International Journal of Chronic Obstructive Pulmonary Disease

Dovepress
Taylor & Francis Group

Open Access Full Text Article

REVIEW

Genetic and Epigenetic Association of Lung Function and Chronic Obstructive Pulmonary Disease-Related Phenotypes in the Korean Genome and Epidemiology Study: A Scoping Review

Sun Young Kang [1,*], Sanghyun Roh [2,*], Jeong-An Gim [1,2]

[1]Institute for Molecular Metabolism Innovation, Soonchunhyang University, Asan, Republic of Korea; [2]Department of Medical Science, Soonchunhyang University, Asan, Republic of Korea

*These authors contributed equally to this work

Correspondence: Jeong-An Gim, Email vitastar@sch.ac.kr

Background: Chronic obstructive pulmonary disease (COPD) is influenced by environmental exposure and genetic susceptibility. Although large-scale genetic studies have identified loci associated with COPD and lung function, most evidence has come from populations of European ancestry, and data from Asian populations remain limited.

Objective: This scoping review aimed to summarize KoGES-based studies on genetic and epigenetic associations with COPD and COPD-related lung function phenotypes.

Methods: PubMed, Scopus, Web of Science, KoreaMed, RISS, KISS, and DBpia were searched from inception to April 27, 2026. Original studies using KoGES data to investigate genetic or epigenetic associations with COPD, airflow limitation, lung function traits, or longitudinal lung function decline were included. The review followed the Arksey and O'Malley framework and PRISMA-ScR guidance.

Results: Of 20 unique records identified for final assessment, one conference abstract/preliminary report corresponding to a subsequently published full-length article was removed before screening. The remaining 19 records were screened, and 12 studies met the eligibility criteria. Most KoGES-based studies focused on spirometric traits, including FEV1, FEV1/FVC, and longitudinal lung function decline, rather than clinically defined COPD. Major approaches included genome-wide association studies (GWAS), exome array analysis, gene–environment interaction, polygenic/genetic risk score (PRS/GRS) analysis, epigenome-wide association studies (EWAS), and heritability analysis. Recurrent signals involved FAM13A and the 6p21/AGER region. Recent studies increasingly addressed longitudinal phenotypes, including rapid FEV1 decline and annual lung function decline.

Conclusion: KoGES-based studies provide population-specific evidence on genetic and epigenetic associations with COPD-related lung function phenotypes in Koreans. However, evidence for clinically defined COPD, externally validated risk prediction, and integrative multi-omics remains limited. Future studies should distinguish spirometric traits, longitudinal decline, and clinically defined COPD while validating KoGES findings in larger Asian and multi-ethnic populations.

Plain Language Summary: Why was this review done? COPD develops through a combination of smoking, environmental exposure, and genetic susceptibility. Most large genetic studies on COPD have been conducted in European populations, so less is known about genetic factors in Koreans. We reviewed studies that used data from the Korean Genome and Epidemiology Study (KoGES) to understand what is already known and what research gaps remain.

What did the researchers do? We searched major Korean and international databases for studies that used KoGES data to investigate genetic or epigenetic factors related to COPD or lung function. We then grouped the studies by research type, such as genome-wide association studies, exome-based studies, polygenic risk score studies, interaction studies, and methylation studies.

https://doi.org/10.2147/COPD.S611602
Received: 24 March 2026
Accepted: 18 May 2026
Published: 27 May 2026

What did the review find? Most KoGES-based studies focused on lung function measurements such as FEV1 and FEV1/FVC rather than clinically diagnosed COPD alone. Repeated evidence was found for several loci, especially FAM13A and the 6p21 region. More recent studies have moved from cross-sectional analyses toward longitudinal outcomes, such as rapid lung function decline and annual change in lung function. Some studies have also begun to evaluate polygenic risk and its interaction with modifiable lifestyle factors.

What do these findings mean? KoGES has helped improve understanding of genetic influences on lung function in Koreans, but important gaps remain. There are still relatively few studies on clinically defined COPD, externally validated genetic risk prediction, and multi-omics integration. More longitudinal and population-specific studies are needed to support future research on early identification, external validation, and prevention-related research questions.

Keywords: COPD, lung function, KoGES, genetics, polygenic risk score, genome-wide association study, DNA methylation

Introduction

Chronic obstructive pulmonary disease (COPD) is a complex and progressive respiratory disorder characterized by persistent airflow limitation associated with chronic airway and/or alveolar abnormalities. According to the Global Initiative for Chronic Obstructive Lung Disease (GOLD), COPD is defined spirometrically by a post-bronchodilator ratio of forced expiratory volume in 1 second (FEV1) to forced vital capacity (FVC) of less than 0.70.[1] COPD remains one of the leading causes of morbidity and mortality worldwide, and its global burden continues to rise. A recent global analysis estimated that the prevalence of COPD among adults aged 30–79 years was 10.3% in 2019, highlighting its major public health impact.[2]

Although cigarette smoking is the most important risk factor for COPD, the disease is influenced by a broad range of environmental and lifestyle-related exposures. In addition to smoking, air pollution, indoor and outdoor pollutant exposure, occupational dust and fumes, respiratory infections, insufficient physical activity, and poor dietary habits have all been associated with COPD incidence and progression.[3–5] The observation that only a proportion of smokers develop COPD suggests that host susceptibility, including genetic predisposition, substantially modifies disease risk.

In recent years, genetic factors have been increasingly recognized as important contributors to COPD pathogenesis. Candidate gene studies and genome-wide association studies (GWAS) have identified multiple genes associated with COPD susceptibility and lung function, including SERPINA1, CHRNA3/5, FAM13A, HHIP, and AGER.[6–9] Variants in CHRNA3/5 and FAM13A, in particular, have shown repeated associations with COPD risk and lung function impairment in previous studies.[8–11] Heritability studies also support a meaningful inherited component of lung function and COPD-related traits, suggesting that genetic susceptibility contributes to inter-individual differences in COPD development and progression.[12]

However, most large-scale genetic studies on COPD have been conducted in populations of European ancestry. As a result, evidence regarding the genetic architecture of COPD in Asian populations remains relatively limited. This is an important gap because allele frequencies, linkage disequilibrium patterns, and gene–environment interactions may differ across ethnic groups. Therefore, population-specific studies are needed to better understand the genetic associations of COPD and lung function in Koreans.

In this context, the Korean Genome and Epidemiology Study (KoGES), a large population-based cohort established to investigate genetic, environmental, and behavioral determinants of chronic diseases in Koreans, provides a valuable platform for COPD-related genetic research. KoGES is particularly suited for this review because it integrates repeated spirometric measurements, genome-wide genotype data, and longitudinal follow-up within a single cohort framework, enabling assessment of both baseline lung function and temporal decline. At the same time, this KoGES-specific focus may limit the generalizability of the findings to broader Korean, Asian, or multi-ethnic populations.[13]

In addition to single-variant association studies, more recent genetic studies increasingly use polygenic risk score (PRS) and genetic risk score (GRS) approaches to quantify cumulative genetic susceptibility based on multiple single nucleotide polymorphisms (SNPs).[14] Furthermore, epigenetic mechanisms such as DNA methylation may provide additional insight into how smoking and other modifiable exposures become biologically embedded and influence COPD-related lung function phenotypes.[15–18]

Recent KoGES-based studies have reported genetic associations with FEV1, FEV1/FVC, rapid decline in FEV1, SNP-by-smoking interaction effects, DNA methylation related to smoking and pulmonary function, and heritability of longitudinal lung function traits. However, despite the growing number of KoGES-based studies, no review has comprehensively summarized research trends in genetic and epigenetic factors related to COPD and COPD-related lung function phenotypes using this cohort. Several repeatedly implicated COPD-related genes have plausible biological relevance to lung function. FAM13A has been associated with lung function and COPD susceptibility and may be involved in airway epithelial responses to cigarette smoke.[19] AGER encodes the receptor for advanced glycation end products and has been linked to pulmonary inflammatory responses, alveolar structure, and lung function.[20] HHIP is involved in hedgehog signaling and lung development, and experimental evidence suggests that Hhip haploinsufficiency may increase susceptibility to emphysema-like lung changes.[21] CHRNA3/5 is related to nicotine dependence and smoking-related COPD risk.[22] Other loci, including PPT2 and TNXB in the 6p21 region, may reflect immune, extracellular matrix, or regional linkage signals, although their direct functional roles in COPD require further investigation.

Although lung function traits such as FEV1 and FEV1/FVC are widely used as intermediate phenotypes for COPD, they are not equivalent to clinically defined COPD.[7–9] Lung function reflects physiological impairment, whereas COPD diagnosis incorporates persistent airflow limitation, symptom burden, exposure history, and clinical context. Longitudinal decline in lung function provides a closer approximation to disease progression than cross-sectional measurements, but it still does not fully represent clinically diagnosed COPD. Therefore, lung function traits, longitudinal decline, and COPD diagnosis should be considered related but distinct phenotypic domains.

Accordingly, this review does not treat spirometric traits as direct substitutes for clinically diagnosed COPD. Rather, it maps how KoGES-based studies have used lung function traits and longitudinal decline as COPD-related intermediate phenotypes to investigate genetic and epigenetic susceptibility in a Korean population. Therefore, this scoping review aimed to systematically summarize KoGES-based studies on the genetic and epigenetic associations of lung function and COPD-related phenotypes, classify the major methodological approaches used, and identify research gaps that may guide future COPD genetics research in Korean populations.

Materials and Methods

Study Design

This study was conducted as a scoping review to map the existing literature on genetic factors associated with COPD and lung function using data from the Korean Genome and Epidemiology Study (KoGES). Scoping reviews are appropriate for examining the breadth of evidence, key concepts, and methodological differences across heterogeneous studies.[23–25] The review followed the five-stage methodological framework proposed by Arksey and O'Malley[23] and was reported in accordance with PRISMA-ScR.[24]

Stage 1: Identifying the Research Question

The first stage of the scoping review was to establish a broad but focused research question. Because scoping reviews aim to capture the range and nature of available evidence rather than answer a narrowly defined hypothesis, the question was formulated to reflect the breadth of KoGES-based COPD genetics research.[23,24] The primary research question of this review was: What are the research trends in genetic and epigenetic studies related to COPD and COPD-related lung function phenotypes using KoGES data, and what gaps remain for future research in Korean populations?

Stage 2: Identifying Relevant Studies

A literature search was performed in PubMed, Scopus, Web of Science, KoreaMed, RISS, KISS, and DBpia. Following the scoping review framework, multiple databases were searched to identify a comprehensive body of potentially relevant studies.[23,24] The search strategy combined three concept blocks and was supplemented by targeted searches using alternative terms for KoGES-related cohorts, respiratory phenotypes, and genetic or epigenetic analytic approaches:

(1) KoGES-related terms ("Korean Genome and Epidemiology Study", "KoGES", "KARE", "Korea Association Resource", "Ansan", "Ansung", and "Ansan–Ansung"); (2) COPD- or lung function-related terms ("COPD", "chronic obstructive pulmonary disease", "chronic lung disease", "lung function", "pulmonary function", "spirometry", "FEV1", "FVC", "FEV1/FVC", "forced expiratory volume", "forced vital capacity", "airflow limitation", "emphysema", "lung function decline", "pulmonary function decline", "rapid decline", and "annual decline"); and (3) genetic or epigenetic terms ("genetic", "genomic", "GWAS", "genome-wide association", "single nucleotide polymorphism", "SNP", "variant", "polymorphism", "exome", "exome array", "polygenic risk score", "PRS", "genetic risk score", "GRS", "heritability", "methylation", "DNA methylation", "epigenome-wide association study", "EWAS", "gene–environment interaction", "gene-by-smoking interaction", and "gene–time interaction").

The search was initially conducted from database inception to March 1, 2026, and updated on April 27, 2026, during revision.

Stage 3: Study Selection

Duplicate records were removed by comparing titles, authors, publication years, and journal information across databases. After duplicate records were removed, titles and abstracts were screened for relevance, and full texts of potentially eligible articles were reviewed according to predefined inclusion and exclusion criteria. Titles/abstracts and full texts were screened independently by two reviewers, and disagreements were resolved through discussion. This stage was conducted in accordance with the study selection process recommended in the original scoping review framework.[23] Studies were included if they

1. Used KoGES or a KoGES-related cohort, including KARE, Ansan–Ansung, or the Healthy Twin Study, as the primary dataset;
2. Investigated eligible genetic, genomic, or epigenetic factors associated with COPD, chronic lung disease, spirometric traits, or longitudinal lung function-related phenotypes;
3. Included outcomes such as COPD, airflow limitation, FEV1, FVC, FEV1/FVC, emphysema, or longitudinal decline in lung function; and
4. Were full-length original research articles published in peer-reviewed journals.

Studies were excluded if they did not use KoGES or a KoGES-related cohort as the primary dataset; did not include eligible genetic, genomic, or epigenetic analyses of COPD- or lung function-related outcomes; did not assess COPD, chronic lung disease, spirometric traits, or longitudinal lung function-related outcomes; were not full-length original research articles; or were outside the predefined analytic scope of this review. Conference abstracts or preliminary reports corresponding to subsequently published full-length articles were removed before screening and were not treated as independent eligible studies. The predefined analytic scope included SNP-based GWAS, exome array analysis, gene–environment interaction, gene–time interaction, heritability analysis, PRS/GRS analysis, and DNA methylation/EWAS. Copy-number variation (CNV)-only studies were not included in the final synthesis.

Stage 4: Charting the Data

The fourth stage involved charting and organizing the data from the included studies.[23] A predefined charting form was used to extract relevant information from each article, including first author, publication year, journal, KoGES subcohort, sample size, study design, phenotype or outcome definition, type of genetic analysis, major loci or genes identified, and key findings. This step enabled a structured comparison of studies with different methodologies and outcomes.

Stage 5: Collating, Summarizing, and Reporting the Results

The final stage involved collating, summarizing, and reporting the results of the included studies. Scoping reviews provide comprehensive overviews of a research field and differ from systematic reviews in that they typically do not perform formal quality assessment to derive pooled effect estimates or narrow evidence-based conclusions.[23,24] In the

present study, the included literature was descriptively summarized and grouped according to the major methodological approaches used in KoGES-based COPD genetics research.

These categories included

1. Genetic loci associated with lung function traits, including GWAS and exome-based studies;
2. Gene–environment interaction studies;
3. Longitudinal decline and gene–time interaction studies;
4. DNA methylation, epigenetic studies, and heritability analyses; and
5. Polygenic/genetic risk score studies.

Because the purpose of this review was to map the breadth of existing evidence rather than to estimate pooled effects, no formal meta-analysis or methodological quality assessment was performed.[23–25]

Supplementary Population-Frequency Mapping of Prioritized Variants

To provide population context for variants repeatedly reported in KoGES-based COPD genetic studies, we conducted a supplementary descriptive analysis of prioritized SNPs identified from the scoping review. Six SNPs were selected to provide representative population context on the basis of repeated appearance in the reviewed KoGES literature, biological relevance to lung function/COPD-related pathways, and coverage of both cross-sectional and longitudinal phenotypes. These included variants in FAM13A, AGER, the 6p21 region, and a rapid decline-associated locus. For each SNP, genotype data from the 1000 Genomes Project were used to calculate selected-allele frequencies across 26 populations (n = 2,504) and five superpopulations. For biallelic variants, the selected allele corresponded to the ALT allele used in genotype recoding; for the multiallelic rs2239688 site, the G allele was used for descriptive population-frequency mapping. Population-specific selected-allele frequencies were summarized and visualized on world maps to illustrate ancestry-related distribution patterns. This analysis was performed to provide descriptive population context for cross-population interpretation and to inform future replication planning; it was not intended as an independent association test.

Results

Study Selection and Characteristics

A total of 20 unique records were identified for final assessment after database searching and duplicate removal. One conference abstract/preliminary report corresponding to a subsequently published full-length article was removed before screening, leaving 19 records for screening and eligibility assessment. Seven records were excluded because they were not based on KoGES or KoGES-related cohorts as the primary dataset (n = 5), were outside the predefined analytic scope as a CNV-only study (n = 1), or did not include eligible genetic/genomic or epigenetic analysis of COPD- or lung function-related outcomes (n = 1). The remaining 12 studies met the eligibility criteria and were included in the final scoping review. The process of study identification, removal before screening, screening, eligibility assessment, and final inclusion is summarized in Figure 1. Most KoGES-based studies focused on lung function traits rather than clinically diagnosed COPD alone. The major outcomes used across studies were FEV1, FEV1/FVC ratio, annual decline in FEV1, rapid decline in FEV1, and longitudinal change in lung function. The main research approaches included GWAS, exome array analysis, polygenic/genetic risk score analysis, genome-wide interaction studies, epigenome-wide association studies, and heritability analyses. The characteristics of the included KoGES-based studies are summarized in Table 1.[10–12,14,18,26–32] As shown in Figure 2, KoGES-based COPD genetics research was limited in the early period and expanded after 2020, with more recent studies shifting from cross-sectional GWAS toward longitudinal, PRS/GRS, and interaction-based approaches. Because several included studies incorporated more than one analytic framework, such as GWAS with gene–environment interaction or GWAS with elastic-net-based GRS construction, the methodological categories were treated as non-mutually exclusive (Figure 3).

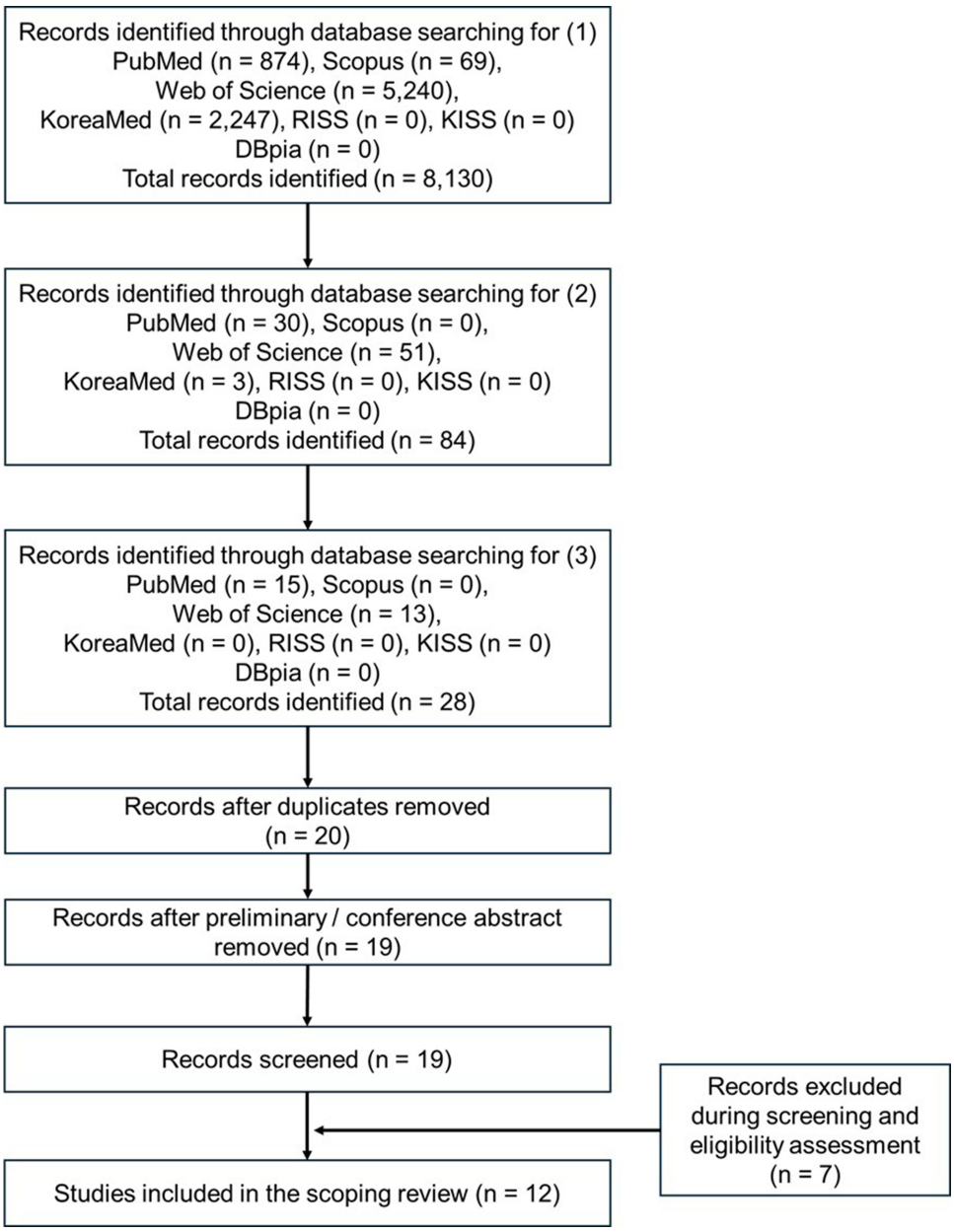

Figure 1 PRISMA flow diagram of study selection. After database searching and duplicate removal, 20 unique records were identified. One conference abstract/preliminary report corresponding to a subsequently published full-length article was removed before screening. The remaining 19 records were screened and assessed for eligibility. Seven records were excluded because they were not based on KoGES or KoGES-related cohorts as the primary dataset (n = 5), were outside the predefined analytic scope of this review as a copy-number variation (CNV)-only study (n = 1), or did not include eligible genetic/genomic or epigenetic analysis of COPD- or lung function-related outcomes (n = 1). Finally, 12 studies met the eligibility criteria and were included in the scoping review.

Genetic Loci Associated with Lung Function Traits

Early KoGES-based studies primarily examined cross-sectional lung function phenotypes. In a GWAS using KARE3 as the discovery cohort and the Healthy Twin Study as the replication set, variants in the 6p21 region and FAM13A were associated with lung function, and PPT2 on chromosome 6p21 was replicated as a locus associated with loss of lung function in Koreans.[11] A subsequent study further demonstrated that FAM13A variants were significantly associated with FEV1/FVC and showed interaction with heavy smoking in the Ansan–Ansung cohort.[10] Later, an exome array study identified coding variants in SMIM29, HMGA1, GIT2, FAM13A, TNXB, AGER, and ARHGEF40 as potential associations of lung function in the Korean population.[28]

Table 1 Characteristics of Included KoGES-Based Studies on Genetic Determinants of Lung Function and COPD-Related Phenotypes

No	First Author	Year	Journal	KoGES sub-Cohort/ Population	Sample Size	Genetic Approach	Phenotype	Key Genes/Loci	Endpoint Type	Study Design Context	Replication/ Validation Status	Main Findings
1	Kim WJ[11]	2014	Respirology	KARE3 discovery + Healthy Twin Study replication	6,223 + 2,730	GWAS	FEV1, FEV1/ FVC	6p21, PPT2, FAM13A	Spirometric trait	Cross-sectional GWAS	KoGES discovery + Korean replication cohort	Identified 6p21 as a lung-function locus in Koreans; PPT2 was replicated and FAM13A showed suggestive association.
2	Kim S[10]	2015	J Hum Genet	Ansan–Ansung	7,993	GWAS + gene–environment interaction	FEV1/FVC	FAM13A	Spirometric trait	Cross-sectional GWAS + interaction	KoGES-internal	FAM13A variants were significantly associated with FEV1/FVC and showed interaction with heavy smoking.
3	Park B[26]	2018	Sci Rep	KoGES/KARE	8,534	Genome-wide gene-by-smoking interaction	FEV1 / COPD-related traits	Genome-wide loci	Spirometric trait / COPD-related phenotype	Genome-wide smoking interaction	KoGES-internal	Performed a genome-wide evaluation of gene-by-smoking interactions affecting COPD-related phenotypes and lung function.
4	Park B[27]	2020	Sci Rep	KARE	7,473	Genome-wide interaction study	FEV1/FVC	6p21	Spirometric trait	Genome-wide interaction study	KoGES-internal	Demonstrated that the effect of the 6p21 region on lung function was modified by smoking exposure.
5	Kwak SY[18]	2020	Clin Nutr Res	KoGES men	209	DNA methylation analysis / EWAS	Pulmonary function	AHRR-related CpG and smoking-associated CpGs	Pulmonary function / methylation	Cross-sectional EWAS	KoGES-internal	Smoking-associated DNA methylation was related to pulmonary function.
6	Lee KS[28]	2021	Korean J Intern Med	KoGES	7,524	Exome array GWAS	FEV1, FEV1/ FVC	SMIM29, HMGA1, GIT2, FAM13A, TNXB, AGER, ARHGEF40	Spirometric trait	Exome array GWAS	KoGES-internal	Identified coding variants associated with lung function in Koreans.
7	Li D[12]	2022	Genes (Basel)	KoGES longitudinal cohort	6,622	Heritability analysis + GWAS	Mean lung function and annual change	Multiple loci	Mean lung function / longitudinal change	Heritability + longitudinal GWAS	KoGES-internal	Showed that both mean lung function and annual decline had heritable components.
8	Kim KS[14]	2023	Nutrients	Ansan–Ansung	8,840	Polygenic risk score + interaction analysis	COPD risk	Lung-related PRS	COPD risk	PRS + lifestyle interaction	KoGES-internal	High PRS was associated with increased COPD risk; PRS interacted with omega-3 fatty acid intake and exercise.
9	Lee SH[29]	2023	J Korean Soc Res Nicotine Tob	KoGES	8,166	Genetic risk score	FEV1/FVC / chronic lung disease	Multiple loci	Spirometric trait / chronic lung disease	GRS + smoking-related analysis	KoGES-internal	GRS for low FEV1/FVC was associated with chronic lung disease risk and evaluated with smoking exposure.

(*Continued*)

https://doi.org/10.2147/COPD.S611602

Table 1 (Continued).

No	First Author	Year	Journal	KoGES sub-Cohort/ Population	Sample Size	Genetic Approach	Phenotype	Key Genes/Loci	Endpoint Type	Study Design Context	Replication/ Validation Status	Main Findings
10	Kim SH[30]	2024	J Korean Med Sci	KoGES	6,516	GWAS	Rapid FEV1 decline	Multiple SNPs	Longitudinal decline	Longitudinal GWAS	KoGES-internal	Identified 15 genome-wide significant signals for rapid decline in FEV1.
11	Bae E[31]	2025	Hum Genomics	Ansan–Ansung prospective cohort	7,357	GWAS + elastic-net + GRS	Annual FEV1 decline	Multiple SNPs	Longitudinal decline	Longitudinal GWAS + GRS	KoGES-internal	Developed a genetic risk model for annual FEV1 decline using elastic-net-based SNP selection.
12	Kim CY[32]	2025	Sci Rep	Ansan–Ansung	8,554	Gene–time interaction study	Lung function decline	FAM13A and other loci	Longitudinal decline	Gene–time interaction study	KoGES-internal	Identified longitudinal genetic signals associated with lung function decline using gene–time interaction models.

Note: Several studies were conducted within related KoGES sub-cohorts, particularly KARE or Ansan–Ansung subsets. Therefore, repeated reporting of specific loci across studies should not be interpreted as fully independent external replication, but rather as recurrent observations within KoGES-based analyses.

Abbreviations: FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; GWAS, genome-wide association study; PRS, polygenic risk score; GRS, genetic risk score; EWAS, epigenome-wide association study.

https://doi.org/10.2147/COPD.S611602

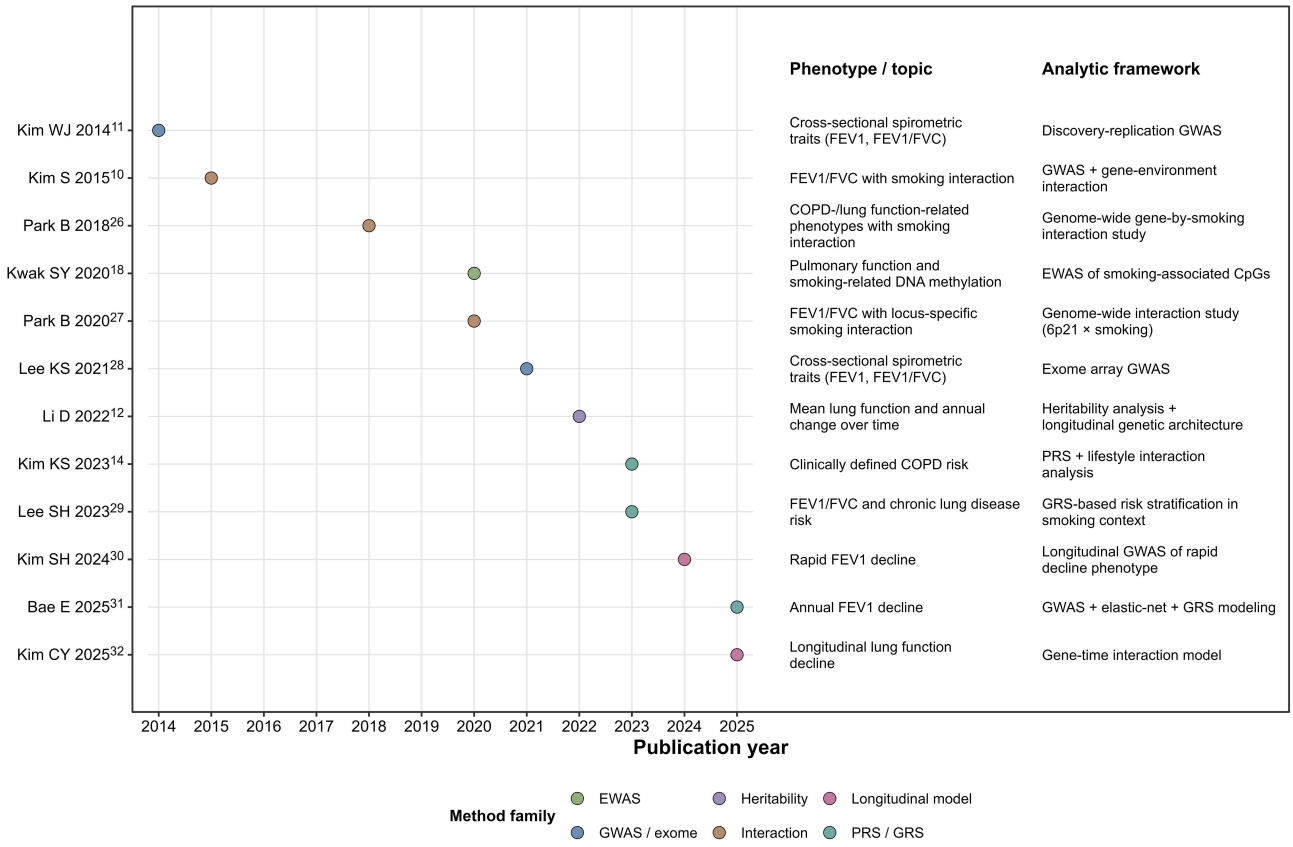

Figure 2 Timeline of KoGES-based genetic and epigenetic studies related to COPD and lung function. The figure aligns publication year, phenotype class, and analytic framework, illustrating the evolution from cross-sectional lung function GWAS to smoking interaction analyses, epigenetic studies, longitudinal decline models, and PRS/GRS-based approaches.

Gene–Environment Interaction Studies

Because smoking is the strongest environmental determinant of COPD, KoGES studies have also examined gene–environment interaction. A 2018 study performed a genome-wide assessment of gene-by-smoking interaction in COPD-related phenotypes.[26] A subsequent 2020 genome-wide interaction study found that the effect of the 6p21 region on FEV1/FVC was modified by smoking exposure.[27] In addition, the 2023 PRS study extended this line of inquiry to lifestyle-related interactions by evaluating omega-3 fatty acid intake and exercise.[14]

Longitudinal Decline and Gene–Time Interaction Studies

More recent KoGES studies have increasingly focused on longitudinal phenotypes, which may better reflect COPD development than single-time spirometric measures. A 2024 GWAS of rapid decline in FEV1 in 6,516 KoGES participants identified 15 genome-wide significant association signals, with different leading SNPs observed according to smoking history.[30] A Human Genomics study used KoGES longitudinal spirometry to assess annual FEV1 decline and constructed a genetic risk score based on SNP selection through elastic-net regression.[31] Another 2025 Scientific Reports study evaluated gene–time interactions and reported that multiple variants, including those in FAM13A, were associated with longitudinal lung function decline.[32]

DNA Methylation, Epigenetic Studies, and Heritability

In addition to SNP-based analyses, KoGES has been used for epigenetic studies relevant to COPD mechanisms. A study in KoGES men showed that smoking-associated DNA methylation was related to pulmonary function, suggesting a potential biological pathway linking smoking exposure and long-term decline in lung function.[18] In parallel, a 2022 study using repeated spirometry in KoGES estimated SNP heritability for both subject-specific mean lung function and

https://doi.org/10.2147/COPD.S611602

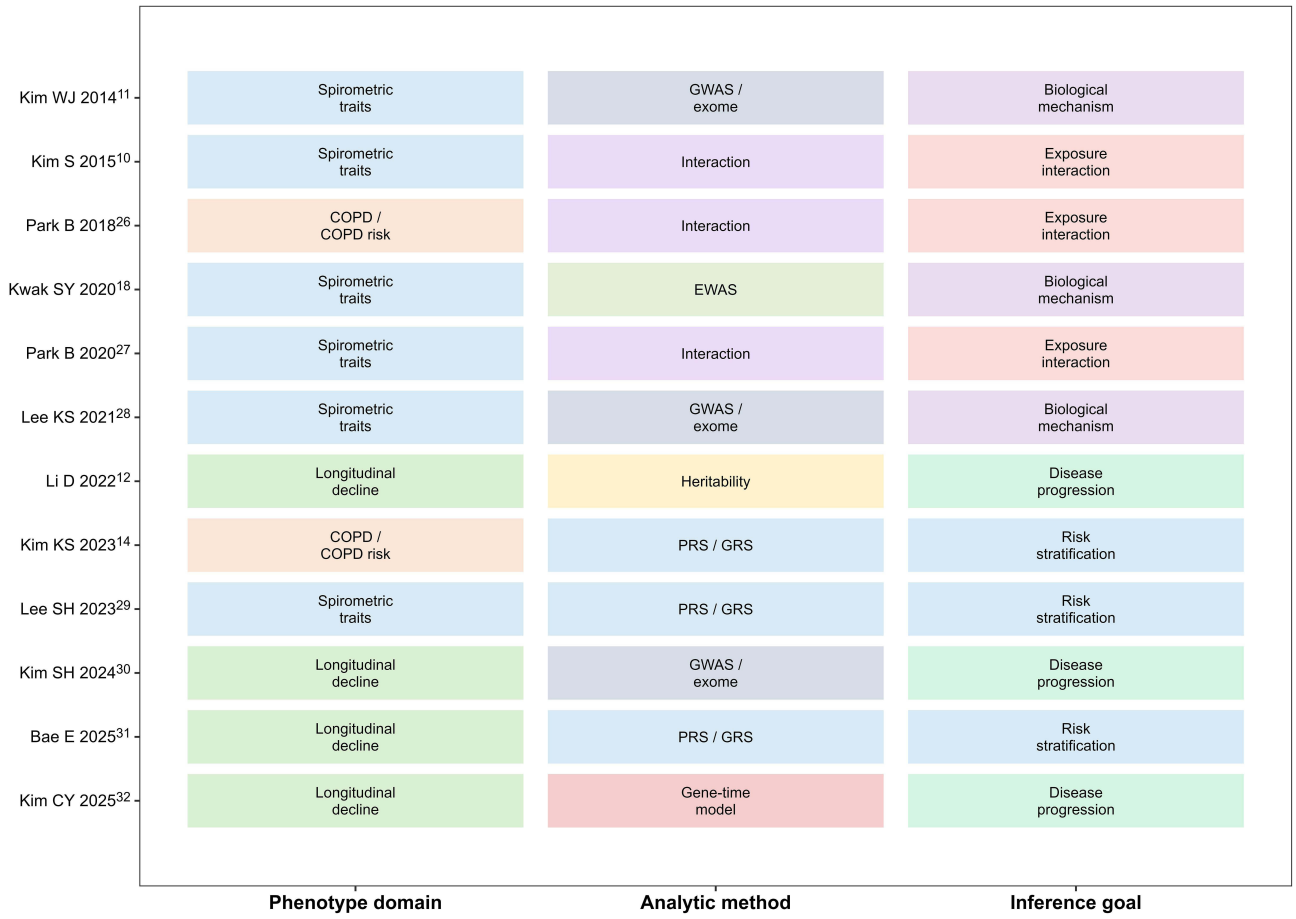

Figure 3 Conceptual framework of KoGES-based COPD-related genetic and epigenetic studies. The framework separates phenotype domain, analytic method, and inference goal to clarify how different study designs contribute to biological interpretation, exposure interaction, disease progression, and future risk-stratification hypotheses.

annual change rates.[12] The current KoGES-based epigenetic evidence remains limited, particularly because the available EWAS included only 209 male participants and focused mainly on smoking-associated CpG sites in peripheral blood. This small sample size limits statistical power and generalizability.

Future epigenetic research may need to extend beyond peripheral blood-cell EWAS toward methylation-based signal readouts that may be more feasible for translational research. In particular, circulating cell-free DNA methylation profiling may provide a future minimally invasive approach for exploring tissue injury-related epigenetic signals. Although cfDNA methylation has not yet been applied in KoGES-based COPD research, circulating cfDNA methylation has been proposed as a minimally invasive framework for capturing disease-related methylation signals from blood. This approach may help future COPD studies explore links among environmental exposure, lung tissue injury, systemic inflammation, and noninvasive stratification research.[33]

Polygenic and Genetic Risk Score Studies

KoGES-based research has also extended from single-variant association studies to polygenic approaches. A 2023 study in the Ansan–Ansung cohort reported that a lung-related polygenic risk score (PRS) was associated with COPD risk and interacted with omega-3 fatty acid intake and exercise, suggesting that PRS-based approaches may generate hypotheses for future risk stratification research.[14] Another study constructed a genetic risk score (GRS) based on FEV1/FVC-associated variants and evaluated its association with chronic lung disease in relation to smoking exposure.[29]

Table 2 Representative KoGES-Related Variants for Cross-Population Interpretation

Variant	Gene/Locus	KoGES-Based Study	Reported Phenotype	Rationale for Prioritization
rs2609264	FAM13A	Kim et al[10]	FEV1/FVC	Representative Korean FAM13A signal
rs7671167	FAM13A	Lee et al[28]	Lung function	Recurrent FAM13A-related signal
rs8192575	6p21 region	Park et al[27]	FEV1/FVC × smoking	Important gene–smoking interaction signal
rs2070600	AGER	Lee et al[28]	FEV1/FVC	Biologically relevant COPD/lung function locus
rs2239688	TNXB/6p21 region	Lee et al[28]	Lung function	Supports 6p21-region relevance
rs16951883	Rapid decline-associated locus	Kim et al[30]	Rapid FEV1 decline	Longitudinal decline-related KoGES signal for descriptive population context

Table 3 Selected-Allele Frequencies of Prioritized KoGES-Related COPD- and Lung Function-Associated SNPs in the 1000 Genomes Project

SNP	Gene/Locus	Selected Allele	East Asian	European	African	South Asian	Admixed American
rs2609264	FAM13A	T	0.58	0.76	0.72	0.55	0.70
rs7671167	FAM13A	T	0.55	0.49	0.58	0.46	0.45
rs8192575	6p21 region	G	0.21	0.05	0.03	0.08	0.07
rs2070600	AGER	T	0.22	0.05	0.01	0.07	0.02
rs2239688	TNXB/6p21 region	G	0.22	0.00	0.00	0.02	0.04
rs16951883	Rapid decline-associated locus	C	0.00	0.00	0.07	0.00	0.01

Note: Values represent selected-allele frequencies summarized by 1000 Genomes superpopulation. For biallelic variants, the selected allele corresponds to the ALT allele used in genotype recoding; for the multiallelic rs2239688 site, the G allele was used for descriptive population-frequency mapping. These frequencies should be interpreted as descriptive population context and not as direct evidence of cross-population association.

Prioritization of Representative SNPs for Cross-Population Interpretation

Based on the scoping review findings, six representative SNPs were prioritized for further contextual interpretation (Table 2). These SNPs do not represent newly discovered associations in the present review but rather summarize recurrent or representative signals identified across KoGES-based studies. These variants were selected based on repeated appearance across KoGES-based studies, biological relevance to lung function or COPD-related pathways, and representation of both cross-sectional and longitudinal phenotypes.

Population-Specific Selected-Allele Frequency Patterns of Prioritized SNPs

Population-specific selected-allele frequencies of the prioritized SNPs were examined using genotype data from the 1000 Genomes Project (Table 3, Figure 4). Selected-allele frequencies varied across continental populations, with variant-specific differences in magnitude and direction. In particular, rs2239688 showed a higher selected-allele frequency in East Asia but low frequencies in other superpopulations, whereas rs16951883 was rare in most superpopulations and showed a relatively higher frequency in African populations. Overall, these findings suggest that baseline allele distributions may differ across ancestries and should be considered when interpreting KoGES-based genetic associations and designing external validation studies.

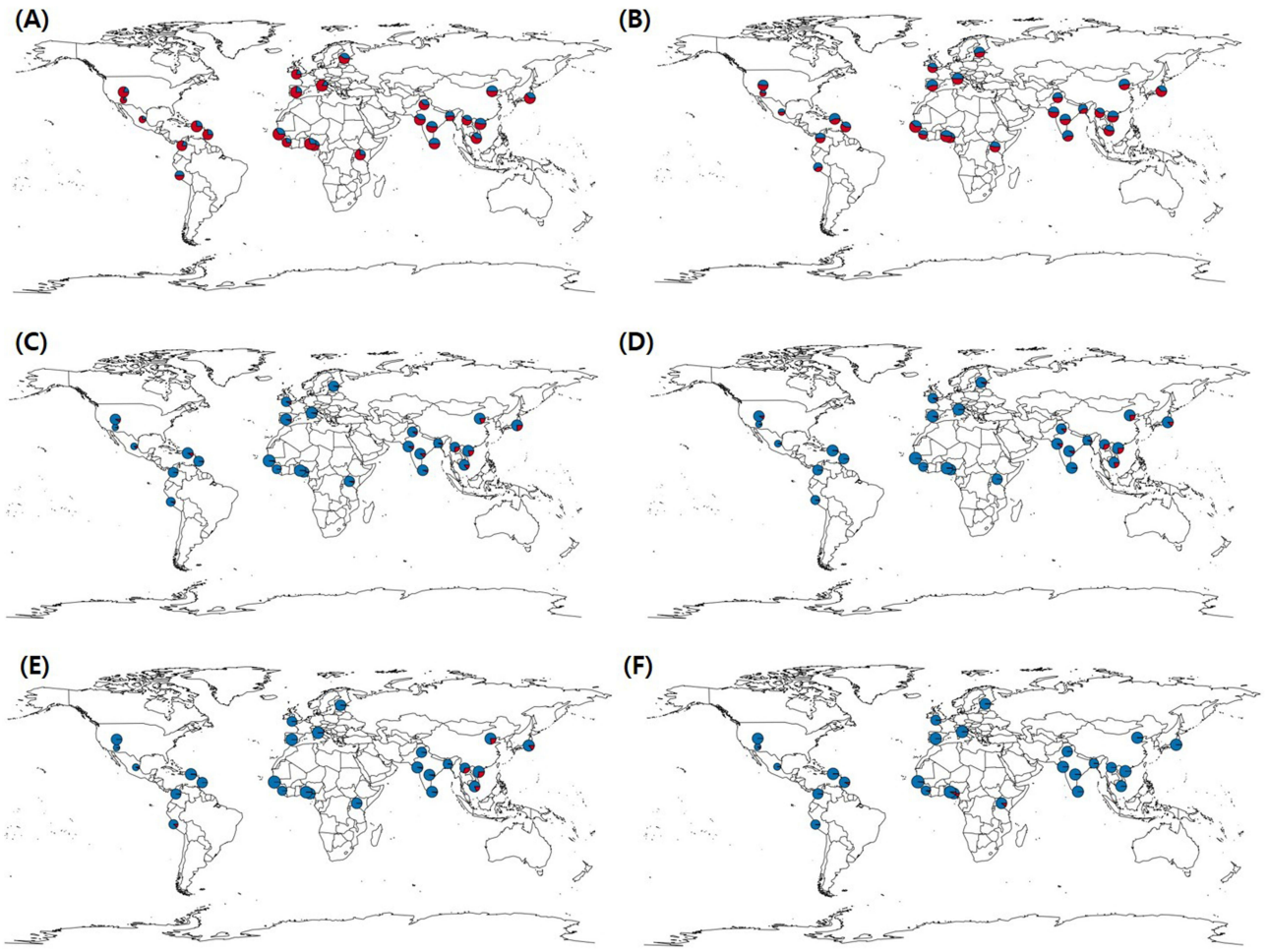

Figure 4 Global distribution of selected-allele frequencies for six prioritized KoGES-related COPD- and lung function-associated SNPs. Selected-allele frequencies were calculated from 1000 Genomes Project genotype data across 26 populations. Each panel represents a single SNP: (**A**) rs2609264, FAM13A, FEV1/FVC, GWAS + interaction; (**B**) rs7671167, FAM13A, lung function, exome array GWAS; (**C**) rs8192575, 6p21 region, FEV1/FVC × smoking, gene-smoking interaction; (**D**) rs2070600, AGER, FEV1/FVC, exome array GWAS; (**E**) rs2239688, TNXB/6p21 region, lung function, exome array GWAS; and (**F**) rs16951883, rapid decline-associated locus, rapid FEV1 decline, longitudinal GWAS. Pie charts indicate the selected-allele frequency for each SNP across populations. Allele-frequency patterns varied across continental populations for several variants, supporting the need to consider ancestry-related allele distributions when interpreting KoGES-based signals and planning external validation studies.

Discussion

This scoping review shows that KoGES-based studies on COPD-related genetic factors have progressed from early cross-sectional genome-wide association studies of lung function to more diverse approaches incorporating longitudinal decline, gene–environment interaction, polygenic risk, epigenetic mechanisms, and heritability.[10–12,14,18,26–32] A notable feature of this literature is that most studies focused on lung function-related phenotypes, such as FEV1, FEV1/FVC, and their longitudinal decline, rather than clinically defined COPD itself. This pattern likely reflects the structure of KoGES as a population-based cohort with repeated spirometric measurements, which is particularly well suited for studying intermediate and preclinical respiratory phenotypes.[13]

Among the repeatedly implicated loci, FAM13A and the 6p21 region appear to be the most prominent in KoGES-based studies.[10,11,27,28,32] FAM13A was identified in early GWAS of lung function, remained relevant in exome-based analyses, and reappeared in more recent gene–time interaction studies. Similarly, the 6p21 region was highlighted in both standard association analyses and smoking-interaction studies. These findings suggest that some loci may influence not only baseline lung function but also longitudinal susceptibility to decline, thereby contributing to COPD-related phenotypes over time.

These patterns are broadly consistent with findings from large international COPD genetics studies, which have identified major susceptibility loci such as FAM13A, HHIP, AGER, and CHRNA3/5, and have demonstrated that COPD-associated variants overlap substantially with loci related to lung function, pulmonary fibrosis, and other respiratory phenotypes.[6–9] In this respect, KoGES-based studies provide valuable Asian population-specific evidence that both complements and extends the predominantly European ancestry-based literature. Because several studies were conducted in related or partially overlapping KoGES sub-cohorts, repeated observation of specific loci should be interpreted as recurrent KoGES-based evidence rather than fully independent external replication.

Another important trend identified in this review is the movement toward longitudinal phenotypes, including annual decline in FEV1 and rapid decline in lung function.[12,30–32] These measures may better capture the dynamic biological processes underlying COPD development than a single spirometric assessment. Recent KoGES studies using rapid decline GWAS, elastic-net-based GRS construction, and gene–time interaction models support the idea that decline-related phenotypes are particularly informative for understanding early disease susceptibility and progression in population-based cohorts. From a translational research perspective, this approach may provide a basis for future hypothesis-generating studies on COPD risk stratification, especially in Asian populations where longitudinal evidence remains limited.

Importantly, KoGES-based COPD genetics has also expanded into polygenic and genetic risk score approaches. The inclusion of the 2023 Nutrients study is particularly meaningful because it showed that a lung-related PRS was associated with COPD risk and interacted with modifiable lifestyle factors such as omega-3 fatty acid intake and exercise.[14] In addition, another KoGES-based study demonstrated that a GRS derived from FEV1/FVC-associated variants was related to chronic lung disease risk in relation to smoking exposure.[29] Together, these findings suggest that KoGES-based research may generate hypotheses for future risk prediction frameworks that integrate genetic background and behavioral exposures, although external validation will be required before clinical application.

KoGES offers a valuable Korean population-based resource for COPD-related genetic epidemiology because it combines repeated spirometry, smoking-related variables, lifestyle information, genome-wide genotype data, and longitudinal follow-up within a single cohort framework. However, its cohort-specific design also limits generalizability to broader Korean, Asian, or multi-ethnic populations.[13] These characteristics make KoGES particularly suitable for examining intermediate respiratory phenotypes, longitudinal lung function decline, and gene–environment interaction in a Korean population. Compared with large European and multi-ancestry COPD genetics studies, however, the number of KoGES-based COPD genetic studies remains relatively small, and clinically defined COPD outcomes are still limited.[6–9]

A cross-population interpretation of KoGES findings suggests both shared and potentially population-specific genetic architecture. FAM13A and the 6p21/AGER region have been repeatedly implicated in KoGES-based studies[10,11,27,28,32] and have also been reported in large European GWAS of lung function and COPD.[6–9] This overlap suggests that part of the genetic architecture underlying airflow limitation and COPD-related traits may be shared across ancestries. However, not all KoGES-prioritized signals can be interpreted as externally replicated, because several findings were observed within overlapping KoGES sub-cohorts and some variants showed marked differences in selected-allele frequency across 1000 Genomes populations. Such differences may reflect ancestry-specific allele distributions, linkage disequilibrium structure, environmental exposure patterns, or differences in phenotype definition. Therefore, KoGES-based findings should be interpreted as population-specific evidence that complements, rather than replaces, findings from larger European and multi-ethnic studies.

Finally, several methodological limitations should be acknowledged. Because this study was conducted as a scoping review, the primary aim was to map the breadth and characteristics of existing evidence rather than to perform a formal quality assessment or pooled quantitative synthesis.[23–25] Therefore, the validity and risk of bias of the included studies were not systematically evaluated. In addition, because the review focused specifically on studies using KoGES as the primary dataset, relevant COPD genetics studies conducted in other Korean or Asian cohorts were not included. Nevertheless, this narrower focus was intentional, as it allowed us to specifically characterize the development of COPD-related genetic research within one major Korean population-based cohort. Future studies should extend this work by incorporating incident COPD, smoking-related methylation scores, PRS × physical activity interactions, and rapid

decline GRS models to develop risk-stratification approaches that can be externally tested in Korean populations. In addition, future KoGES-based studies should evaluate whether blood-cell and circulating cell-free DNA methylation profiles can complement genetic risk scores by capturing exposure-related and tissue injury-related pathways involved in COPD development, while recognizing that cfDNA methylation remains a future translational direction rather than an established KoGES-based COPD approach.[33]

Conclusion

In conclusion, KoGES-based studies have provided valuable evidence on the genetic and epigenetic associations of lung function and COPD-related phenotypes in Koreans. The literature has evolved from early cross-sectional GWAS of lung function to studies incorporating polygenic and genetic risk scores, longitudinal analyses of lung function decline, gene–environment interaction, and epigenetic mechanisms.[10–12,14,18,26–32]

However, studies directly addressing clinically defined COPD, externally validated risk prediction models, and integrative multi-omics remain limited. KoGES findings show partial overlap with European COPD genetics, particularly for FAM13A and the 6p21/AGER region, but allele-frequency differences and differences in cohort-specific phenotype definitions suggest that some signals may be population-specific and require external validation.

Future KoGES-based research should build on these findings by incorporating incident COPD, detailed smoking exposure, lifestyle-related modifiers, external validation, and integrated genomic approaches to support hypothesis-driven research on risk stratification in Korean and broader Asian populations.

Use of Artificial Intelligence

ChatGPT (OpenAI) was used to assist with language editing and manuscript structuring. The authors reviewed, edited, and verified all AI-assisted outputs and take full responsibility for the final content.

Data Sharing Statement

No new individual-level datasets were generated in this study. The review was based on previously published literature, and the supplementary population-frequency mapping used publicly available 1000 Genomes Project genotype data to calculate selected-allele frequencies.

Ethics Approval and Informed Consent

Not applicable. This study was a review of previously published literature and did not involve direct analysis of individual participant data.

Acknowledgments

The authors thank the participants and staff of the Korean Genome and Epidemiology Study (KoGES) for their valuable contributions.

Author Contributions

The first and second authors, Sun Young Kang and Sanghyun Roh, contributed equally to this work as co-first authors. All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This research was supported by the Global Learning & Academic Research Institution for Master's–PhD Students and Postdocs (G-LAMP) Program of the National Research Foundation of Korea (NRF), funded by the Ministry of Education (No. RS-2026-25441283).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Agustí A, Celli BR, Criner GJ, et al. Global initiative for chronic obstructive lung disease 2023 report: GOLD executive summary. *Am J Respir Crit Care Med*. 2023;207(7):819–837. doi:10.1164/rccm.202301-0106PP
2. Adeloye D, Song P, Zhu Y, Campbell H, Sheikh A, Rudan I. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. *Lancet Respir Med*. 2022;10(5):447–458. doi:10.1016/S2213-2600(21)00511-7
3. Pando-Sandoval A, Ruano-Ravina A, Candal-Pedreira C, et al. Risk factors for chronic obstructive pulmonary disease in never-smokers: a systematic review. *Clin Respir J*. 2022;16(4):261–275. doi:10.1111/crj.13479
4. Bhatt SP, Kim Y-I, Harrington KF, et al. Smoking duration alone provides stronger risk estimates of chronic obstructive pulmonary disease than pack-years. *Thorax*. 2018;73(5):414–421. doi:10.1136/thoraxjnl-2017-210722
5. Ambrosino N, Bertella E. Lifestyle interventions in prevention and comprehensive management of COPD. *Breathe*. 2018;14(3):186–194. doi:10.1183/20734735.018618
6. Wain LV, Shrine N, Artigas MS, et al. Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat Genet*. 2017;49(3):416–425. doi:10.1038/ng.3787
7. Shrine N, Guyatt AL, Erzurumluoglu AM, et al. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat Genet*. 2019;51(3):481–493. doi:10.1038/s41588-018-0321-7
8. Sakornsakolpat P, Prokopenko D, Lamontagne M, et al. Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nat Genet*. 2019;51(3):494–505. doi:10.1038/s41588-018-0342-2
9. Hobbs BD, de Jong K, Lamontagne M, et al. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet*. 2017;49(3):426–432. doi:10.1038/ng.3752
10. Kim S, Kim H, Cho N, et al. Identification of FAM13A gene associated with the ratio of FEV1 to FVC in Korean population by genome-wide association studies including gene–environment interactions. *J Hum Genet*. 2015;60(3):139–145. doi:10.1038/jhg.2014.118
11. Kim WJ, Lee MK, Shin C, et al. Genome-wide association studies identify locus on 6p21 influencing lung function in the Korean population. *Respirology*. 2014;19(3):360–368. doi:10.1111/resp.12230
12. Li D, Kim W, An J, Kim S, Lee S, Do A. Heritability analyses uncover shared genetic effects of lung function and change over time. *Genes*. 2022;13(7):1261. doi:10.3390/genes13071261
13. Kim Y, Han B-G. KoGES group. Cohort profile: the Korean Genome and Epidemiology Study (KoGES) consortium. *Int J Epidemiol*. 2017;46(2):e20. doi:10.1093/ije/dyv316
14. Kim K-S, Park S. Impact of lung-related polygenic risk scores on chronic obstructive pulmonary disease risk and their interaction with w-3 fatty acid intake in middle-aged and elderly individuals. *Nutrients*. 2023;15(13):3062. doi:10.3390/nu15133062
15. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet*. 2003;33 (S3):245–254. doi:10.1038/ng1089
16. Światowy WJ, Drzewiecka H, Kliber M, et al. Physical activity and DNA methylation in humans. *Int J Mol Sci*. 2021;22(23):12989. doi:10.3390/ijms222312989
17. Kadayifci FZ, Zheng S, Pan Y-X. Molecular mechanisms underlying the link between diet and DNA methylation. *Int J Mol Sci*. 2018;19(12):4055. doi:10.3390/ijms19124055
18. Kwak S-Y, Park CY, Shin M-J. Smoking may affect pulmonary function through DNA methylation: an epigenome-wide association study in Korean men. *Clin Nutr Res*. 2020;9(2):134–145. doi:10.7762/cnr.2020.9.2.134
19. Chen Q, de Vries M, Nwozor KO, et al. A protective role of FAM13A in human airway epithelial cells upon exposure to cigarette smoke extract. *Front Physiol*. 2021;12:690936. doi:10.3389/fphys.2021.690936
20. Oczypok EA, Perkins TN, Oury TD. All the "RAGE" in lung disease: the receptor for advanced glycation endproducts (RAGE) is a major mediator of pulmonary inflammatory responses. *Paediatr Respir Rev*. 2017;23:40–49. doi:10.1016/j.prrv.2017.03.012
21. Lao T, Jiang Z, Yun J, et al. Hhip haploinsufficiency sensitizes mice to age-related emphysema. *Proc Natl Acad Sci U S A*. 2016;113(32):E4681–E4687. doi:10.1073/pnas.1602342113
22. Lassi G, Taylor AE, Timpson NJ, et al. The CHRNA5–A3–B4 Gene cluster and smoking: from discovery to therapeutics. *Trends Neurosci*. 2016;39 (12):851–861. doi:10.1016/j.tins.2016.10.005
23. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19–32. doi:10.1080/1364557032000119616
24. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169 (7):467–473. doi:10.7326/M18-0850
25. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci*. 2010;5(1):69. doi:10.1186/1748-5908-5-69
26. Park B, Koo S-M, An J, et al. Genome-wide assessment of gene-by-smoking interactions in COPD. *Sci Rep*. 2018;8(1):9319. doi:10.1038/s41598-018-27463-5
27. Park B, An J, Kim W, et al. Effect of 6p21 region on lung function is modified by smoking: a genome-wide interaction study. *Sci Rep*. 2020;10 (1):13075. doi:10.1038/s41598-020-70092-0
28. Lee K-S, Kim KH, Oh Y-M, Han B, Kim WJ. A genome wide association study for lung function in the Korean population using an exome array. *Korean J Intern Med*. 2021;36(Suppl 1):S142–S150. doi:10.3904/kjim.2019.204
29. Lee SH, Baek JW, Jee SH. Interaction between smoking and genetic risk of FEV 1 /FVC ratio in chronic lung disease: a cross-sectional study. *J Korean Soc Res Nicotine Tob*. 2023;14(1):1–8. doi:10.25055/JKSRNT.2023.14.1.1
30. Kim SH, Lee H, Jo YS, Yoo J, Choi JY. Genome-wide association analysis of rapid decline in lung function: analysis from the Korean genome and epidemiology study. *J Korean Med Sci*. 2024;39(42):e275. doi:10.3346/jkms.2024.39.e275

31. Bae E, Oh S, Lee J-K, Heo EY, Kim DK, Lee HW. Genome-wide association study of lung function decline using elastic-net regression: a prospective analysis in Korean cohort. *Hum Genomics*. 2025;20(1):3. doi:10.1186/s40246-025-00862-2
32. Kim CY, Park B, Jung JY, et al. Genome wide interaction study of genetic variants associated with lung function decline. *Sci Rep*. 2025;15(1):9824. doi:10.1038/s41598-025-93147-6
33. Dong W, Lau C-H, Li J, et al. Pan-cancer methylation analysis of circulating cell-free DNA. *Cancer Genetics*. 2025;296-297:182–195. doi:10.1016/j.cancergen.2025.07.014