Check for updates

DR. MICHAEL FRICKER (Orcid ID : 0000-0002-8587-1774)
DR. KIAN FAN CHUNG (Orcid ID : 0000-0001-7101-1426)

Article type : Original Article

Correspoding author mail id :- Michael.fricker@newcastle.edu.au

An altered sputum macrophage transcriptome contributes to the neutrophilic asthma endotype

Manuscript Acceptance Date: 21-Aug-2021

Running title: Transcriptomic features of sputum macrophages in asthma

Authors: Michael Fricker[1,2,3], Ling Qin[4], Stephany Sánchez-Ovando[1,3], Jodie L Simpson[1,3,5], Katherine J Baines[1,3], Carlos Riveros[6], Hayley A Scott[3,7], Lisa G Wood[3,7], Peter AB Wark[1,3,5], Nazanin Z Kermani[8,9], Kian Fan Chung[8,9], Peter G Gibson[1,2,3,5].

Author affiliations:

[1] School of Medicine and Public Health, Faculty of Health and Medicine and Priority Research Centre for Healthy Lungs, The University of Newcastle, NSW, Australia

[2] National Health and Medical Research Council Centre for Excellence in Severe Asthma, Newcastle, NSW, Australia

[3] Hunter Medical Research Institute, Newcastle, NSW, Australia

[4] Department of Respiratory Medicine, Department of Pulmonary and Critical Care Medicine, Xiangya Hospital, Central South University, Changsha, China

[5] Department of Respiratory and Sleep Medicine, John Hunter Hospital, Newcastle, NSW, Australia

[6] Statistical services (CReDITSS), Hunter Medical Research Institute, Newcastle, NSW, Australia

[7] School of Biomedical Sciences and Pharmacy, Faculty of Health and Medicine, Priority Research Centre for Healthy Lungs, The University of Newcastle, NSW, Australia

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/ALL.15087

This article is protected by copyright. All rights reserved

Accepted Article

[8] Data Science Institute, Imperial College London, London, UK.
[9] National Heart and Lung Institute, Imperial College London, Dovehouse St, London, SW3 6LY, UK.

Name and address of corresponding author:
Michael Fricker
Address: Priority Research Centre for Healthy Lungs, Hunter Medical Research Institute, Lot 1 Kookaburra Circuit, New Lambton Heights, NSW, 2305, Australia.
Phone: +61 2 404 20207 Fax: +61 2 404 20046
Email: michael.fricker@newcastle.edu.au

Author contributions: M.F. designed, performed and supervised laboratory studies, designed and performed data analysis, interpreted data and wrote the manuscript. L.Q. performed laboratory studies and edited the manuscript. S.S-O designed and performed data analysis and edited the manuscript. J.L.S. supervised clinical studies and edited the manuscript. K.J.B. supervised clinical studies, designed analysis and edited the manuscript. C.R. designed and performed data analysis, interpreted data and edited the manuscript. H.A.S. supervised clinical studies and edited the manuscript. L.G.W. supervised clinical studies and edited the manuscript. P.A.B.W., N.Z.K., and K.F.C. supervised data analysis and edited the manuscript. P.G.G. designed laboratory studies, supervised clinical studies, designed data analysis, interpreted data and edited the manuscript.

Acknowledgments
The authors would like to thank the clinical and laboratory staff who conducted study visits and sample processing: Netsanet Negewo, Penny Chan, Catherine Delahunty, Lakshitha Gunawardhana, Bridgette Donati, Melanie Erriah, Amber Smith. We would also like to thank Nicole Cole and Jessica Weaver for technical advice.

Funding: This study was supported by the National Health and Medical Research Council (NHMRC) Centre of Research Excellence in Severe Asthma (project identifier 1078579), NHMRC project grants, Asthma Australia, and the John Hunter Hospital Charitable Trust. M.F. was supported by a Thoracic Society of Australia and New Zealand-AstraZeneca Respiratory

This article is protected by copyright. All rights reserved

Accepted Article

Research Fellowship. P.G.G. was supported by an NHMRC Practitioner Fellowship. K.J.B. was supported by an Australian Lung Foundation COPD Research Fellowship.

Conflict of Interest statement: Dr. Fricker reports grants from NHMRC, grants from Thoracic Society of Australia and New Zealand, during the conduct of the study, and grants from GlaxoSmithKline outside the submitted work. Dr. Baines reports grants from NHMRC, grants from Australian Lung Foundation, grants from HMRI, outside the submitted work. Dr. Gibson reports personal fees from AstraZeneca, GlaxoSmithKline, Novartis, grants from AstraZeneca, GlaxoSmithKline, outside the submitted work. L.Q., S.S-O., J.L.S., C.R., H.A.S. and L.W. have nothing to disclose.

Accepted Article

This article is protected by copyright. All rights reserved

Accepted Article

Abstract

Background: Neutrophilic asthma (NA) is a clinically important asthma phenotype, the cellular and molecular basis of which is not completely understood. Airway macrophages are long-lived immune cells that exert important homeostatic and inflammatory functions which are dysregulated in asthma. Unique transcriptomic programmes reflect varied macrophage phenotypes *in vitro*. We aimed to determine whether airway macrophages are transcriptomically altered in NA.

Methods: We performed RNASeq analysis on flow cytometry-isolated sputum macrophages comparing NA (n=7) and non-neutrophilic asthma (NNA, n=13). qPCR validation of RNASeq results was performed (NA n=13, NNA n=23). Pathway analysis (*PANTHER, STRING*) of differentially expressed genes (DEGs) was performed. Gene set variation analysis (GSVA) was used to test for enrichment of NA macrophage transcriptomic signatures in whole sputum microarray (cohort 1 - controls n=16, NA n=29, NNA n=37; cohort 2 U-BIOPRED - controls n=16, NA n=47, NNA n=57).

Results: Flow cytometry sorting significantly enriched sputum macrophages (99.4% post-sort, 44.9% pre-sort, $p<0.05$). RNASeq analysis confirmed macrophage purity and identified DEGs in NA macrophages. Selected DEGs (*SLAMF7, DYSF, GPR183, CSF3, PI3, CCR7,* all $p<0.05$ NA vs NNA) were confirmed by qPCR. Pathway analysis of NA macrophage DEGs was consistent with responses to bacteria, contribution to neutrophil recruitment and increased expression of phagocytosis and efferocytosis factors. GSVA demonstrated neutrophilic macrophage gene signatures were significantly enriched in whole sputum microarray in NA vs NNA and controls in both cohorts.

Conclusions: We demonstrate a pathophysiologically-relevant sputum macrophage transcriptomic programme in NA. The finding that there is transcriptional activation of inflammatory programmes in cell types other than neutrophils supports the concept of NA as a specific endotype.

Keywords

Asthma; Endotype; Macrophage; Neutrophil; Transcriptome

Abbreviations

AHR	Airways hyperresponsiveness

This article is protected by copyright. All rights reserved

ES	Enrichment score
FEV_1	Forced expiratory volume in 1 second
GCSF	Granulocyte colony stimulating factor
GO	Gene ontology
GSVA	Gene set variation analysis
IFN	Interferon
IL	Interleukin
LPS	Lipopolysaccharide
NA	Neutrophilic asthma
NNA	Non-neutrophilic asthma
QC	Quality control
qPCR	Quantitative polymerase chain reaction
RNASeq	RNA sequencing
TNF	Tumour necrosis factor

Accepted Article

This article is protected by copyright. All rights reserved

Accepted Article

Introduction

Airway inflammation causes asthma pathophysiology and symptoms through promotion of airway remodelling, airways hyperresponsiveness and asthma flare-ups. Airway inflammation in asthma can be categorized into eosinophilic/non-eosinophilic or type 2 high/type-2 low based on the presence or absence of airway and systemic biomarkers[1,2]. Eosinophilic airway is more responsive to inhaled corticosteroid treatment in mild-moderate asthma, and can be targeted with biologicals blocking type 2 cytokine signalling in severe asthma to reduce flare-ups[1,3]. In contrast, effective targeted anti-inflammatory treatment for non-eosinophilic asthma is lacking[2]. Greater understanding of the cellular and molecular inflammatory pathways present in non-eosinophilic asthma are required to identify new therapeutic targets.

Non-eosinophilic asthma can be sub-categorized by presence or absence of elevated sputum neutrophils. Neutrophilic asthma (NA) is associated with corticosteroid resistance, asthma flare-ups, airflow limitation and airway dysbiosis[4-6]. The cellular and molecular pathways that promote airway neutrophilia and clinical features of NA are incompletely understood. Transcriptomic profiling of whole sputum samples demonstrated alterations in NA[7] or a transcriptomic cluster with elevated neutrophils[8], including enrichment of gene expression relating to IL-1β, TNF and IFNγ signalling. It is unclear if sputum transcriptomic alterations in NA reflect contribution of altered proportions of distinct immune cell types to the total RNA pool, and/or reflect a genuine transcriptional activation of inflammatory programmes in specific cell types. Better understanding of cell-specific dysregulation of inflammatory pathways and their contribution to airway inflammation in asthma may inform development of new therapeutics.

Macrophages are a key immune cell of the airway lumen that mount functionally diverse responses and exert important homeostatic and regulatory functions[9]. Macrophage functions are dysregulated in asthma and may relate to NA[10]. *In vitro* studies have demonstrated that transcriptomic alterations underpin functional differences in macrophages exposed to varying type 1 (yielding M1 macrophages) and type 2 (yielding M2 macrophages) inflammatory stimuli[10,11]. Sputum macrophages constitute a major lumenal immune cell population across asthma airway inflammatory phenotypes. Their longevity as well as phenotypic and functional plasticity prompt the concept that they may play context-specific roles in promoting/regulating asthma airway inflammatory phenotypes. We hypothesized that sputum macrophages would display

This article is protected by copyright. All rights reserved

transcriptomic alterations in NA, and that these alterations would also be reflected in previously described transcriptomic data derived from whole sputum samples. To address this, we developed a flow cytometry-based method to isolate highly pure preparations of sputum macrophages and performed an exploratory transcriptomic analysis of sputum macrophages using a bulk RNAseq approach to compare macrophage transcriptomes in NA vs non-neutrophilic asthma (NNA).

Accepted Article

This article is protected by copyright. All rights reserved

Methods

Study population

Clinical studies were approved by the Hunter New England Health Human Research Ethics Committee (16/04/20/3.0; 17/02/15/3.04; 15/03/18/3.04; 17/04/12/4.03). All participants provided written informed consent. Participants were aged 18 or over and had a doctor diagnosis of asthma with documented evidence of variable airflow limitation in the past 10 years in the form of: AHR (PD_{15} < 15 mL to hypertonic saline OR mannitol); bronchodilator response (change post-bronchodilator FEV_1 > 12% OR 200 mL) or; FEV_1 variability > 12% (two values measured within 2 months of each other) or; Peak Flow variability > 12% over at least one week of monitoring. At time of visit, participants had stable asthma (no respiratory infection, asthma exacerbation or change in asthma maintenance therapy in the past month). Participants were never smokers or had no recent history of smoking (prior 6 months), were not pregnant or breastfeeding, and had no current lung cancer or other blood, lymphatic or solid organ malignancy. For U-BIOPRED cohort-based microarray analysis, severe asthma was defined as previously described[12]. For all other analyses, severe asthma was defined according to ERS/ATS guidelines[13].

Spirometry and sputum induction

Spirometry and sputum induction with hypertonic saline (4.5%) was performed and sputum cell suspensions and cytospins prepared as previously described (see online supplement)[14]. Samples were categorized into inflammatory subtypes based on sputum differential cell count using a threshold ≥61% neutrophils for NA vs <61% NNA[14]. Additional cell suspension was immediately processed for flow cytometry-mediated isolation of sputum macrophages.

Flow cytometry-mediated sputum macrophage isolation

General sputum flow cytometry methods including fluorescence minus one controls have been described[15], for details see online supplement. Sputum macrophages were identified as $CD45^+$, HLA-DR^+, $CD14^{intermediate}$, CD16 high, $CD206^+$, side scatter (SSC) high events (Figure 1), consistent with prior reports[16,17]. Sputum cell suspensions were run on a BD FACSAria III cell sorter, using a 70 micron filter and macrophages sorted directly into PBS or TRIzol LS (*ThermoFisher*) chilled at 4 °C for preparation of cytospins or RNA preparation respectively. Cytospins of sorted sputum macrophages were stained with May-Grünwald and Giemsa and conventional differential cell count performed on 400 cells.

This article is protected by copyright. All rights reserved

RNA isolation, cDNA synthesis, qPCR

RNA was isolated from macrophages lysed in TRIzol LS (see online supplement). 200-500 pg of sputum macrophage RNA was reverse transcribed to produce cDNA using a SuperScript™ IV VILO™ Master Mix (*ThermoFisher*) following manufacturer's instructions. qPCR was run using standard TaqMan reagents and probesets (*ThermoFisher*) as previously described[7]. Probesets used were: *SLAMF7* – Hs00904275_m1; *DYSF* – Hs01002513_m1; *GPR183* – Hs00953886_m1; *MERTK* – Hs01031979_m1; *CSF3* – Hs00738432_g1; *CCR7* - Hs01013469_m1*; PI3* - Hs00160066_m1. Relative mRNA abundance of target to the housekeeper *ACTB* was calculated ($2^{-\Delta Ct}$).

RNASeq library preparation and sequencing

Library preparation, library QC and sequencing service was provided by Ramaciotti Centre for Genomics (*University of New South Wales, Sydney, Australia*). cDNA library synthesis was conducted using a SMARTer® Stranded Total RNA-Seq kit v2 (Takara Bio), following the manufacturer's instructions. This library synthesis kit was selected due to its compatibility with picogram inputs of total RNA of variable quality. Between 500-1000 pg total RNA was used as input for cDNA library preparation. All samples processed successfully yielded libraries based on confirmation of expected DNA concentration and lack of dimer peaks. Libraries were sequenced on the NovaSeq 6000 platform (single NovaSeq S1 2 x 100bp lane), generating an average of 45.5 M paired end reads per sample (lowest was 36.3 M).

Bioinformatics processing of RNASeq data and differential gene expression analysis

Paired read RNASeq data analysis consisted of paired adapter and quality trimming (TrimGalore 0.6.3, cutadapt 1.18), followed by sequence alignment (HISAT2 2.0.2[18]) and count (subread 1.6.4[19]) against the Homo sapiens GRCH38 v97 transcriptomic reference plus decoys. Quality control was performed at every step with FastQC and Multiqc[20], mapping efficiency was above 90% and 72% of paired end reads had unique mappings on average. An average of 25% of unique reads were assigned to transcripts. Differential expression analysis was performed in R with the DESeq2 package[21] and included logarithmic fold change and significance p-value estimation (s-value) with the approximate posterior estimation for generalised linear model (apeglm) package[22],

This article is protected by copyright. All rights reserved

which outperforms the normal shrinkage estimators. A detailed report on bioinformatics analysis pipelines is provided in the supplemental data (Figure E1).

Gene ontological analysis

Gene ontological (GO) analysis was performed using a PANTHER Overrepresentation Test (www.pantherdb.org) using the GO Biological Process annotation data set and fisher's exact test with correction for false discovery rate. STRING functional protein association network tool (https://string-db.org/, version 11) was used to investigate protein-protein interactions of DEGs. A minimum required interaction score of high confidence (>0.7) was applied and only the 1st shell query proteins were included in the analysis for the maximum number of interactions. Network edges were marked as confidence where line thickness indicates strength of data support. Disconnected nodes and interactions between only 2 proteins were removed from the analysis.

Gene set variation analysis (GSVA)

GSVA was used to identify differences in expression of predefined NA macrophage gene sets in a two previously published whole sputum microarray datasets comparing non-asthma controls, NA and NNA: the first dataset primarily included participants with mild-moderate asthma[7,23,24], the second being the U-BIOPRED sputum microarray sub-cohort which primarily included participants with severe asthma[8,25]. An enrichment score (ES) ranging from -1 to +1 was calculated for the expression of each gene set for each sample across the whole sample population.

Statistics

Mann-Whitney test was used for two group comparisons and Kruskal-Wallis with Dunn's multiple comparisons for multiple group comparisons of non-parametric data. P-values <0.05 were considered statistically significant.

This article is protected by copyright. All rights reserved

Results

Flow cytometry isolation of sputum macrophages

Induced sputum cell suspensions were rapidly processed for flow cytometry-mediated sorting of sputum macrophages. Sputum macrophages were identified as a distinct population on the basis of CD45+, HLA-DR+, CD14+, CD16+, CD206+ with a high side scatter profile (Figure 1A). Other immune cells, including monocytes, visible in subpanel 7 of Figure 1A as a population with low SSC and low CD206 signal, were excluded through gating. Differential cell count of unsorted sputum cell suspensions vs flow cytometry sorted macrophages (Figure 1B) confirmed the enrichment high level of purity of macrophages (mean 44.9% macrophages pre-sort, 99.4% macrophages post-sort, $p = 0.008$ Figure 1C).

Bulk transcriptomic analysis of isolated sputum macrophages by RNASeq

RNA sequencing was performed on 20 macrophage samples from participants with asthma (Table 1). Of the 20 samples, 7 were NA and 13 were of NNA airway inflammatory phenotype. Clinical features did not differ between NA and NNA participants, and sputum neutrophil and macrophage proportions differed as expected. Across all 20 samples high expression of macrophage signature genes and low to undetectable expression of signature genes representing lung neutrophils, monocytes, dendritic cells, mast cells, T- and NK cells and B cells [26] was observed (Figure E2), consistent with flow-cytometry isolation of sputum macrophages to a high degree of purity. PANTHER gene ontology (GO) analysis of the 200 most highly expressed RNA species across all macrophage samples revealed statistical overrepresentation of transcripts relating to cellular iron homeostasis, T cell mediated immunity, positive regulation of the adaptive immune response, wound healing, granulocyte chemotaxis, leukocyte chemotaxis, response to LPS, protein folding and actin filament organization, consistent with known macrophage roles in maintaining airway homeostasis and orchestration of immune responses (Table E1).

Differential gene expression analysis reveals transcriptomic alteration of sputum macrophages in neutrophilic asthma

We conducted an exploratory analysis to address our hypothesis that sputum macrophages would display inflammatory specific transcriptomic changes in NA. Analysis of NA (n=7) vs NNA (n=13) samples returned 44 RNA species with s-values (equivalent of p value) below 0.05, the

This article is protected by copyright. All rights reserved

majority of which were increased in NA vs NNA (Table E2). Thus, NA was associated with a specific transcriptomic alteration of sputum macrophages.

qPCR validation of NA macrophage DEGs

We ranked genes in terms of level of differential expression between NA and NNA using the ratio of s-value to (absolute) log fold change, with smaller values representing greater differential expression, generating a list of the 200 most differentially expressed genes, which included the 44 RNA species with s values below 0.05 (Table E2). To verify this list as truly differentially expressed between NA and NNA, we performed qPCR in a larger sample set of 36 flow cytometry-isolated sputum macrophage samples, 14 of which were included in the RNASeq cohort (NA n=13, NNA n=23, Table E3). For validation, we selected genes from the top 200 DEG list with known immune functions and with a range of average expression in the RNASeq dataset (Table E4: *MERTK, SLAMF7, DYSF, GPR183, CSF3, PI3, CCR7*). Expression was normalized to the housekeeper gene *ACTB* (β-actin), which was highly expressed across all samples and was not differentially expressed between NA vs NNA in the RNASeq dataset (log2 fold change -0.008, s value 0.98). All genes tested by qPCR were significantly increased in NA vs NNA in accordance with the RNASeq data, except *MERTK* which showed a similar trend (Figure 2A-G). Thus, qPCR confirmed increased expression of DEGs identified by RNASeq as increased in NA vs NNA sputum macrophages.

Pathway analysis of NA macrophage DEGs

We next performed GO and protein network analyses to characterise the altered transcriptomic landscape of macrophages in NA vs NNA. GO analysis using PANTHER of the top 200 DEGS from NA vs NNA macrophages revealed overrepresentation of genes involved in positive regulation of Janus kinase activity, regulation of granulocyte chemotaxis, regulation of neutrophil migration, positive regulation of leukocyte chemotaxis, regulation of the inflammatory response, response to bacterium, lymphocyte activation and defense response (Table 2). The STRING protein interaction network tool identified 163 protein coding transcripts amongst the top 200 DEGs, and three networks containing three or more members, including a major node containing 38 genes (Figure 3A-C), with a significant PPI enrichment value ($p<1x10^{13}$), indicating that the proteins identified were at least partially biologically connected as a group. STRING also

This article is protected by copyright. All rights reserved

identified statistical functional enrichment in numerous GO biological process categories, many of which were common with PANTHER GO analysis (Table E5).

Enrichment of the macrophage-derived NA signature in whole sputum samples

We performed GSVA analysis to test for enrichment of NA macrophage gene signatures in a whole sputum microarray dataset, which primarily featured participants with mild-moderate asthma (non-asthma controls n=16, NA n=29, NNA n=37, Table E6)[7,23]. Of the 200 most differentially expressed RNA species between NA and NNA macrophages, 150 were also present in the microarray dataset (gene set 1, Table E7). Gene set 1 was significantly enriched in NA vs NNA and non-asthma controls (Fig 4A). A significantly increased enrichment score was also observed in NA vs NNA when gene set 1 was restricted to genes which had s values below 0.05 (Figure 4B, gene set 2, Table E7). These results were then validated through analysis of whole sputum microarray data in a second independent cohort, the U-BIOPRED cohort, primarily composed of participants with severe asthma[8,25]. Both gene set 1 and 2 produced increased enrichment scores in NA vs NNA and non-asthma controls (Fig 4C-D). When the U-BIOPRED microarray cohort was further categorized into those with mild-moderate and those with severe asthma, we observed increased enrichment scores in NA irrespective of asthma severity, and no difference between mild-moderate and severe asthma enrichment scores in NA or NNA (Fig 4E-F).

Clinical correlations of macrophage-derived NA signature

Enrichment scores for gene set 1 (Fig 4A) were tested for correlation with clinical characteristics in participants with asthma (n=66). Increased expression of gene set 1 was significantly associated with increased age, worse lung function, poorer asthma control and increased ICS dose (Table 3).

This article is protected by copyright. All rights reserved

Accepted Article

Discussion

Macrophages are important, functionally diverse, innate immune cells of the airways whose dysregulation has been implicated in asthma. Altered transcriptomes are considered to underpin macrophage functional plasticity in response to varying extracellular queues, however it is unclear from transcriptomic studies of airway samples including sputum whether macrophage transcriptomic identity is altered in asthma. Here we performed bulk RNASeq analysis of flow cytometry-isolated sputum macrophages and demonstrate an altered macrophage transcriptome in NA. Transcriptomic alterations in NA macrophages identified by RNASeq were confirmed by qPCR in a larger sample set. Bioinformatic analysis suggested NA macrophage transcriptomic alteration may reflect response to an altered extracellular environment including bacteria, potential contribution to neutrophil recruitment and clearance, and an immunomodulatory role. Finally, we demonstrate that NA macrophage gene signatures are enriched in whole sputum microarray from NA in both mild-moderate and severe asthma and correlate with important clinical characteristics including worse lung function and poorer asthma control, demonstrating a clinically-relevant macrophage-specific element of the previously described whole sputum NA transcriptome.

Bioinformatic analyses of DEGs in sputum macrophages isolated from NA identified enrichment of genes related to processes with pathophysiological relevance in NA. Among these were GO categories relating to regulation of inflammatory responses and responses to bacteria. Airway dysbiosis is evident in NA[6], and our data suggest macrophage transcriptomes and consequently function are altered in response to bacterial queues. Many of the DEGs identified in our analysis are inducible by the bacterial ligand LPS in macrophages or monocytes. These included genes that promote LPS inflammatory responses (e.g. signalling adaptor *TRAF1*[27], transcription factor *POU2F2/*Oct2[28]) but also genes that suppress LPS responses through mechanisms including LPS binding (*CRISPLD2*[29]), suppression of LPS-induced pro-inflammatory gene transcription (*IL10*[30]) and modulation of intracellular signalling (*MARCKS*[31]). This suggests sputum macrophages initiate and regulate inflammatory responses to the altered airway bacterial environment in NA. Also consistent with an immune modulatory/homeostatic role in NA was upregulation of several genes involved in efferocytosis (*MERTK, STAB2, SIGLEC10, SLAMF* members)[32,33] and phagocytosis (*FAIM3, HRH2, STAB2, MARCKS*)[34-36], increasing cellular components that facilitate macrophage uptake of host cells and microbes. This is consistent with an M2-like phenotype where increased uptake of dead/dying cells and bacteria promotes inflammatory

This article is protected by copyright. All rights reserved

resolution. Defective efferocytosis by airway and monocyte-derived macrophages has been demonstrated in non-eosinophilic asthma, suggesting any potential defect in efferocytosis by NA macrophages may be post-transcriptional[37,38]. *CSF3* and *CSF3R*, encoding granulocyte-colony stimulating factor (GCSF) and its receptor, were increased in NA macrophages. CSF3R is expressed on mature neutrophils but also monocytes, and GCSF, a *POU2F2*/Oct2-inducible gene[28], can induce altered differentiation of monocytes towards an immunosuppressive M2-like macrophage phenotype[39]. Aberrant GCSF signalling has recently been linked to NA[40] and severe asthma[41]. *CSF3* and *CSF3R* clustered with several other notable DEGs using the STRING tool. These included receptors for IL1β (*IL1R*) and IL6 (*IL6R*), two cytokines postulated to play important roles in neutrophilic and severe asthma[42,43], as well as the IL23 subunit *IL23A*, which has links to airway obstruction in IL17-mediated neutrophilic airways disease[44]. Thus, macrophages display numerous transcriptomic alterations consistent with both inflammatory promotion and resolution in NA, which may reflect functional plasticity and diversity amongst the population studied.

Whole sputum microarray studies show alteration of the transcriptome in NA[7] or severe asthma with elevated sputum neutrophils[8]. Use of immune cell specific gene signatures implicates altered contribution of specific immune cell gene pools to altered sputum transcriptomes of type 2 high vs type 2 low asthma[45]. A recent study employing GSVA of *in vitro* derived macrophage gene signatures[11] in whole sputum microarray was suggestive of a contribution of altered macrophage transcriptomes to altered sputum transcriptomes in severe and neutrophilic asthma[46]. Here we demonstrate enrichment of a NA macrophage gene signature in whole sputum microarray data from participants with NA. This is the first direct evidence that altered macrophage transcriptomes contribute to the overall gene expression differences observed in whole sputum microarray in NA, and imply altered macrophage function is a significant facet of the airway inflammatory milieu in NA. Analysis of the U-BIOPRED sputum microarray dataset demonstrated similar alterations in mild-moderate vs severe NA, implying the altered macrophage transcriptome is a feature of NA irrespective of asthma severity. Importantly, increased expression of the NA macrophage gene signature correlated with clinical features including worse lung function and poorer asthma control. NA is a clinically important phenotype with limited therapeutic options for those with corticosteroid resistant disease symptoms, and neutrophilia appears unaffected by corticosteroid treatment. However trials of CXCR2 antagonists that reduce neutrophil recruitment to the airways

This article is protected by copyright. All rights reserved

have failed to show clinical benefit in severe asthma[2]. Our present data and those above suggest therapeutic modulation of immune cells such as macrophages could be considered for further investigation in the context of neutrophilic airway inflammation. These data also suggest the term 'neutrophilic' could be considered something of a misnomer[47], and perhaps a more precise description encompassing cellular and molecular alterations beyond the neutrophil may be enabled by this and other studies.

This study has limitations. Our RNASeq analysis should be considered exploratory due to the relatively low sample numbers used, although we confirmed DEGS using qPCR in a larger cohort. Low sample number in the RNASeq cohort prevented meaningful analysis of relationship of identified gene signatures to other clinical characteristics. Our analysis is limited to transcriptomic alterations and further studies should seek to confirm these observations using protein and functional assays, however altered macrophage transcriptomes underpin macrophage functional plasticity in response to differing extracellular queues[11]. Future studies should aim to test whether smoking or prior smoking alters macrophage phenotype in asthma[48]. We performed bulk RNASeq on isolated sputum macrophage populations, and thus our dataset lacks the resolution to identify specific transcriptomic and functional subsets of macrophages that may be present. We cannot exclude that contaminating cells such as neutrophils contributed to our macrophage RNA pool, although all analyses indicated that we achieved highly pure macrophage preparations, DEGs identified did not include lung neutrophil signature genes[26], and the majority of DEGs appeared plausibly derived from macrophages based on literature searches. We were also careful to exclude monocytes from our macrophage preparations in order to be able to specifically evaluate the macrophage niche independently of monocyte influx[49].

In sum, this work reveals inflammatory context-specific modulation of airway macrophage phenotype in NA. Bioinformatic analyses indicated that these transcriptomic alterations relate to both pathophysiologically-relevant inflammation promoting and resolving responses to an altered extracellular milieu, including response to bacteria, promotion of neutrophil recruitment and enhanced clearance of bacteria and host cells. This work demonstrates transcriptomic activation of inflammatory programmes in a cell type other than the neutrophil in NA, supporting the concept of neutrophilic asthma as a discrete asthma endotype.

This article is protected by copyright. All rights reserved

Accepted Article

References

1. Peters MC, Wenzel SE. Intersection of biology and therapeutics: type 2 targeted therapeutics for adult asthma. *Lancet.* 2020;395(10221):371-383.
2. Sze E, Bhalla A, Nair P. Mechanisms and therapeutic strategies for non-T2 asthma. *Allergy.* 2020;75(2):311-325.
3. Fricker M, Heaney LG, Upham JW. Can biomarkers help us hit targets in difficult-to-treat asthma? *Respirology.* 2017;22(3):430-442.
4. Green RH, Brightling CE, Woltmann G, Parker D, Wardlaw AJ, Pavord ID. Analysis of induced sputum in adults with asthma: identification of subgroup with isolated sputum neutrophilia and poor response to inhaled corticosteroids. *Thorax.* 2002;57(10):875-879.
5. Moore WC, Hastie AT, Li X, et al. Sputum neutrophil counts are associated with more severe asthma phenotypes using cluster analysis. *J Allergy Clin Immunol.* 2014;133(6):1557-1563 e1555.
6. Taylor SL, Leong LEX, Choo JM, et al. Inflammatory phenotypes in patients with severe asthma are associated with distinct airway microbiology. *J Allergy Clin Immunol.* 2018;141(1):94-103 e115.
7. Baines KJ, Simpson JL, Wood LG, Scott RJ, Gibson PG. Transcriptional phenotypes of asthma defined by gene expression profiling of induced sputum samples. *J Allergy Clin Immunol.* 2011;127(1):153-160, 160 e151-159.
8. Kuo CS, Pavlidis S, Loza M, et al. T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum transcriptomics in U-BIOPRED. *Eur Respir J.* 2017;49(2).
9. Puttur F, Gregory LG, Lloyd CM. Airway macrophages as the guardians of tissue repair in the lung. *Immunol Cell Biol.* 2019;97(3):246-257.
10. Fricker M, Gibson PG. Macrophage dysfunction in the pathogenesis and treatment of asthma. *Eur Respir J.* 2017;50(3).
11. Xue J, Schmidt SV, Sander J, et al. Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity.* 2014;40(2):274-288.
12. Shaw DE, Sousa AR, Fowler SJ, et al. Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. *Eur Respir J.* 2015;46(5):1308-1321.

This article is protected by copyright. All rights reserved

13. Chung KF, Wenzel SE, Brozek JL, et al. International ERS/ATS guidelines on definition, evaluation and treatment of severe asthma. *Eur Respir J.* 2014;43(2):343-373.
14. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology.* 2006;11(1):54-61.
15. Fricker M, Qin L, Niessen N, et al. Relationship of sputum mast cells with clinical and inflammatory characteristics of asthma. *Clin Exp Allergy.* 2020;50(6):696-707.
16. Lay JC, Peden DB, Alexis NE. Flow cytometry of sputum: assessing inflammation and immune response elements in the bronchial airways. *Inhal Toxicol.* 2011;23(7):392-406.
17. Brooks CR, van Dalen CJ, Hermans IF, Douwes J. Identifying leukocyte populations in fresh and cryopreserved sputum using flow cytometry. *Cytometry B Clin Cytom.* 2013;84(2):104-113.
18. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37(8):907-915.
19. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923-930.
20. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics.* 2016;32(19):3047-3048.
21. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
22. Zhu A, Ibrahim JG, Love MI. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics.* 2019;35(12):2084-2092.
23. Wang G, Baines KJ, Fu JJ, et al. Sputum mast cell subtypes relate to eosinophilia and corticosteroid response in asthma. *Eur Respir J.* 2016;47(4):1123-1133.
24. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013;14:7.
25. Kermani NZ, Song WJ, Badi Y, et al. Sputum ACE2, TMPRSS2 and FURIN gene expression in severe neutrophilic asthma. *Respir Res.* 2021;22(1):10.
26. Vieira Braga FA, Kar G, Berg M, et al. A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat Med.* 2019;25(7):1153-1163.

Accepted Article

This article is protected by copyright. All rights reserved

Accepted Article

27. Oyoshi MK, Barthel R, Tsitsikov EN. TRAF1 regulates recruitment of lymphocytes and, to a lesser extent, neutrophils, myeloid dendritic cells and monocytes to the lung airways following lipopolysaccharide inhalation. *Immunology.* 2007;120(3):303-314.
28. Chou YY, Gao JI, Chang SF, Chang PY, Lu SC. Rapamycin inhibits lipopolysaccharide induction of granulocyte-colony stimulating factor and inducible nitric oxide synthase expression in macrophages by reducing the levels of octamer-binding factor-2. *FEBS J.* 2011;278(1):85-96.
29. Wang ZQ, Xing WM, Fan HH, et al. The novel lipopolysaccharide-binding protein CRISPLD2 is a critical serum protein to regulate endotoxin function. *J Immunol.* 2009;183(10):6646-6656.
30. Murray PJ. The primary mechanism of the IL-10-regulated antiinflammatory response is to selectively inhibit transcription. *Proc Natl Acad Sci U S A.* 2005;102(24):8686-8691.
31. Mancek-Keber M, Bencina M, Japelj B, et al. MARCKS as a negative regulator of lipopolysaccharide signaling. *J Immunol.* 2012;188(8):3893-3902.
32. Poon IK, Lucas CD, Rossi AG, Ravichandran KS. Apoptotic cell clearance: basic biology and therapeutic potential. *Nat Rev Immunol.* 2014;14(3):166-180.
33. Chen J, Zhong MC, Guo H, et al. SLAMF7 is critical for phagocytosis of haematopoietic tumour cells via Mac-1 integrin. *Nature.* 2017;544(7651):493-497.
34. Lang KS, Lang PA, Meryk A, et al. Involvement of Toso in activation of monocytes, macrophages, and granulocytes. *Proc Natl Acad Sci U S A.* 2013;110(7):2593-2598.
35. Fultz R, Engevik MA, Shi Z, et al. Phagocytosis by macrophages depends on histamine H2 receptor signaling and scavenger receptor 1. *Microbiologyopen.* 2019;8(10):e908.
36. Johnston DGW, Kearney J, Zaslona Z, Williams MA, O'Neill LAJ, Corr SC. MicroRNA-21 Limits Uptake of Listeria monocytogenes by Macrophages to Reduce the Intracellular Niche and Control Infection. *Front Cell Infect Microbiol.* 2017;7:201.
37. Simpson JL, Gibson PG, Yang IA, et al. Impaired macrophage phagocytosis in non-eosinophilic asthma. *Clin Exp Allergy.* 2013;43(1):29-35.
38. Erriah M, Pabreja K, Fricker M, et al. Galectin-3 enhances monocyte-derived macrophage efferocytosis of apoptotic granulocytes in asthma. *Respir Res.* 2019;20(1):1.
39. Hollmen M, Karaman S, Schwager S, et al. G-CSF regulates macrophage phenotype and associates with poor overall survival in human triple-negative breast cancer. *Oncoimmunology.* 2016;5(3):e1115177.

This article is protected by copyright. All rights reserved

Accepted Article

40. Kim YM, Kim H, Lee S, et al. Airway G-CSF identifies neutrophilic inflammation and contributes to asthma progression. *Eur Respir J.* 2020;55(2).
41. Wang H, FitzPatrick M, Wilson NJ, et al. CSF3R/CD114 mediates infection-dependent transition to severe asthma. *J Allergy Clin Immunol.* 2019;143(2):785-788 e786.
42. Kim RY, Pinkerton JW, Essilfie AT, et al. Role for NLRP3 Inflammasome-mediated, IL-1beta-Dependent Responses in Severe, Steroid-Resistant Asthma. *Am J Respir Crit Care Med.* 2017;196(3):283-297.
43. Peters MC, McGrath KW, Hawkins GA, et al. Plasma interleukin-6 concentrations, metabolic dysfunction, and asthma severity: a cross-sectional analysis of two cohorts. *Lancet Respir Med.* 2016;4(7):574-584.
44. Moermans C, Damas K, Guiot J, et al. Sputum IL-25, IL-33 and TSLP, IL-23 and IL-36 in airway obstructive diseases. Reduced levels of IL-36 in eosinophilic phenotype. *Cytokine.* 2021;140:155421.
45. Peters MC, Ringel L, Dyjack N, et al. A Transcriptomic Method to Determine Airway Immune Dysfunction in T2-High and T2-Low Asthma. *Am J Respir Crit Care Med.* 2019;199(4):465-477.
46. Tiotiu A, Zounemat Kermani N, Badi Y, et al. Sputum macrophage diversity and activation in asthma: Role of severity and inflammatory phenotype. *Allergy.* 2020.
47. Nair P, Surette MG, Virchow JC. Neutrophilic asthma: misconception or misnomer? *Lancet Respir Med.* 2021;9(5):441-443.
48. Takahashi K, Pavlidis S, Ng Kee Kwong F, et al. Sputum proteomics and airway cell transcripts of current and ex-smokers with severe asthma in U-BIOPRED: an exploratory analysis. *Eur Respir J.* 2018;51(5).
49. Niessen NM, Baines KJ, Simpson JL, et al. Neutrophilic asthma features increased airway classical monocytes. *Clin Exp Allergy.* 2020.

This article is protected by copyright. All rights reserved

Tables

Table 1. Clinical characteristics and sputum differential cell count of sputum macrophage RNASeq cohort

	All (n=20)	NNA (n=13)	NA (n=7)
Female†	14 (70)	9 (69)	5 (71)
Age‡	59.1 (13.1)	58.3 (13.1)	60.5 (13.9)
BMI#	28.8 (24.7, 39.2)	32.6 (26.7, 39.5)	26 (23.8, 28.9)
Pre-β2 FEV1 (% pred)‡	69.5 (14.7)	69.1 (14.5)	70.2 (16.1)
Pre-β2 FEV1/FVC‡	64.2 (11.2)	65.6 (9.6)	61.5 (14.1)
Ex-smoker†	5 (25)	4 (30.8)	1 (14.3)
ACQ6§	0.8 (0.5, 1.7)	0.7 (0.5, 2.1)	1 (0.2, 1.2)
ICS treated†	18 (90)	12 (92)	6 (86)
LABA treated†	17 (85)	11 (85)	6 (86)
ICS dose (fluticasone equivalent)§	500 (500, 937.5)	500 (375, 1000)	500 (500, 500)
OCS treated†	0 (0)	-	-
LAMA treated†	7 (35)	4 (31)	3 (43)
Severe asthma†	4 (20)	2 (15)	2 (29)
Sputum TCC (x10^6/mL)§	6.4 (4.1, 13.2)	4.7 (2.8, 8,5)	15 (7.1, 23.9)*
Sputum macrophage (%)§	44.1 (26.1, 63)	58.1 (44.1, 70.6)	18.3 (10.8, 28)***
Sputum neutrophil (%)§	45.9 (29.6, 67.8)	37.5 (20.4, 45.9)	74.7 (64, 83.8)***
Sputum eosinophil (%)§	1.4 (0.6, 3.6)	2 (0.7, 4.0)	1 (0.3, 2)
Sputum lymphocyte (%)§	1.5 (0.8, 2.4)	2 (1.3, 2.5)	0.5 (0, 1.8)*
Sputum macrophage (x10^4/mL)§	229 (172, 425)	279 (186, 411)	228 (76, 456)
Sputum neutrophil (x10^4/mL)§	267 (137, 564)	148 (95, 358)	1323 (559, 1526)***
Sputum eosinophil (x10^4/mL)	9 (3, 23)	9 (4, 18)	17 (2, 31)
Sputum lymphocyte (x10^4/mL)	9 (2, 17)	10 (5, 18)	2 (0, 17)

† = *Number (%),*‡ = *Mean (SD),* § = *Median (IQR).* * = p<0.05 NA vs NNA. *** = p<0.001 NA vs NNA. *BMI = body mass index. FEV1 = forced expiratory volume in 1 second. FVC = forced vital*

This article is protected by copyright. All rights reserved

capacity. ACQ = asthma control questionnaire. ICS = inhaled corticosteroid. TCC = total cell count.

Accepted Article

This article is protected by copyright. All rights reserved

Table 2. PANTHER gene ontology analysis of top 200 DEGs from NA vs NNA macrophages

PANTHER GO biological process	Fold Enrichment	raw P value	FDR
positive regulation of activation of Janus kinase activity	62.48	5.55E-05	2.00E-02
↳positive regulation of receptor signalling pathway via JAK-STAT	8.06	1.52E-04	4.13E-02
↳regulation of response to stimulus	1.59	5.39E-05	2.00E-02
↳positive regulation of receptor signalling pathway via STAT	7.81	1.79E-04	4.46E-02
↳positive regulation of protein phosphorylation	2.40	1.52E-04	4.20E-02
↳positive regulation of phosphorylation	2.42	6.59E-05	2.22E-02
↳positive regulation of phosphate metabolic process	2.28	1.29E-04	3.84E-02
↳positive regulation of phosphorus metabolic process	2.28	1.29E-04	3.91E-02
↳regulation of activation of Janus kinase activity	41.65	1.31E-04	3.75E-02
regulation of granulocyte chemotaxis	13.15	1.24E-05	9.44E-03
↳regulation of leukocyte chemotaxis	7.09	2.85E-05	1.49E-02
↳regulation of response to external stimulus	2.38	5.26E-05	1.99E-02
↳regulation of locomotion	2.79	2.89E-06	5.49E-03
↳regulation of leukocyte migration	5.11	4.12E-05	1.84E-02
↳regulation of immune system process	2.11	4.75E-05	1.95E-02
↳regulation of cell migration	2.74	1.52E-05	1.05E-02
↳regulation of cell motility	2.80	4.37E-06	4.73E-03
↳regulation of cellular component movement	2.66	6.53E-06	5.83E-03
↳regulation of localization	1.91	4.33E-06	5.05E-03
regulation of neutrophil migration	12.62	8.12E-05	2.62E-02
positive regulation of leukocyte chemotaxis	8.33	3.51E-05	1.67E-02
↳positive regulation of leukocyte migration	6.17	7.16E-05	2.36E-02
leukocyte chemotaxis	6.82	1.20E-05	9.60E-03

This article is protected by copyright. All rights reserved

↳cell chemotaxis	4.78	1.63E-04	4.27E-02
↳chemotaxis	3.31	4.00E-05	1.84E-02
↳response to stimulus	1.38	3.50E-05	1.72E-02
↳taxis	3.29	4.23E-05	1.84E-02
↳locomotion	2.56	3.92E-06	4.96E-03
↳response to external stimulus	2.19	1.39E-07	7.03E-04
↳cell migration	2.85	8.40E-06	7.09E-03
↳cell motility	2.61	2.17E-05	1.22E-02
↳movement of cell or subcellular component	2.38	3.78E-06	5.74E-03
↳localization of cell	2.61	2.17E-05	1.27E-02
↳localization	1.48	4.33E-05	1.83E-02
↳leukocyte migration	4.61	3.68E-06	6.20E-03
↳immune system process	1.79	1.18E-04	3.64E-02
regulation of inflammatory response	4.04	8.86E-07	3.36E-03
↳regulation of defense response	2.58	1.30E-04	3.79E-02
response to bacterium	3.46	3.88E-06	5.36E-03
↳response to other organism	2.50	6.05E-06	6.13E-03
↳response to external biotic stimulus	2.50	6.05E-06	5.74E-03
↳response to biotic stimulus	2.43	1.33E-05	9.65E-03
lymphocyte activation	3.34	1.95E-04	4.78E-02
↳cell activation	2.37	5.66E-05	2.00E-02
regulation of system process	3.06	1.59E-04	4.24E-02
↳regulation of multicellular organismal process	1.68	1.72E-04	4.35E-02
cation transport	2.72	6.23E-05	2.15E-02
cell adhesion	2.69	4.84E-05	1.93E-02
↳biological adhesion	2.67	5.16E-05	2.01E-02
defense response	2.57	2.37E-06	5.15E-03
immune response	2.18	1.59E-05	1.05E-02

Accepted Article

This article is protected by copyright. All rights reserved

regulation of biological quality	1.58	1.64E-04	4.22E-02
multicellular organismal process	1.43	1.47E-04	4.12E-02
RNA processing	< 0.01	8.28E-05	2.62E-02
↳RNA metabolic process	.17	2.28E-05	1.23E-02
↳nucleic acid metabolic process	.20	1.44E-06	3.64E-03
↳nucleobase-containing compound metabolic process	.20	4.27E-08	6.48E-04
↳cellular nitrogen compound metabolic process	.28	7.95E-08	6.04E-04
↳organic cyclic compound metabolic process	.38	1.84E-05	1.17E-02
↳heterocycle metabolic process	.29	1.28E-06	3.88E-03
↳cellular aromatic compound metabolic process	.38	2.87E-05	1.45E-02
↳gene expression	.23	2.11E-05	1.28E-02

FDR = False Discovery Rate (Benjamini-Hochberg). GO = gene ontology.

Accepted Article

This article is protected by copyright. All rights reserved

Table 3. Correlation of NA sputum macrophage gene signature (gene set 1) GSVA enrichment scores with clinical characteristics

	Spearman r value
Age	0.364**
BMI	0.003
ACQ6	0.284*
Pre-β2 FEV1 (% pred)‡	-0.539***
Pre-β2 FEV1/FVC‡	-0.471***
ICS dose (fluticasone equivalent)	0.254*

** $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$*

Accepted Article

This article is protected by copyright. All rights reserved

Figure legends

Figure 1. Flow cytometry-mediated sorting and isolation of highly pure sputum macrophage preparations. (**A**) Flow cytometry sorting strategy for isolation of sputum macrophages. Macrophages were gated as CD45+, HLA-DR+, CD14+, CD16+, CD206+ with a high side scatter (SSC) profile. (**B**) Representative May Grünwald Giemsa-stained cytospin images of sputum cell suspensions pre flow cytometry sort (*left hand panel*), and post flow cytometry sort for macrophages (*right hand panel*). (**C**) Differential cell count of cytospins of pre- and post- flow cytometry sort sputum samples (n=5). Mann-Whitney test. * = $p<0.05$, ** = $p<0.01$.

Figure 2. qPCR validation of neutrophilic asthma sputum macrophage DEGs. RNA was isolated from flow cytometry-sorted sputum macrophages and cDNA prepared (n=36). qPCR was run measuring relative abundance of *MERTK* (**A**)*, SLAMF7* (**B**)*, DYSF* (**C**)*, GPR183* (**D**)*, CSF3* (**E**)*, PI3* (**F**) and *CCR7* (**G**) relative to the housekeeper gene *ACTB*. Relative mRNA abundance of target mRNA to the housekeeper *ACTB* was calculated ($2^{-\Delta Ct}$) and compared between NA (n=13) and NNA (n=23), Mann-Whitney test. * = $p<0.05$, *** = $p<0.001$. Data is plotted as median ± interquartile range.

Figure 3. STRING database protein interaction network of top 200 DEGs identified in NA vs NNA sputum macrophages. Of the top 200 DEGs, STRING identified 167 encoded protein-coding genes. Modules with 2 or more interacting genes displayed (labelled **A-C**).

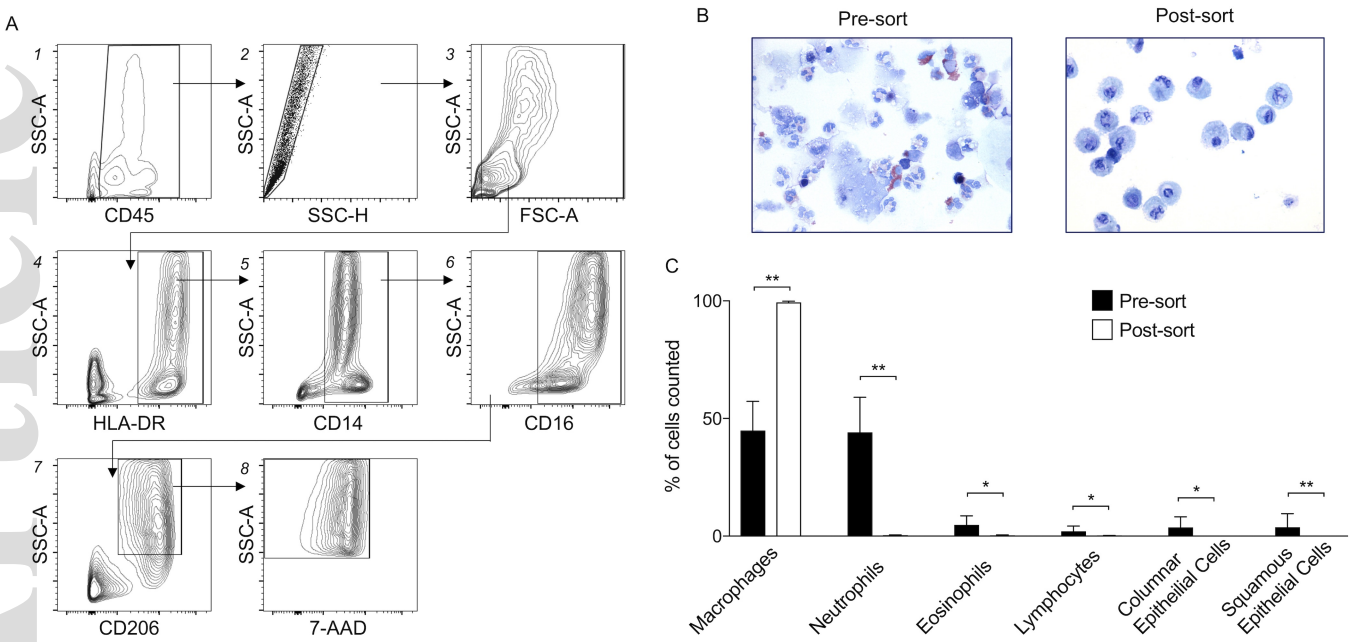
Figure 4. Gene set variation analysis demonstrates enrichment of NA sputum macrophage gene signatures in whole sputum microarray. Gene sets tested using GSVA were the 150 genes present in both the top 200 DEGs of NA vs NNA macrophages and the whole sputum microarray dataset (gene set 1, **A**), and those genes within gene set 1 that had an s value below 0.05 in NA vs NNA RNASeq differential expression comparison (gene set 2, **B**). Gene set lists are provided in the online supplement. Gene set enrichment scores were calculated across 82 whole sputum microarray samples (non-asthma controls n=16, NA n=29, NNA n=37; **A, B**). Gene set 1 and 2 were also tested by GSVA in the U-BIOPRED sputum microarray dataset (non-asthma controls n=16, NA n=47, NNA n=57; **C, D**), and comparing mild-moderate vs severe asthma by neutrophilic phenotype (mild-moderate NNA n=12, mild-moderate NA n=8, severe NNA n=45,

This article is protected by copyright. All rights reserved

severe NA n=39; **E, F**). Kruskal-Wallis test with Dunn's multiple comparisons test. * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$.

Accepted Article

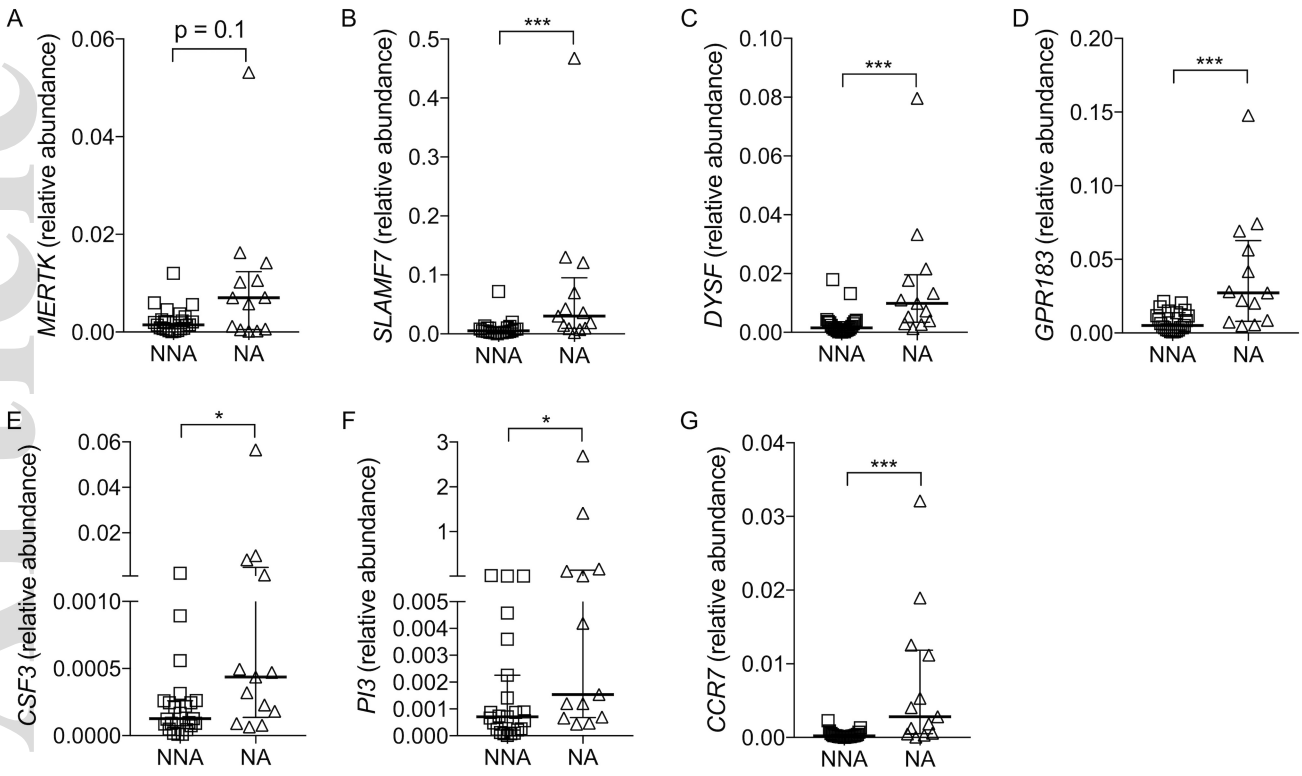
This article is protected by copyright. All rights reserved



all_15087_f1.jpg

Accepted Article

This article is protected by copyright. All rights reserved

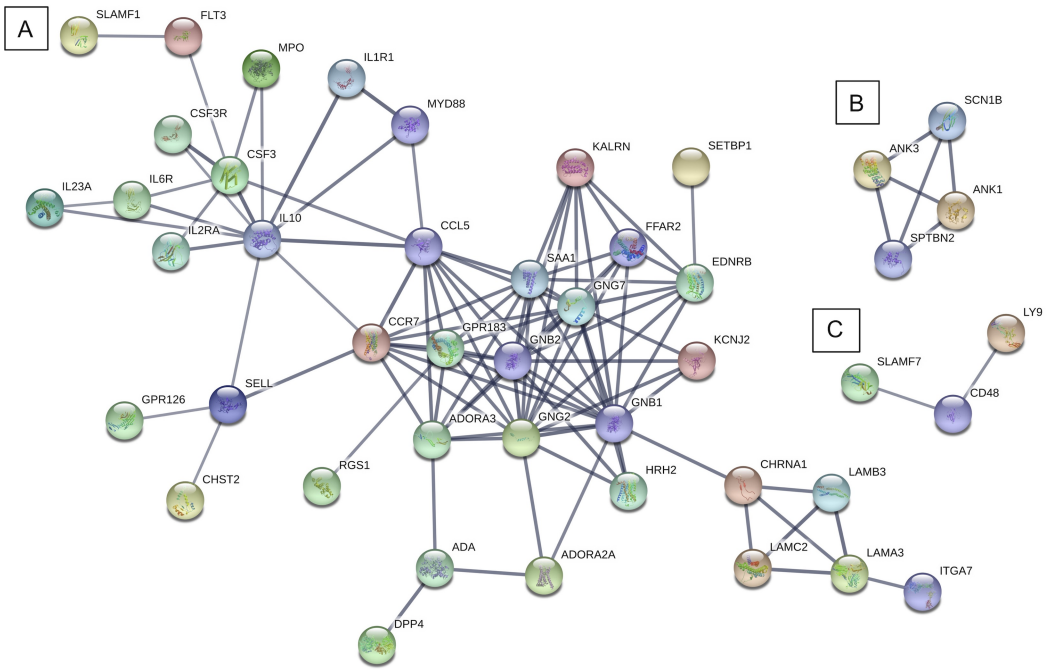


all_15087_f2.jpg

Accepted Article

This article is protected by copyright. All rights reserved

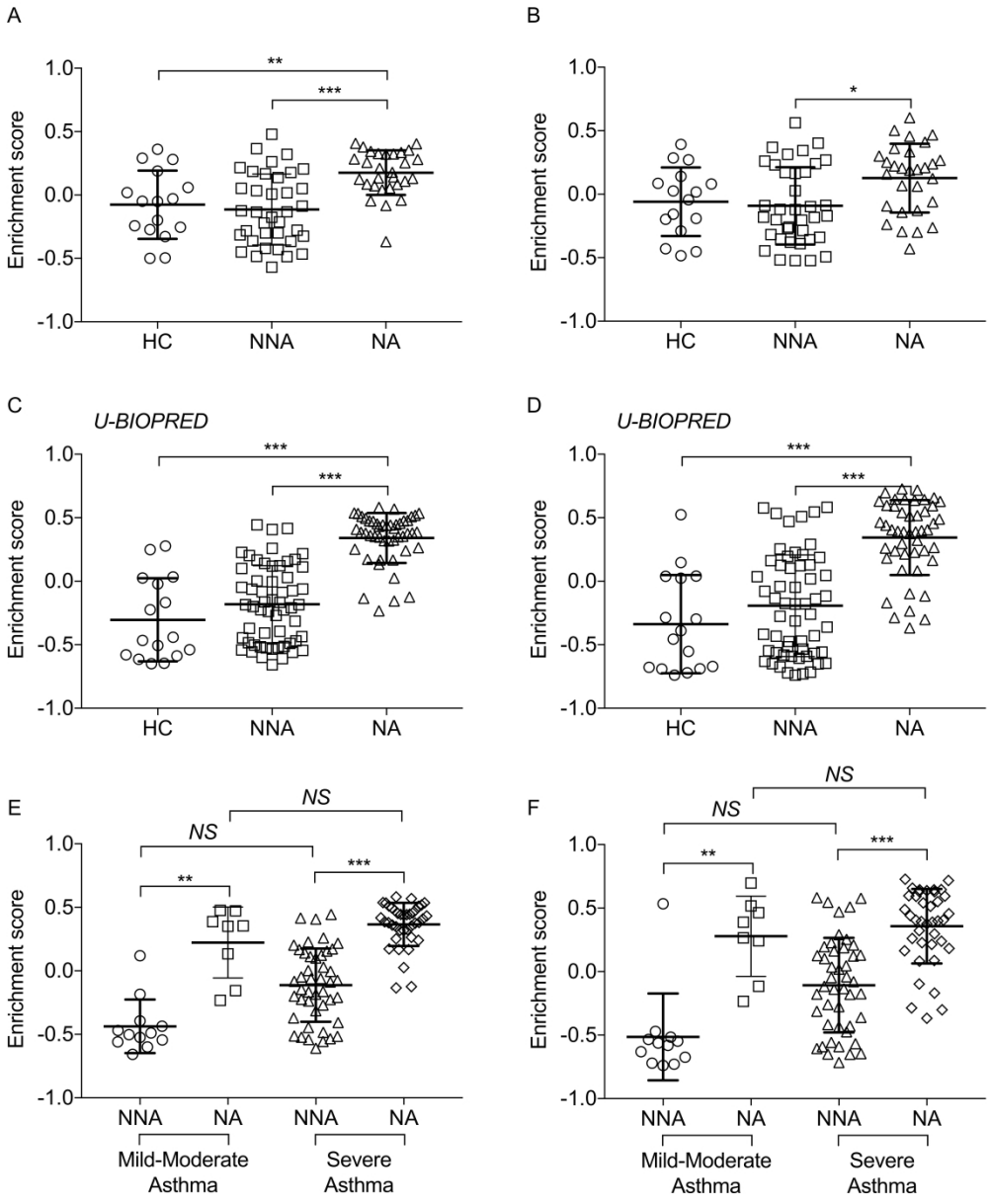
Accepted Article



all_15087_f3.jpg

This article is protected by copyright. All rights reserved

Accepted Article



all_15087_f4.jpg

This article is protected by copyright. All rights reserved

An altered sputum macrophage transcriptome contributes to the neutrophilic asthma endotype

Michael Fricker, Ling Qin, Stephany Sánchez-Ovando, Jodie L Simpson, Katherine J Baines, Carlos Riveros, Hayley A Scott, Lisa G Wood, Peter AB Wark, Nazanin Z Kermani, Ian M Adcock, Kian Fan Chung and Peter G Gibson

6 May 2021

Abstract

Bioinformatics data analysis pipeline

Contents

1 RNA Seq data pipeline description 2
1.1 Pipeline steps program version and processing parameters 2
1.2 Initial raw reads QC report (step 2) 2
1.3 Trimming QC report (step 3) 7
1.4 Alignment report (step 5) 12
1.5 Feature counting report (step 6) 13
2 Differential expression analysis 15
2.1 `DESeq` analysis workflow 15
3 R Session Info 16
3.1 `DESeq2` analysis 16
References 18

1 RNA Seq data pipeline description

Data analysis starts from demultiplexed 100bp paired end reads RNA Seq fastq files for 20 samples.

The data analysis workflow consists of:

1. fastqc quality control analysis of raw reads
2. multiqc aggregate raw quality control of individual `fastqc` analysis
3. TrimGalore adapter and quality trimming in paired-end mode, with trimmed `fastqc` analysis
4. multiqc aggregate trimmed data quality control
5. hisat2 paired-end alignment to GRCh38 human genome
6. featureCounts read counting to transcriptome
7. multiqc aggregate alignment and read counting quality control
8. Differential expression, annotation, and gene set variance downstream analysis in R / rstudio.

1.1 Pipeline steps program version and processing parameters

Program	Version	Notes	Reference
`fastqc`	0.11.7	K-mer counting	Andrews (2010)
`multiqc`	1.7		Ewels et al. (2016)
`TrimGalore`	0.6.3	Paired-end mode, `--nextseq 20`	Krueger (2015)
`cutadapt`	1.18	Implicitly used by `TrimGalore`	
`hisat2`	2.1.0	paired-end mode, `grch38` Homo Sapiens genome plus decoys	Kim et al. (2019)
`samtools`	1.9	Sorting of BAM files	
`subRead` (`featureCounts`)	1.6	Paired end, no multimapping, no overlapping, primary only. Ensembl GRCh38 v97 GTF transcriptome annotation.	Liao, Smyth, and Shi (2014)

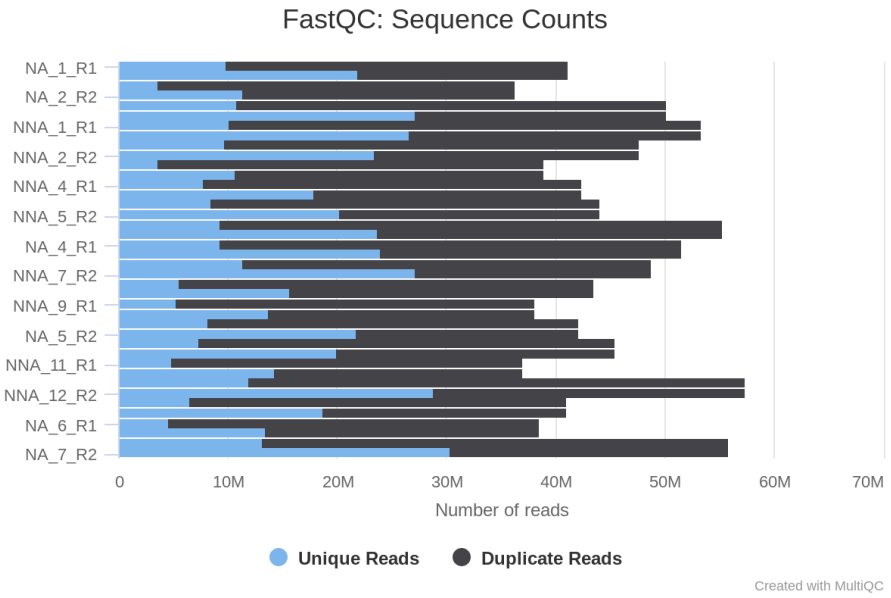
1.2 Initial raw reads QC report (step 2)

Sample Name	% Dups	% GC	K Seqs
NA_1_R1	76.0%	52%	41113.0
NA_1_R2	46.6%	53%	41113.0
NA_2_R1	89.9%	55%	36347.0
NA_2_R2	68.6%	56%	36347.0
NA_3_R1	78.5%	51%	50209.7
NA_3_R2	46.0%	52%	50209.7
NNA_1_R1	81.2%	52%	53417.6
NNA_1_R2	50.2%	53%	53417.6
NNA_2_R1	79.6%	53%	47679.0
NNA_2_R2	50.9%	54%	47679.0

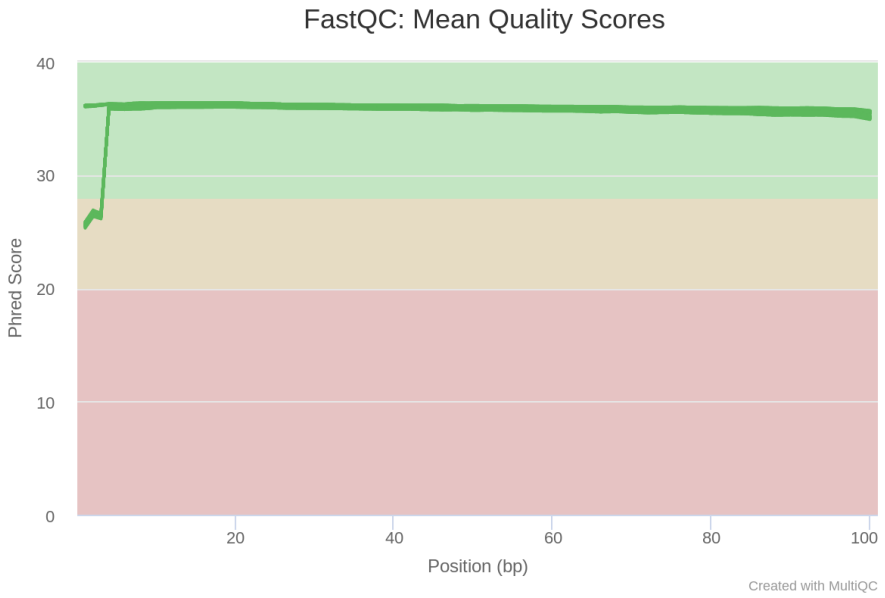
Sample Name	% Dups	% GC	K Seqs
NNA_3_R1	90.9%	52%	39018.5
NNA_3_R2	72.7%	53%	39018.5
NNA_4_R1	81.7%	54%	42441.3
NNA_4_R2	57.8%	56%	42441.3
NNA_5_R1	80.8%	53%	44016.9
NNA_5_R2	54.1%	54%	44016.9
NNA_6_R1	83.1%	52%	55296.6
NNA_6_R2	57.2%	53%	55296.6
NA_4_R1	81.9%	53%	51568.9
NA_4_R2	53.4%	54%	51568.9
NNA_7_R1	76.6%	51%	48829.9
NNA_7_R2	44.3%	52%	48829.9
NNA_8_R1	87.1%	54%	43461.9
NNA_8_R2	63.8%	55%	43461.9
NNA_9_R1	86.3%	59%	38187.4
NNA_9_R2	64.1%	60%	38187.4
NA_5_R1	80.6%	49%	42080.7
NA_5_R2	48.3%	49%	42080.7
NNA_10_R1	83.8%	53%	45499.1
NNA_10_R2	56.1%	54%	45499.1
NNA_11_R1	86.9%	52%	37069.3
NNA_11_R2	61.3%	53%	37069.3
NNA_12_R1	79.3%	52%	57412.9
NNA_12_R2	49.8%	54%	57412.9
NNA_13_R1	84.0%	52%	41045.2
NNA_13_R2	54.6%	53%	41045.2
NA_6_R1	88.3%	53%	38525.0
NA_6_R2	65.2%	54%	38525.0
NA_7_R1	76.4%	52%	55919.5
NA_7_R2	45.6%	54%	55919.5

An altered sputum macrophage transcriptome contributes to the neutrophilic asthma endotype

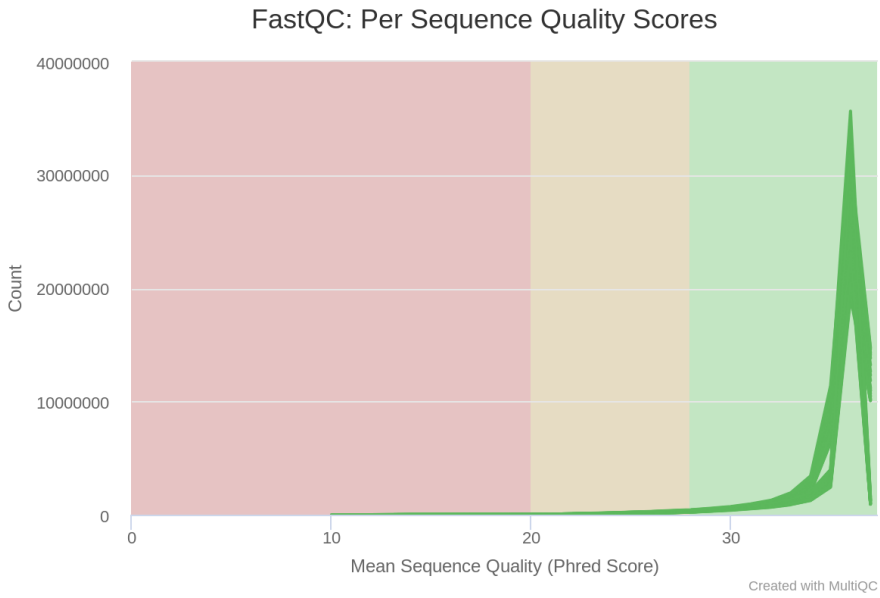
Sequence counts



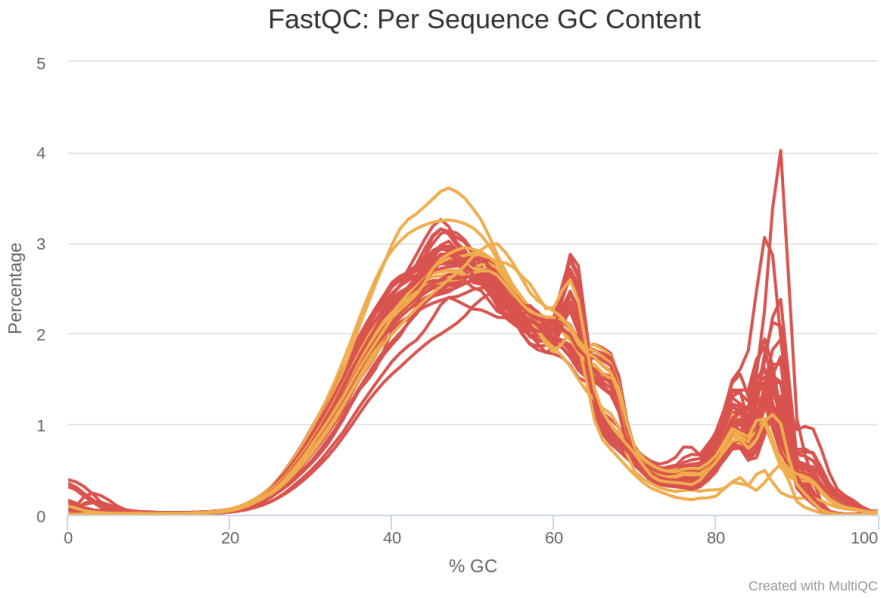
Sequence quality histograms



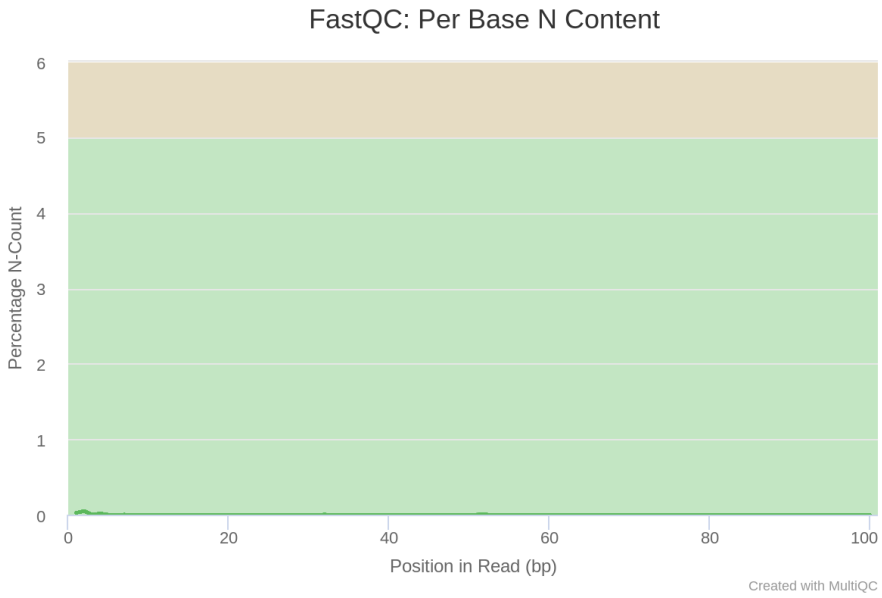
Per sequence quality scores



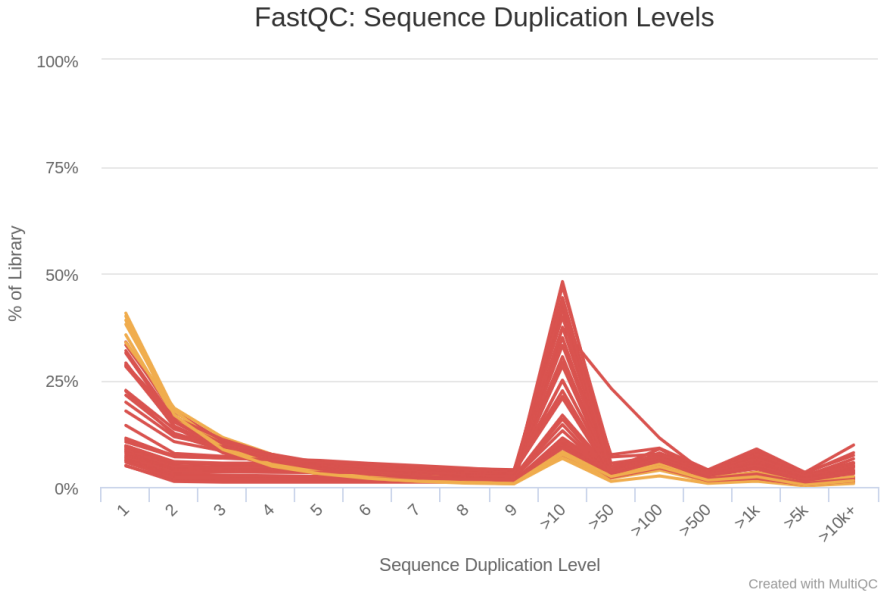
Per sequence GC content



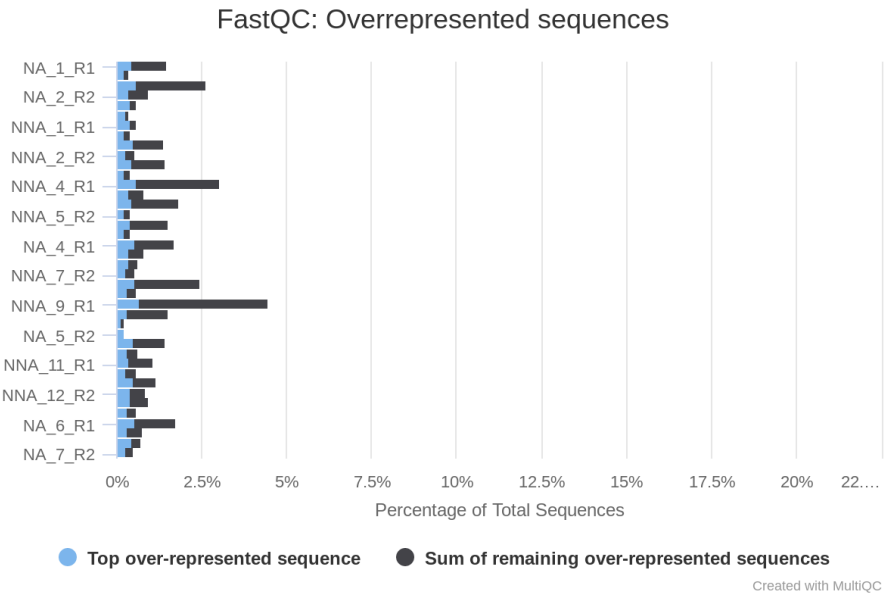
Per base N content



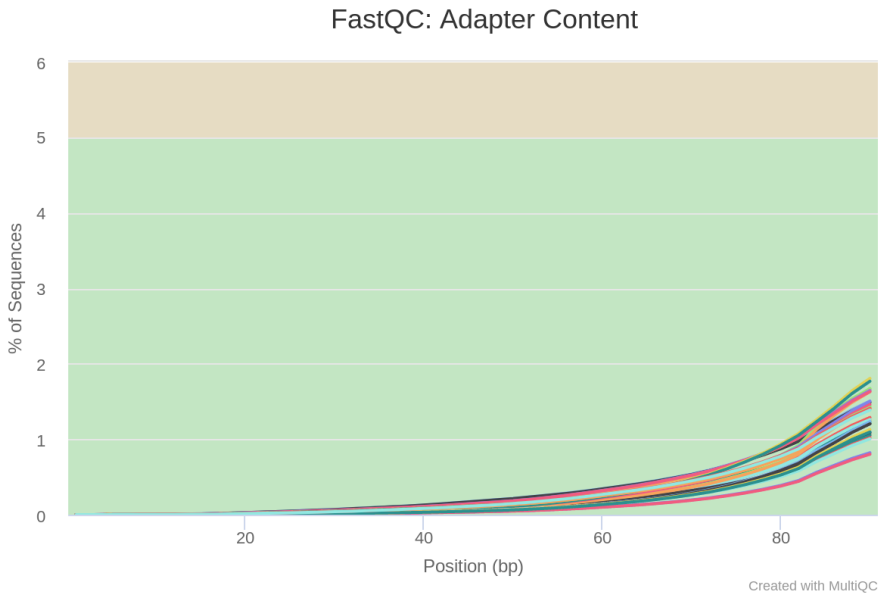
Sequence duplication levels



Overrepresented sequences



Adapter content

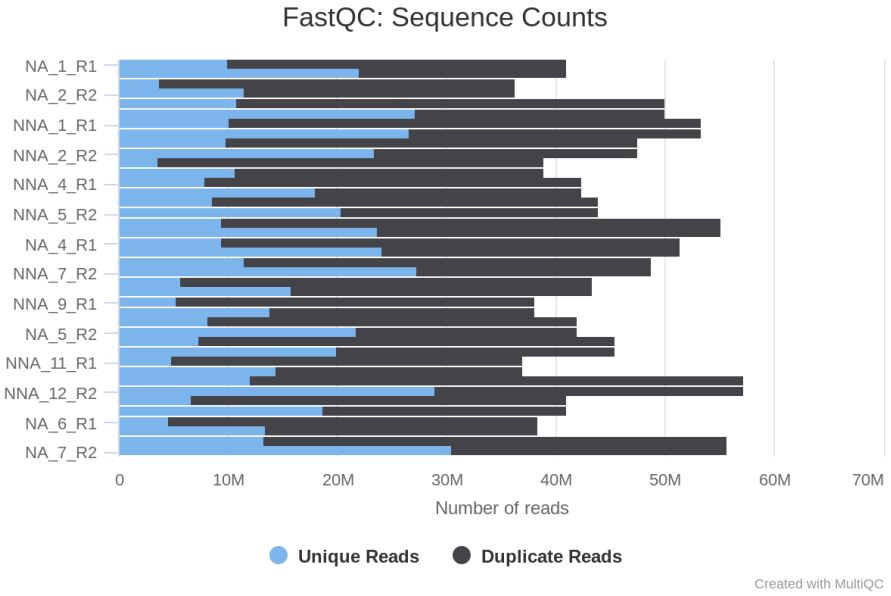


1.3 Trimming QC report (step 3)

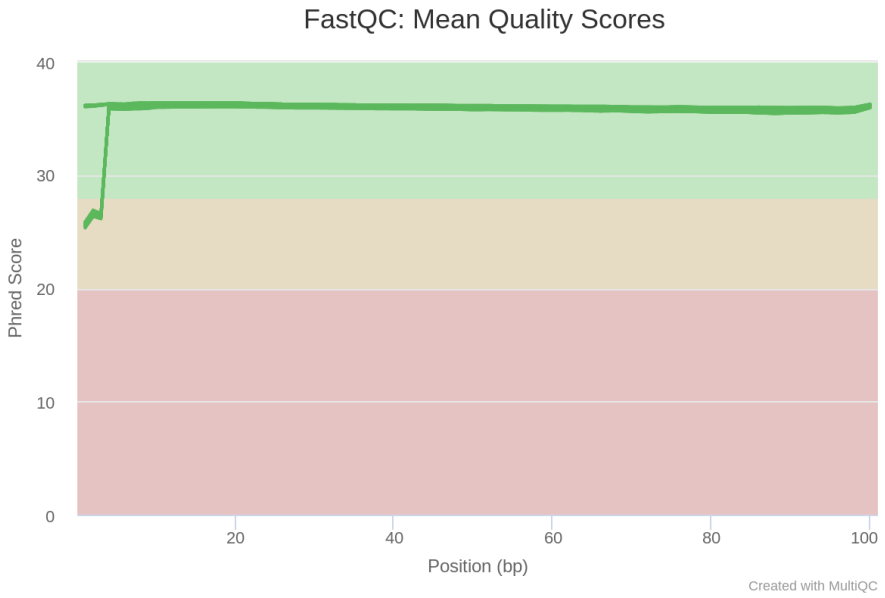
The analysis of trimmed reads reveals a high level of duplication and overrepresented sequences from a few Kmer patterns.

An altered sputum macrophage transcriptome contributes to the neutrophilic asthma endotype

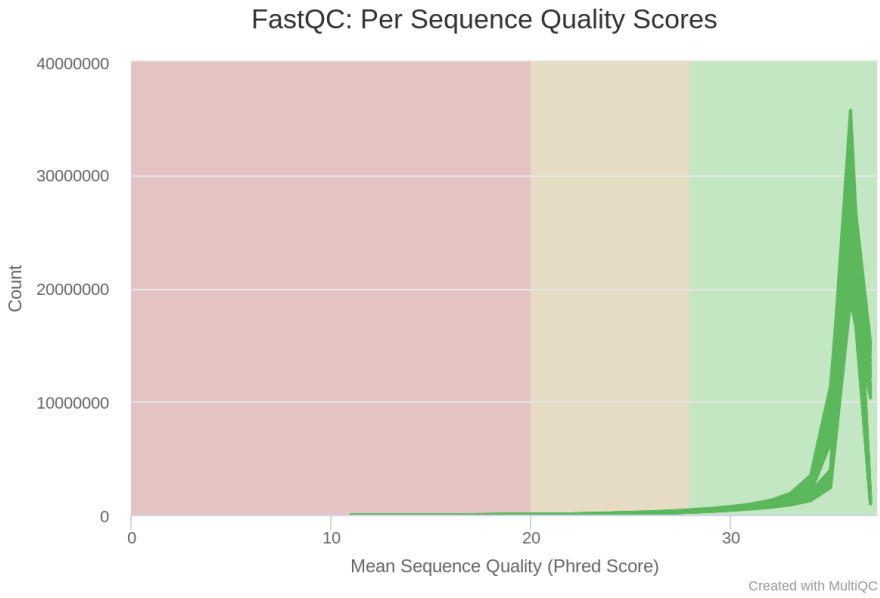
Sequence counts



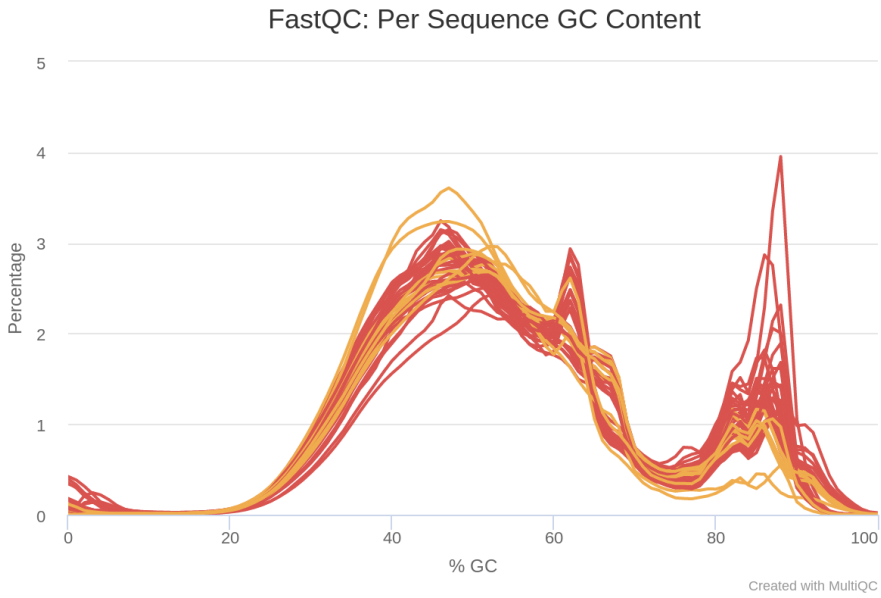
Sequence quality histograms



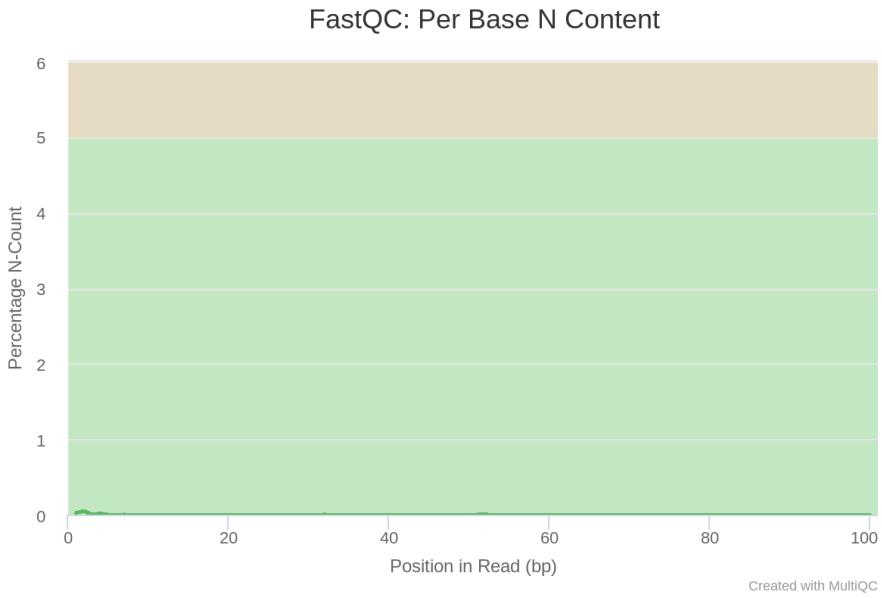
Per sequence quality scores



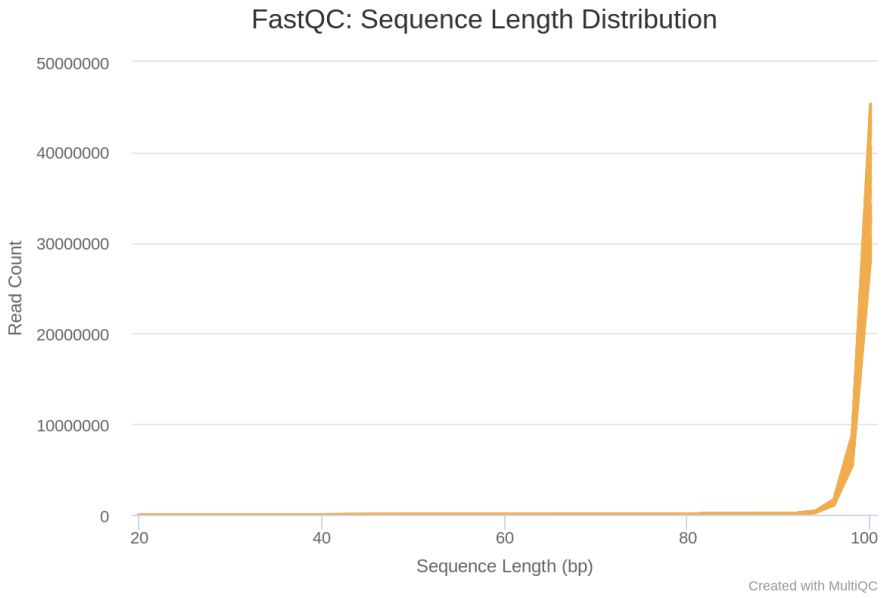
Per sequence GC content



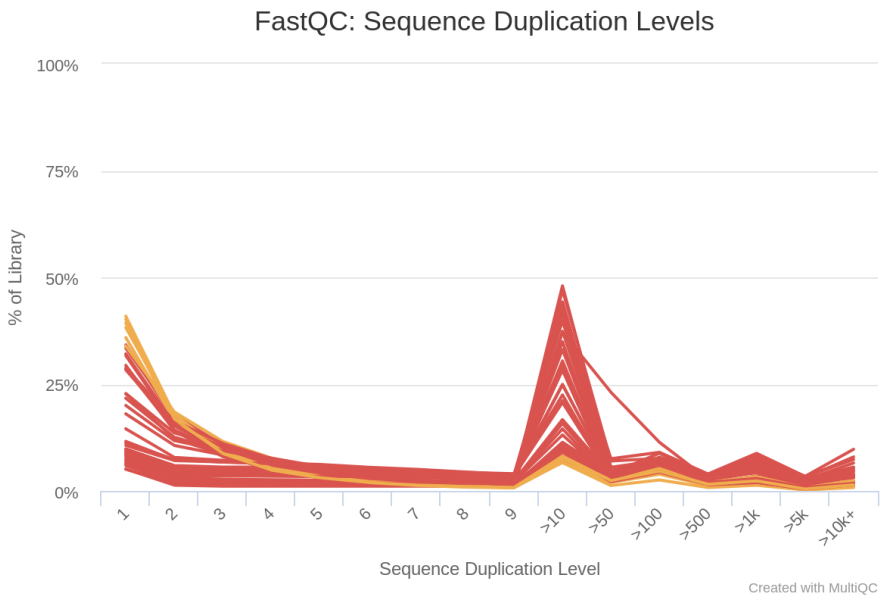
Per base N content



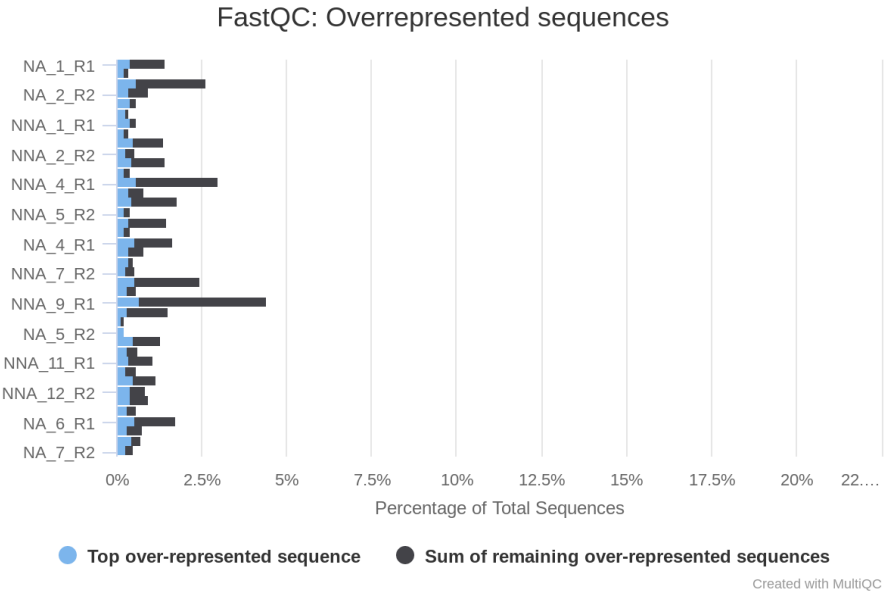
Sequence length distribution



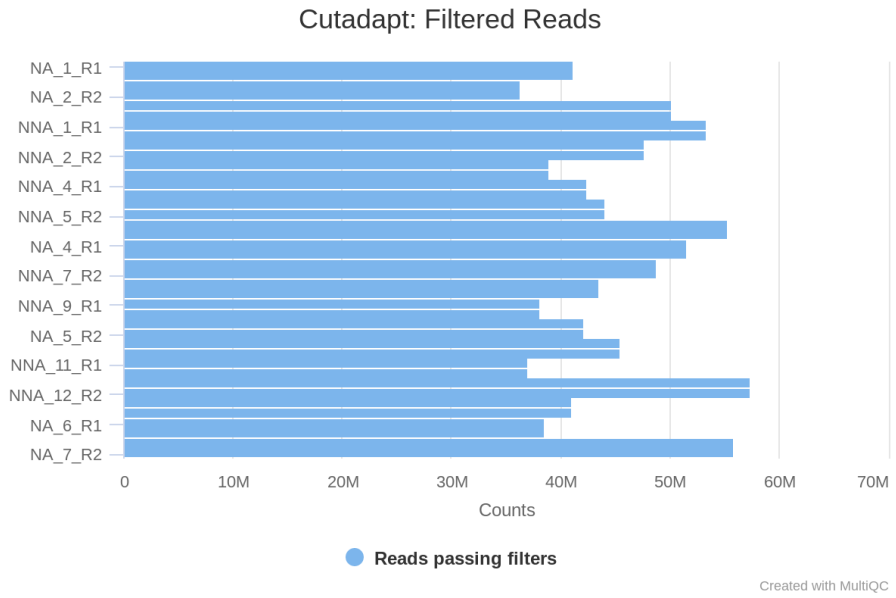
Sequence duplication levels



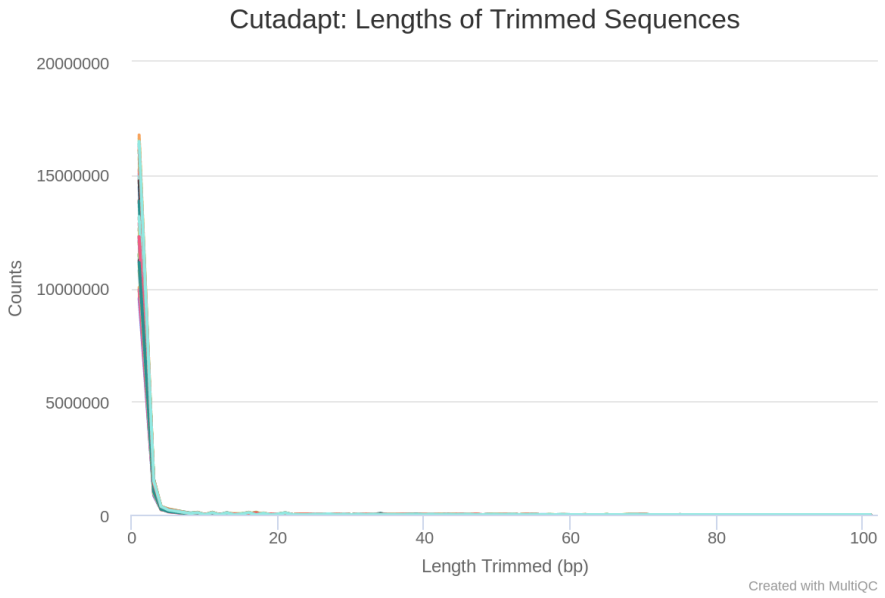
Overrepresented sequences



Filtered reads



Trimmed sequence lengths

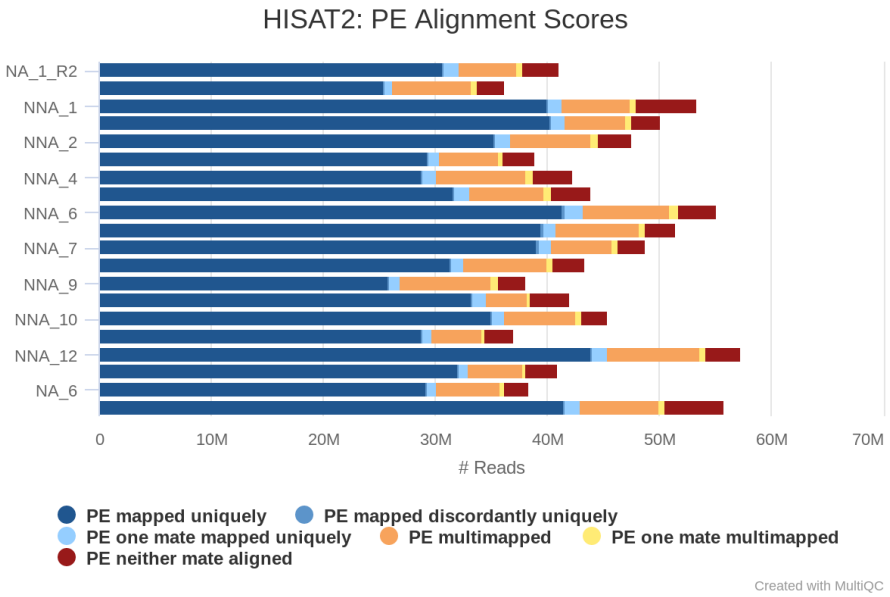


1.4 Alignment report (step 5)

Sample Name	% Aligned
NA_1	2.3%
NA_2	3.0%
NNA_1	0.0%

Sample Name	% Aligned
NA_3	4.9%
NNA_2	3.7%
NNA_3	2.6%
NNA_4	1.6%
NNA_5	1.9%
NNA_6	3.8%
NA_4	4.9%
NNA_7	5.0%
NNA_8	3.6%
NNA_9	3.6%
NA_5	1.8%
NNA_10	5.0%
NNA_11	3.2%
NNA_12	4.7%
NNA_13	3.2%
NA_6	4.3%
NA_7	0.6%

Number of reads aligned



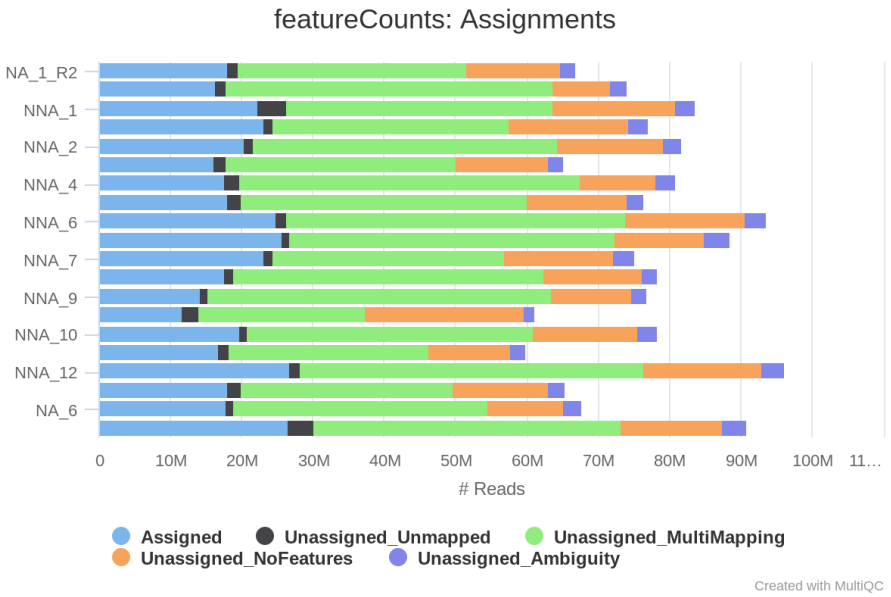
Mapping shows a relatively high level of uniquely mapped pairs.

1.5 Feature counting report (step 6)

Sample Name	% Assigned	K Assigned
NA_1	26.9%	17987.5
NA_2	22.1%	16369.9
NNA_1	26.7%	22330.5

Sample Name	% Assigned	K Assigned
NA_3	30.1%	23185.5
NNA_2	24.8%	20292.5
NNA_3	24.8%	16153.4
NNA_4	21.8%	17611.7
NNA_5	23.5%	17958.0
NNA_6	26.6%	24914.9
NA_4	29.0%	25685.3
NNA_7	30.9%	23239.8
NNA_8	22.4%	17532.2
NNA_9	18.5%	14193.7
NA_5	19.2%	11731.5
NNA_10	25.2%	19751.1
NNA_11	28.1%	16804.6
NNA_12	27.9%	26775.4
NNA_13	27.7%	18138.8
NA_6	26.2%	17735.6
NA_7	29.3%	26604.4

Number of reads assigned



There is a high proportion of multimapped reads in the feature counting which can not be unambiguously assigned to transcripts. This is related to the high number of duplicated and overrepresented sequences detected earlier.

To investigate the possible source of these fragments, all overrepresented sequences were aligned and organised in clusters, after removing duplicated sequences and poly-T.

The largest groups have homology with ribosomal RNA and cytoplasmic RNA, suggesting rRNA depletion had low efficiency during library preparation. Only uniquely mapped and assigned reads were used for the downstream differential expression and enrichment analysis.

2 Differential expression analysis

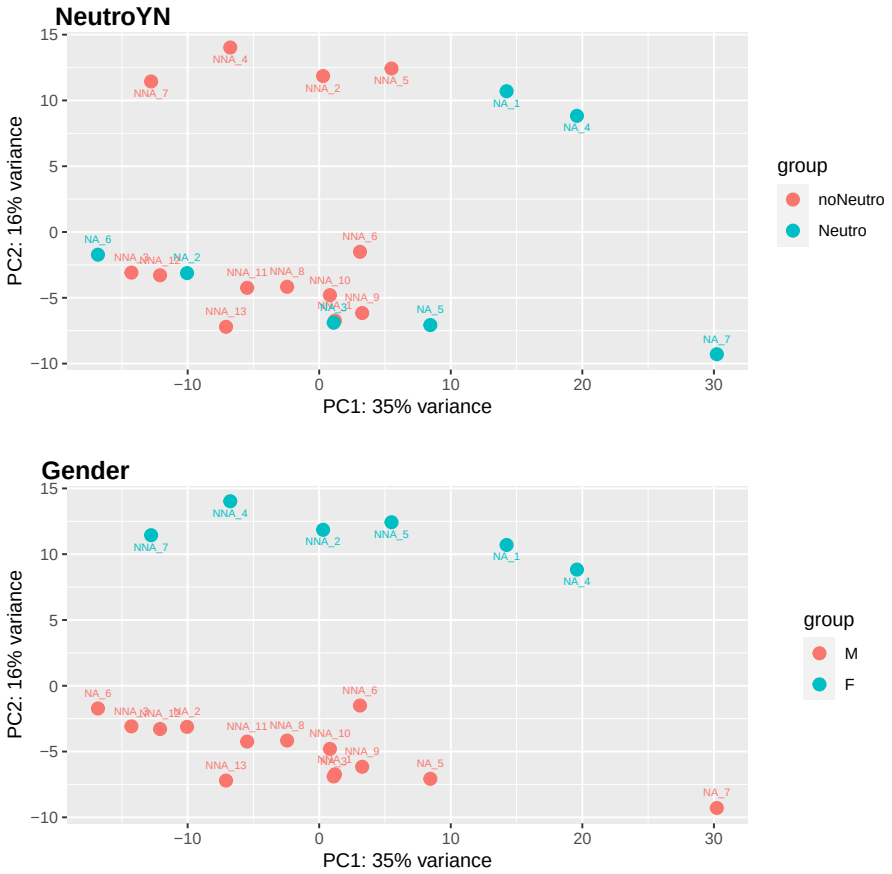
The differential expression analysis of the data was performed with the *R* package DESeq2 (Love, Huber, and Anders (2014)), with the APE Glm log fold change shrinkage estimator method (Zhu, Ibrahim, and Love (2019)).

2.1 `DESeq` analysis workflow

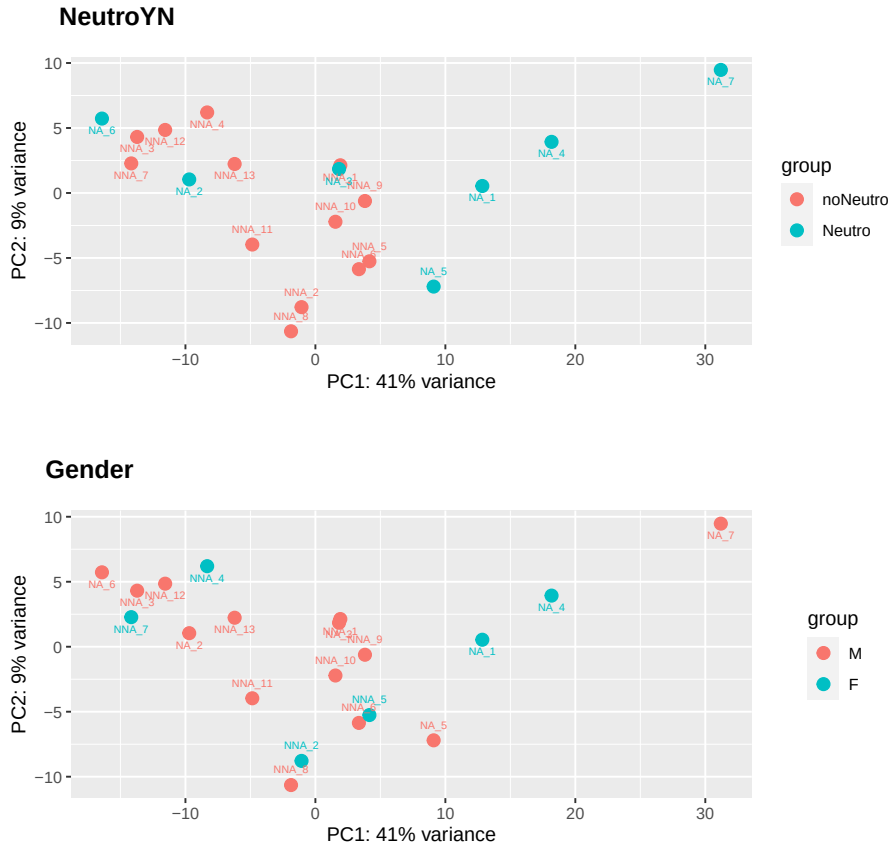
The outline of the differential expression analysis is:

1. Estimation of library size factors and preliminary `rlog` transformation for inspection PCA analysis.
2. PCA analysis and review of phenotype variables.
3. Remove genes in X and Y chromosomes.
4. Re-estimate `rlog` transformation and PCA inspection of reduced data.
5. `DESEq2` analysis on Neutrophilic vs. non-Neutrophilic asthma phenotype:

Initial PCA



PCA after remove X, Y chromosome genes



3 R Session Info

3.1 `DESeq2` analysis

```
R version 3.5.3 (2019-03-11)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Fedora 29 (Workstation Edition)

Matrix products: default
BLAS: /opt/microsoft/ropen/3.5.3/lib64/R/lib/libRblas.so
LAPACK: /opt/microsoft/ropen/3.5.3/lib64/R/lib/libRlapack.so

locale:
 [1] LC_CTYPE=en_AU.UTF-8       LC_NUMERIC=C
 [3] LC_TIME=en_AU.UTF-8        LC_COLLATE=en_AU.UTF-8
 [5] LC_MONETARY=en_AU.UTF-8    LC_MESSAGES=en_AU.UTF-8
 [7] LC_PAPER=en_AU.UTF-8       LC_NAME=C
 [9] LC_ADDRESS=C               LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_AU.UTF-8 LC_IDENTIFICATION=C
```

```
attached base packages:
 [1] grid      parallel  stats4    stats     graphics  grDevices utils
 [8] datasets  methods   base

other attached packages:
 [1] nVennR_0.2.1                rsvg_1.3
 [3] grImport2_0.1-4             vsn_3.50.0
 [5] cowplot_1.0.0               ggplot2_3.2.1
 [7] futile.logger_1.4.3         openxlsx_4.1.0.1
 [9] DESeq2_1.22.2               SummarizedExperiment_1.12.0
[11] DelayedArray_0.8.0          BiocParallel_1.16.6
[13] matrixStats_0.55.0          Biobase_2.42.0
[15] GenomicRanges_1.34.0        GenomeInfoDb_1.18.2
[17] IRanges_2.16.0              S4Vectors_0.20.1
[19] BiocGenerics_0.28.0         data.table_1.12.4
[21] RevoUtils_11.0.3            RevoUtilsMath_11.0.0

loaded via a namespace (and not attached):
 [1] bitops_1.0-6           bit64_0.9-7            RColorBrewer_1.1-2
 [4] httr_1.4.1             numDeriv_2016.8-1.1    tools_3.5.3
 [7] backports_1.1.5        R6_2.4.0               affyio_1.52.0
[10] rpart_4.1-15           Hmisc_4.2-0            DBI_1.0.0
[13] lazyeval_0.2.2         colorspace_1.4-1       apeglm_1.4.2
[16] nnet_7.3-12            withr_2.1.2            tidyselect_0.2.5
[19] gridExtra_2.3          bit_1.1-14             compiler_3.5.3
[22] preprocessCore_1.44.0  formatR_1.7            htmlTable_1.13.2
[25] labeling_0.3           scales_1.0.0           checkmate_1.9.4
[28] hexbin_1.27.3          genefilter_1.64.0      affy_1.60.0
[31] stringr_1.4.0          digest_0.6.21          foreign_0.8-72
[34] rmarkdown_1.16         XVector_0.22.0         jpeg_0.1-8
[37] base64enc_0.1-3        pkgconfig_2.0.3        htmltools_0.4.0
[40] bbmle_1.0.20           limma_3.38.3           htmlwidgets_1.5.1
[43] rlang_0.4.0            rstudioapi_0.10        RSQLite_2.1.2
[46] jsonlite_1.6           acepack_1.4.1          dplyr_0.8.3
[49] zip_2.0.4              RCurl_1.95-4.12        magrittr_1.5
[52] GenomeInfoDbData_1.2.0 Formula_1.2-3          Matrix_1.2-17
[55] Rcpp_1.0.2             munsell_0.5.0          yaml_2.2.0
[58] stringi_1.4.3          MASS_7.3-51.4          zlibbioc_1.28.0
[61] plyr_1.8.4             blob_1.2.0             crayon_1.3.4
[64] lattice_0.20-38        splines_3.5.3          annotate_1.60.1
[67] locfit_1.5-9.1         zeallot_0.1.0          knitr_1.25
[70] pillar_1.4.2           geneplotter_1.60.0     futile.options_1.0.1
[73] XML_3.98-1.20          glue_1.3.1             evaluate_0.14
[76] latticeExtra_0.6-28    lambda.r_1.2.4         BiocManager_1.30.7
[79] vctrs_0.2.0            png_0.1-7              gtable_0.3.0
[82] purrr_0.3.2            assertthat_0.2.1       emdbook_1.3.11
[85] xfun_0.10              xtable_1.8-4           coda_0.19-3
[88] viridisLite_0.3.0      survival_2.44-1.1      tibble_2.1.3
[91] AnnotationDbi_1.44.0   memoise_1.1.0          cluster_2.1.0
```

References

Andrews, Simon. 2010. *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Ewels, Philip, Måns Magnusson, Sverker Lundin, and Max Käller. 2016. "MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report." *Bioinformatics* 32 (19): 3047–8. https://doi.org/10.1093/bioinformatics/btw354.

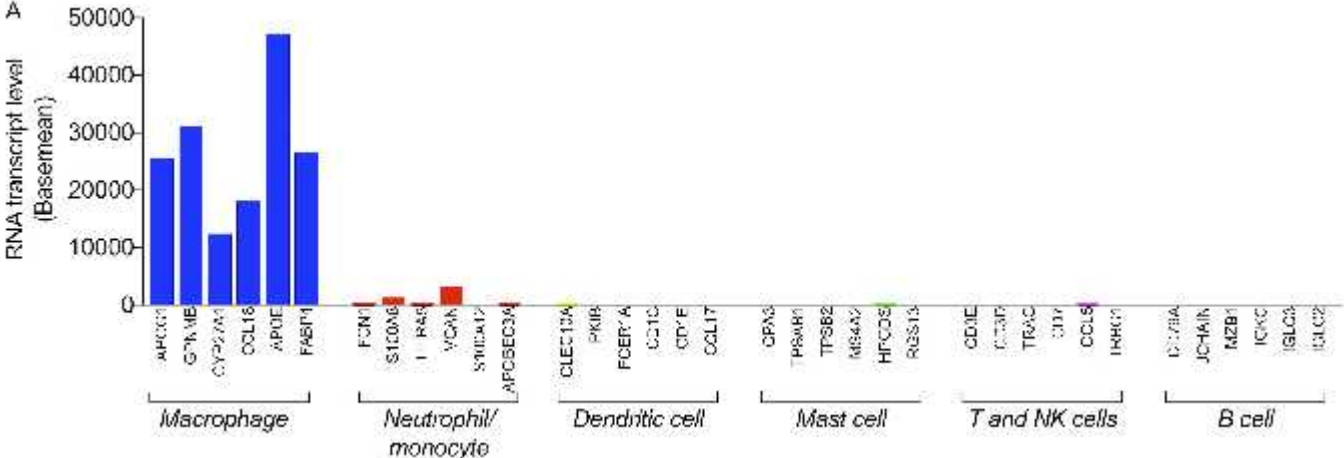
Kim, Daehwan, Joseph M. Paggi, Chanhee Park, Christopher Bennett, and Steven L. Salzberg. 2019. "Graph-Based Genome Alignment and Genotyping with HISAT2 and HISAT-Genotype." *Nature Biotechnology* 37 (8): 907–15. https://doi.org/10.1038/s41587-019-0201-4.

Krueger, Felix. 2015. "Trim Galore." *A Wrapper Tool Around Cutadapt and FastQC to Consistently Apply Quality and Adapter Trimming to FastQ Files*.

Liao, Yang, Gordon K. Smyth, and Wei Shi. 2014. "featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features." *Bioinformatics* 30 (7): 923–30. https://doi.org/10.1093/bioinformatics/btt656.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2." *Genome Biology* 15 (12): 550. https://doi.org/10.1186/s13059-014-0550-8.

Zhu, Anqi, Joseph G. Ibrahim, and Michael I. Love. 2019. "Heavy-Tailed Prior Distributions for Sequence Count Data: Removing the Noise and Preserving Large Differences." *Bioinformatics* 35 (12): 2084–92. https://doi.org/10.1093/bioinformatics/bty895.


A
RNA transcript level (Basemean)
50000
40000
30000
20000
10000
0
APOC1
GPNMB
CYP27A1
CCL18
APOE
FABP4
FCN1
S100A9
LYZ
VCAN
SNHG12
APOBEC3A
CLEC10A
PKIB
FCER1A
CD1C
CD1E
CCL17
CPA3
TPSAB1
TPSB2
MS4A2
HPGDS
RGS13
CD3E
CD3D
TRAC
CD7
CCL5
IKZF3
CD79A
JCHAIN
MZB1
IGKC
IGLC3
IGLC2
Macrophage
Neutrophil/ monocyte
Dendritic cell
Mast cell
T and NK cells
B cell

An altered sputum macrophage transcriptome contributes to the neutrophilic asthma endotype

Authors: Michael Fricker; Ling Qin; Stephany Sánchez-Ovando; Jodie L Simpson; Katherine J Baines; Carlos Riveros; Hayley A Scott; Lisa G Wood; Peter AB Wark; Nazanin Z Kermani; Kian Fan Chung; Peter G Gibson D Med.

SUPPORTING INFORMATION

Methods

Spirometry and sputum induction

Spirometry and sputum induction with hypertonic saline (4.5%) was performed as previously described[1]. Participants were asked to breathe nebulized hypertonic saline for consecutive intervals of 30 seconds, 1 minute, 2 minutes, 4 minutes, 4 minutes and 4 minutes for total maximum time of 15.5 minutes. After each time interval spirometry measurements were taken. Selected sputum portions were disaggregated using dithiothreitol and total cell count performed, with viability assessed through exclusion of trypan blue. Cytospins were prepared, stained with May-Grünwald and Giemsa and 400 non-squamous cells quantified for differential cell count. Samples were categorized into inflammatory subtypes based on differential cell count as previously described[1]: paucigranulocytic (PGA) <3% eosinophils and <61% neutrophils; eosinophilic (EA) ≥3% eosinophils and <61% neutrophils; neutrophilic (NA) <3% eosinophils and ≥61% neutrophils; mixed granulocytic (MGA) ≥3% eosinophils and ≥61% neutrophils. For two-way comparison of eosinophilic vs non-eosinophilic, samples were categorised as eosinophilic based on ≥3% eosinophils alone.

Sputum flow cytometry

A minimum of 2 x 10^5 sputum cells (viability > 40%) was resuspended in flow cytometry buffer (phosphate buffered saline, 1% foetal calf serum, 2 mM Ethylenediaminetetraacetic acid, 0.01% Sodium Azide) at 4 °C. Human Fc block (BD Biosciences) was added following manufacturer guidelines (10 minutes, 4 °C), cells pelleted by centrifugation (400 *g*, 8 minutes, 4 °C), supernatant discarded and cells washed in flow cytometry buffer prior to repeat centrifugation. Cell pellet was resuspended in flow cytometry buffer containing optimized titrated amounts of antibodies directed against CD45 (BD Biosciences, clone HI30), CD14 (BD Biosciences, clone M5E2), CD16 (BD Biosciences, clone 3G8), HLA-DR

(BD Biosciences, clone G46-6) and CD206 (BD Biosciences, clone 19.2) and incubated in the dark for 20 minutes, 4 °C, followed by a further wash in flow cytometry buffer. Cells were resuspended in flow cytometry buffer, viability 7-AAD (BD Biosciences) added, incubated on ice for 10 minutes and then immediately underwent flow cytometry-mediated cell sorting for isolation of macrophages. Macrophages were sorted directly into TRIzol LS which was chilled at 4°C. Sorting typically took between 10-30 minutes. As soon as flow cytometry cell sorting of macrophages in a sputum sample was complete, TRIzol LS containing sorted macrophages was thoroughly mixed by vortexing to ensure complete lysis of cells, and then stored at -80°C prior to subsequent processing steps.

Macrophage RNA isolation and quantification

Up to 1 mL of sample in TRIzol LS was thawed on ice, pulse vortexed then centrifuged (12,000 x *g*, 10 minutes, 4 °C). 250 µL chloroform was added to supernatant, sample pulse-vortexed and then incubated at room temperature for 10 minutes prior to centrifugation (12,000 x *g*, 15 minutes, 4 °C). The upper aqueous phase was transferred to fresh tubes containing 2 µL of 5 µg/µL glycogen (*Ambion*), pulse vortexed, and 500 µL of ice-cold isopropanol was added followed by further pulse vortexing and storage at -20 °C overnight. The next day samples were pulse vortexed, centrifuged (12,000 x *g*, 10 minutes, 4 °C) and supernatant removed. RNA pellet was washed in ice cold 75% EtOH, pulse-vortexed and centrifuged (7,500 *g*, 5 minutes, 4 °C), and EtOH wash step repeated. RNA pellet was resuspended in nuclease-free water. Following DNase (TURBO DNase, *Thermofisher*, manufacturer's instructions) treatment, 1/10 volume of 3M NaOAc, pH 5.2, 5 volumes EtOH and 0.5 µL glycogen (5 µg/µL) was added to RNA, and samples incubated overnight at -80 °C. The following day samples were centrifuged (12,000 x *g*, 20 minutes, 4 °C) and the RNA pellet washed twice with ice cold 80% EtOH. Wash was removed, pellet air dried and

resuspended in nuclease-free water. RNA quantity and quality was assessed using a Bioanalyzer RNA pico kit (*Agilent*).

Supplemental references

1. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology.* 2006;11(1):54-61.
2. Vieira Braga FA, Kar G, Berg M, et al. A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat Med.* 2019;25(7):1153-1163.

Supplemental tables

Table E1 – Gene ontology analysis of top 200 most highly expressed genes in isolated sputum macrophages

PANTHER GO Biological Process of top 200 most highly expressed genes in sputum macrophages	Fold Enrichment	raw P value	FDR
cellular iron ion homeostasis (GO:0006879)	33.83	1.79E-05	1.23E-02
↳iron ion homeostasis (GO:0055072)	20.30	9.13E-05	1.71E-02
T cell mediated immunity (GO:0002456)	27.68	3.35E-05	1.38E-02
↳adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains (GO:0002460)	16.55	2.69E-05	1.38E-02
↳adaptive immune response (GO:0002250)	12.69	8.25E-05	2.12E-02
↳immune response (GO:0006955)	4.23	9.02E-05	1.85E-02
↳response to stimulus (GO:0050896)	1.83	7.10E-05	2.08E-02
↳immune system process (GO:0002376)	3.20	5.08E-04	3.17E-02
↳lymphocyte mediated immunity (GO:0002449)	14.50	2.80E-04	2.51E-02
↳leukocyte mediated immunity (GO:0002443)	12.69	4.40E-04	3.12E-02
positive regulation of adaptive immune response (GO:0002821)	25.37	4.61E-06	9.48E-03
↳regulation of response to stimulus (GO:0048583)	2.47	4.85E-04	3.32E-02
wound healing (GO:0042060)	13.24	3.81E-04	2.80E-02
↳response to wounding (GO:0009611)	12.18	5.05E-04	3.25E-02
granulocyte chemotaxis (GO:0071621)	11.89	1.08E-04	1.71E-02
↳myeloid leukocyte migration (GO:0097529)	9.76	2.51E-04	2.46E-02
↳leukocyte migration (GO:0050900)	8.27	5.05E-04	3.35E-02
↳cell migration (GO:0016477)	4.09	2.42E-04	2.49E-02
↳cell motility (GO:0048870)	3.99	1.48E-04	1.90E-02
↳locomotion (GO:0040011)	3.57	2.00E-04	2.42E-02
↳localization of cell (GO:0051674)	3.99	1.48E-04	2.03E-02
↳leukocyte chemotaxis (GO:0030595)	10.29	2.01E-04	2.29E-02
↳cell chemotaxis (GO:0060326)	7.32	8.45E-04	4.96E-02
↳response to chemical (GO:0042221)	2.51	2.75E-04	2.57E-02
↳cellular response to chemical stimulus (GO:0070887)	2.84	3.60E-04	2.74E-02
response to lipopolysaccharide (GO:0032496)	9.28	3.10E-04	2.55E-02
↳response to oxygen-containing compound (GO:1901700)	5.27	8.29E-05	1.89E-02
↳response to molecule of bacterial origin (GO:0002237)	9.06	3.44E-04	2.72E-02
↳response to organic substance (GO:0010033)	3.49	1.25E-05	1.28E-02
↳response to bacterium (GO:0009617)	7.14	2.90E-04	2.48E-02
↳response to lipid (GO:0033993)	8.07	4.38E-05	1.50E-02
protein folding (GO:0006457)	6.12	2.19E-04	2.37E-02
actin filament organization (GO:0007015)	4.51	5.54E-04	3.35E-02

↳supramolecular fiber organization (GO:0097435)	4.14	1.07E-04	1.83E-02
cellular response to organic substance (GO:0071310)	3.27	1.44E-04	2.11E-02

FDR = False Discovery Rate (Benjamini-Hochberg). GO = gene ontology.

Table E2 – Top 200 most differentially expressed genes in NA vs NNA ranked by ratio of s value to absolute (log2) fold change

gene name	basemean	log2 fold change (log FC)	s value	s value:log FC
ALS2CL	18.3308059	6.23702503	**4.68E-05**	7.50E-06
AC016739.1	39.3857545	3.89333632	**7.34E-05**	1.88E-05
PCAT1	15.3416111	4.94087738	**0.00083966**	0.00016994
MEIS2	15.0060612	4.07088106	**0.00155269**	0.00038141
CT75	5.12772434	6.28413555	**0.00333884**	0.00053131
KCNN1	12.1364537	4.06041588	**0.00231619**	0.00057043
CHRNA1	36.4039276	2.69822201	**0.00286392**	0.00106141
TNNT3	10.7682381	3.4749123	**0.00468312**	0.00134769
FP236383.7	6.5809419	3.43735738	**0.0073903**	0.00214999
MTCL1	4.77697416	3.8814654	**0.00888854**	0.00229
HIST2H2AA3	19.5196043	-2.1912641	**0.00596279**	0.00272116
ASIC4	9.45349547	4.08705745	**0.01168584**	0.00285923
CSF3	27.3603882	4.36780292	**0.0145193**	0.00332416
FP671120.4	97.029638	2.66239212	**0.01023485**	0.00384423
STAB2	15.9178548	3.43163807	**0.01573109**	0.00458413
GNG2	64.0255601	2.43708813	**0.01317257**	0.00540504
PI3	131.920889	3.09240412	**0.01784574**	0.00577083
ARNT2	15.5004995	3.62168066	**0.02193361**	0.0060562
SEMA5B	16.2044673	4.31806254	**0.02728612**	0.00631907
TESPA1	9.23897809	3.092004	**0.01990585**	0.00643785
SETBP1	17.0794076	2.93198469	**0.01891659**	0.0064518
GJB2	59.7538374	2.42267114	**0.01681126**	0.00693914
GNG7	12.1567058	2.96063542	**0.02091907**	0.00706574
NWD1	10.4237459	3.6399842	**0.02846776**	0.00782085
TMC2	8.07393929	3.431363	**0.0295698**	0.00861751
VWCE	11.4902075	2.82614801	**0.02489335**	0.00880823
CHST2	51.527035	2.94930476	**0.02609103**	0.0088465
HS3ST2	484.14142	2.51249253	**0.02286962**	0.00910236
WDR64	10.9970895	3.28579957	**0.03192119**	0.00971489
AC107223.1	13.1016113	3.3988442	**0.03585101**	0.010548
CACNG8	1.79248996	3.14997115	**0.03329827**	0.01057098
AL136309.3	7.05832363	3.44127742	**0.03845797**	0.01117549
ADORA2A	6.20729089	3.0552116	**0.03715963**	0.0121627
SCIN	99.0877689	2.30952908	**0.03059922**	0.01324912
OR2R1P	9.56125912	3.2118045	**0.04374637**	0.0136205
LINC01315	8.74292193	2.89542696	**0.03976286**	0.01373299
BACH2	14.0182056	3.11025347	**0.04502978**	0.01447785
GPR183	464.554089	1.63464929	**0.02388133**	0.01460945
AC091117.2	7.25864356	3.25627773	**0.04774075**	0.01466114
ANK3	12.0029747	2.10720816	**0.03460782**	0.01642354
PHYHD1	6.50095779	2.7543454	0.05043706	0.01831182
NKG7	105.044052	2.77129656	0.05554	0.02004116
DRAIC	16.597115	2.10603896	**0.0424374**	0.02015034
FLT3	19.8874968	2.66889361	0.05430375	0.02034692
CLEC10A	109.462126	2.19659688	**0.04635112**	0.02110133
IL2RA	39.2988354	2.26323583	**0.04911949**	0.02170321
BAALC	30.2656793	1.88685086	**0.04107111**	0.02176701
PARD3	16.9434622	2.76799988	0.06223009	0.02248197
ADGRG6	18.3835451	2.46192056	0.05946466	0.02415377
SAA1	34.2753828	2.32187581	0.05677343	0.02445154
LINC00511	12.2261977	2.00311399	0.05174927	0.02583441
AC112206.2	11.5383643	2.7259777	0.07137848	0.02618454
CAPN8	8.41744422	2.74340177	0.07262621	0.02647305
SCUBE1	9.12337648	2.93206835	0.08106071	0.02764625
NIPAL4	52.4580041	2.59446223	0.07512077	0.02895427
S100A14	12.3635852	3.00917974	0.08771177	0.02914807
BCL2L15	8.46470978	2.67525002	0.08661304	0.03237568
SCN1B	63.8629254	1.87610007	0.06086916	0.03244452
SGPP2	92.5863261	1.78529227	0.05811936	0.03255454
HELLPAR	77.5972584	1.62604305	0.05301631	0.03260449
CRISPLD2	34.0105144	2.13689642	0.07013445	0.03282071
EFR3B	21.3210027	2.30801793	0.0763134	0.03306447
ROCK1P1	16.0675184	2.03229299	0.06758111	0.03325363
CCR7	143.248739	2.20858995	0.07389245	0.03345684
TRIM67	5.72048978	2.4026925	0.08219917	0.03421127
SPTBN2	23.2911697	2.26051448	0.07752132	0.03429366
MCTP2	58.5375979	1.8994159	0.06627106	0.03489023
RYR3	54.1934019	1.80495438	0.06355218	0.03520985
RHBDF1	9.43815998	2.3698967	0.08444339	0.03563167
GSG1L	17.2003593	2.38553022	0.08879202	0.03722109

SLC26A4-AS1	4.61951268	2.56472246	0.09816633	0.03827561
C5orf64	12.6572257	2.03586846	0.07990401	0.03924812
CERS4	14.5253052	2.43863827	0.10023798	0.04110408
ATP8B5P	10.9223233	2.52663679	0.10532519	0.04168592
CUBN	13.4266911	1.98274927	0.08333159	0.04202831
CYBRD1	111.397025	1.53608696	0.06491661	0.04226103
EBI3	35.8227126	2.64029354	0.11210049	0.04245759
SLC18B1	131.185287	1.61327342	0.0688746	0.04269246
GGT5	11.1542109	2.48572229	0.10927733	0.043962
CHST10	13.0851255	2.1980081	0.10329352	0.04699415
AMPH	23.7809918	2.01331815	0.09503899	0.04720515
SH3RF3	128.586072	1.95728093	0.09297342	0.04750131
AKAP6	14.3501966	2.36109124	0.11308704	0.04789609
ERICH6B	3.23731223	2.5376029	0.12195575	0.04805943
AC091100.1	8.43924766	2.1958042	0.10730531	0.04886834
ST18	132.872097	1.55232829	0.07871424	0.05070721
PLXNC1	1183.53107	1.63943251	0.08553169	0.05217152
NT5DC2	45.4184327	1.84736542	0.09712668	0.05257578
TMIGD3	146.722497	2.16477417	0.11513777	0.05318697
LY9	101.6663	1.99279773	0.10632265	0.05335346
CCDC26	5.54738546	2.11766383	0.11973457	0.05654088
GOLGA8VP	13.1976025	2.13633153	0.12084896	0.05656845
KANK1	101.852395	1.57205691	0.08986191	0.05716199
SLAMF7	1594.74823	1.60412567	0.09193703	0.05731286
MTCO3P12	123.104762	1.58590272	0.09090436	0.05732026
HULC	19.3436471	1.68682692	0.10125688	0.06002802
TDRD6	350.399238	1.56412681	0.09401884	0.06010948
OAF	31.0892626	1.88325399	0.11405564	0.06056307
ITGA7	54.4817732	1.81116427	0.11023578	0.0608646
SRGAP3	64.5614415	1.62247299	0.09920459	0.06114406
ANK1	19.3909889	2.02546211	0.12414344	0.06129141
MYO1D	111.10786	2.00709615	0.12525844	0.06240779
WTIP	8.40887311	2.22829216	0.1393611	0.06254167
LYPLAL1-DT	10.9736705	2.05374614	0.13428264	0.06538424
TRAF1	456.24369	1.50701687	0.10225858	0.06785497
AC099489.1	239.680235	1.41211644	0.09609397	0.06804961
SYNE2	77.5281564	1.53214914	0.10431988	0.06808729
SLAMF1	121.069294	1.76353165	0.12749017	0.07229253
CTBP2P4	6.3110456	1.80070387	0.13311708	0.07392503
SNCA	115.817019	1.46903821	0.11117756	0.07568051
ZC3H12C	361.017483	1.41539343	0.10830074	0.07651635
AC022784.1	6.38477887	1.75448482	0.13553082	0.07724821
SIGLEC10	506.237746	1.58208149	0.12305117	0.07777802
DYSF	1318.88828	1.50141218	0.11745609	0.07823041
DPP4	63.4139703	1.84722554	0.1459559	0.07901358
AC010300.1	24.9979797	1.45576422	0.11630549	0.07989308
OBSL1	17.20077	1.81816847	0.14727133	0.08099983
PSAT1	19.3945054	1.50791537	0.13081352	0.08675123
IL10	96.1124828	1.56068719	0.13680567	0.08765733
IL1R1	465.929592	1.45235689	0.12971958	0.0893166
FFAR2	146.10423	1.29524902	0.11860249	0.09156733
TRPV6	30.8502174	1.60883709	0.14856933	0.09234579
OLFML2B	890.746677	1.54736398	0.1432927	0.09260439
LRP4	17.6867409	1.95043374	0.1838877	0.09428041
MPO	21.815474	1.65439744	0.15895958	0.09608306
AC013457.1	640.635203	-1.3143297	0.12638325	0.09615795
KIAA0319	25.4857401	1.47351208	0.14194005	0.09632771
LAMB3	137.708319	1.3479703	0.13194179	0.09788182
DSEL	12.4367334	1.71994709	0.1697449	0.09869193
LRRC49	22.1028629	1.50739967	0.14984777	0.09940812
AKR1C1	193.571458	-1.2760576	0.128612	0.10078855
PRG4	10.8658404	1.58490599	0.16159798	0.10196061
PRLR	39.5083147	1.52635915	0.15767175	0.10329925
NEURL3	13.4142969	1.65255033	0.17254026	0.10440848
KALRN	37.1800684	1.31419486	0.14064592	0.1070206
ALOX15B	605.151109	1.52639158	0.16427295	0.10762176
SDS	988.139968	1.32964222	0.14462921	0.10877303
BASP1	1412.81311	1.24553267	0.13808406	0.11086346
EDNRB	16.1032035	1.63053622	0.18102392	0.1110211
HRH2	187.335102	1.4963571	0.16693958	0.111564
SH3RF1	64.8782741	1.35021319	0.15247412	0.11292596
SLC38A1	18.5887388	1.64023302	0.18685475	0.11391964
G0S2	1731.86255	1.3150516	0.15377118	0.11693167
RGL1	874.818368	1.3211555	0.15637539	0.11836259
ADAM28	127.313709	1.26134875	0.1511684	0.11984663

LAMA3	52.6918914	1.34248112	0.16294297	0.1213745
LINC01480	16.3752188	1.35165533	0.1683363	0.12454085
HS3ST1	126.084076	1.24250213	0.15507995	0.12481263
FBLIM1	34.7709187	1.361933	0.17114839	0.12566579
MYLK	38.9768949	1.38384436	0.17394077	0.12569388
EFCAB6	24.3253054	1.32463015	0.17676841	0.13344737
TCF4	227.373268	1.14895288	0.16023945	0.13946564
TFF3	75.2630459	-1.170518	0.1656085	0.14148309
RASSF2	771.607852	1.22720728	0.17816534	0.1451795
LSP1P4	16.9006776	1.25485013	0.18244628	0.14539289
PLA2G7	1666.99617	1.18334356	0.17536085	0.14819099
LAMC2	19.8306085	1.27758418	0.18986789	0.14861478
AC034199.1	46.0825877	1.2268738	0.18538007	0.15109955
MCOLN3	65.1118196	1.17522615	0.17959663	0.15281879
RELN	7.21268123	1.20997873	0.19285444	0.15938664
MIR4458HG	44.7118171	1.11531093	0.18836489	0.16889002
MERTK	704.662421	1.15211552	0.19582797	0.16997252
IL23A	60.395182	1.10948687	0.1913585	0.17247478
ADA	113.993821	1.13513614	0.20200753	0.17795886
LINC01482	48.4206447	1.10758899	0.1973272	0.17815923
SRGAP1	191.049696	1.11091732	0.19886394	0.17900877
AZIN1-AS1	226.715733	1.08696781	0.20041561	0.18438044
AC005972.3	312.493579	-1.0458488	0.19434948	0.18582943
ZC3H12D	100.547671	1.02881501	0.20365173	0.19794786
TMEM154	61.9654032	0.97897893	0.20542146	0.20983236
FAM20A	888.543681	0.94643405	0.20920093	0.22104121
SELL	74.602287	0.92940626	0.2111248	0.22716094
GAREM1	33.4404202	0.87944615	0.20729871	0.23571506
SERPINB9	333.265042	0.91620986	0.21736843	0.23724742
AC025822.1	32.6273375	0.88913233	0.21317027	0.23975089
CCL5	140.238767	0.88743456	0.21525253	0.24255594
CD48	370.050431	0.89568414	0.21964715	0.24522835
ATP8A1	99.9631522	0.86944068	0.22462849	0.25835976
RNVU1-7	414.75192	0.87761727	0.23472409	0.26745609
TCN2	2353.7491	0.78399241	0.23726748	0.30264002
CDCA4	82.7538857	0.76151173	0.24485084	0.32153259
CSF3R	1339.1169	0.79528655	0.26003252	0.32696708
NCR3LG1	320.614233	0.76946158	0.25746161	0.3345997
PFKFB3	2725.10972	0.78313759	0.26273995	0.33549654
ST8SIA4	286.512149	0.67977472	0.2321838	0.34155993
SGTB	663.104033	0.7746128	0.27083251	0.34963598
NEB	131.35929	0.70495043	0.24735674	0.3508853
MARCKS	2442.53097	0.62013821	0.24232971	0.39076726
FCMR	306.507988	0.57016593	0.22714723	0.39838794
IL6R	922.52502	0.71291033	0.29306369	0.41108071
SLC25A37	374.316325	0.69675862	0.28744518	0.41254628
RARRES1	208.907019	0.54749861	0.22964785	0.41944919
FNIP2	3509.89201	0.67757484	0.29025287	0.4283702
CYB5RL	107.905516	0.63753493	0.28466717	0.44651226
KCNJ2	131.135779	0.60281606	0.27357633	0.45383053
POU2F2	178.074218	0.61182527	0.27905328	0.45609963
GPR153	305.551329	-0.6543106	0.30717805	0.46946824
KSR1	909.124966	-0.6355308	0.30150711	0.4744178
FCGR3B	116.841618	0.51679367	0.24986272	0.48348642
RGS1	4667.25496	0.56874386	0.2763026	0.48581201

Basemean indicates average reads per sample across all samples. S value is estimated p value in the apeglm differential expression package. Bolded underlined RNA species were also tested by qPCR in a larger sample set (Figure 2). S values below 0.05 are bolded.

Table E3 – Sputum macrophage qPCR validation cohort clinical characteristics

	All (n=36)	NNA (n=23)	NA (n=13)
Female[†]	17 (47.2%)	10 (45.5%)	7 (50%)
Age[§]	60.9 (45.7, 69.3)	56.2 (39.1, 69.9)	62.6 (52.9, 66.5)
BMI[§]	28.6 (25.7, 37.6)	30.9 (26.2, 36.3)	27.7 (24.0, 38.4)
Pre-β2 FEV1 (% pred)[‡]	76.0 (18.8)	75.8 (18.7)	76.3 (19.7)
Pre-β2 FVC (% pred)[‡]	87.6 (15.8)	86.5 (17.9)	89.4 (12.2)
Pre-β2 FEV1/FVC	67.1 (12.4)	68.2 (10.8)	65.5 (14.9)
Ex-smoker[†]	12 (33%)	6 (27.2%)	6 (42.9%)
ACQ6[§]	0.75 (0.33, 1.67)	0.5 (0.33, 1.54)	1.1 (0.3, 2.17)
ICS treated[†]	30 (83.3%, n=35)	19 (86.4%)	11 (84.6%)
ICS dose (fluticasone equivalent)[§]	500 (500,500)	500 (250, 625)	500 (500, 625)
Sputum TCC (x10^6/mL)[§]	6.5 (4.3, 12.5)	4.9 (3.7, 8.4)	13.6 (6.7, 21.1)[*]
Sputum macrophage (%)[§]	42.6 (19.4, 60.7)	59.1 (44.1, 67.8)	12.3 (4.6, 26.1)[*]
Sputum neutrophil (%)[§]	50.6 (31.8, 68)	36.9 (24.3, 50.3)	81.4 (63.9, 89.4)[*]
Sputum eosinophil (%)[§]	1.5 (0.8, 4.2)	1.5 (0.8, 3.3)	1.5 (0.4, 5.9)
Sputum lymphocyte (%)[§]	1.3 (0.5, 2)	1.6 (0.9, 2.2)	0.6 (0, 2.4)
Sputum macrophage (x10^4/mL)[§]	208 (138, 322)	289 (187, 483)	124 (65, 228)[***]
Sputum neutrophil (x10^4/ mL)[§]	267 (152, 612)	178 (132, 278)	1216 (503, 1692)[***]
Sputum eosinophil (x10^4/ mL)[§]	9 (4, 24)	8 (4, 18)	15 (2, 63)
Sputum lymphocyte (x10^4/ mL)[§]	9 (2, 19)	9 (4, 18)	6 (0, 22)

[†] = Number (%),[‡] = Mean (SD), [§] = Median (IQR). NA = Neutrophilic asthma. NNA = Non-neutrophilic asthma. BMI = body mass index. FEV_1 = forced expiratory volume in 1 second. FVC = forced vital capacity. ICS = inhaled corticosteroids. ACQ6 = asthma control questionnaire, 6 item. TCC = total cell count.

** = $p<0.05$ vs NNA. Comparisons are NA vs NNA, Mann-Whitney test for non-parametric variables, Student's T-test for parametric variables, Fisher's exact for categorical variables.*

Table E4 – differentially expressed genes selected for qPCR validation

Gene symbol	Coded protein	Basemean	Function
MERTK	Mer tyrosine kinase	704.6	Efferocytosis receptor
SLAMF7	SLAMF7/CD319	1594.7	Macrophage phagocytosis promoter
DYSF	Dysferlin	1318.9	Cell-cell adhesion
GPR183	G-protein coupled receptor 183/EBI2	464.6	Chemotactic receptor
CSF3	Granulocyte colony stimulating factor	27.3	Promote neutrophil survival and function
PI3	Elafin/peptidase inhibitor 3	131.9	Serine antiprotease/ antimicrobial/immune modulator
CCR7	C-C chemokine receptor 7	143.2	Chemokine receptor/immune signaling

Table E5 – STRING GO categories enriched in top 200 NA sputum macrophage DEGs

GO term ID	GO term description	observed gene count	background gene count	FDR
GO:0006952	defense response	30	1234	0.00068
GO:0050727	regulation of inflammatory response	15	338	0.00068
GO:0002682	regulation of immune system process	31	1391	0.00087
GO:0001775	cell activation	25	1024	0.0019
GO:0030595	leukocyte chemotaxis	9	130	0.0019
GO:0046649	lymphocyte activation	14	358	0.0019
GO:0006954	inflammatory response	16	482	0.0022
GO:0032879	regulation of localization	43	2524	0.0022
GO:0002376	immune system process	41	2370	0.0024
GO:0006955	immune response	31	1560	0.0024
GO:0030334	regulation of cell migration	20	753	0.0024
GO:0040012	regulation of locomotion	22	881	0.0024
GO:0009605	response to external stimulus	38	2152	0.0025
GO:0050900	leukocyte migration	12	296	0.0027
GO:0051239	regulation of multicellular organismal process	45	2788	0.0027
GO:0031347	regulation of defense response	18	676	0.0043
GO:0051270	regulation of cellular component movement	21	886	0.0044
GO:0045321	leukocyte activation	21	894	0.0048
GO:0007155	cell adhesion	20	843	0.0055
GO:0042327	positive regulation of phosphorylation	22	984	0.0055
GO:0045937	positive regulation of phosphate metabolic process	23	1052	0.0055
GO:0050896	response to stimulus	92	7824	0.0055
GO:0010647	positive regulation of cell communication	30	1631	0.0057
GO:0022409	positive regulation of cell-cell adhesion	10	238	0.0057
GO:0051240	positive regulation of multicellular organismal process	29	1551	0.0057
GO:0051249	regulation of lymphocyte activation	13	401	0.0057
GO:0001934	positive regulation of protein phosphorylation	21	941	0.006
GO:0002694	regulation of leukocyte activation	14	470	0.0061
GO:0010536	positive regulation of activation of Janus kinase activity	3	7	0.0064
GO:0050863	regulation of T cell activation	11	302	0.0066
GO:0016477	cell migration	19	812	0.0067
GO:0019935	cyclic-nucleotide-mediated signaling	7	115	0.0067
GO:0043270	positive regulation of ion transport	10	252	0.0067
GO:0048583	regulation of response to stimulus	54	3882	0.0071
GO:0006928	movement of cell or subcellular component	26	1355	0.0072
GO:0033632	regulation of cell-cell adhesion mediated by integrin	3	8	0.0072
GO:0055065	metal ion homeostasis	15	555	0.0073
GO:0023056	positive regulation of signaling	29	1638	0.0096
GO:0002687	positive regulation of leukocyte migration	7	127	0.0097
GO:0065008	regulation of biological quality	50	3559	0.0097
GO:0002685	regulation of leukocyte migration	8	175	0.0108

GO:0002690	positive regulation of leukocyte chemotaxis	6	91	0.011
GO:0006812	cation transport	19	866	0.011
GO:0009967	positive regulation of signal transduction	27	1493	0.011
GO:1903037	regulation of leukocyte cell-cell adhesion	10	278	0.011
GO:0007166	cell surface receptor signaling pathway	35	2198	0.0118
GO:0002684	positive regulation of immune system process	19	882	0.0123
GO:0032101	regulation of response to external stimulus	20	955	0.0123
GO:0048518	positive regulation of biological process	68	5459	0.0123
GO:0072511	divalent inorganic cation transport	10	291	0.0134
GO:0031581	hemidesmosome assembly	3	12	0.0135
GO:0048584	positive regulation of response to stimulus	33	2054	0.0142
GO:0002696	positive regulation of leukocyte activation	10	298	0.0153
GO:0006935	chemotaxis	13	491	0.0156
GO:0006936	muscle contraction	9	244	0.0156
GO:0007528	neuromuscular junction development	4	36	0.0156
GO:0009617	response to bacterium	14	555	0.0156
GO:0019933	cAMP-mediated signaling	6	107	0.0156
GO:0022407	regulation of cell-cell adhesion	11	366	0.0156
GO:0023052	signaling	64	5108	0.0156
GO:0030155	regulation of cell adhesion	15	623	0.0156
GO:0031401	positive regulation of protein modification process	22	1149	0.0156
GO:0032740	positive regulation of interleukin-17 production	3	13	0.0156
GO:0033555	multicellular organismal response to stress	5	67	0.0156
GO:0040011	locomotion	22	1144	0.0156
GO:0048246	macrophage chemotaxis	3	14	0.0156
GO:0050801	ion homeostasis	16	708	0.0156
GO:0050862	positive regulation of T cell receptor signaling pathway	3	13	0.0156
GO:0050870	positive regulation of T cell activation	8	193	0.0156
GO:0072507	divalent inorganic cation homeostasis	12	429	0.0156
GO:0080134	regulation of response to stress	24	1299	0.0156
GO:0097529	myeloid leukocyte migration	6	103	0.0156
GO:0045785	positive regulation of cell adhesion	11	375	0.0161
GO:0051707	response to other organism	22	1173	0.0162
GO:0050670	regulation of lymphocyte proliferation	8	205	0.0163
GO:0051251	positive regulation of lymphocyte activation	9	260	0.0163
GO:0046427	positive regulation of receptor signaling pathway via JAK-STAT	5	73	0.0171
GO:0001819	positive regulation of cytokine production	11	390	0.0181
GO:0019221	cytokine-mediated signaling pathway	15	655	0.0181
GO:0019932	second-messenger-mediated signaling	9	266	0.0181
GO:0030335	positive regulation of cell migration	12	452	0.0181
GO:0032501	multicellular organismal process	76	6507	0.0181
GO:0044057	regulation of system process	13	516	0.0181
GO:0048522	positive regulation of cellular process	61	4898	0.0181

GO:1902533	positive regulation of intracellular signal transduction	19	959	0.0181
GO:0050728	negative regulation of inflammatory response	6	117	0.0182
GO:0030001	metal ion transport	15	664	0.0186
GO:0051179	localization	64	5233	0.0186
GO:0051050	positive regulation of transport	18	892	0.0192
GO:0032026	response to magnesium ion	3	18	0.0197
GO:0050729	positive regulation of inflammatory response	6	120	0.0197
GO:0055074	calcium ion homeostasis	11	402	0.0214
GO:0002407	dendritic cell chemotaxis	3	19	0.0217
GO:0010646	regulation of cell communication	45	3327	0.0217
GO:0042110	T cell activation	8	225	0.0217
GO:0098657	import into cell	14	609	0.0217
GO:1900273	positive regulation of long-term synaptic potentiation	3	19	0.0217
GO:0030098	lymphocyte differentiation	8	226	0.0221
GO:0071622	regulation of granulocyte chemotaxis	4	47	0.0231
GO:0042592	homeostatic process	25	1491	0.0233
GO:0070838	divalent metal ion transport	9	287	0.0237
GO:2000107	negative regulation of leukocyte apoptotic process	4	48	0.0243
GO:0023051	regulation of signaling	45	3360	0.0244
GO:2000106	regulation of leukocyte apoptotic process	5	85	0.0245
GO:0007154	cell communication	63	5219	0.0251
GO:0031348	negative regulation of defense response	7	180	0.0251
GO:0046634	regulation of alpha-beta T cell activation	5	86	0.0251
GO:0050671	positive regulation of lymphocyte proliferation	6	130	0.0251
GO:0050731	positive regulation of peptidyl-tyrosine phosphorylation	7	180	0.0251
GO:0090026	positive regulation of monocyte chemotaxis	3	21	0.0251
GO:0002335	mature B cell differentiation	3	22	0.027
GO:2000026	regulation of multicellular organismal development	29	1876	0.0272
GO:0034329	cell junction assembly	6	135	0.0278
GO:0007165	signal transduction	58	4738	0.0291
GO:0050920	regulation of chemotaxis	7	188	0.0292
GO:0019226	transmission of nerve impulse	4	53	0.0293
GO:0032655	regulation of interleukin-12 production	4	53	0.0293
GO:0006810	transport	52	4130	0.0295
GO:0050776	regulation of immune response	17	873	0.0295
GO:0050808	synapse organization	7	189	0.0295
GO:0098655	cation transmembrane transport	15	720	0.0295
GO:0001932	regulation of protein phosphorylation	23	1370	0.0297
GO:0008284	positive regulation of cell population proliferation	17	878	0.0299
GO:0009966	regulation of signal transduction	41	3033	0.0299
GO:0048731	system development	52	4144	0.0308
GO:0042325	regulation of phosphorylation	24	1465	0.0312
GO:0051046	regulation of secretion	15	728	0.0312

GO:0007204	positive regulation of cytosolic calcium ion concentration	8	251	0.0317
GO:0002521	leukocyte differentiation	9	313	0.0322
GO:0034695	response to prostaglandin E	3	25	0.0322
GO:0043410	positive regulation of MAPK cascade	12	512	0.0322
GO:0046640	regulation of alpha-beta T cell proliferation	3	25	0.0322
GO:0032649	regulation of interferon-gamma production	5	97	0.0326
GO:0019220	regulation of phosphate metabolic process	26	1657	0.0342
GO:0032272	negative regulation of protein polymerization	4	58	0.0343
GO:0030183	B cell differentiation	5	99	0.0345
GO:0034112	positive regulation of homotypic cell-cell adhesion	2	6	0.035
GO:0060405	regulation of penile erection	2	6	0.035
GO:0071286	cellular response to magnesium ion	2	6	0.035
GO:0032270	positive regulation of cellular protein metabolic process	24	1496	0.0372
GO:0042531	positive regulation of tyrosine phosphorylation of STAT protein	4	60	0.0372
GO:0007275	multicellular organism development	57	4726	0.0377
GO:0051247	positive regulation of protein metabolic process	25	1587	0.0377
GO:0042129	regulation of T cell proliferation	6	151	0.038
GO:0042127	regulation of cell population proliferation	25	1594	0.0391
GO:0048878	chemical homeostasis	18	995	0.0391
GO:0001508	action potential	5	104	0.0398
GO:0030048	actin filament-based movement	5	105	0.0412
GO:0043279	response to alkaloid	5	106	0.0426
GO:0001817	regulation of cytokine production	13	615	0.0438
GO:0001662	behavioral fear response	3	30	0.045
GO:0048265	response to pain	3	30	0.045
GO:0098662	inorganic cation transmembrane transport	13	618	0.045
GO:0014068	positive regulation of phosphatidylinositol 3-kinase signaling	4	65	0.0451
GO:0001776	leukocyte homeostasis	4	66	0.0473
GO:0035637	multicellular organismal signaling	5	110	0.0473
GO:0030832	regulation of actin filament length	6	161	0.0476
GO:0006941	striated muscle contraction	5	111	0.0486
GO:0015889	cobalamin transport	2	8	0.0488
GO:0002250	adaptive immune response	8	280	0.0496
GO:1903706	regulation of hemopoiesis	10	412	0.0496
GO:0060306	regulation of membrane repolarization	3	32	0.0498
GO:0071624	positive regulation of granulocyte chemotaxis	3	32	0.0498
GO:0098660	inorganic ion transmembrane transport	14	707	0.0498

Table E6 – Whole sputum microarray cohort clinical characteristics

	Control (n=16)	NNA (n=37)	NA (n=29)
Age[†]	63.1 (27.1, 72.4)	59.3 (47.4, 68.5)	63.3 (53.8, 69.7)
Gender (female, %)	56.25	62.16	44.83
BMI[†]	26.3 (23.7, 28.1)	29.0 (23.5, 35.2)	30.4 (27.3, 33.9)*
Ex/current smoker (%)	43.75	43.24	58.62
Pack years[†]	39.5 (11.8, 54)	7.4 (0.6, 22.4)*	6.8 (3.9, 27.5)
Pre-bronchodilator FEV_1 (% predicted)[‡]	101.2 (11.2)	77.3 (19.5)*	68.9 (20.2)*
Pre-bronchodilator FVC (% (predicted)[‡]	103.9 (12.5)	92.5 (17.9)	87.8 (16.1)*
Pre-bronchodilator FEV_1/FVC[‡]	79.2 (5.8)	67.8 (10.5)	64.6 (10.4)
Taking ICS (%)	-	62.2	82.8
ICS dose (Budesonide equivalent)[†]	-	400 (0, 1000)	1000 (200, 2000)[#]
ACQ6[†]		0.3 (0, 1.2)	0.8 (0.2, 1.5)
Sputum TCC ($x10^6$/mL)[†]	3.7 (2.1, 5.7)	3.0 (2.0, 4.8)	5.5 (2.9, 10.7)[#]
Macrophage %[†]	42.2 (33, 68.8)	55 (36.4, 69.1)	22 (13.3, 24.3)*,[#]
Neutrophil %[†]	44.7 (13.8, 60.8)	35 (16, 44.9)	75.5 (67.9, 82.3)*,[#]
Eosinophil %[†]	0.25 (0, 0.75)	4.5 (0, 22.75)*	1 (0.38, 4.8)
Lymphocyte %[†]	1 (0.25, 3)	0.5 (0.25, 1.5)	0.25 (0, 1)

[†] = *Median (IQR),*[‡] = *Mean (SD).* NA = Neutrophilic asthma. NNA = Non-neutrophilic asthma. BMI = body mass index. FEV_1 = forced expiratory volume in 1 second. FVC = forced vital capacity. ICS = inhaled corticosteroids. ACQ6 = asthma control questionnaire, 6 item. TCC = total cell count. * = $p<0.05$ vs control. [#] = $p<0.05$ vs NNA. Kruskal-Wallis with Dunn's multiple comparison test for non-parametric variables, ANOVA with Tukey's multiple comparisons for parametric variables.

Table E7 – neutrophilic macrophage gene sets for GSVA analysis

Gene set and description	Gene IDs
Gene set 1: most differentially expressed genes NA vs NNA macrophages	*ALS2CL, MEIS2, KCNN1, CHRNA1, TNNT3, KIAA0802, HIST2H2AA3, ACCN4, CSF3, STAB2, GNG2, PI3, ARNT2, SEMA5B, SETBP1, GJB2, GNG7, NWD1, TMC2, VWCE, CHST2, HS3ST2, WDR64, CACNG8, ADORA2A, SCIN, BACH2, EBI2, ANK3, PHYHD1, NKG7, FLT3, CLEC10A, IL2RA, BAALC, PARD3, ADGRG6, SAA1, SCUBE1, S100A14, BCL2L15, SCN1B, SGPP2, CRISPLD2, CCR7, TRIM67, SPTBN2, MCTP2, RYR3, RHBDF1, GSG1L, LASS4, CUBN, CYBRD1, EBI3, C6ORF192, GGT5, CHST10, AMPH, AKAP6, ST18, PLXNC1, NT5DC2, LY9, CCDC26, KANK1, SLAMF7, TDRD6, OAF, ITGA7, SRGAP3, ANK1, MYO1D, WTIP, TRAF1, SYNE2, SLAMF1, SNCA, SIGLEC10, DYSF, DPP4, PSAT1, IL10, IL1R1, FFAR2, TRPV6, OLFML2B, LRP4, MPO, KIAA0319, LAMB3, DSEL, LRRC49, AKR1C1, PRLR, KALRN, ALOX15B, SDS, BASP1, EDNRB, HRH2, SH3RF1, SLC38A1, G0S2, RGL1, ADAM28, LAMA3, HS3ST1, FBLIM1, MYLK, EFCAB6, TCF4, TFF3, RASSF2, PLA2G7, LAMC2, MCOLN3, RELN, MERTK, IL23A, ADA, SRGAP1, TMEM154, FAM20A, SELL, GAREM1, SERPINB9, CCL5, CD48, ATP8A1, TCN2, CDCA4, CSF3R,, DKFZP686O24166, PFKFB3, ST8SIA4, SGTB, NEB, MARCKS, FAIM3, IL6R, SLC25A37, RARRES1, LOC606495, KCNJ2, POU2F2, GPR153, KSR1, FCGR3B, RGS1*
Gene set 2: genes from gene set 1 with s value <0.05	*ALS2CL, MEIS2, KCNN1, CHRNA1, TNNT3, HIST2H2AA3, KIAA0802, ACCN4, GNG2, CSF3, STAB2, GJB2, PI3, SETBP1, GNG7, ARNT2, HS3ST2, EBI2, VWCE, CHST2, SEMA5B, NWD1, TMC2, SCIN, WDR64, CACNG8, ANK3, ADORA2A, BAALC, BACH2, CLEC10A, IL2RA*

Supplemental figures and captions

Figure E1. Bioinformatics report outlining analysis pipeline and QC for sputum macrophage RNASeq dataset.

Figure E2. Average expression of lung immune cell signature genes in across asthma sputum macrophage RNASeq samples. Representative gene signatures were identified by Vieira-Braga et al. [2]. Data expressed as basemean, the average number of reads per gene across all sputum macrophage samples with RNASeq data (n=20). Due to a potential miscall of monocytes as neutrophils in the Vieira-Braga study, we label here as neutrophil/monocyte.

Figure E2:

