
Peer reviewed version

Link to published version (if available):
10.1111/apa.14027

Link to publication record in Explore Bristol Research

PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at https://onlinelibrary.wiley.com/doi/full/10.1111/apa.14027. Please refer to any applicable terms of use of the publisher.

**University of Bristol - Explore Bristol Research**

**General rights**

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/
DNA methylation profiles between airway epithelium and proxy tissues in children

Rossa Brugha1, Robert Lowe2, A. John Henderson3, John W. Holloway4, Vardhman Rakyan2, Eva Wozniak5, Nadiya Mahmud5, Kay Seymour6, Jonathan Grigg2 Seif O. Shaheen2.

1National Heart and Lung Institute, Imperial College London, London, UK.

2Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary, University of London, London, UK

3School of Social and Community Medicine, University of Bristol, Bristol, UK.

4Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK.

5Genome Centre, Barts and The London School of Medicine and Dentistry, London, UK.

6Ear, Nose and Throat Surgery, Barts Health NHS Trust, London, UK.

Correspondence and reprint requests to: Prof Seif Shaheen, Department of Primary Care and Public Health, Blizard Institute, Barts and The London School of Medicine and Dentistry, 58 Turner Street, London E1 2AB.

Email: s.shaheen@qmul.ac.uk

Tel: +44 (0)20 7882 2480; Fax: +44 (0)20 7882 2552

Key words: Epigenetics, asthma and early wheeze

Short title: Airway DNA methylation Word count: 2396
Abstract

Background

Epidemiological studies of DNA methylation in airway disease have largely been conducted using blood or buccal samples. However, given tissue specificity of DNA methylation, these surrogate tissues may not allow reliable inferences about methylation in the lung.

Objective

To compare the pattern of DNA methylation in blood, buccal and nasal epithelial cells to that in airway epithelial cells from children.

Methods

Samples of blood, and buccal, nasal and airway epithelium were obtained from six children undergoing elective anaesthesia for adenotonsillectomy. DNA methylation was assessed at 450,000 sites using the Illumina HumanMethylation450 array.

Results

Eighteen samples from all sites were suitable for analysis. Hierarchical clustering demonstrated that the methylation profile in nasal epithelium was most representative of that in airway epithelium; the profile in buccal cells was moderately similar; and that in blood was least similar.

Conclusion

DNA methylation in blood poorly reflects methylation in airway epithelium. Future epidemiological studies of DNA methylation and airway diseases should consider measurement of methylation either in buccal cells or, preferably, in nasal epithelial cells.
Article summary

DNA methylation is tissue specific, and primary tissue may not be available for large scale studies of epigenetic markers in airways disease. We sought to compare nasal, buccal, airway and blood DNA methylation in order to establish a best proxy for airway cells. We demonstrate that DNA methylation in nasal epithelium is the most representative of airway DNA methylation.


Introduction

There is growing interest in the role of epigenetic control of gene expression in the lung [1], and in the likely importance of epigenetic mechanisms in the aetiology of asthma [2-4]. Consequently there have been calls to include epigenetics in epidemiological studies of asthma [5] and evidence has been gradually accumulating to implicate altered DNA methylation, the most readily studied epigenetic process in population-based studies, in childhood asthma. For example, differential methylation has been reported in relation to childhood wheezing and asthma [6, 7], and to risk factors such as maternal smoking in pregnancy [8, 9], although epidemiological data confirming whether the associations between early life risk factors and asthma are mediated through altered DNA methylation are still sparse. Experimental evidence in mouse models has provided further evidence for a potential role of altered DNA methylation in asthma [10-12].

Epidemiological studies of epigenetics and asthma to date have tended to measure DNA methylation predominantly using DNA extracted from peripheral blood leukocytes [6, 7, 9], and sometimes buccal samples [8], as these sources of DNA are readily accessible and direct lower airway sampling is generally not feasible in population-based studies. A potential difficulty with this approach is that DNA methylation tends to be highly tissue-specific [13] and even cell-type specific [14], and hence these tissues may not allow reliable inferences about methylation in the lung. It has been argued that methylation measured in buccal DNA samples may better reflect methylation in the lung than methylation measured in peripheral blood leukocyte DNA [15], and we recently showed that buccal samples are likely to be a more informative surrogate tissue than
blood for epigenome-wide association studies (EWAS) of non-blood based diseases/phenotypes, especially those linked to epithelial disorders such as paediatric eosinophilic oesophagitis [16]. However, the most informative proxy tissue may be nasal epithelium, given that the nose and airways form a continuous tract, and that nasal epithelial cells have been used as surrogates for bronchial epithelial cells in studies of airway inflammation [17]. To date, no comparison has been made between the DNA methylation profiles of airway and nasal epithelial cells, buccal cells and blood to confirm the best proxy tissue for epidemiological studies of airway diseases. We have therefore sought to establish which proxy tissue exhibits the closest pattern of epigenome-wide DNA methylation to that of airway epithelial cells. Our a priori hypothesis was that methylation profiles from nasal epithelial cells would show a greater similarity than those from buccal samples and blood.
**Methods**

**Participants**

In order to obtain uncontaminated airway cells, children requiring elective endotracheal intubation were recruited to the study. Six children (four female, two male) aged 5-13 years undergoing elective tonsillectomy or adenotonsillectomy were recruited from the Ear, Nose and Throat clinic at the Royal London Hospital, three of whom had doctor diagnosed asthma. Children with a respiratory condition other than asthma were excluded. Ethical approval was obtained from the National Research Ethics Service Committee London - Riverside (REC 11/LO/1579), and parents or carers consented for study procedures.

**Sampling**

Samples from 4 tissue sites (airway, nasal, buccal and blood) were obtained. Buccal cells were collected pre-operatively using swabs (Isohelix SK-2S, Cell Projects, Kent, UK). Immediately following intubation, tracheal epithelial cells were collected by performing blind brushings via the endotracheal tube using a 2mm cytology brush (BC-202D-5010, Olympus KeyMed, Essex, UK). This has been shown to be an effective and safe method in children when bronchoscopy is not feasible [18, 19]. Nasal epithelial cells were obtained by directly observed curettage of the inferior nasal turbinates (Rhino-Pro® Nasal Mucosal Curette, Arlington Scientific, Utah, USA). Peripheral blood was obtained by venesection and mixed with EDTA (BD vacutainer, Becton Dickinson, Oxford, UK).
Measurement of DNA methylation

DNA Extraction and Quality control

Upon receipt, buccal swabs were treated as per the Isohelix DNA Isolation Kit DDK-50 protocol Part A, until appropriate numbers for extraction had been received. Nasal mucosal samples and airway brushings were immersed in 0.9% sodium chloride and frozen at -80°C. After thawing, tubes were vortexed to loosen collected cells, probes and brushes were removed, and the resulting cell suspension was then centrifuged at 3,500 RCF for 10 minutes in a bench-top centrifuge to pellet the cells. The supernatant was removed and DNA extraction of buccal swabs, nasal and airway samples was carried out as in the Isohelix protocol (beginning at Part B for buccal swabs), with the optional step 6 included, and the elution volume used was 30µl. Blood samples were stored at -20°C until extraction. Extraction was performed using a salting out procedure [20]. Briefly, blood samples were defrosted overnight before 40mls of cold water were added to lyse red cells, followed by centrifugation at 3000 RCF for 20 minutes at 4°C to precipitate white cells. Supernatant was discarded and the pellet re-suspended in 40mls solution 1 (0.32M Sucrose, 0.01M Tris-Cl, pH 7.6, 0.005M MgCl₂, 1% TRitonX100, 0.02 Sodium Azide). Samples were centrifuged at 3000 RCF for 20 minutes at 4°C. The supernatant was discarded and the pellet re-suspended in 30mls solution 1. Tubes were centrifuged at 3000 RCF for 20 minutes at 4°C and the pellet re-suspended in 11mls of Solution 2 (0.05M Tris-Cl, pH 8, 0.02M EDTA, pH 8, 2% SDS). Samples were digested with 1mg Proteinase K overnight at 37°C followed by protein precipitation with saturated 4 mls NaCl solution. Protein was precipitated by centrifugation at 3500 RCF for 20min at 4°C and DNA was precipitated with 3x volume 100% Ethanol. DNA quantity and purity were measured using a Nanodrop8000 spectrophotometer (Thermofisher Scientific, Delaware, USA). A Qubit® 2.0 fluorometer
LifeTechnologies, Paisley, United Kingdom) with the Qubit® dsDNA HS kit was used to quantify intact double-stranded DNA. Samples were also checked for RNA contamination by an Agilent RNA 6000 Nano kit on the Agilent 2100 Bioanalyser (Agilent Technologies, California, USA).

**Bisulphite conversion and Quality control**

Any sample with less than 500ng total DNA quantity was excluded from further analysis. 500ng of DNA were bisulphite converted using the Zymo EZ-96 DNA Methylation Kit D5004 (Zymo Research Corp, California, USA). Bisulphite conversion was performed as per manufacturer’s instructions but with incubation conditions recommended by Illumina for the Infinium® Methylation arrays. The resulting quantity of single-stranded, converted sample was measured using the Nanodrop with the RNA setting. An initial QC of the bisulfite conversion efficiency suggested one of the samples had a particularly low level and hence this was removed before normalisation.

**Illumina Infinium® Methylation 450K arrays**

Samples with a converted concentration of less than 10ng/µl were excluded from array processing. CpG site methylation was analysed at 485,577 sites using the Illumina Infinium HumanMethylation 450 Beadchip (450k). This has been validated [21] and recently used to demonstrate differential methylation in relation to maternal smoking in pregnancy [8, 9]. 4µl of bisulphite converted DNA was used and samples were processed according to the Illumina Infinium®HD Meth450 Protocol (Part No 15019522) and scanned on the Illumina iScan (San
Diego, CA). Initial data QC was performed using the Genome Studio analysis package (Illumina (San Diego, CA)).

**Statistical analysis**

Raw intensity values extracted from the IDAT files were quantile normalised separately for each colour channel as well as probe type using a custom script. We remapped the genomic sequence of all probes to hg19 and removed those probes which mapped with >90% similarity to more than one genomic location as well as those probes which mapped to either chromosome X or Y. Furthermore we removed those probes which recorded intensity levels around background by using a detection p-value <0.01. This left 441,946 probes which were used for further analysis.

We calculated beta values using the following equation [21] and used this as a measure of methylation across the 441,946 probes:

\[
\beta = \frac{I_m}{I_m + I_u + 100}
\]

where \(I_m\) is the intensity in the methylated channel and \(I_u\) is the intensity in the unmethylated channel. To ascertain how similar different tissue types are in terms of their methylation state we performed an unsupervised hierarchical clustering. We first calculated the Euclidean distance (d) between each of the samples by summing the square of the beta differences across all of the probes. Then we clustered the samples using a mean or average linkage. To further identify how close nasal, buccal and blood are in terms of methylation state to airway epithelium we first averaged the beta values of the samples for each tissue and then calculated the Euclidean
distance by summing the square of the differences between the average beta value of airway
epithelium samples and each of the other tissues.
Results

Characteristics of the children are shown in Table 1. One nasal sample and three buccal samples were insufficient for analysis, one blood sample was misplaced, and one sample failed bisulphite conversion, leaving 18 samples (plus 5 duplicate samples, i.e. 23 arrays in total) which were suitable for analysis.

Figure 1 demonstrates the hierarchical clustering of the samples. Taking into account the 441,946 probes for each sample, the diagram represents the difference in methylation at the same site in the DNA sequence between tissues. The greater the distance (represented by height on the diagram), the greater the difference between tissues. This shows that the methylation profile of nasal epithelial cells was most similar to that of airway epithelial cells; the profile of buccal cells was moderately similar; that of blood was least similar. When we quantified this by calculating a standard Euclidean distance of the average methylation beta values for each tissue from the average value for airway epithelial cells, the distances for nasal, buccal and blood were, respectively, 46.1, 72.3, and 91.4. To further evaluate this we carried out a Wilcoxon rank sum test on the Euclidean distances for each sample of nasal, buccal and blood against the average methylation for airway epithelial cells. Both nasal (p-value = 0.004329) and buccal (p-value = 0.009524) were significantly closer to airway epithelial cells than blood.

Figure 2 demonstrates the extent of methylation for the 20,000 most variable probes (defined by having the largest variance in beta values across all of our samples), where yellow represents hypomethylation and blue represents hypermethylation; this gives a visual representation of the difference in methylation across the separate tissues, with blood clearly separate from the
respiratory tract samples. We analysed the blood cell type data for the same 20,000 variable probes analysed in Figure 2. We found that nasal, airway and buccal cells clustered together as before, while whole blood clustered with the various blood subsets (data not shown), confirming that our conclusions above are not affected by blood cell type.

Figure 3 demonstrates the hierarchical clustering of the samples when we limited the analyses to children with asthma (n=3). This confirmed a similar hierarchy of methylation profiles across tissues to that seen in all children. Euclidean distances of the average methylation beta values for each tissue from the average value for airway epithelial cells, were 42.5, 84.2 and 92.8 for nasal, buccal and blood respectively.
Discussion

Asthma is primarily an epithelial disorder [22] and hence increasing our understanding of epigenetic changes in airway epithelium may shed light on asthma aetiology. In epidemiological studies of the epigenetics of asthma, in which airway epithelial cells cannot be obtained, we therefore need to focus efforts on obtaining the most representative surrogate tissue. In this study we have shown that nasal epithelial cells are a better proxy tissue than buccal cells for making inferences about epigenome-wide methylation patterns in airway epithelial cells, with blood being the worst proxy tissue. Our findings suggest that, whilst studies to date which have measured DNA methylation in blood may have obtained some useful information, future studies of the epigenetics of asthma are likely to prove more informative if buccal cells, or preferably nasal epithelial cells, are collected. We acknowledge, though, that blood DNA methylation profiling may still be informative for understanding epigenetic influences on the immune (atopic) component of asthma, as suggested by previous studies of Th1/Th2 gene methylation [23-25] or possibly in relation to DNA methylation changes induced by exposures that occur in utero or are inherited as a result of trans-generational transmission of DNA methylation.

Our conclusion that blood methylation profiles may not be informative for making inferences about DNA methylation in the airway is supported by a recent study of children which compared DNA methylation profiles between airway epithelial cells (AECs) and peripheral blood mononuclear cells (PBMCs)[26]. The authors demonstrated differential methylation of CpG sites in AECs compared with PBMCs, irrespective of disease phenotype, but nasal epithelial cells and buccal cells were not studied. Although we did not separate out PBMCs before DNA extraction in our study, and we did not take into account the differential white cell count in our analysis of DNA methylation in blood, it has recently been shown that polymorphonuclear leucocytes and
PBMCs do not differ in terms of methylation profiles [9]. Whilst it was recently proposed that nasal cell DNA may be a more appropriate and accessible proxy tissue for studying the epigenetics of asthma [27], airway epithelial cells and other proxy tissues were not compared in that study.

A potential limitation of our study is that the sample size was small, as we experienced a number of refusals from parents because of concerns about taking airway brushings and prolonging the anaesthetic time. As our study includes children from multiple ethnic backgrounds, it cannot be excluded that ethnic differences may have influenced these findings. Despite this, by taking multiple samples from the same subject, we were able to demonstrate that a clear hierarchy of similarity between the methylation profiles of different proxy tissues and that of airway epithelial cells exists within individuals. Furthermore, a similar hierarchy was seen in children with asthma. Our nasal samples were acquired using a small curette which is an uncomfortable procedure when awake. Whilst tolerated well by adults it would be challenging to perform in younger children when not anaesthetised. Soft brushing of the nasal epithelium using a swab, as described by Baccarelli et al [27], offers an alternative technique for acquiring nasal epithelial cells in epidemiological studies, and may be more suitable especially in young children.

In summary, we have shown that the best surrogate tissue for capturing methylation marks relevant to airway epithelium is the nasal epithelium, followed by buccal cells. DNA methylation profiling of blood samples does not accurately reflect the pattern of methylation seen in the airway. Future epidemiological studies of the role of DNA methylation in airways diseases
should take this into account, and existing or forthcoming data from studies based on methylation in blood should be interpreted with caution.

Acknowledgements

We would like to thank the children and families who took part, and the nursing, theatre and anaesthetic staff who facilitated sampling.

Competing interests

No competing interests to declare

**Funding:** Asthma UK, Barts and The London Charity
References


**Legends for figures**

**Figure 1:** Hierarchical cluster dendrogram of DNA methylation by tissue type. The respiratory tract tissues cluster together, separately from blood. The dendrogram is calculated using agglomerative hierarchical clustering. The root of the dendrogram is at the top of the figure and its y-axis represents the Euclidean distance between clusters at each splitting point. A greater distance indicates greater differences between tissues.

**Figure 2:** To visually inspect the data we took 20,000 probes with the highest variability across all samples. The rows of the heatmap represent an individual probe while the columns represent an individual sample. The beta value as recorded from the array is plotted from yellow to blue with yellow representing low methylation, green fractional methylation and blue high methylation. In the top part of the heatmap the respiratory tract tissues contain cytosines in the genome that are hypomethylated compared to blood. In the lower half of the heatmap the majority of probes are specifically hypomethylated in blood.

**Figure 3:** Hierarchical cluster dendrogram of DNA methylation by tissue type restricted to children with asthma.