	**Mother** Caucasian (51.2 %) Non-caucasian (48.8 %) **Father** Caucasian (44.1 %) Non-caucasian (55.9 %)
Smoking exposure	**Yes** 4.76 % (n=4) **No** 95.54 % (n=80)
Maternal asthma	**Maternal asthma history** 11.91 % (n=10) **No maternal asthma history** 88.09 % (n=74)
Gestational age (mean ± SD)	39.12 ± 1.28 weeks
Delivery Mode	**Vaginal delivery** 47.62% (n=40) **Caesarean delivery** 47.62% (n=40) ***NA*** 4.76% (n=4)

567 ***SD****: standard deviation;* ***NA***: data not available

568

569 **Table 2. Sequencing statistics for 84 matched samples from neonatal nasal and maternal**

570 **amnion.** Values represent the mean ± standard deviation.

	Nasal samples	Amnion samples
M seq (in million)	168.67 ± 33.50	183 ± 28.90
% GC	27 ± 0.01	27% ± 0.01
% Duplication	15% ± 0.03	18% ± 0.03
Insert Size	225.35 bp ± 21.02	212.95 bp ± 9.98
% Target Bases 10X	99% ± 0.00	99% ± 0.00
% Target Bases 30X	93% ± 5.73	93% ± 8.55

571 bp: base pair

572

573 **Table 3. Top 10 Most Differentially Methylated Positions (DMPs) in Maternal Asthma**

574 **History.**

Position	Symbol*	Log_2FC	Adj. p-value
chr16:89629809	*DPEP1*	2.98	0.049936
chr17:74802258	*TMEM104*	2.36	0.02597409
chr10:13502863	*BEND7*	2.34	0.02597409
chr15:60787177	*RORA*	2.28	0.03648546
chr15:60787176	*RORA*	2.28	0.02597409
chr10:3332958		-1.45	0.00505511
chr6:169655096	*WDR27*	-1.46	0.04672635
chr18:45723869	SLC14A1; LOC105372092; LOC105372093	-2.22	0.01668038
chr18:45723868	SLC14A1; LOC105372092; LOC105372093	-2.27	0.02282406
chrX:70149858	*IGBP1*	-2.61	0.01668038

575 *CpG sites without a gene symbol may be located in regions with incomplete or less characterized
576 gene annotations.
577

578

579 **Table 4. Top 10 Most Differentially Methylated Positions (DMPs) in Maternal Smoking**

580 **Exposure.**

Position	Symbol*	Log_2FC	Adj. p-value
chr8:143251263	*ZFP41*	4.99	0.02767552
chr6:15500540	*JARID2*	4.77	0.019276332
chr11:17779027	*KCNC1*; LOC107984317	4.74	0.015961482
chr8:143229735		4.73	0.012421924
chr14:74362734		4.66	5.60224E-07
chr7:158796784	*ESYT2*	-4.14	0.0415906
chr6:65566339	*EYS*	-4.20	0.008385907
chr12:31887708	LINC02422	-4.20	0.028869893
chr12:31887707	LINC02422	-4.47	0.011409155
chr10:129774152		-4.84	0.002327469

581 *CpG sites without a gene symbol may be located in regions with incomplete or less characterized
582 gene annotations.
583

584

585 **Figure Legends**
586
587 **Figure 1. Comparative analysis of DNA methylation landscapes in amnion and nasal tissues.**
588 (A) Density plot illustrating CpG enrichment around Transcriptional Start Sites (TSSs). (B)
589 Methylation levels of CpG islands with ± 2 kb. CpG islands are rescaled to the same width from
590 Start to End vertical lines. (C) Principal component analysis (PCA) of genome-wide DNA
591 methylation tissue profiles reveals distinct clustering of Amnion (red) and Nasal (blue) samples.
592 Principal Component 1 (PC1) accounts for 39.93% of the variance.
593
594 **Figure 2. Regions differentially methylated between amniotic and nasal samples are enriched**
595 **in cell development pathways.** (A) Volcano plot displaying differentially methylated regions
596 (DMRs) between amnion and nasal tissues. The x-axis represents the mean methylation difference,
597 and the y-axis represents the $\log_2$-transformed region size. Genes associated with the most
598 significant DMRs are labelled in red. (B) Genomic tracks showing methylation loss or gain in
599 specific regions of Chromosomes 13 (*NAXD*) and 15 (Intergenic region), respectively. CpG islands
600 are marked in green. The line graph summarizes the methylation difference between amnion (red)
601 and nasal (blue) tissues across the region. (C) Gene Ontology (GO) enrichment was calculated
602 using genes near DMRs between amniotic and nasal tissues, followed by a similarity matrix
603 analysis. The similarity heatmap clusters significant GO terms into related Biological Process
604 (BP), with colour intensity representing the significance of enrichment. Similarity of related terms
605 within the same cluster is shown in the heatmap.
606
607 **Figure 3. Conserved methylation landscapes were enriched in cell differentiation and**
608 **morphogenesis pathways.** (A) Scatter plot comparing DNA methylation levels between amnion
609 and nasal tissues. Each dot represents a genomic region, with its position indicating the
610 methylation level in each tissue. The strong positive correlation (Pearson's $R \geq 0.8$, p-value =
611 0.000) suggests a high degree of conservation in methylation patterns between these tissues. (B)
612 Heatmap illustrating the expression levels of lung- and epithelial-specific genes across various
613 tissues. Each row represents a gene, and each column represents a tissue. The colour intensity
614 indicates the level of expression, with red representing high expression and blue representing low
615 expression. Data are derived from the GTEx catalogue. (C) Gene Ontology (GO) enrichment was
616 calculated using genes near CpG sites from the conserved methylation signature, followed by a
617 similarity matrix analysis. The similarity heatmap displays GO terms associated with Biological
618 Process (BP), with colour intensity representing the significance of enrichment.
619
620 **Figure 4. Methylation changes associated with maternal asthma in the conserved signature.**
621 (A) Principal Component Analysis (PCA) of DNA methylation variation in newborn amnion and
622 nasal tissues, showing the first two principal components (PC1 and PC2). Samples are coloured
623 based on maternal asthma history, with asthma in orange and non-asthma in blue. (B) Bar plot of
624 loading scores from Dimension 1 (Dim1) of a Partial Least Squares Discriminant Analysis (PLS-
625 DA), highlighting the top 15 differentially methylated CpG sites associated with maternal asthma.
626 (C) Violin plots of the top three CpG sites identified by PLS-DA, showing the distribution of
627 methylation levels (expressed as M-values) at features associated with maternal asthma in newborn
628 amnion and nasal tissues, further stratified by maternal asthma history during pregnancy status.
629

Figure 5. Methylation changes associated with maternal smoking exposure in the conserved signature. (A) Principal Component Analysis (PCA) of DNA methylation variation in newborn amnion and nasal tissues, showing the first two principal components (PC1 and PC2). Samples are coloured based on maternal smoking exposure, with smoking exposure in orange and non-smoking exposure in blue. (B) Bar plot of loading scores from Dimension 1 (Dim1) of a Partial Least Squares Discriminant Analysis (PLS-DA), highlighting the top 15 differentially methylated CpG sites associated with maternal smoking exposure. (C) Violin plots of the top three CpG sites identified by PLS-DA, showing the distribution of methylation levels (expressed as M-values) at features associated with maternal smoking exposure in newborn amnion and nasal tissues, further stratified by prenatal smoking status.

Supplemental Data

Table E1. Tissue-Specific Differentially Methylated Regions (DMRs) and Positions (DMPs). Results of differential methylation analysis comparing amniotic and nasal epithelium.

Table E2. Conserved CpG Methylation Across Lung-Specific Genes. List of 1,493,976 CpG loci showing conserved methylation patterns across tissues, with genomic feature annotations.

Table E3. CpG Sites Associated with Maternal Asthma. Binary logistic regression analysis, adjusted for multiple comparisons (FDR ≤ 0.05), identified significant associations at CpG sites with maternal asthma (MA) and smoking exposure (SE).

Figure E1. Analysis Workflow. This study was conducted in three main phases: Sample processing and sequencing (green), Bioinformatics analysis (blue), and Statistical analysis (pink). The statistical analysis phase was further divided into four key steps: Step 1 – Dimensional reduction using Principal Component Analysis (PCA); Step 2 – Differential methylation analysis at the regional level (differentially methylated regions, DMRs) associated with tissue differences; Step 3 – Differential methylation analysis at the CpG site level (differentially methylated positions, DMPs) associated with tissue differences; Step 4 – Identification of a conserved DMP signature, which was then used to assess associations with gestational exposures (maternal asthma and maternal smoking). MD: Mean Difference; FDR: False Discovery Rate; FC: Fold Change.

Figure E2. Conserved methylation landscape across tissues. (A) Histogram counts of CpGs annotated to genomic features for the conserved methylome. (B) Significantly enriched biological processes within the conserved methylome for CpGs annotate to nearest transcriptional start sites. Point size represents count of number of genes relative to total genes within the pathway (C) Raw sequencing coverage across the SFTPA1 genet coding region (top panel) and the SCGB3A1 coding region. Ticks in red are methylated reads and blue represents unmethylated reads. Height of the grey silhouettes represents sequencing coverage.

Figure E3. Pathway enrichment of maternal asthma history. (A) Significantly enriched biological processes within the conserved methylome for CpGs associated with maternal asthma. Point size represents count of number of genes relative to total genes within the pathway. (B) Gene Ontology (GO) enrichment was calculated using genes near CpG sites from the maternal asthma signature, followed by a similarity matrix analysis. The similarity heatmap displays GO terms

676 associated with Biological Process (BP), with colour intensity representing the significance of
677 enrichment.
678
679 **Figure E4. Pathway enrichment of maternal smoking exposure.** (A) Significantly enriched
680 biological processes within the conserved methylome for CpGs associated with maternal smoking.
681 Point size represents count of number of genes relative to total genes within the pathway. (B) Gene
682 Ontology (GO) enrichment was calculated using genes near CpG sites from the maternal smoking
683 signature, followed by a similarity matrix analysis. The similarity heatmap displays GO terms
684 associated with Biological Process (BP), with colour intensity representing the significance of
685 enrichment.

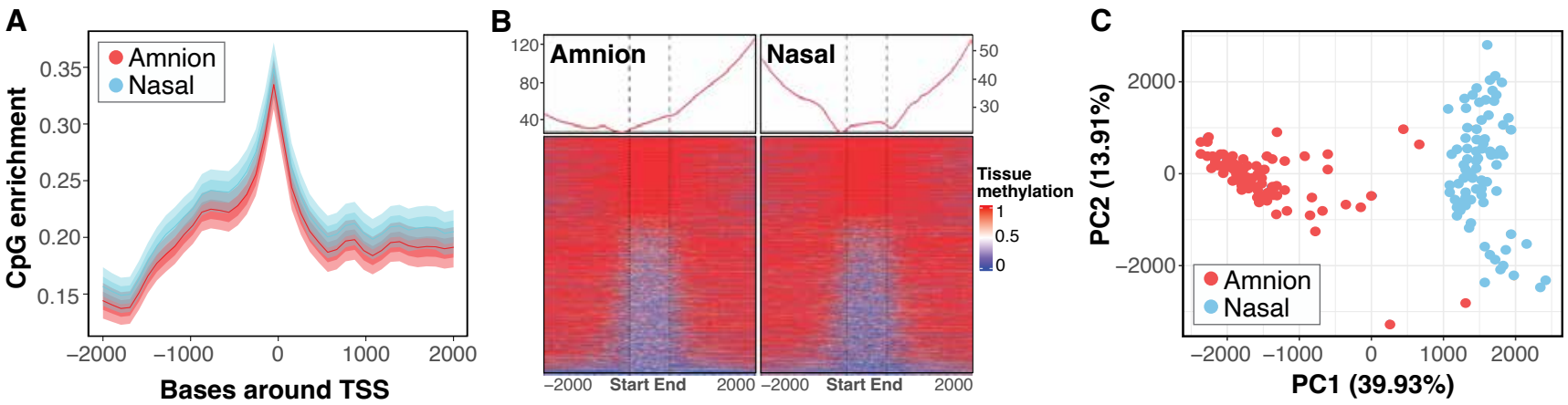
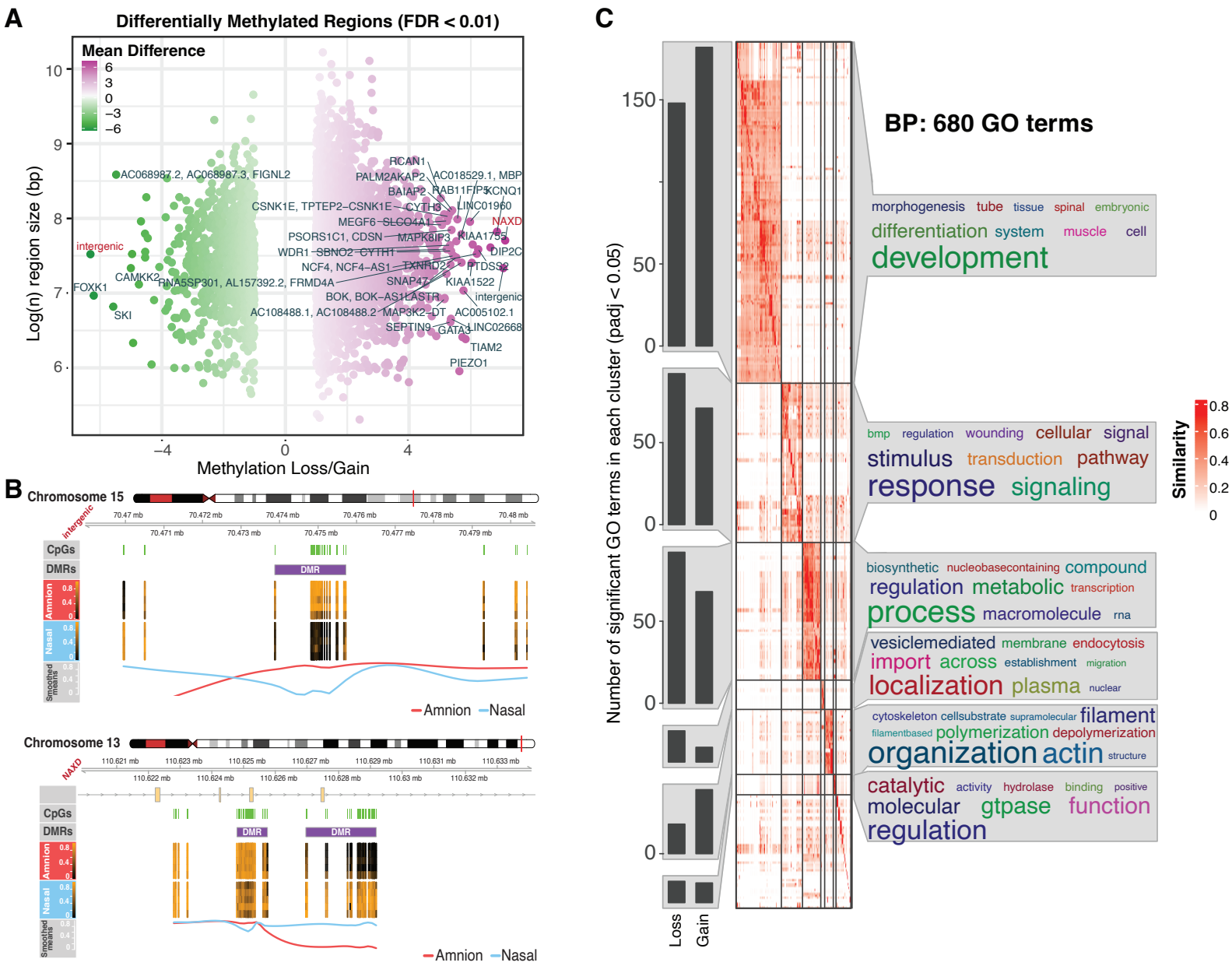
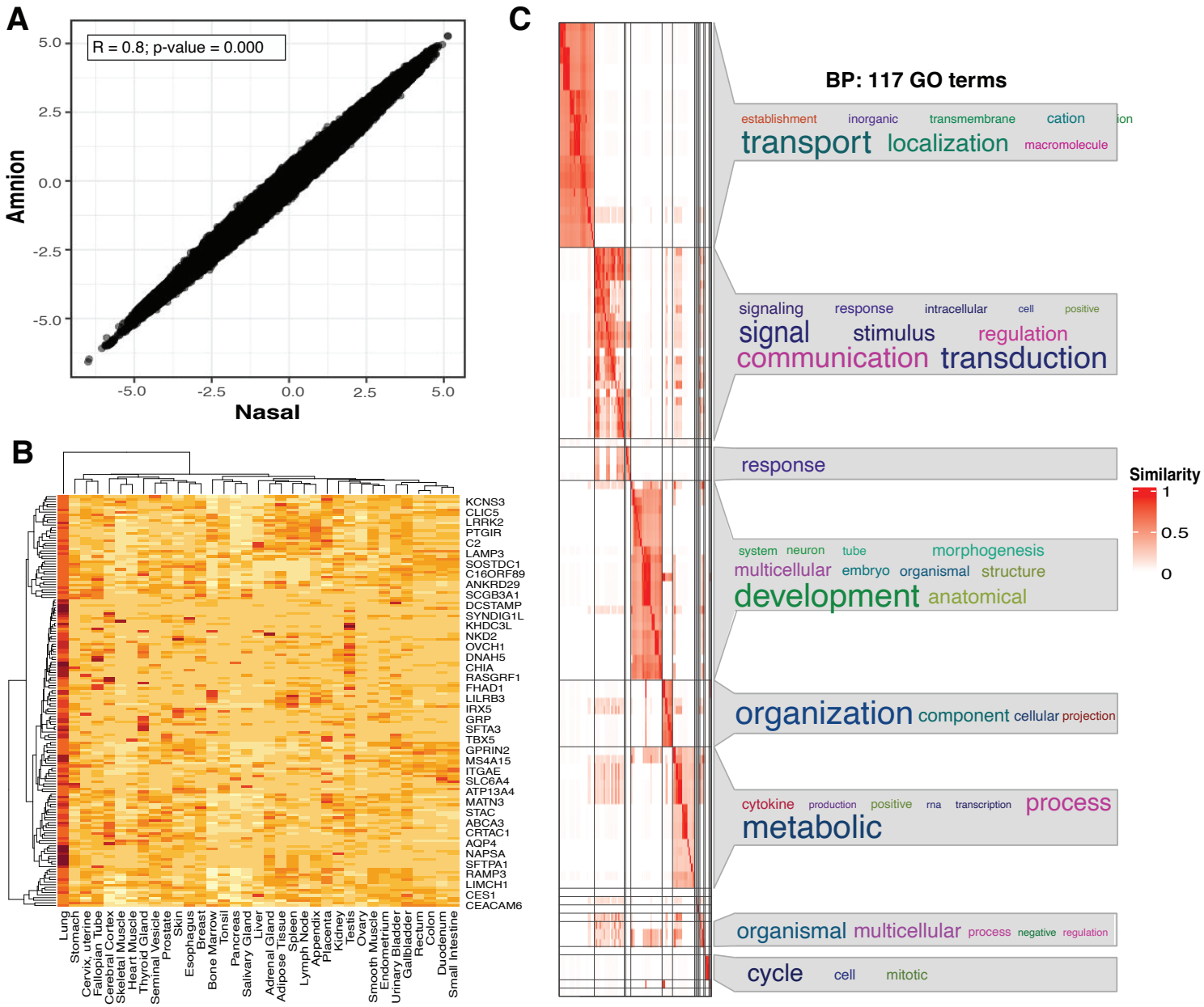


Figure 1. Comparative analysis of DNA methylation landscapes in amnion and nasal tissues. (A) Density plot illustrating CpG enrichment around Transcriptional Start Sites (TSSs). (B) Methylation levels of CpG islands with ± 2 kb. CpG islands are rescaled to the same width from Start to End vertical lines. (C) Principal component analysis (PCA) of genome-wide DNA methylation tissue profiles reveals distinct clustering of Amnion (red) and Nasal (blue) samples. Principal Component 1 (PC1) accounts for 39.93% of the variance.


A
Differentially Methylated Regions (FDR < 0.01)
Mean Difference
6
3
0
−3
−6
Log(n) region size (bp)
Methylation Loss/Gain
AC068987.2, AC068987.3, FIGNL2
intergenic
FOXK1
SKI
CAMKK2
RNA5SP301, AL157392.2, FRMD4A
CSNK1E, TPTEP2−CSNK1E
PSORS1C1, CDSN
WDR1
NCF4, NCF4−AS1
BOK, BOK−AS1
AC108488.1, AC108488.2
RCAN1
PALM2AKAP2
AC018529.1, MBP
BAIAP2
RAB11FIP5
KCNQ1
CYTH3
LINC01960
MEGF6
SLCO4A1
NAXD
MAPK8IP3
KIAA1755
DIP2C
TXNRD2
PTDSS2
SNAP47
KIAA1522
LASTR
AC005102.1
MAP3K2−DT
SEPTIN9
GATA3
LINC02668
TIAM2
PIEZO1
B
Chromosome 15
Chromosome 13
CpGs
DMRs
DMR
Amnion
Nasal
Smoothed means
Amnion
Nasal
C
BP: 680 GO terms
Number of significant GO terms in each cluster (padj < 0.05)
Loss
Gain
morphogenesis tube tissue spinal embryonic
differentiation system muscle cell
development
bmp regulation wounding cellular signal
stimulus transduction pathway
response signaling
biosynthetic nucleobasecontaining compound
regulation metabolic transcription
process macromolecule rna
vesiclemediated membrane endocytosis
import across establishment migration
localization plasma nuclear
cytoskeleton cellsubstrate supramolecular filament
filamentbased polymerization depolymerization
organization actin structure
catalytic activity hydrolase binding positive
molecular gtpase function
regulation
Similarity
0.8
0.6
0.4
0.2
0

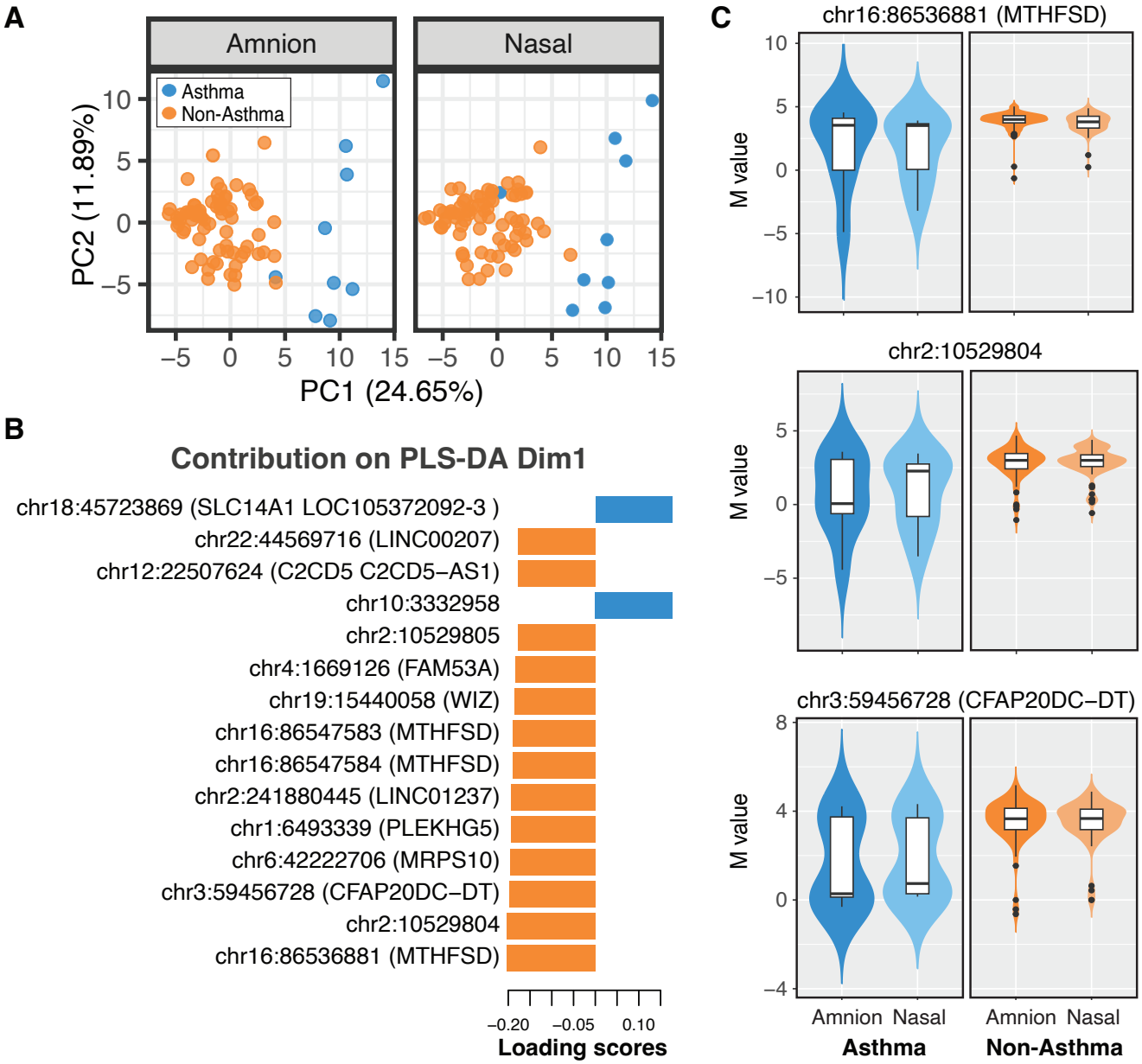
693

694 **Figure 2. Regions differentially methylated between amniotic and nasal samples are enriched in cell development pathways.** (A)
695 Volcano plot displaying differentially methylated regions (DMRs) between amnion and nasal tissues. The x-axis represents the mean
696 methylation difference, and the y-axis represents the $\log_2$-transformed region size. Genes associated with the most significant DMRs
697 are labelled in red. (B) Genomic tracks showing methylation loss or gain in specific regions of Chromosomes 13 (*NAXD*) and 15
698 (Intergenic region), respectively. CpG islands are marked in green. The line graph summarizes the methylation difference between
699 amnion (red) and nasal (blue) tissues across the region. (C) Gene Ontology (GO) enrichment was calculated using genes near DMRs
700 between amniotic and nasal tissues, followed by a similarity matrix analysis. The similarity heatmap clusters significant GO terms into
701 related Biological Process (BP), with colour intensity representing the significance of enrichment. Similarity of related terms within the
702 same cluster is shown in the heatmap.
703


A
R = 0.8; p-value = 0.000
Amnion
Nasal
5.0
2.5
0.0
-2.5
-5.0
B
KCNS3
CLIC5
LRRK2
PTGIR
C2
LAMP3
SOSTDC1
C16ORF89
ANKRD29
SCGB3A1
DCSTAMP
SYNDIG1L
KHDC3L
NKD2
OVCH1
DNAH5
CHIA
RASGRF1
FHAD1
LILRB3
IRX5
GRP
SFTA3
TBX5
GPRIN2
MS4A15
ITGAE
SLC6A4
ATP13A4
MATN3
STAC
ABCA3
CRTAC1
AQP4
NAPSA
SFTPA1
RAMP3
LIMCH1
CES1
CEACAM6
Lung
Stomach
Cervix, uterine
Fallopian Tube
Cerebral Cortex
Skeletal Muscle
Heart Muscle
Thyroid Gland
Seminal Vesicle
Prostate
Skin
Esophagus
Breast
Bone Marrow
Tonsil
Pancreas
Salivary Gland
Liver
Adrenal Gland
Adipose Tissue
Spleen
Lymph Node
Appendix
Placenta
Kidney
Testis
Ovary
Smooth Muscle
Endometrium
Urinary Bladder
Gallbladder
Rectum
Colon
Duodenum
Small Intestine
C
BP: 117 GO terms
establishment inorganic transmembrane cation
transport localization macromolecule
signaling response intracellular cell positive
signal stimulus regulation
communication transduction
response
system neuron tube morphogenesis
multicellular embryo organismal structure
development anatomical
organization component cellular projection
cytokine production positive rna transcription process
metabolic
organismal multicellular process negative regulation
cycle cell mitotic
Similarity
1
0.5
0

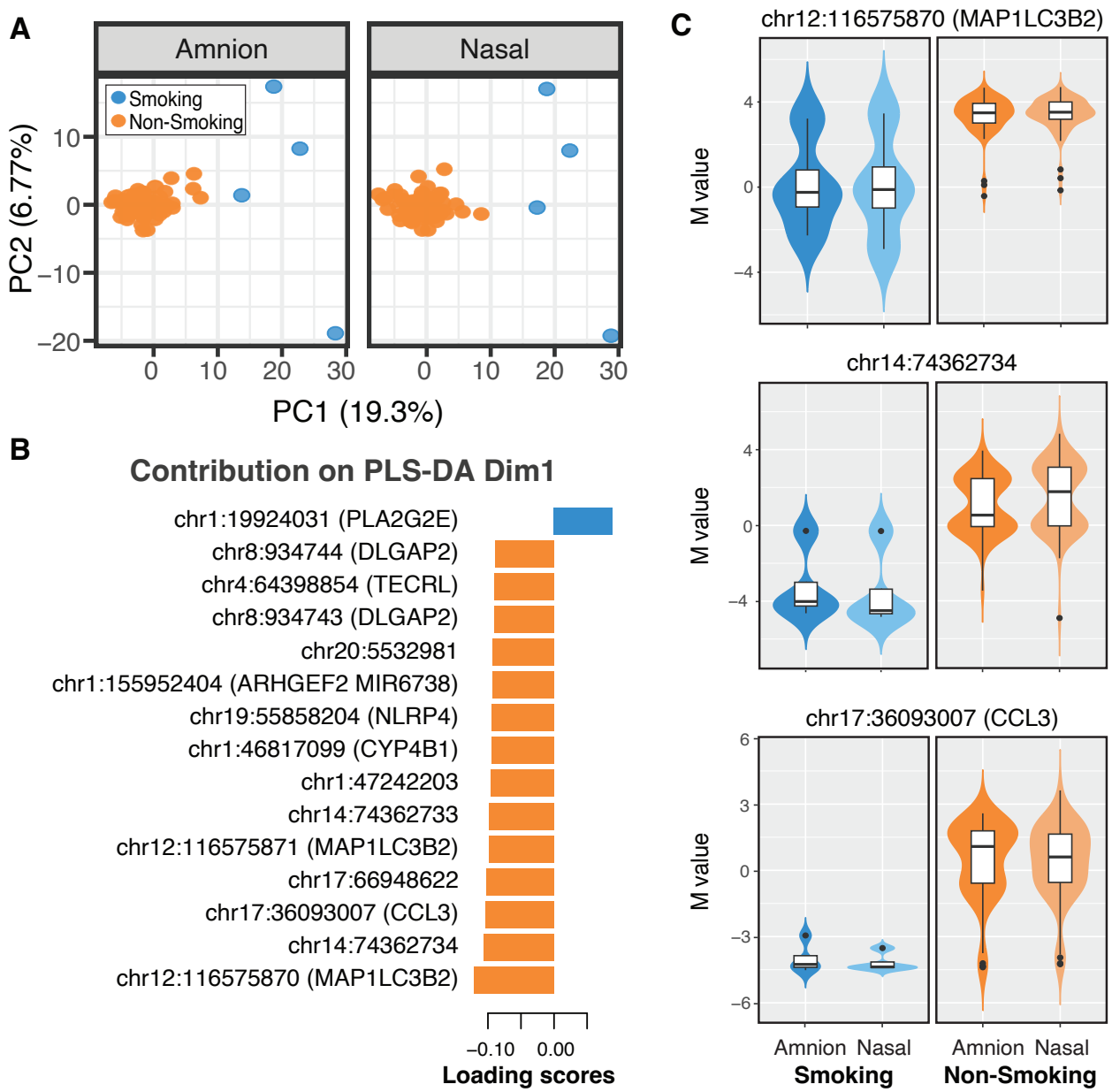
704

705 **Figure 3. Conserved methylation landscapes were enriched in cell differentiation and morphogenesis pathways.** (A) Scatter plot
706 comparing DNA methylation levels between amnion and nasal tissues. Each dot represents a genomic region, with its position indicating
707 the methylation level in each tissue. The strong positive correlation (Pearson's $R \geq 0.8$, p-value = 0.000) suggests a high degree of
708 conservation in methylation patterns between these tissues. (B) Heatmap illustrating the expression levels of lung- and epithelial-specific
709 genes across various tissues. Each row represents a gene, and each column represents a tissue. The colour intensity indicates the level
710 of expression, with red representing high expression and blue representing low expression. Data are derived from the GTEx catalogue.
711 (C) Gene Ontology (GO) enrichment was calculated using genes near CpG sites from the conserved methylation signature, followed by
712 a similarity matrix analysis. The similarity heatmap displays GO terms associated with Biological Process (BP), with colour intensity
713 representing the significance of enrichment.


A
Amnion
Nasal
Asthma
Non-Asthma
PC2 (11.89%)
PC1 (24.65%)
B
Contribution on PLS-DA Dim1
chr18:45723869 (SLC14A1 LOC105372092-3)
chr22:44569716 (LINC00207)
chr12:22507624 (C2CD5 C2CD5–AS1)
chr10:3332958
chr2:10529805
chr4:1669126 (FAM53A)
chr19:15440058 (WIZ)
chr16:86547583 (MTHFSD)
chr16:86547584 (MTHFSD)
chr2:241880445 (LINC01237)
chr1:6493339 (PLEKHG5)
chr6:42222706 (MRPS10)
chr3:59456728 (CFAP20DC–DT)
chr2:10529804
chr16:86536881 (MTHFSD)
–0.20
–0.05
0.10
Loading scores
C
chr16:86536881 (MTHFSD)
chr2:10529804
chr3:59456728 (CFAP20DC–DT)
M value
Amnion
Nasal
Asthma
Non-Asthma

714

715 **Figure 4. Methylation changes associated with maternal asthma in the conserved signature.** (A) Principal Component Analysis
716 (PCA) of DNA methylation variation in newborn amnion and nasal tissues, showing the first two principal components (PC1 and PC2).
717 Samples are coloured based on maternal asthma history, with asthma in orange and non-asthma in blue. (B) Bar plot of loading scores
718 from Dimension 1 (Dim1) of a Partial Least Squares Discriminant Analysis (PLS-DA), highlighting the top 15 differentially methylated
719 CpG sites associated with maternal asthma. (C) Violin plots of the top three CpG sites identified by PLS-DA, showing the distribution
720 of methylation levels (expressed as M-values) at features associated with maternal asthma in newborn amnion and nasal tissues, further
721 stratified by maternal asthma history during pregnancy status.


A
Amnion
Nasal
Smoking
Non-Smoking
PC2 (6.77%)
PC1 (19.3%)
B
Contribution on PLS-DA Dim1
chr1:19924031 (PLA2G2E)
chr8:934744 (DLGAP2)
chr4:64398854 (TECRL)
chr8:934743 (DLGAP2)
chr20:5532981
chr1:155952404 (ARHGEF2 MIR6738)
chr19:55858204 (NLRP4)
chr1:46817099 (CYP4B1)
chr1:47242203
chr14:74362733
chr12:116575871 (MAP1LC3B2)
chr17:66948622
chr17:36093007 (CCL3)
chr14:74362734
chr12:116575870 (MAP1LC3B2)
Loading scores
C
chr12:116575870 (MAP1LC3B2)
chr14:74362734
chr17:36093007 (CCL3)
M value
Amnion Nasal
Smoking
Non-Smoking

722

723 **Figure 5. Methylation changes associated with maternal smoking exposure in the conserved signature.** (A) Principal Component
724 Analysis (PCA) of DNA methylation variation in newborn amnion and nasal tissues, showing the first two principal components (PC1
725 and PC2). Samples are coloured based on maternal smoking exposure, with smoking exposure in orange and non-smoking exposure in
726 blue. (B) Bar plot of loading scores from Dimension 1 (Dim1) of a Partial Least Squares Discriminant Analysis (PLS-DA),
727 highlighting the top 15 differentially methylated CpG sites associated with maternal smoking exposure. (C) Violin plots of the top
728 three CpG sites identified by PLS-DA, showing the distribution of methylation levels (expressed as M-values) at features associated
729 with maternal smoking exposure in newborn amnion and nasal tissues, further stratified by prenatal smoking status.

730
731