

Cross-Ancestry Whole-Blood Transcriptomic Profiling Reveals Bidirectional Immune and Pharmacogenomic Relevant Differences in Severe Asthma: A Sex-Adjusted Analysis

Miriam Watafua

UREKA Biotec/OmicsLogic Africa

Adekunle Oluwatosin Adeluwoye

adekunle.a@urekabiotec.org

UREKA Biotec/OmicsLogic Africa

Katherine Olajumoke David

UREKA Biotec/OmicsLogic Africa

Lawal Jamiu

UREKA Biotec/OmicsLogic Africa

Ayanfeoluwa Alabetutu

UREKA Biotec/OmicsLogic Africa

Research Article

Keywords: asthma, severe asthma, transcriptomics, ancestry, whole blood, differential expression, curated asthma genes, pharmacogenomics, sex-adjusted analysis, reproducibility

Posted Date: October 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11369877/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Marked disparities in asthma morbidity occur across racialised populations, but bulk-blood molecular differences that accompany these disparities remain incompletely characterised and may reflect both gene regulation and circulating-cell composition.

Objective

To identify whole-blood expression differences associated with the ancestry grouping recorded in GSE69683, while adjusting for sex, and to examine whether a curated asthma-gene panel showed biologically coherent bidirectional differences in severe asthma.

Methods

We analysed RMA-normalised, log₂-transformed microarray profiles from 24 U-BIOPRED participants with severe asthma (12 recorded as black_african and 12 as white_caucasian; eight females and four males per group; GPL13158, 54,715 probe sets). The primary probe-level model regressed expression on ancestry group with sex as a covariate; Benjamini-Hochberg correction was applied across all probes. Gene-level analyses included a 128-gene measurable curated asthma panel tested under one correction family and a reciprocal ancestry-adjusted sex contrast.

Results

Ancestry group was the primary source of variation in the selected samples (PC1 22.6%; PC1-PC3 41.3%). The adjusted model identified 2,134 probes at FDR < 0.05; 1,020 additionally exceeded |adjusted log₂FC| > 0.5, with 612 higher in the African-ancestry group and 408 higher in the European-ancestry group. Thirteen of 128 measurable curated asthma genes remained significant after one correction family, spanning immune, remodelling, epithelial-barrier, oxidative/metabolic and treatment-relevant biology in opposing directions. The sex contrast yielded 97 FDR-significant probes and 73 dual-threshold probes, dominated by expected sex-linked transcripts; no curated asthma gene was significant after correction.

Conclusions

Whole-blood expression differences between the two recorded ancestry groups were bidirectional rather than uniformly inflammatory in one group. The curated-gene pattern, together with a strong ACKR1/Duffy-related signal, is consistent with a mixture of disease-related transcriptional differences

and variation in circulating-cell composition. Unavailable clinical variables, including age, treatment, smoking and exacerbation history, leave residual confounding possible. Independent replication in larger ancestry-diverse and cell-resolved cohorts with relevant clinical and treatment-response data is required before these associations can be interpreted as ancestry-specific mechanisms or predictors of therapeutic response.

1. Introduction

Asthma morbidity is not distributed evenly across populations. In the United States, Black populations have experienced higher asthma prevalence in several age groups and disproportionate emergency attendance, hospitalisation and mortality, patterns that arise within wider inequities in environmental exposure, access to high-quality care and social conditions [1–5]. Race is a social classification rather than a biological mechanism, and genetic ancestry does not substitute for the social, structural and environmental processes that shape these outcomes [5]. Molecular studies are therefore most informative when ancestry-associated biology is interpreted alongside, rather than instead of, the social determinants of asthma disparity.

Asthma is also biologically heterogeneous. Molecular and clinical work has established Type 2-high and Type 2-low inflammatory states, but these categories are not exhaustive and do not map cleanly onto a single clinical phenotype [6–10]. Severe and difficult-to-treat asthma further reflects variable contributions from airway inflammation, structural change, treatment exposure, adherence, comorbidity and environmental triggers [11]. This heterogeneity makes population-level comparisons difficult: a difference in bulk-blood expression can arise from disease biology, treatment, cell composition, unmeasured clinical characteristics or several of these processes at once.

Genetic studies nevertheless show that asthma susceptibility and exacerbation risk are not identical across ancestries. Genome-wide association studies in African-ancestry and multi-ethnic cohorts have identified both shared and population-dependent risk architecture [12, 13], while admixture-based analyses have associated a higher proportion of African genetic ancestry with asthma or total IgE in some African-admixed populations [14, 15]. Associations with exacerbation risk have also varied by sex and cohort [16], and recent larger African American studies continue to identify loci not fully represented by European-ancestry discovery efforts [17]. These findings support greater ancestry diversity in asthma genomics, while also distinguishing race/ethnicity categories from direct genotype-derived ancestry measurements.

The distinction between source-recorded race/ethnicity categories and genotype-derived ancestry remains important in transcriptomic studies. Human genomics remains markedly under-representative of global diversity [18, 19], and analytical guidance for diverse cohorts emphasises careful reporting of ancestry, admixture and race/ethnicity rather than treating these concepts as interchangeable [20–22]. Gene regulation itself varies across tissues and populations, and tissue-specific expression quantitative trait loci complicate extrapolation from genotype to blood expression [23, 24]. In the present study, the

GEO metadata record race/ethnicity categories supplied with the original U-BIOPRED samples; genetic ancestry was not independently re-estimated in this secondary analysis. We therefore use the terms African-ancestry group and European-ancestry group as concise labels for the recorded categories and avoid attributing the observed expression differences to genotype-derived ancestry alone.

Whole blood offers an accessible view of systemic immune activity, but it is a cellular mixture. This matters for the Duffy-null ACKR1 variant (rs2814778), which is common in many populations of African descent and is associated with lower circulating neutrophil counts [25, 26]. Mechanistic work links erythroid ACKR1 expression to haematopoiesis and neutrophil phenotype [27], and clinical studies show how Duffy-null-associated low neutrophil counts can be misinterpreted when ancestry is used without genotype or Duffy-status information [28, 29]. Consequently, an ancestry-associated whole-blood expression signal at ACKR1 or neutrophil-enriched genes can reflect cell abundance, cell-intrinsic transcription or both.

Treatment-relevant pathways provide another reason to examine ancestry-associated expression cautiously. Variation in ALOX5 and other leukotriene-pathway genes has been associated with response to anti-leukotriene treatment [30, 31]. ADRB2 pharmacogenetic studies have examined heterogeneity in response to long-acting beta2-agonists [32], while beta-arrestin 2 regulates allergic airway responses and constrains beta2-adrenergic receptor signalling in experimental systems [33, 34]. Glucocorticoid-response pharmacogenetics likewise demonstrates substantial inter-individual variation [35]. Expression differences in these pathways may therefore identify biologically plausible treatment-related hypotheses, but without individual treatment-response measurements they cannot establish pharmacogenomic effects.

U-BIOPRED provides one of the most deeply characterised severe-asthma resources, with clinical, inflammatory and multi-omic profiling across blood and airway compartments [36–40]. We used a sex-balanced 24-participant subset of the public whole-blood microarray series GSE69683 to ask whether the ancestry groups differed in genome-wide expression after adjustment for sex. We then examined a pre-specified curated asthma-gene panel and quantified the reciprocal sex contrast as an internal check on covariate handling. Systems-level GSEA, ssGSEA and co-expression analyses of the same cohort are reported separately in a companion systems-level report [43–46]. The central interpretive question of this study was not whether one ancestry group was globally more inflammatory, but whether the gene-level data showed a coherent, bidirectional pattern that could be separated from obvious sex effects and interpreted with appropriate attention to whole-blood cell composition.

2. Materials and Methods

2.1 Study design and data source

This cross-sectional secondary analysis used publicly available whole-blood expression data from GEO accession GSE69683, generated within U-BIOPRED [36, 37]. The series was produced on the Affymetrix

HT HG-U133 + PM Array Plate (GPL13158), which contains 54,715 probe sets. GPL13158 and GPL570 share probe-set identifiers after removal of the GPL13158 '_PM_' infix, but they are distinct platforms; GPL570 annotation was used only as a cross-platform identifier resource. Expression values in the public series were supplied RMA-normalised and log2-transformed, consistent with the robust multi-array average framework [41].

2.2 Cohort definition and terminology

GSE69683 contains 498 samples. We parsed ancestry-group and sex variables programmatically from the GEO sample-characteristics fields. A version-controlled sample manifest defined 24 participants with severe asthma: 12 recorded as race: black_african and 12 as race: white_caucasian. Each group comprised eight females and four males. The white_caucasian comparison group was frequency-matched on sex. The public record did not support matching on age, treatment, smoking, exacerbation history or other clinical covariates for this selected subset; residual confounding by these factors therefore remains possible.

The source GEO metadata record the groups as race: black_african and race: white_caucasian. Genetic ancestry was not independently re-estimated in this secondary analysis because genotype data were not available for the selected samples. We use African-ancestry group and European-ancestry group as concise labels for these recorded source categories, while avoiding claims that require direct genetic-ancestry measurement.

2.3 Quality control and exploratory analysis

We checked the selected matrix for missing values, expression range and sample-level distributions. Box plots, density plots and a Pearson sample-correlation matrix were used to identify gross anomalies. These procedures assess data consistency but do not establish absence of batch effects. No array-plate or processing-batch variable was available for explicit modelling. Principal component analysis (PCA) was fitted to the sample-by-probe matrix. t-SNE was computed from PCA scores with perplexity 8, PCA initialisation and random seed 42 and was treated only as exploratory visualisation. Hierarchical clustering used average linkage on correlation distance (1 - Pearson r).

2.4 Differential expression

For each probe, the primary ancestry model was $\text{expression} \sim \text{ancestry group} + \text{sex}$, fitted by ordinary least squares. The ancestry coefficient was coded as African minus European expression with sex held constant. Because the input matrix is already on a log2 scale, the coefficient is reported as an adjusted log2 expression difference; no second logarithmic transformation was applied. The reciprocal model $\text{expression} \sim \text{sex} + \text{ancestry group}$ was used for the sex contrast. P values were adjusted across all 54,715 probes using the Benjamini-Hochberg false-discovery-rate procedure [42]. We summarised results as (i) $\text{FDR} < 0.05$ and (ii) $\text{FDR} < 0.05$ with $|\text{adjusted log2FC}| > 0.5$. Directional counts refer only to the dual-threshold subset. Unadjusted Welch tests were retained as sensitivity analyses.

2.5 Probe-to-gene mapping

Probe identifiers were cross-mapped to GPL570 annotation after removal of the GPL13158 '_PM_' infix. Probes were classified as unambiguous single-gene mappings, ambiguous multi-gene mappings or unmapped probes. Probe-level results remained the primary genome-wide product. For gene-labelled plots, the probe with the smallest ancestry FDR was selected per unambiguous gene. For curated-gene summaries and repository-compatible gene-level outputs, expression was averaged across unambiguous probes per gene, producing 21,653 gene-level profiles. Ambiguous and unmapped probes were retained in supplementary audit files.

2.6 Curated asthma-gene panel

Ten literature-curated gene sets represented Immune Response, Airway Remodelling, Oxidative Stress, Bronchial Hyperresponsiveness, Epithelial Barrier Function, Innate Immunity/TLR Signalling, Type 17/Neutrophilic Inflammation, Hypoxia/Metabolic Response, Steroid Response and Drug Metabolism. The definitions comprised 140 pathway-gene memberships, 129 unique genes and 128 genes measurable in the collapsed matrix. A single tab-separated source generated the corresponding pathway definitions to avoid divergence between formats. These sets define the curated asthma-gene panel; the companion systems analysis applies GSEA and sample-level enrichment frameworks to the same definitions and reports those pathway-level results separately [43–46].

2.7 Curated-gene testing

The 128 measurable unique curated genes were tested with the sex-adjusted ancestry model, followed by one Benjamini-Hochberg correction across the entire curated-gene family. Per-pathway correction was generated as a supplementary comparison. The primary interpretation is based on the single-family correction, so that the reported 13 genes arise from one consistent multiplicity framework.

2.8 Software and reproducibility

Analyses were performed in Python using pandas, NumPy, SciPy, statsmodels, scikit-learn, seaborn, matplotlib and GEOparse. The analysis notebook records the input checksum, output checksums, software environment and random seed, and writes data products, tables and figures to persistent repository folders.

Analysis provenance. All numerical results, figures and summary counts derive from a single fully executed notebook,

Cross_Ancestry_Asthma_Transcriptomics_PUBLICATION_MASTER_EXECUTED_v1.ipynb (execution started 30 July 2026 at 14:27:12 UTC). The run used GSE69683 on GPL13158 and random seed 42; it indexed 72 generated files and passed all programmed validation checks.

3. Results

3.1 Cohort and quality assessment

The analysed cohort contained 24 participants with severe asthma, with 12 participants in each ancestry group and eight females and four males in each group (Table 1). No expression values were missing across 54,715 probe sets. Values ranged from 1.62 to 13.90, with a mean of 5.20 and median of 4.60. Sample distributions were comparable and no gross outlier was identified. Pairwise Pearson correlations were high (mean 0.974; range 0.926–0.988), indicating technical consistency within the selected matrix but not excluding unmeasured batch structure (see Fig. 1).

Table 1
Characteristics of the analysed cohort.

Characteristic	African-ancestry group	European-ancestry group	Total
Participants	12	12	24
Female, n (%)	8 (66.7)	8 (66.7)	16 (66.7)
Male, n (%)	4 (33.3)	4 (33.3)	8 (33.3)
GEO race value	black_african	white_caucasian	Not applicable
Dataset	GSE69683	GSE69683	GSE69683
Platform	GPL13158	GPL13158	54,715 probes
Tissue	Whole blood	Whole blood	Whole blood

3.2 Ancestry group is the primary source of variation in the selected samples

PC1 explained 22.6% of expression variance, PC2 11.1% and PC3 7.6%, giving 41.3% cumulatively across the first three components. Samples separated primarily by ancestry group along PC1, whereas female and male samples were interspersed within both groups. t-SNE using perplexity 8 produced a visually compatible separation, but because of the small sample and the non-inferential nature of t-SNE, it was used only as an exploratory representation (Fig. 2). Correlation-distance clustering likewise grouped samples mainly by ancestry group.

3.3 Genome-wide ancestry-associated expression

The sex-adjusted ancestry model identified 2,134 probes at FDR < 0.05. Of these, 1,020 also exceeded an absolute adjusted log2FC of 0.5: 612 were higher in the African-ancestry group and 408 in the European-ancestry group (Table 2). The unadjusted Welch sensitivity analysis identified 1,445 FDR-significant probes and 751 dual-threshold probes. The larger adjusted set indicates that modelling sex removed residual sex-linked variation, but the analysis cannot exclude confounding by other unavailable clinical or technical covariates.

Table 2
Differential-expression counts for the primary, sensitivity and sex-contrast models.

Contrast and model	Probes tested	FDR < 0.05	FDR < 0.05 and adjusted log2FC > 0.5	Higher in first group	Higher in second group
Ancestry: African vs European, OLS adjusted for sex	54,715	2,134	1,020	612	408
Ancestry: unadjusted Welch sensitivity	54,715	1,445	751	407	344
Sex: male vs female, OLS adjusted for ancestry	54,715	97	73	33	40

ACKR1 (DARC) was among the strongest ancestry-associated genes, with an adjusted African-minus-European log2FC of approximately - 1.69. MMP9 and several other immune-related transcripts were also higher in the European-ancestry group. Given the known relationship between Duffy-null status and neutrophil counts, the prominence of ACKR1 is biologically coherent but is not interpretable as proof of a cell-intrinsic ancestry effect (see Fig. 3).

3.4 Curated asthma genes show bidirectional differences

Thirteen of 128 measurable unique curated genes were significant after one correction family (Table 3). Nine showed higher expression in the European-ancestry group: MMP9, CXCL1, CXCL8, ALOX5, ARRB2, TIMP2, TGFB1, LTC4S and CLDN18. Four were higher in the African-ancestry group: GATA3, LDHA, CD28 and PRDX1. The opposing directions across immune, remodelling, metabolic and treatment-relevant genes argue against a simple model in which one group has uniformly greater inflammatory expression (Fig. 4).

Table 3. Curated asthma genes significantly associated with ancestry group after sex adjustment and one Benjamini-Hochberg correction across 128 genes.

Gene	Pathway	Mean African	Mean European	Adjusted log2FC	FDR	Higher group
ALOX5	BHR	7.089	7.695	-0.606	0.00045	European
ARRB2	Drug metabolism	9.938	10.389	-0.452	0.00160	European
MMP9	Airway remodelling	7.077	8.809	-1.733	0.00160	European
CD28	Immune response	5.165	4.858	+0.308	0.00304	African
TIMP2	Airway remodelling	9.863	10.233	-0.371	0.00338	European
LDHA	Hypoxia/metabolic	11.051	10.698	+0.353	0.00884	African
CXCL1	Type 17/neutrophilic	6.299	7.278	-0.979	0.00884	European
CLDN18	Epithelial barrier	3.107	3.255	-0.148	0.00884	European
TGFB1	Airway remodelling	6.167	6.568	-0.401	0.01012	European
PRDX1	Oxidative stress	10.179	9.887	+0.293	0.02023	African
GATA3	Immune response	7.276	6.738	+0.538	0.02603	African
LTC4S	BHR	3.976	4.222	-0.246	0.03599	European
CXCL8	Innate/TLR	5.138	5.994	-0.856	0.04217	European

Abbreviations: BHR, bronchial hyperresponsiveness; TLR, Toll-like receptor.

3.5 Sex contrast as an internal validation analysis

The ancestry-adjusted sex contrast yielded 97 probes at FDR < 0.05 and 73 at the dual threshold, with 33 higher in males and 40 higher in females. The strongest signals were dominated by expected X- and Y-linked transcripts, including XIST, RPS4Y1, DDX3Y, KDM5D and UTY (see Figure 5). No curated asthma gene was significant after correction, indicating that the ancestry-associated curated-gene findings were not explained by the deliberately balanced sex distribution in this cohort.

4. Discussion

The main finding is not a unidirectional inflammatory difference between the two ancestry groups. Instead, the data show a bidirectional whole-blood pattern in which airway-remodelling, chemokine and leukotriene-related transcripts were higher in the European-ancestry group, while GATA3, CD28, LDHA and PRDX1 were higher in the African-ancestry group. The opposing directions across the 13 curated genes are more informative than any single transcript in isolation, but the whole-blood design means that both cell abundance and cell-intrinsic regulation remain plausible contributors.

4.1 Cell composition constrains interpretation

The ACKR1 result provides an internal biological consistency check and, at the same time, a reason for caution. The Duffy-null regulatory variant is strongly associated with lower circulating neutrophil counts in populations in which it is common [25-29]. Several transcripts that were higher in the European-ancestry group, including MMP9, CXCL1, CXCL8 and ALOX5, are expressed prominently in neutrophils or other myeloid compartments. A lower bulk-blood signal in the African-ancestry group could therefore arise from fewer circulating neutrophils, altered transcription within cells, treatment effects, or a combination of these processes. Computational deconvolution can estimate cell fractions from bulk expression data [47], but no validated deconvolution result or differential blood count was available for these 24 samples. We therefore do not interpret the ACKR1-linked pattern as evidence that ancestry itself directly regulates the observed genes.

4.2 Remodelling- and epithelial-barrier-related genes

Several of the curated genes that differed between groups are directly relevant to airway structural biology. MMP9, TGFB1 and TIMP2 were all higher in the European-ancestry group. TGF-beta1 coordinates airway smooth-muscle, fibroblast and extracellular-matrix responses relevant to structural change [48], while MMP9 has been linked to extracellular-matrix turnover and subepithelial fibrosis in asthma [49,50]. Airway remodelling is a composite process involving epithelium, smooth muscle, matrix deposition and inflammatory injury rather than a single molecular programme [51]. CLDN18 was also higher in the European-ancestry group, providing a gene-level connection to epithelial-barrier biology, where junctional integrity and barrier disruption shape both innate and adaptive airway responses [52,53]. Because the present samples were whole blood rather than airway tissue, these transcript differences should be interpreted as systemic associations with remodelling- and barrier-related genes rather than direct evidence of structural airway change.

4.3 Oxidative and metabolic findings add a second direction to the signal

PRDX1 and LDHA were higher in the African-ancestry group even though several neutrophil- and remodelling-associated genes were higher in the European-ancestry group. Oxidative stress and antioxidant imbalance are well described in asthma [54,55], making PRDX1 biologically plausible as part of an antioxidant response. LDHA is also increasingly linked to immunometabolic control of inflammatory cells; recent experimental work has implicated LDHA-dependent glycolysis in neutrophilic asthma mechanisms [56]. These associations should not be converted into a metabolic endotype from this dataset. Rather, they show that the ancestry-associated profile contains opposing biological directions that would be missed by describing one group as globally more inflammatory.

4.4 Treatment-related expression warrants mechanistic follow-up, not clinical inference

ALOX5 and LTC4S were higher in the European-ancestry group, and ALOX5 has long-standing pharmacogenetic relevance to leukotriene-modifier response [30,31]. ARRB2 was also higher in the European-ancestry group. Beta-arrestin 2 regulates allergic airway responses and beta2-adrenergic receptor desensitisation in experimental systems [33,34], while ADRB2 and glucocorticoid-response pharmacogenetics demonstrate that treatment response can vary with genotype [32,35]. This GEO subset does not provide individual bronchodilator, corticosteroid or leukotriene-modifier response measures. These genes therefore identify treatment-relevant candidates for follow-up in cohorts where genotype, exposure and longitudinal response are measured.

4.5 Sex adjustment behaved as an internal validation marker rather than a substantive sex analysis

The reciprocal model recovered XIST, RPS4Y1, DDX3Y, KDM5D and UTY, a pattern consistent with established sex-biased peripheral-blood expression [57,58]. Sex differences in asthma prevalence, inflammatory biology and treatment response are clinically important and interact with hormonal state across the life course [59-62]. In this selected cohort, however, there were only eight males and the design was deliberately balanced on sex within ancestry group. The absence of significant curated asthma genes in the sex contrast should therefore not be interpreted as evidence that sex is unimportant in severe asthma; it simply reduces concern that the reported ancestry-group differences are a consequence of gross sex imbalance.

4.6 Replication should separate ancestry, cell composition and clinical state

Severe asthma can contain neutrophilic and inflammasome-related biology that is not captured by a simple Type 2-high/low distinction [63]. Recent whole-blood transcriptomic work in non-Hispanic Black participants has also identified state-dependent expression changes across asthma establishment and exacerbation [64], illustrating both the value of ancestry-diverse blood datasets and the importance of clinical context. A convincing next step would therefore test these 13 curated genes and the remodelling- and barrier-related gene pattern in an independent ancestry-diverse cohort with independent ancestry characterisation where available, Duffy status or complete blood counts, detailed treatment and exacerbation data, and preferably paired airway or cell-resolved transcriptomics. Such replication could determine which signals persist after cell-composition adjustment and whether the treatment-related genes predict actual therapeutic response.

4.7 Strengths and limitations

The study combines programmatic GEO metadata validation, a version-controlled sample manifest, explicit GPL13158 platform checking, a single pre-specified sex-adjusted model, transparent probe-mapping audits and checksum-backed outputs. Its main constraint is the small, selected 24-participant subset, which limits precision and leaves age, treatment, smoking, exacerbation history and other clinical factors unavailable for adjustment; these factors may contribute to residual confounding. The grouping follows the race/ethnicity categories recorded in GEO and was not independently re-estimated genetically in this secondary analysis, while whole blood is a heterogeneous cell mixture without differential counts or validated deconvolution, so cell-composition effects cannot be separated from cell-intrinsic transcription. The absence of an independent replication cohort and a batch covariate further limits causal, mechanistic and treatment-response interpretation.

4.8 Future directions

Independent replication should prioritise ancestry-diverse cohorts with independent ancestry characterisation where available, Duffy status, complete clinical covariates and differential blood counts or validated transcriptomic deconvolution. Paired airway tissue, single-cell or cell-sorted transcriptomics would help distinguish systemic compositional signals from airway-specific or cell-intrinsic regulation. Treatment-response analyses should then test whether *ARRB2*, *ALOX5*, *LTC4S* and related transcript patterns relate prospectively to bronchodilator, leukotriene-modifier or corticosteroid outcomes. These steps are necessary before expression differences observed in this small cross-sectional dataset can inform ancestry-aware treatment hypotheses.

5. Conclusion

The recorded ancestry grouping was the primary source of transcriptomic variation in this selected severe-asthma whole-blood cohort, but the associated expression pattern was bidirectional rather than uniformly inflammatory. The 13-gene curated panel included opposing immune, remodelling, epithelial-barrier, oxidative/metabolic and treatment-relevant differences, while the strong *ACKR1* signal indicates that circulating-cell composition probably contributes materially to the bulk-blood pattern. Residual confounding from unavailable clinical variables also cannot be excluded. The present data therefore support a focused gene-level set of findings for independent replication without establishing ancestry-specific endotypes or treatment responses.

Declarations

Acknowledgements. We thank the U-BIOPRED participants, study investigators and data-generating teams, and the original authors for making the GSE69683 dataset publicly available.

Funding. This study did not receive specific funding. The analyses presented here were conducted using the publicly available GSE69683 dataset from the U-BIOPRED study. The original U-BIOPRED project was supported by its own funding sources, as acknowledged in the primary publications. No additional financial support was obtained for this secondary analysis.

Competing interests. The authors declare no competing interests.

Author contributions. W.M: Investigation, Data curation, Formal analysis, Visualisation, Writing - original draft, Writing - review and editing. A.O.A: Conceptualisation, Methodology, Supervision, Project administration, Formal analysis, Interpretation, Writing - original draft, Writing - review and editing. K.D: Validation, Literature curation, Visualisation, Writing - review and editing. L.J: Validation, Reproducibility checks, Writing - review and editing. A.A: Methodology, Software, Validation, Formal analysis, Writing - review and editing. W.M and A.O.A contributed equally to this work and share first authorship. All authors read the final manuscript and approved it before submission.

Ethical approval. This secondary analysis used publicly available, de-identified data and did not involve new participant recruitment, intervention or access to identifiable information. No additional ethical approval was required for this secondary analysis. The original U-BIOPRED study obtained ethical approval at participating sites.

Informed consent. The original U-BIOPRED study obtained informed consent from participants according to its approved protocols. No new consent was required for this analysis of publicly available de-identified data.

AI use disclosure. Generative AI tools were used during manuscript preparation for language editing, grammar refinement and structural suggestions. All analyses, numerical results, and scientific interpretations were performed and verified by the authors, and the authors remain responsible for the scientific interpretation and final wording. AI was not used to generate or alter primary data or statistical outputs.

Data availability. All analysed expression data are publicly available from NCBI Gene Expression Omnibus under accession GSE69683 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE69683>). The version-controlled analysis code, sample manifest, curated pathway definitions, machine-readable result tables, figures, validation reports and checksums are maintained in the project repository. The public GitHub repository is available at <https://github.com/UREKA-Biotec-OmicsLogic-Africa/cross-ancestry-asthma-transcriptomics-public>, and a version-specific Zenodo DOI is available at <https://doi.org/10.5281/zenodo.23227701>.

References

1. Akinbami OJ. Trends in asthma prevalence, health care use, and mortality in the United States, 2001–2010. US Department of Health and Human Services, Centers for Disease Control and

- Prevention, National Center for Health Statistics; 2012.
2. Celedón JC, Roman J, Schraufnagel DE, Thomas A, Samet J. Respiratory health equality in the United States. The American thoracic society perspective. *Annals Am Thorac Soc*. 2014;11(4):473–9. 10.1513/AnnalsATS.201402-059PS.
 3. Ford JG, McCaffrey L. Understanding disparities in asthma outcomes among African Americans. *Clinics in chest medicine*. 2006;27(3):423–30. 10.1016/j.ccm.2006.05.001
 4. Oh SS, Galanter J, Thakur N, Pino-Yanes M, Barcelo NE, White MJ, et al. Diversity in clinical and biomedical research: a promise yet to be fulfilled. *PLoS Med*. 2015;12(12):e1001918. 10.1371/journal.pmed.1001918.
 5. Borrell LN, Elhawary JR, Fuentes-Afflick E, Witonsky J, Bhakta N, Wu AH, et al. Race and genetic ancestry in medicine—a time for reckoning with racism. *Mass Medical Soc*; 2021. pp. 474–80. 10.1056/NEJMms2029562.
 6. Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med*. 2012;18(5):716–25. 10.1038/nm.2678.
 7. Lambrecht BN, Hammad H. The immunology of asthma. *Nat Immunol*. 2015;16(1):45–56. 10.1038/ni.3049.
 8. Fahy JV. Type 2 inflammation in asthma—present in most, absent in many. *Nat Rev Immunol*. 2015;15(1):57–65. 10.1038/nri3786.
 9. Woodruff PG, Modrek B, Choy DF, Jia G, Abbas AR, Ellwanger A, et al. T-helper type 2–driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med*. 2009;180(5):388–95. 10.1164/rccm.200903-0392OC.
 10. Hinks TS, Levine SJ, Brusselle GG. Treatment options in type-2 low asthma. *Eur Respir J*. 2021;57(1):2000528. 10.1183/13993003.00528-2020.
 11. Israel E, Reddel HK. Severe and difficult-to-treat asthma in adults. *N Engl J Med*. 2017;377(10):965–76. 10.1056/NEJMra1608969.
 12. Mathias RA, Grant AV, Rafaels N, Hand T, Gao L, Vergara C, et al. A genome-wide association study on African-ancestry populations for asthma. *J allergy Clin Immunol*. 2010;125(2):336–46. 10.1016/j.jaci.2009.08.031. e4.
 13. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet*. 2011;43(9):887–92. 10.1038/ng.888.
 14. Flores C, Ma S-F, Pino-Yanes M, Wade MS, Pérez-Méndez L, Kittles RA, et al. African ancestry is associated with asthma risk in African Americans. *PLoS ONE*. 2012;7(1):e26807. 10.1371/journal.pone.0026807.
 15. Vergara C, Murray T, Rafaels N, Lewis R, Campbell M, Foster C, et al. African ancestry is a risk factor for asthma and high total IgE levels in African admixed populations. *Genet Epidemiol*. 2013;37(4):393–401. 10.1002/gepi.21702.

16. Rumpel JA, Ahmedani BK, Peterson EL, Wells KE, Yang M, Levin AM, et al. Genetic ancestry and its association with asthma exacerbations among African American subjects with asthma. *J allergy Clin Immunol*. 2012;130(6):1302–6. 10.1016/j.jaci.2012.09.001.
17. Chang X, March M, Mentch F, Qu H, Liu Y, Glessner J, et al. Genetic architecture of asthma in African American patients. *J Allergy Clin Immunol*. 2023;151(4):1132–6. 10.1016/j.jaci.2022.09.001.
18. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*. 2016;538(7624):161–4. 10.1038/538161a.
19. Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. *Cell*. 2019;177(1):26–31. 10.1016/j.cell.2019.02.048.
20. Peterson RE, Kuchenbaecker K, Walters RK, Chen C-Y, Popejoy AB, Periyasamy S, et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. *Cell*. 2019;179(3):589–603. 10.1016/j.cell.2019.08.051.
21. Mersha TB, Abebe T. Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. *Hum Genomics*. 2015;9(1):1. 10.1186/s40246-014-0023-x.
22. Khan AT, Gogarten SM, McHugh CP, Stilp AM, Sofer T, Bowers ML, et al. Recommendations on the use and reporting of race, ethnicity, and ancestry in genetic research: Experiences from the NHLBI TOPMed program. *Cell genomics*. 2022;2(8):100155. 10.1016/j.xgen.2022.100155.
23. Gamazon ER, Segrè AV, Van De Bunt M, Wen X, Xi HS, Hormozdiari F, et al. Using an atlas of gene regulation across 44 human tissues to inform complex disease-and trait-associated variation. *Nat Genet*. 2018;50(7):956–67. 10.1038/s41588-018-0154-4.
24. Consortium G. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*. 2020;369(6509):1318–30. 10.1126/science.aaz1776.
25. Reich D, Nalls MA, Kao WL, Akylbekova EL, Tandon A, Patterson N, et al. Reduced neutrophil count in people of African descent is due to a regulatory variant in the Duffy antigen receptor for chemokines gene. *PLoS Genet*. 2009;5(1):e1000360. 10.1371/journal.pgen.1000360.
26. Merz LE, Story CM, Osei MA, Jolley K, Ren S, Park HS, et al. Absolute neutrophil count by Duffy status among healthy Black and African American adults. *Blood Adv*. 2023;7(3):317–20. 10.1182/bloodadvances.2022007679.
27. Duchene J, Novitzky-Basso I, Thiriot A, Casanova-Acebes M, Bianchini M, Etheridge SL, et al. Atypical chemokine receptor 1 on nucleated erythroid cells regulates hematopoiesis. *Nat Immunol*. 2017;18(7):753–61. 10.1038/ni.3763.
28. Van Driest SL, Abul-Husn NS, Glessner JT, Bastarache L, Nirenberg S, Schildcrout JS, et al. Association between a common, benign genotype and unnecessary bone marrow biopsies among African American patients. *JAMA Intern Med*. 2021;181(8):1100–5. 10.1001/jamainternmed.2021.3108.
29. Rappoport N, Simon AJ, Amariglio N, Rechavi G. The Duffy antigen receptor for chemokines, ACKR 1, –‘Jeanne DARC’ of benign neutropenia. *Br J Haematol*. 2019;184(4):497–507. 10.1111/bjh.15730.

30. Drazen JM, Yandava CN, Dubé L, Szczerback N, Hippensteel R, Pillari A, et al. Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment. *Nat Genet.* 1999;22(2):168–70. 10.1038/9680.
31. Lima JJ, Zhang S, Grant A, Shao L, Tantisira KG, Allayee H, et al. Influence of leukotriene pathway polymorphisms on response to montelukast in asthma. *Am J Respir Crit Care Med.* 2006;173(4):379–85. 10.1164/rccm.200509-1412OC.
32. Bleecker ER, Meyers DA, Bailey WC, Sims A-M, Bujac SR, Goldman M, et al. ADRB2 polymorphisms and budesonide/formoterol responses in COPD. *Chest.* 2012;142(2):320–8. 10.1378/chest.11-1655.
33. Walker JK, Fong AM, Lawson BL, Savov JD, Patel DD, Schwartz DA, et al. β -Arrestin-2 regulates the development of allergic asthma. *J Clin Investig.* 2003;112(4):566–74. 10.1172/JCI17265.
34. Deshpande DA, Theriot BS, Penn RB, Walker JK. β -Arrestins specifically constrain β 2-adrenergic receptor signaling and function in airway smooth muscle. *FASEB journal: official publication Federation Am Soc Experimental Biology.* 2008;22(7):2134. 10.1096/fj.07-102459.
35. Tantisira KG, Lasky-Su J, Harada M, Murphy A, Litonjua AA, Himes BE, et al. Genomewide association between GLCCI1 and response to glucocorticoid therapy in asthma. *N Engl J Med.* 2011;365(13):1173–83. 10.1056/NEJMoa0911353.
36. Shaw DE, Sousa AR, Fowler SJ, Fleming LJ, Roberts G, Corfield J, et al. Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. *Eur Respir J.* 2015;46(5):1308–21. 10.1183/13993003.00779-2015.
37. Bigler J, Boedigheimer M, Schofield JP, Skipp PJ, Corfield J, Rowe A, et al. A severe asthma disease signature from gene expression profiling of peripheral blood from U-BIOPRED cohorts. *Am J Respir Crit Care Med.* 2017;195(10):1311–20. 10.1164/rccm.201604-0866OC.
38. Wilson SJ, Ward JA, Sousa AR, Corfield J, Bansal AT, De Meulder B, et al. Severe asthma exists despite suppressed tissue inflammation: findings of the U-BIOPRED study. *Eur Respir J.* 2016;48(5):1307–19. 10.1183/13993003.01129-2016.
39. Kuo C-HS, Pavlidis S, Loza M, Baribaud F, Rowe A, Pandis I, et al. A transcriptome-driven analysis of epithelial brushings and bronchial biopsies to define asthma phenotypes in U-BIOPRED. *Am J Respir Crit Care Med.* 2017;195(4):443–55. 10.1164/rccm.201512-2452OC.
40. Hoda U, Pavlidis S, Bansal AT, Takahashi K, Hu S, Ng Kee Kwong F, et al. Clinical and transcriptomic features of persistent exacerbation-prone severe asthma in U-BIOPRED cohort. *Clin translational Med.* 2022;12(4):e816. 10.1002/ctm2.816.
41. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics.* 2003;4(2):249–64. 10.1093/biostatistics/4.2.249.
42. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc: Ser B (Methodol).* 1995;57(1):289–300. 10.1111/j.2517-6161.1995.tb02031.x.

43. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the national academy of sciences*. 2005;102(43):15545–50. 10.1073/pnas.0506580102
44. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*. 2009;462(7269):108–12. 10.1038/nature08460.
45. Hänzelmann S, Castelo R, Guinney J. *BMC Bioinformatics*. 2013;14(1):7. 10.1186/1471-2105-14-7. GSVA: gene set variation analysis for microarray and RNA-seq data.
46. Reimand R, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g: Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc*. 2019;14(2):482–517. 10.1038/s41596-018-0103-9.
47. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–7. 10.1038/nmeth.3337.
48. Ojiaku CA, Yoo EJ, Panettieri RA Jr. Transforming growth factor β 1 function in airway remodeling and hyperresponsiveness. The missing link? *Am J Respir Cell Mol Biol*. 2017;56(4):432–42. 10.1165/rcmb.2016-0307TR.
49. Hoshino M, Nakamura Y, Sim J, Shimojo J, Isogai S. Bronchial subepithelial fibrosis and expression of matrix metalloproteinase-9 in asthmatic airway inflammation. *J allergy Clin Immunol*. 1998;102(5):783–8. 10.1016/s0091-6749(98)70018-1.
50. Ohbayashi H, Shimokata K. Matrix metalloproteinase-9 and airway remodeling in asthma. *Curr Drug Targets-Inflammation Allergy*. 2005;4(2):177–81. 10.2174/1568010053586246.
51. Fehrenbach H, Wagner C, Wegmann M. Airway remodeling in asthma: what really matters. *Cell Tissue Res*. 2017;367(3):551–69. 10.1007/s00441-016-2566-8.
52. Georas SN, Rezaee F. Epithelial barrier function: at the front line of asthma immunology and allergic airway inflammation. *J allergy Clin Immunol*. 2014;134(3):509–20. 10.1016/j.jaci.2014.05.049.
53. Heijink I, Nawijn M, Hackett TL. Airway epithelial barrier function regulates the pathogenesis of allergic asthma. *Clin Experimental Allergy*. 2014;44(5):620–30. 10.1111/cea.12296.
54. Comhair SA, Erzurum SC. Redox control of asthma: molecular mechanisms and therapeutic opportunities. *Antioxid Redox Signal*. 2010;12(1):93–124. 10.1089/ars.2008.2425.
55. Sahiner UM, Birben E, Erzurum S, Sackesen C, Kalayci Ö. Oxidative stress in asthma: Part of the puzzle. *Pediatr Allergy Immunol*. 2018;29(8):789–800. 10.1111/pai.12965.
56. Su G, Peng J, Quan J, Feng Z, Zhong Y, Xiong Z, et al. SIRT6-mediated immunometabolic reprogramming of macrophages drives neutrophilic asthma via LDHA-dependent glycolysis. *Cell Rep*. 2026;45(5):117294. 10.1016/j.celrep.2026.117294.
57. Jansen R, Batista S, Brooks AI, Tischfield JA, Willemsen G, Van Grootheest G, et al. Sex differences in the human peripheral blood transcriptome. *BMC Genomics*. 2014;15(1):33. 10.1186/1471-2164-15-33.

58. Staedtler F, Hartmann N, Letzkus M, Bongiovanni S, Scherer A, Marc P, et al. Robust and tissue-independent gender-specific transcript biomarkers. *Biomarkers*. 2013;18(5):436–45. 10.3109/1354750X.2013.811538.
59. Yung JA, Fuseini H, Newcomb DC. Hormones, sex, and asthma. *Annals of Allergy. Asthma Immunol*. 2018;120(5):488–94. 10.1016/j.anai.2018.01.016.
60. Fuseini H, Newcomb DC. Mechanisms driving gender differences in asthma. *Curr Allergy Asthma Rep*. 2017;17(3):19. 10.1007/s11882-017-0686-1.
61. Chowdhury NU, Guntur VP, Newcomb DC, Wechsler ME. Sex and gender in asthma. *Eur Respiratory Rev*. 2021;30(162):210067. 10.1183/16000617.0067-2021.
62. Chiarella SE, Cardet JC, Prakash Y, editors. Sex, cells, and asthma. *Mayo Clinic Proceedings*; 2021: Elsevier. 10.1016/j.mayocp.2020.12.007
63. Lachowicz-Scroggins ME, Dunican EM, Charbit AR, Raymond W, Looney MR, Peters MC, et al. Extracellular DNA, neutrophil extracellular traps, and inflammasome activation in severe asthma. *Am J Respir Crit Care Med*. 2019;199(9):1076–85. 10.1164/rccm.201810-1869OC.
64. Wu B, Yang M, Hu D, Hochstadt S, Lee BH, Lanfear DE, et al. Distinct blood transcriptomic profiling underlying asthma establishment and exacerbation. *J Allergy Clin Immunology: Global*. 2026;100644. 10.1016/j.jacig.2026.100644.

Figures

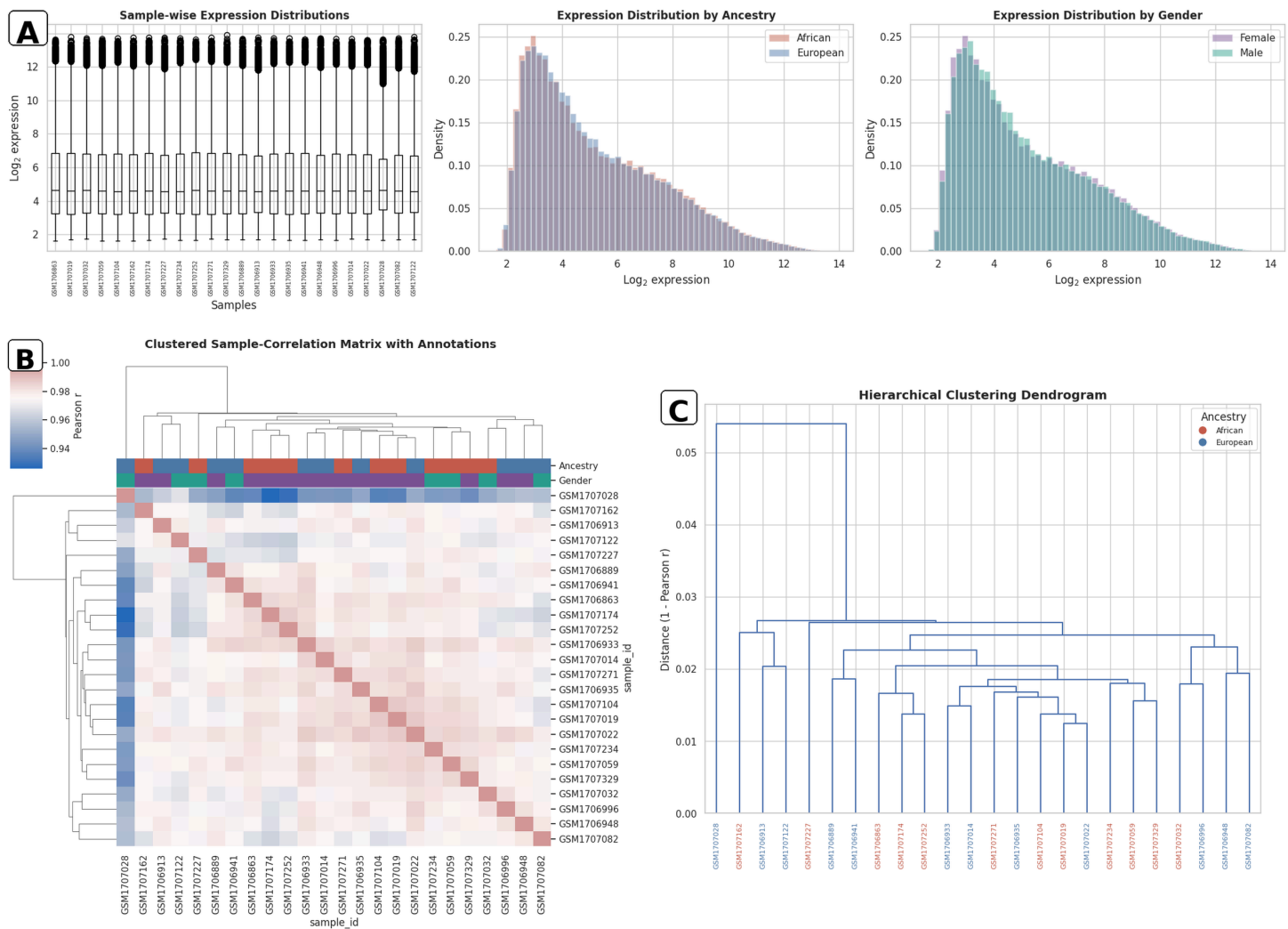


Figure 1

Sample quality and clustering. (A) Sample-wise expression distributions across 24 samples. (B) Clustered Pearson correlation matrix with ancestry-group and sex annotation bars. (C) Hierarchical clustering on correlation distance. These quality-control displays identify gross anomalies but do not establish absence of batch effects.

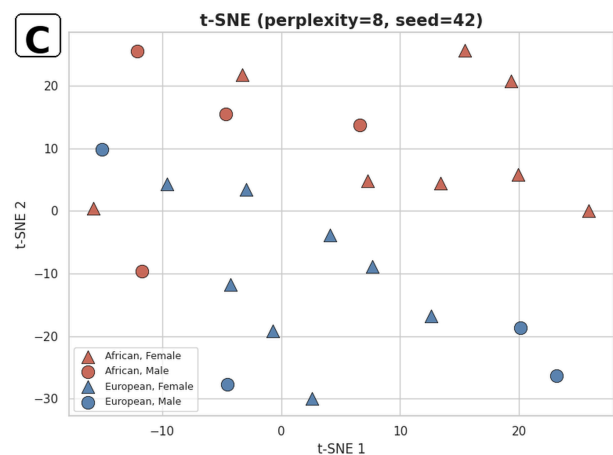
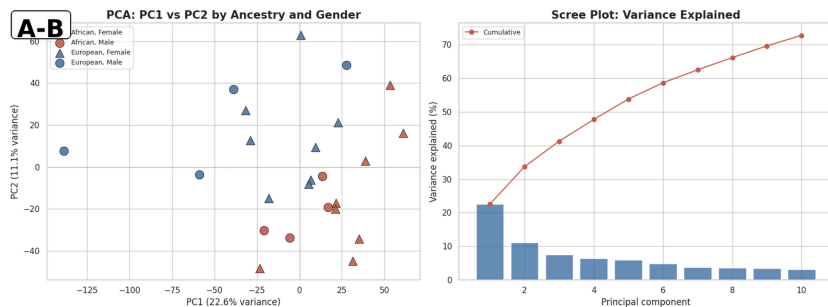


Figure 2

Exploratory transcriptomic structure. (A) PCA of the sample-by-probe matrix, coloured by ancestry group and shaped by sex. (B) Scree and cumulative-variance plot. (C) t-SNE of PCA scores using perplexity 8, PCA initialisation and seed 42. t-SNE is shown as visualisation only.

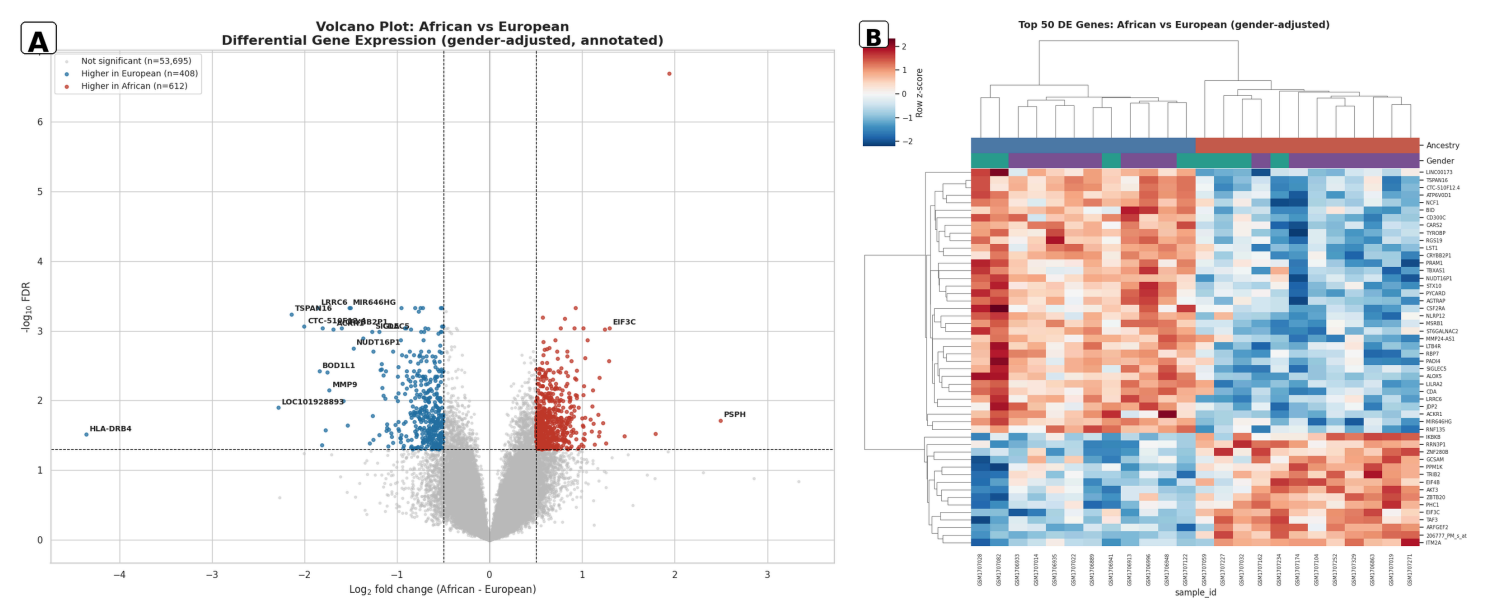


Figure 3

Genome-wide ancestry-associated expression. (A) Volcano plot of the sex-adjusted ancestry coefficient against $-\log_{10}$ FDR for 54,715 probes; dual-threshold probes are coloured by direction. (B) Row-standardised heatmap of the top 50 dual-threshold probes.

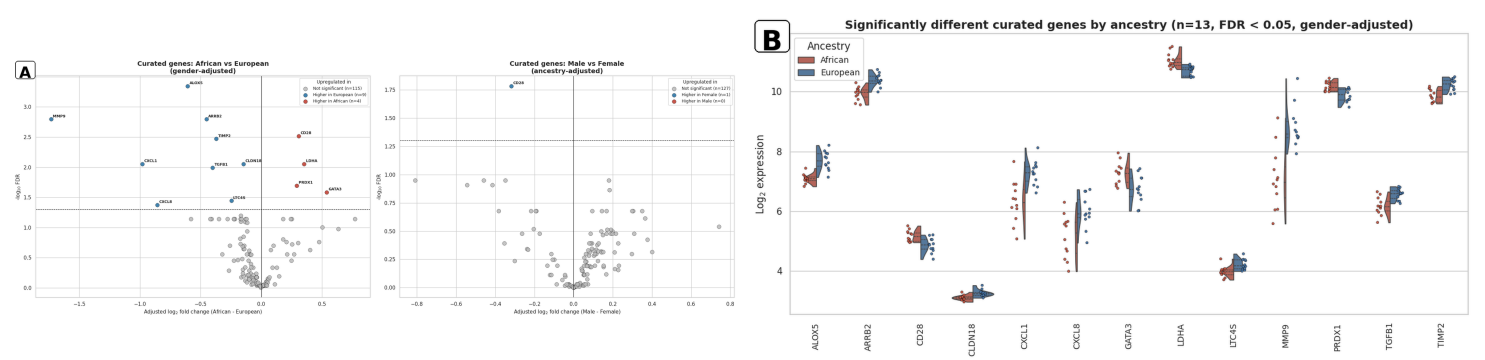


Figure 4

Curated asthma genes. (A) Volcano plot of 128 measurable curated genes after one Benjamini-Hochberg correction family. (B) Split violin plots for the 13 significant genes, with ancestry groups displayed using the pre-specified accessible palette.

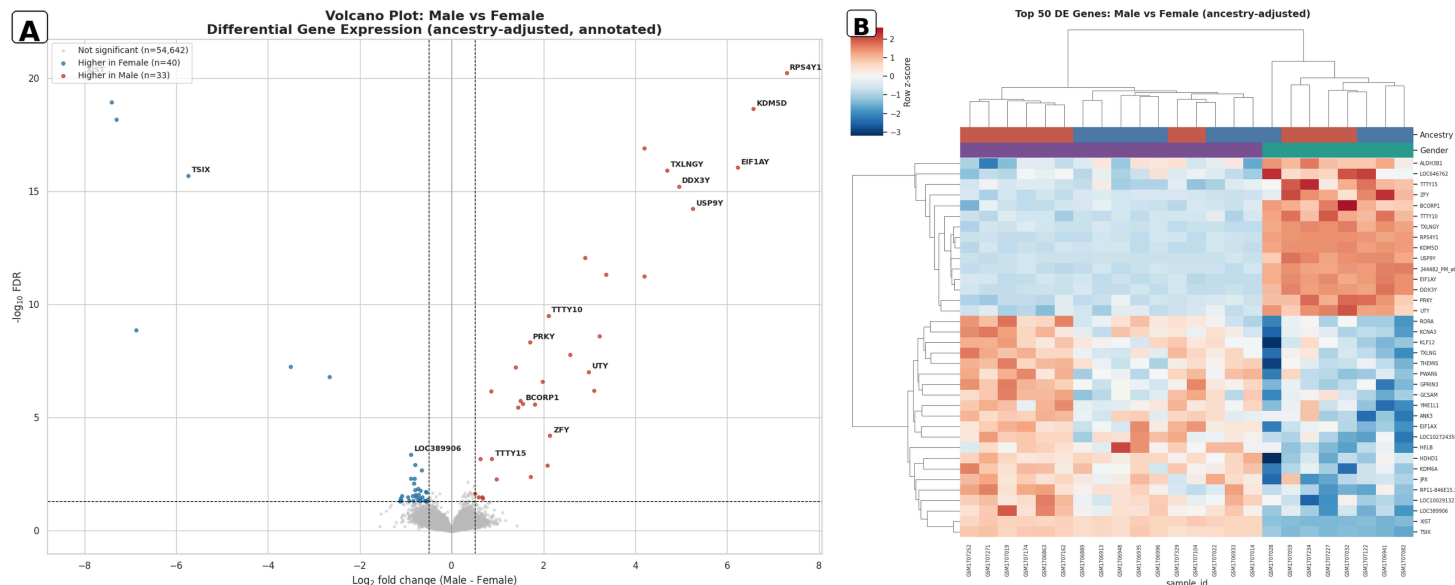


Figure 5

Sex contrast adjusted for ancestry group. (A) Volcano plot of male-minus-female expression. (B) Heatmap of the top sex-associated probes. The signal is dominated by expected sex-linked transcripts; no curated asthma gene is significant after multiple-testing correction.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.zip](#)
- [Additionalfile2.zip](#)
- [Additionalfile3.png](#)