Monaldi Archives for Chest Disease

eISSN 2532-5264 https://www.monaldi-archives.org/

Monaldi Arch Chest Dis 2024 [Online ahead of print]

To cite this Article:
Hassan M, Ali AS, Zubairi ABS, et al. **Gene polymorphisms and risk of idiopathic pulmonary fibrosis: a systematic review and meta-analysis.** *Monaldi Arch Chest Dis* doi: 10.4081/monaldi.2024.2952

Gene polymorphisms and risk of idiopathic pulmonary fibrosis: a systematic review and meta-analysis

Maryam Hassan,[1*] Akbar Shoukat Ali,[1*] Ali Bin Sarwar Zubairi,[1,2*] Zahra Ali Padhani,[3] Salman Kirmani,[4] Huzaifa Ahmad,[1] Zafar Fatmi,[5] Jai K Das[4,6]

[1]Department of Medicine, Aga Khan University Hospital, Karachi, Pakistan; [2]Department of Medicine, Southern Illinois University School of Medicine, Springfield, IL, USA; [3]Faculty of Health and Medical Sciences, School of Public Health, University of Adelaide, Australia; [4]Department of Pediatrics and Child Health, Aga Khan University Hospital, Karachi, Pakistan; [5]Department of Community Health Sciences, Aga Khan University Hospital, Karachi, Pakistan; [6]Institute of Global Health and Development, Aga Khan University, Karachi, Pakistan

*Contributed equally as co-first authors

Correspondence: Ali Bin Sarwar Zubairi, Department of Medicine, Aga Khan University Hospital, Karachi, Pakistan. Tel: 092-300-2339962. E-mail: ali.zubairi@aku.edu

Contributions: MH, ABSZ, HA, ZF, concept and design of the study; ASA, JKD, ZAP, data analysis; MH, ASA, ABSZ, JKD, ZAP, SK, data interpretation and drafting the manuscript; ABSZ, JKD, SK, ZF, critical revision for intellectual content. All authors have approved the final version to be published and are jointly accountable for all aspects of the work.

Conflict of interest: the authors declare that there is no conflict of interest.

Ethics approval and consent to participate: not applicable.

Patient consent for publication: not applicable.

Funding: none.

Availability of data and materials: all data generated or analyzed during this study are included in this published article.

Abstract

Idiopathic pulmonary fibrosis (IPF) has been widely hypothesized to occur as a result of an interplay between a nexus of environmental and genetic risk factors. However, not much is known about the genetic aspect of this disease. The objective of this review was to identify the genetic polymorphisms associated with the risk of developing IPF. We searched PubMed, EBSCO CINAHL Plus, Web of Science, and Wiley Cochrane Library databases for studies on risk factors of IPF published between March 2000 and November 2023. Studies with an IPF diagnosis based only on the American Thoracic Society and the European Respiratory Society guidelines were included. Thirty-one case-control studies were included with 3997 IPF and 20,925 non-IPF subjects. Two of the studies enrolled biopsy-proven IPF patients; 13 studies diagnosed IPF on the basis of clinical and high-resolution computed tomography (HRCT) findings; and 14 studies diagnosed based on both biopsy and clinical and HRCT findings. 16 studies with *MUC5B rs35705950*, *IL-4 rs2243250*, *IL-4 rs2070874*, and *tumor necrosis factor α* (*TNFα)-308* were eligible for meta-analysis. The allele contrast model (T *versus* G) for *MUC5B rs35705950* revealed statistically significant association of T allele with the risk of IPF [odds ratio (OR) 3.84, 95% confidence interval (CI) 3.20 to 4.61, adjusted p<0.0001), as was the allele contrast model for Asian (OR 2.83, 95% CI 1.51 to 5.32, adjusted p=0.009) and Caucasian (OR 4.11, 95% CI 3.56 to 4.75, adjusted p<0.0001). The allele contrast models for *IL-4 rs2243250*, *IL-4 rs2070874*, and *TNFα-308* did not demonstrate any significant association with IPF. This review suggests an association of *MUC5B rs35705950* T allele with the risk of developing IPF. To our knowledge, this study is the first to aggregate several genetic polymorphisms associated with IPF.

Key words: idiopathic pulmonary fibrosis, interstitial lung disease, *MUC5B*, *rs35705950*, gene polymorphism.

Additional information - study protocol registration: PROSPERO Registration Number CRD4202018170.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive, and fibrotic lung disease of unknown etiology; it is the most common, and most lethal type of ILD. It has an estimated global prevalence of 13 and 20 cases per 100,000 among adult females and males, respectively [1]. Most patients present between 50 to 70 years of age with gradually progressive dyspnea and non-productive cough [2] and the post-diagnosis median survival time is 2-4 years, mainly due to the relentless progression of the disease leading to respiratory failure.

The clinical course of IPF is variable, and therefore, makes the diagnosis of the disease a perplexing task. In 2000, the American College of Chest Physicians (ACCP), the American Thoracic Society along with the European Respiratory Society (ATS/ERS) recognized IPF as a separate clinical entity and issued an international consensus on the diagnosis and treatment of IPF [2] and this consensus statement was revised in 2011 and 2018 [3,4]. The most recent evidence-based guidelines recommend a combination of clinical, radiologic, and/or histopathologic findings to diagnose IPF. Based on current criteria, patients are diagnosed with IPF after exclusion of known causes of ILD and either the presence of high-resolution computed tomography (HRCT) pattern of usual interstitial pneumonitis (UIP) or specific combination of HRCT and histopathology in patients subjected to lung biopsy [3].

Our understanding of the pathogenesis of the disease is still evolving. Recent studies have suggested that the epithelial-mesenchymal pathway may contribute to fibrosis by disrupting the normal regeneration of alveolar epithelium [4]. Ample evidence suggests that chronic inflammation plays a major role in the development of IPF. Higher levels of IL-14 have shown to increase the risk of IPF [5]; inversely, higher levels of circulating adipokines decreased the risk of IPF [6].

Though the etiology of IPF is largely unknown, it is widely hypothesized that it occurs as a result of an interplay between a nexus of environmental and genetic risk factors. Several risk factors such as cigarette smoking, gastroesophageal reflux disease (GERD), environmental and occupational exposures have been identified [1]. A systematic review and meta-analyses conducted by Park *et al*. 2021 found that environmental exposure to wood dust, metal dust and pesticides increased the risk of developing IPF [7]. Moreover, farmers or those who worked in agriculture also had an increased the risk of IPF and smoking, an already established risk factor, was also further cemented as a social risk factor.

In addition to the above mentioned environmental and occupational risk factors, recent studies are trending towards analyzing genetic risk factors associated with IPF. Significant evidence has established a causal link between genetics and the development of IPF [8]. Genetics have

also been shown to lead to different patterns in HRCT as compared to sporadic IPF, suggesting that genetics may not only play an important role in the underlying pathogenesis but also in determining the prognosis and treatment of the disease [9]. Genetic risk factors that have been identified for familial IPF so far are *MUC5B* polymorphism [10], surfactant protein C [11], and telomerase proteins [12]. In fact, a recent systematic review and meta-analysis by Wu *et al*. 2021 confirmed the association between *MUC5B* polymorphism rs35705950 and risk of developing IPF [13].

To the best of our knowledge, our study is the first to adhere to stringent criteria, based on studies that diagnosed IPF using the ATS/ERS guidelines. Previous studies have focused on one gene, this study is the first to collate all the genetic polymorphisms associated with IPF. The main objective of this systematic review is to systematically identify all the genetic polymorphisms that may be associated with the risk of developing IPF (diagnosed only on the ATS/ERS guidelines).

Methods

This review adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement [14], and its protocol is filed with PROSPERO (www.crd.york.ac.uk/prospero/), the international prospective register of systematic reviews, under registration number CRD42020181703 [15].

Eligibility criteria

Types of studies

We included case-control studies that assessed the genetic polymorphisms associated with the risk of developing IPF and included studies that had diagnosed IPF based only on the ATS/ERS guidelines published in 2000 along with subsequent revisions [2,3,16-19].

Main criteria

Main criteria for inclusion were: Studies which assessed any genetic polymorphisms associated with development of IPF, reported baseline demographic and clinical characteristics of the IPF patients, examined at least 10 IPF patients, involving adults (18) of all racial backgrounds and ethnicities with IPF (cases), including either healthy or diseased adults (18) subjects as controls. Studies that included patients with acute exacerbation of IPF (AE-IPF), focusing on predictors of progression and mortality, and with incomplete or missing genetic polymorphism data (after consulting with corresponding authors of respective articles) were excluded.

ATS/ERS guidelines

The ATS/ERS guidelines regarding clinical management of IPF have evolved from being consensus-based in 2000 to evidence-based in 2011 [17]. The first ever evidence-based guideline for the clinical diagnosis and management of IPF was issued in 2011 [17], with updated recommendations in 2015 [19]. In 2018, a collaborative multidisciplinary effort of IPF specialists from leading respiratory societies (ATS, ERS, Japanese Respiratory Society and Latin American Thoracic Society) restructured the IPF diagnostic criteria, and made recommendations for diagnosis of IPF [3].

Systematic literature search

We searched PubMed, EBSCO CINAHL Plus, Web of Science, and Wiley Cochrane Library databases for studies assessing the genetic risk factors associated with IPF and published between March 2000 and November 2023. We used the combination of MeSH and key words for terms of "idiopathic pulmonary fibrosis", "IPF", "usual interstitial pneumonia", "UIP", "risk factors", "association", "genetic", "environmental", and "occupational" to find relevant articles. Detailed search strategy can be accessed through (*Supplementary Table 1*). Article search was restricted to those written in the English language only.

Data collection and analysis

Selection of studies

The titles and abstracts of all the records identified as a result of the search strategy were independently screened by two reviewers (MH and ASA) on EndNote and the duplicates were removed. Discrepancies were resolved by discussion with a third reviewer (ABSZ). The full texts were also screened in duplicate, and the final list of included studies was finalized with consensus.

Data extraction

Two reviewers (MH and ASA) extracted the following data on an excel sheet: first author's name, publication year, study population, sample size, methodology, genes with SNPs of interest, genotyping, quality assessment parameters, and statistical methods. Any disagreement between the two reviewers, during the review process was discussed with a third reviewer (ABSZ) to reach unanimity.

Quality assessment of included studies

To evaluate the quality of included studies, we used Newcastle-Ottawa Scale (NOS) for case-control studies [20]. The following components were used in the scale: selection (adequacy of case definition, representativeness of the cases, selection of controls and definition of controls), comparability and exposure (ascertainment of exposure, same method for ascertainment of cases and controls and non-response rate).

Statistical analyses

All the quantitative statistical analyses, including Hardy-Weinberg Equilibrium (HWE) calculations, were executed using the MetaGenyo: Meta-Analysis of Genetic Association Studies tool specifically designed for meta-analysis of genetic studies [21]. A meta-analysis for risk of IPF was conducted for each SNP where 2 studies documented on the same genetic variant, and ORs with their respective 95% CIs were reported. For all the SNPs included in meta-analysis (*MUC5B rs35705950*, *IL-4 rs2243250*, *IL-4 rs2070874*, and *TNFα -308*), we estimated the association under six different genetic models. For all studies, we estimated the association under five different types of ORs, namely the allele contrast, recessive, dominant, over-dominant, homozygote co-dominant, and heterozygote co-dominant models. An I^2 statistical measure of heterogeneity was used with value of 25% was regarded as low, 50% as moderate and 75% as high heterogeneity. Based on significant heterogeneity among studies, a random effect model was deemed acceptable. We did a descriptive analysis of all genotypes where a meta-analysis could not be performed. A sub-group analysis for ethnicity was performed for *MUC5B rs35705950*. The Egger's test and inverted funnel plots were also executed for SNPs with 10 studies; *MUC5B rs35705950*.

Results

Results of the literature search

We retrieved 13,747 publications through search on electronic databases (Figure 1). After omitting duplicates, 13,597 articles were screened by title and abstract, and 13,453 records were excluded. One hundred forty-four full-text articles were assessed for eligibility and a total of 113 articles were removed based on the following reasons: non-genetic studies (n=61), GWAS studies (n=8), exome sequencing (n=3) and others specified in the PRISMA flow diagram (n=41). Finally, 31 studies were included, of which 16 were eligible for quantitative analyses (meta-analysis).

Characteristics of the included studies

The characteristics of the included studies are presented in (*Supplementary Table 2*). A total of 31 studies were included [10,22-49] which were published between 2003 and 2020 and included IPF (n=3997) and non-IPF subjects (n=20,925). Two of the studies enrolled biopsy-proven IPF patients [28,44], 13 studies diagnosed IPF on the basis of clinical and high-resolution computed tomography (HRCT) findings [22-26,30,34,39,40,42,43,45,46], and 14 diagnosed IPF based on both biopsy and clinical and HRCT findings [10,27,29,31-33,35-38,41,47-49]. A total of 11 studies used the ATS/ERS guidelines published in 2000 [24,25,27-29,33,35,36,39,40,45], 13 adhered to the 2001 guidelines [10,24-26,31,32,37,38,41,44,45,47,48], nine followed the 2011 guidelines [23,24,26,30,34,42,46,49], two 2013 guidelines [26,43], and one followed 2018 guideline [42]. Thirteen studies performed genetic association analysis in European populations such as France, Czech Republic, Greece, Germany, UK, Spain and Italy [10,22,24-27,31,32,39,41,42,45], five studies were from United States [34,43,47,48], one from Australia [35], and four studies each were from South Korea [28,37,38,44], Mexico [29,30,33,36], and China [23,40,46,49]. *MUC5B rs35705950* was the most frequently reported polymorphism [10,22-26,34,42,43,45-48]. Genotyping methods varied among studies and included TaqMan, Allele-specific Matrix Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF), Polymerase Chain Reaction Restriction Fragment Length Polymorphism (PCR-RFLP), RFLP, DNA sequencing, Single Base Extension (SBE) and electrophoresis, PCR Sequence Specific Amplification (PCR-SSP) and Reference Strand-mediated Conformation Analyses (RSCA) genotyping methods. Key factors adjusted among studies for the multivariate analysis were age, gender, and smoking [25,28,34,36-39,44,47]. *Supplementary Table 3* depicts association between genetic variants and risk of IPF reported.

Quality assessment of the included studies

Quality assessment of the included case-control studies was examined using NOS. As depicted in *Supplementary Table 4*, total stars ranged from 6 to 8, suggesting moderate-to-high quality studies and minimal risk of bias. Case and control definitions, representativeness of the cases, selection of controls, ascertainment of exposure, method for ascertainment of cases and controls were adequate in all included studies. On comparability part, only 11 studies sufficiently controlled for most important (smoking) and additional confounders (age and gender) [25,28,34,36-39,44,47].

Meta-analysis of MUC5B rs35705950, IL-4 rs2243250, IL-4 rs2070874,* and *TNFα -308

A meta-analysis was performed since 2 studies evaluated the association of risk of IPF with *MUC5B rs35705950* [10,22-26,34,42,43,45-48], *IL-4 rs2243250* [24,27], *IL-4 rs2070874* [24,27], and *TNFα -308* [35,49] polymorphism. Genotype distributions with unadjusted and adjusted HWE *P* values are displayed in (*Supplementary Tables 5 and 6*). We performed an allele contrast, recessive, dominant, over-dominant, homozygote co-dominant and heterozygote co-dominant models for *MUC5B rs35705950*, *IL-4 rs2243250*, *IL-4 rs2070874*, and *TNFα -308*.

The allele contrast model (T vs. G) for *MUC5B rs35705950* revealed statistically significant association of T allele with that of risk of IPF (OR 3.84, 95% CI 3.20 to 4.61, adjusted *P*<0.0001). However, *IL-4 rs2243250* allele contrast model T vs. C (OR 1.88, 95% CI 0.27 to 12.87, adjusted *P*=1.000), *IL-4 rs2070874* allele contrast model T vs. C (OR 1.63, 95% CI 0.41 to 6.49, adjusted *P*=1.000), and *TNFα -308* allele contrast model A vs. G (OR 1.91, 95% CI 0.69 to 5.32, adjusted *P*=1.000) did not demonstrate a statistically significant association with risk of IPF. The subgroup analysis of the *MUC5B rs35705950* demonstrated: Asian (OR 2.83, 95% CI 1.51 to 5.32, adjusted *P*=0.009) and Caucasian (OR 4.11, 95% CI 3.56 to 4.75, adjusted *P*<0.0001) (Figures 2 and 3). Forest plots of genotypic models of *MUC5B rs35705950*, *IL-4 rs2243250*, *IL-4 rs2070874*, and *TNFα -308* are shown in *Supplementary Figure 1*.

Descriptive analysis of genetic polymorphisms other than MUC5B rs35705950, IL-4 rs2243250, IL-4 rs2070874,* and *TNFα -308

Besides *MUC5B rs35705950*, *IL-4 rs2243250*, *IL-4 rs2070874*, and *TNFα -308* polymorphisms, several other genetic polymorphisms were identified in the included studies. *TNF-α* [35], *TGF-b_1 T869C* [50], *ACE -5538* [37], *COX2.3050* and *COX2.8473* [39], *HLA-A*3, HLA-B*14*, *HLA-B*15*, *HLA-B*40*, *HLA-A2B15*, *HLA-A2B40*, *HLA-A11B15*, *HLA-A24B58* and *HLA-A30B40* [40], *HLA-A*02-DRB1*04* [19], *CR1 -5507 e33* [41], *IL-8 rs4073* [28], *MICA* [29], *FcγRIIIb CD16b NA1* allele heterozygotes and homozygotes [31], *FCGR3B* copy number [32], and *IVD rs2034650* [34] polymorphisms were found to associated with risk of IPF based on individual studies. Interestingly, IPF subjects with history of smoking had higher *SP-B B1580_C* [36] and *MMP-1 -755* [33] polymorphisms compared with non-smokers and therefore might increase the risk of development of IPF. Interestingly, two polymorphisms; *ADAM33 rs628977* [38],

HSPA1B rs1061581, *HSPA1L rs2227956* and *HSPA1 rs1043618* [30], and *FcγRIIIb CD16b NA2* allele [31] favored decreased risk of IPF (*Supplementary Table 3*).

Discussion

In this systematic review and meta-analyses, we have assessed the genetic risk factors associated with IPF. After reviewing a total of 13,597 articles, 31 studies met our inclusion criteria and of these studies only 16 studies were included for meta-analysis.

Our study of the underlying genetic risk factors of IPF revealed a diverse data set differing in population size and research methodologies. While some of the included studies boasted data sets in the thousands, others utilized a relatively smaller data set. This variation in sample size reflects the logistical issues in conducting genetic studies in a large study population while also highlighting the need for more collaborative research efforts. All included studies were case-control of moderate-to-high quality that analyzed a wide variety of ethnicities ranging from Europeans, Mexicans, Japanese, Han Chinese, North Americans, and Koreans. We also captured data on different races which included, Asians, Caucasians, and Mixed. Hence, this study has collated varied ethnic and racial data and further aided in understanding the genetic predisposition of IPF in different populations. However, we found that data on South Asian, African, and Middle Eastern populations was lacking. While age is a well-established risk factor for IPF, our dataset revealed a variation in age group where, in some studies, the age was greater in the cases than controls while the inverse was true in others. This variability in age reflects the natural history of patients who have a genetic predisposition as these patients may have an earlier onset of disease with a worse prognosis and severity as compared to the general population.

The genetic markers studied encompass a wide array of SNPs involving genes implicated in the pathophysiology the disease by impacting pathways involved in inflammation, fibrosis and tissue repair. Various genotyping methods such as, TaqMan assays, PCR, DNA sequencing, and PCR-RFLP were used to accurately detect and characterize genetic variations. *MUC5B rs35705950* [10,22-26,34,42,43,45-48], *IL-4 rs2243250* [24, 27], *IL-4 rs2070874* [24,27], and *TNFα -308* [35,49] were the most frequently reported polymorphisms among these studies. Although several genes were found to be associated with IPF, this is the first study to collate several genes into one review. Our study also showed that Asians and Caucasians having minor alleles T were vulnerable to develop IPF. Several causal genetic risk factors for IPF have been

well established through GWAS and meta-analyses. These genetic risk factors are further stratified according to whether the disease is sporadic or familial in nature [51].

There is extensive literature on the environmental and occupational risk factors for IPF [52]. In a systematic review and meta-analyses on occupational and environmental risk factors that lead to the development of IPF, Park *et al*. 2021 found that exposure to wood dust, metal dust, and pesticides increased the risk of developing IPF. In terms of occupations, a history of farming and agriculture were associated with an increased risk of developing the disease. Furthermore, smoking was also established as a risk factor associated with IPF [7].

Our review of the literature on risk factors of IPF found evidence of genetic polymorphism associated with the disease. The function of each of these genes is diverse - some of them are linked to the expression of cytokines, telomerase proteins (*TERT*), metalloproteinases (*MMP-1*), and toll-like receptors (*TOLLIP*); while others are linked to the production of mucous (*MUC5B*), pulmonary surfactant (surfactant protein), and the angiotensin converting enzyme (*ACE*).

Previous meta-analyses conducted on genetic polymorphisms associated with IPF limited the scope of their studies to a single gene unlike ours which included several. For example, Lee MG *et al*. 2015 [53] and Wu *et al*. 2021 [13] conducted a meta-analysis exploring the association between the MUC5B rs35705950 polymorphism and susceptibility to IPF. Similar to our study Lee MG *et al*. and Wu *et al*. found that an increased expression of the T allele was associated with increased susceptibility. Furthermore, unlike our study they performed a meta-analysis to discern the frequency of the allele expression in different ethnic groups. However, ethnically their study was limited to Europeans and Asians; whereas our study captured a diverse population that included Asian, Caucasians, and Mixed. In stark contrast to our study, Lee *et al*. did not have a strict diagnostic criterion for IPF and did not specifically utilize the ATS/ERS guidelines for the diagnoses of IPF [53]. Zhu *et al*. 2015 also conducted a meta-analysis investigating the association between MUC5B promoter polymorphism rs35705950 and IPF [23]. Their findings were similar to Lee *et al*., in that they reiterated the association between minor T allele expression and IPF; they also found that the expression of the minor T allele was increased in Caucasians as compared to Asians. However, they also did not use a specific criteria or guidelines to diagnose IPF.

These genes have also been identified by previous studies that established their association with IPF via GWAS studies. Noth *et al*. 2013 [54], Fingerlin *et al*. 2013 [50], and Allen RJ *et al*. 2017 [55] conducted GWAS studies that established the association between *MUC5B* and IPF. Furthermore, the meta-analyses conducted by Lee *et al*. 2015 showed that the expression of *MUC5B* increased the risk of IPF four-fold [53]. We conducted a meta-analysis on the SNPs

that were reported by two or more studies. These included the *MUC5B rs35705950* [10,22-26,34,42,43,45-48], *IL-4 rs2243250* [24, 27], *IL-4 rs2070874* [24,27], and *TNFα -308* [35,49]. The meta-analyses revealed a statistically significant association between the *MUC5B* T allele and the development of IPF for and a significant association between the dominant model (TT + TG vs. GG). However, no statistically significant association between IPF and *IL-4 rs2243250*, *IL-4 rs2070874*, and *TNFα -308* was found in any of the models.

MUC5B encodes the mucin family of proteins which form integral components of mucous secreted by the bronchial glands and play an integral part in the safe guarding of the airway [23]. The role of MUC5B rs35705950 polymorphism in the development of IPF still remains unclear; however, there are several working hypotheses. Seibold et al. hypothesized that the MUC5B mutation leads to the overproduction of mucin, which in turn leads to airway remodeling and fibrotic changes in lung tissue [56]. According to Chen et al., the overproduction of MUC5B leads to an aberrant form of mucin by altering its rheological properties, which results in mucociliary dysfunction leading to impaired clearing of mucin. The retained mucin alters lung tissue architecture via chronic inflammation and tissue damage resulting in fibrosis [57]. O'Dwyer et al. theorize that this retained mucous acts as a reservoir for pro-inflammatory cytokines and growth factors leading to the characteristic fibrosis and scarring seen in IPF [58].

Increased expression of MUC5B mRNA in the lungs was associated with the presence of T allele rs35705950, therefore it is hypothesized that it leads to a functional protein [59]. The T allele also demonstrated a 37-fold increase in the expression of *MUC5B i*n the lungs of unaffected subjects and 17-fold increase in patients with IPF as compared to controls. Furthermore, the minor T allele frequency of this polymorphism was found in 30-40% patients with IPF when compared to controls who demonstrated a frequency of 9-10%. Immunohistochemistry analysis have also revealed an increased expression of *MUC5B* in areas with honeycombing in the lungs of patients with IPF [10]. Based on this it can be extrapolated that a link between *MUC5B* and disease pathogenesis exists. It is important to analyze this link as *MUC5B* can potentially serve as a target for drug therapy. Studies done on the role of MUC5B polymorphism and the response to existing treatment modalities such as anti-fibrotic drugs (Pirfenidone and Nintedanib) have shown inconsistent responses, hence more studies are needed to elucidate its prognostic role in relation to treatment [54,60]. However, the identification of MUC5B polymorphisms as a potential biomarker of disease may play a role in a personalized approach to the treatment of IPF; where patients are stratified based on whether

or not they carry the mutation and given a more personalized treatment plan to optimize clinical outcomes and prognosis [61,62].

MUC5B has been extensively studied as a potential risk factor for the development of IPF in several populations including the European, South American, North American, Japanese and Chinese. Interestingly, *MUC5B* polymorphisms demonstrated contrasting changes in pulmonary function testing (PFT) in different ethnic populations. Jiang *et al*. 2015 reported decreased FVC and DLCO along with shorter overall survival in the Chinese Han population; while Stock *et al*. 2013 reported a slower decline in forced vital capacity (FVC) in patients with *MUC5B* polymorphisms among Caucasians in the UK. Similarly, T allele carriage demonstrated differing overall survival. In the Chinese Han population it lead to decreased overall survival and it improved survival among Hispanics [34]. It is possible that the disease severity and survival can be attributed to the different genetic and ethnic makeup of these populations. For example, Hispanic populations demonstrated a greater degree of *MUC5B* polymorphisms when compared to Koreans [34]. However, results in South-Asian, Oceanic, and African populations is lacking. This is important as some studies have shown an absence of *MUC5B* polymorphisms in African patients, although this pool of patients is small [63].

Research on the role of MUC5B polymorphisms on other ILDs is evolving. A strong association between MUC5B polymorphisms and usual interstitial pneumonia (UIP) patterns has been established while the association with non-specific Interstitial pneumonia (NSIP) has shown to be weak [64,65]. Emerging evidence suggest that there may be a link with Connective Tissue Disease-associated ILD (CTD-ILD) and Chronic Hypersensitivity Pneumonitis (CHP) [66]. These studies have mostly been conducted in Caucasian populations particularly North American and European populations. Hence, more larger scale studies with diverse ethnic cohorts are needed to elucidate the role of MUC5B in other ILDs [25,59].

Limitations

Our review has several limitations since only a few studies could be included in the meta-analysis, even though our review identified a wide variety of SNPs. While there was overlap between some of the SNPs, most of them showed little overlap because of which the results could not be replicated. Given that our results were based on genetic epidemiological studies, our outcomes, patient populations, genotyping and conclusions showed great variability as is common for such studies. In addition, the sample size of most of these studies was small. There is a need for more studies with adequate sample sizes exploring the various genes in different context in order to ascertain the actual association of the genetic polymorphism that are

associated with the risk of IPF. These would greatly help in understanding the origins of such conditions and proposing ways to diagnose and manage properly.

Conclusions

Our study marks a significant advancement in the field as it represents the first comprehensive effort to consolidate multiple genetic polymorphisms linked to IPF. By systematically reviewing and synthesizing data from a wide range of studies, we have been able to provide a more comprehensive understanding of the genetic landscape of IPF. This inclusive approach allows for a more nuanced analysis of the complex genetic factors contributing to the development and progression of the disease. Through our meta-analyses, we have identified key genetic markers, such as MUC5B rs35705950, IL-4 rs2243250, IL-4 rs2070874, and TNFα -308, shedding light on their potential roles in IPF susceptibility and pathogenesis. By pooling data from diverse populations and ethnicities, our study not only enhances our understanding of the genetic risk factors associated with IPF but also underscores the importance of considering genetic variability across different demographic groups. Overall, our study represents a significant step forward in elucidating the genetic basis of IPF and lays the groundwork for future research aimed at unraveling the underlying mechanisms of this debilitating disease.

References

1. Nakamura Y, Suda T. Idiopathic pulmonary fibrosis: diagnosis and clinical manifestations. Clin Med Insights Circ Respir Pulm Med 2015;9:163-71.
2. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). Am J Respir Crit Care Med 2000;161:646-64.
3. Raghu G, Remy-Jardin M, Myers JL, et al. Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. Am J Respir Crit Care Med 2018;198:e44-68.
4. Salton F, Ruaro B, Confalonieri P, Confalonieri M. Epithelial–mesenchymal transition: a major pathogenic driver in idiopathic pulmonary fibrosis? Medicina 2020;56:608.
5. Jia Q, Lei Y, Chen S, et al. Circulating inflammatory cytokines and risk of idiopathic pulmonary fibrosis: a Mendelian randomization study. BMC Pulm Med 2023;23:369.
6. Huang D, Gong L, Wu Z, et al. Genetic association of circulating adipokines with risk of idiopathic pulmonary fibrosis: a two-sample Mendelian randomization study. Lung 2023;201:355-62.
7. Park Y, Ahn C, Kim TH. Occupational and environmental risk factors of idiopathic pulmonary fibrosis: a systematic review and meta-analyses. Sci Rep 2021;11:4318.
8. Steele MP, Speer MC, Loyd JE, et al. Clinical and pathologic features of familial interstitial pneumonia. Am J Respir Crit Care Med 2005;172:1146-52.
9. Baratella E, Ruaro B, Giudici F, et al. Evaluation of correlations between genetic variants and high-resolution computed tomography patterns in idiopathic pulmonary fibrosis. Diagnostics 2021;11:762.
10. Borie R, Crestani B, Dieude P, et al. The MUC5B variant is associated with idiopathic pulmonary fibrosis but not with systemic sclerosis interstitial lung disease in the European Caucasian population. PloS One 2013;8:e70621.
11. Moorsel CHMv, Oosterhout MFMv, Barlo NP, et al. Surfactant protein C mutations are the basis of a significant portion of adult familial pulmonary fibrosis in a Dutch cohort. Am J Respir Crit Care Med 2010;182:1419-25.
12. Alder JK, Chen JJ-L, Lancaster L, et al. Short telomeres are a risk factor for idiopathic pulmonary fibrosis. Proc Natl Acad Sci U S A 2008;105:13051-6.
13. Wu X, Li W, Luo Z, Chen Y. The minor T allele of the MUC5B promoter rs35705950 associated with susceptibility to idiopathic pulmonary fibrosis: a meta-analysis. Sci Rep 2021;11:24007.

14. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
15. Zubairi ABS, Hassan M, Ali AS, et al. Risk factors associated with idiopathic pulmonary fibrosis: a systematic review and meta-analyses. PROSPERO. 2020.
16. American Thoracic Society, European Respiratory Society. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. Am J Respir Crit Care Med 2002;165:277-304.
17. Raghu G. Idiopathic pulmonary fibrosis: guidelines for diagnosis and clinical management have advanced from consensus-based in 2000 to evidence-based in 2011. Eur Respir J 2011;37:743-6.
18. Wells AU. The revised ATS/ERS/JRS/ALAT diagnostic criteria for idiopathic pulmonary fibrosis (IPF)--practical implications. Respir Res 2013;14 Suppl 1:S2.
19. Raghu G, Rochwerg B, Zhang Y, et al. An official ATS/ERS/JRS/ALAT clinical practice guideline: treatment of idiopathic pulmonary fibrosis. An update of the 2011 clinical practice guideline. Am J Respir Crit Care Med 2015;192:e3-19.
20. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available from: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
21. Martorell-Marugan J, Toro-Dominguez D, Alarcon-Riquelme ME, Carmona-Saez P. MetaGenyo: a web tool for meta-analysis of genetic association studies. BMC Bioinformatics 2017;18:563.
22. Horimasu Y, Ohshimo S, Bonella F, et al. MUC 5 B promoter polymorphism in J apanese patients with idiopathic pulmonary fibrosis. Respirology 2015;20:439-44.
23. Jiang H, Hu Y, Shang L, et al. Association between MUC5B polymorphism and susceptibility and severity of idiopathic pulmonary fibrosis. Int J Clin Exp Pathol 2015;8:14953-8.
24. Kishore A, Žižková V, Kocourková L, et al. Association study for 26 candidate loci in idiopathic pulmonary fibrosis patients from four European populations. Front Immunol 2016;7:274.

25. Stock CJ, Sato H, Fonseca C, et al. Mucin 5B promoter polymorphism is associated with idiopathic pulmonary fibrosis but not with development of lung fibrosis in systemic sclerosis or sarcoidosis. Thorax 2013;68:436-41.
26. van der Vis JJ, Snetselaar R, Kazemier KM, et al. Effect of M uc5b promoter polymorphism on disease predisposition and survival in idiopathic interstitial pneumonias. Respirology 2016;21:712-7.
27. Vasakova M, Striz I, Slavcev A, et al. Th1/Th2 cytokine gene polymorphisms in patients with idiopathic pulmonary fibrosis. Tissue Antigens 2006;67:229-32.
28. Ahn MH, Park BL, Lee SH, et al. A promoter SNP rs4073T> A in the common allele of the interleukin 8 gene is associated with the development of idiopathic pulmonary fibrosis via the IL-8 protein enhancing mode. Respir Res 2011;12:73.
29. Aquino-Galvez A, Pérez-Rodríguez M, Camarena Á, et al. MICA polymorphisms and decreased expression of the MICA receptor NKG2D contribute to idiopathic pulmonary fibrosis susceptibility. Hum Genet 2009;125:639-48.
30. Aquino-Gálvez A, González-Ávila G, Pérez-Rodríguez M, et al. Analysis of heat shock protein 70 gene polymorphisms Mexican patients with idiopathic pulmonary fibrosis. BMC Pulm Med 2015;15:129.
31. Bournazos S, Bournazou I, Murchison JT, et al. Fcγ receptor IIIb (CD16b) polymorphisms are associated with susceptibility to idiopathic pulmonary fibrosis. Lung 2010;188:475-81.
32. Bournazos S, Bournazou I, Murchison JT, et al. Copy number variation of FCGR3B is associated with susceptibility to idiopathic pulmonary fibrosis. Respiration 2011;81:142-9.
33. Checa M, Ruiz V, Montaño M, et al. MMP-1 polymorphisms and the risk of idiopathic pulmonary fibrosis. Hum Genet 2008;124:465-72.
34. Peljto AL, Selman M, Kim DS, et al. The MUC5B promoter polymorphism is associated with idiopathic pulmonary fibrosis in a Mexican cohort but is rare among Asian ancestries. Chest 2015;147:460-4.
35. Riha R, Yang I, Rabnott G, et al. Cytokine gene polymorphisms in idiopathic pulmonary fibrosis. Intern Med J 2004;34:126-9.
36. Selman M, Lin H-M, Montaño M, et al. Surfactant protein A and B genetic variants predispose to idiopathic pulmonary fibrosis. Hum Genet 2003;113:542-50.
37. Uh ST, Kim TH, Shim EY, et al. Angiotensin-converting enzyme (ACE) gene polymorphisms are associated with idiopathic pulmonary fibrosis. Lung 2013;191:345-51.

38. Uh ST, Jang AS, Park SW, et al. ADAM33 gene polymorphisms are associated with the risk of idiopathic pulmonary fibrosis. Lung 2014;192:525-32.
39. Xaubet A, Fu W, Li M, et al. A haplotype of cyclooxygenase-2 gene is associated with idiopathic pulmonary fibrosis. Sarcoidosis Vasc Diffuse Lung Dis 2010;27:121-30.
40. Zhang J, Xu D, Xu K, et al. HLA-A and HLA-B gene polymorphism and idiopathic pulmonary fibrosis in a Han Chinese population. Respir Med 2012;106:1456-62.
41. Zorzetto M, Ferrarotti I, Trisolini R, et al. Complement receptor 1 gene polymorphisms are associated with idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2003;168:330-4.
42. Bonella F, Campo I, Zorzetto M, et al. Potential clinical utility of MUC5B und TOLLIP single nucleotide polymorphisms (SNPs) in the management of patients with IPF. Orphanet J Rare Dis 2021;16:111.
43. Helling BA, Gerber AN, Kadiyala V, et al. Regulation of MUC5B expression in idiopathic pulmonary fibrosis. Am J Respir Cell Mol Biol 2017;57:91-9.
44. Son JY, Kim SY, Cho SH, et al. TGF-β1 T869C polymorphism may affect susceptibility to idiopathic pulmonary fibrosis and disease severity. Lung 2013;191:199-205.
45. Stock CJ, Conti C, Montero-Fernandez Á, et al. Interaction between the promoter MUC5B polymorphism and mucin expression: is there a difference according to ILD subtype? Thorax 2020;75:901-3.
46. Wang C, Zhuang Y, Guo W, et al. Mucin 5B promoter polymorphism is associated with susceptibility to interstitial lung diseases in Chinese males. PloS One 2014;9:e104919.
47. Wei R, Li C, Zhang M, et al. Association between MUC5B and TERT polymorphisms and different interstitial lung disease phenotypes. Transl Res 2014;163:494-502.
48. Zhang Y, Noth I, Garcia JG, Kaminski N. A variant in the promoter of MUC5B and idiopathic pulmonary fibrosis. N Engl J Med 2011;364:1576-7.
49. Zhang H-P, Zou J, Xie P, et al. Association of HLA and cytokine gene polymorphisms with idiopathic pulmonary fibrosis. Kaohsiung J Med Sci 2015;31:613-20.
50. Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. Nat Genet 2013;45:613-20.
51. Barros A, Oldham J, Noth I. Genetics of idiopathic pulmonary fibrosis. Am J Med Sci 2019;357:379-83.
52. Nalysnyk L, Cid-Ruzafa J, Rotella P, Esser D. Incidence and prevalence of idiopathic pulmonary fibrosis: review of the literature. Eur Respir Rev 2012;21:355-61.

53. Lee MG, Lee YH. A meta-analysis examining the association between the MUC5B rs35705950 T/G polymorphism and susceptibility to idiopathic pulmonary fibrosis. Inflamm Res 2015;64:463-70.
54. Noth I, Zhang Y, Ma SF, et al. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. Lancet Respir Med 2013;1:309-17.
55. Allen RJ, Porte J, Braybrooke R, et al. Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of European ancestry: a genome-wide association study. Lancet Respir Med 2017;5:869-80.
56. Seibold MA, Smith RW, Urbanek C, et al. The idiopathic pulmonary fibrosis honeycomb cyst contains a mucocilary pseudostratified epithelium. PloS One 2013;8:e58658.
57. Chen G, Ribeiro CM, Sun L, et al. XBP1S regulates MUC5B in a promoter variant–dependent pathway in idiopathic pulmonary fibrosis airway epithelia. Am J Respir Crit Care Med 2019;200:220-34.
58. O'Dwyer DN, Ashley SL, Moore BB. Influences of innate immunity, autophagy, and fibroblast activation in the pathogenesis of lung fibrosis. Am J Physiol Lung Cell Mol Physiol 2016;311:L590-601.
59. Seibold MA, Wise AL, Speer MC, et al. A common MUC5B promoter polymorphism and pulmonary fibrosis. N Engl J Med 2011;364:1503-12.
60. Raghu G, van den Blink B, Hamblin MJ, et al. Long-term treatment with recombinant human pentraxin 2 protein in patients with idiopathic pulmonary fibrosis: an open-label extension study. Lancet Respir Med 2019;7:657-64.
61. Hewson T, McKeever TM, Gibson JE, et al. Timing of onset of symptoms in people with idiopathic pulmonary fibrosis. Thorax 2017. doi: 10.1136/thoraxjnl-2017-210177.
62. Cottin V, Hirani NA, Hotchkin DL, et al. Presentation, diagnosis and clinical course of the spectrum of progressive-fibrosing interstitial lung diseases. Eur Respir Rev 2018;27:180076.
63. Putman RK, Gudmundsson G, Araki T, et al. The MUC5B promoter polymorphism is associated with specific interstitial lung abnormality subtypes. Eur Respir J 2017;50:1700537.
64. Kinder BW, Brown KK, McCormack FX, et al. Serum surfactant protein-A is a strong predictor of early mortality in idiopathic pulmonary fibrosis. Chest 2009;135:1557-63.
65. Todd NW, Scheraga RG, Galvin JR, et al. Lymphocyte aggregates persist and accumulate in the lungs of patients with idiopathic pulmonary fibrosis. J Inflamm Res 2013;6:63-70.
66. Kojima M, Kojima T, Suzuki S, et al. Depression, inflammation, and pain in patients with rheumatoid arthritis. Arthritis Rheum 2009;61:1018-24.

Online supplementary material:

Supplementary Table 1. Search strategy for systematic review.

Supplementary Table 2. Characteristics of the included studies.

Supplementary Table 3. Association between SNPs and IPF mentioned in the included studies.

Supplementary Table 4. Quality assessment of studies using Newcastle-Ottawa scale

Supplementary Table 5. Genotypic distribution and HWE of IPF and non-IPF subjects for a) *MUC5B rs35705950*, b) *IL-4 rs2243250*, c) *IL-4 rs2070874*, and d) *TNFa -308*

Supplementary Table 6. Association findings of a) *MUC5B rs35705950*, b) *IL-4 rs2243250*, c) *IL-4 rs2070874*, and d) *TNFa* -308 and IPF using random effects model

Supplementary Figure 1. Forest Plot demonstrating association between risk of IPF genotypic models of a) *MUC5B rs35705950*, b) *IL-4 rs2243250*, c) *IL-4 rs2070874*, and d) *TNFa -308*.

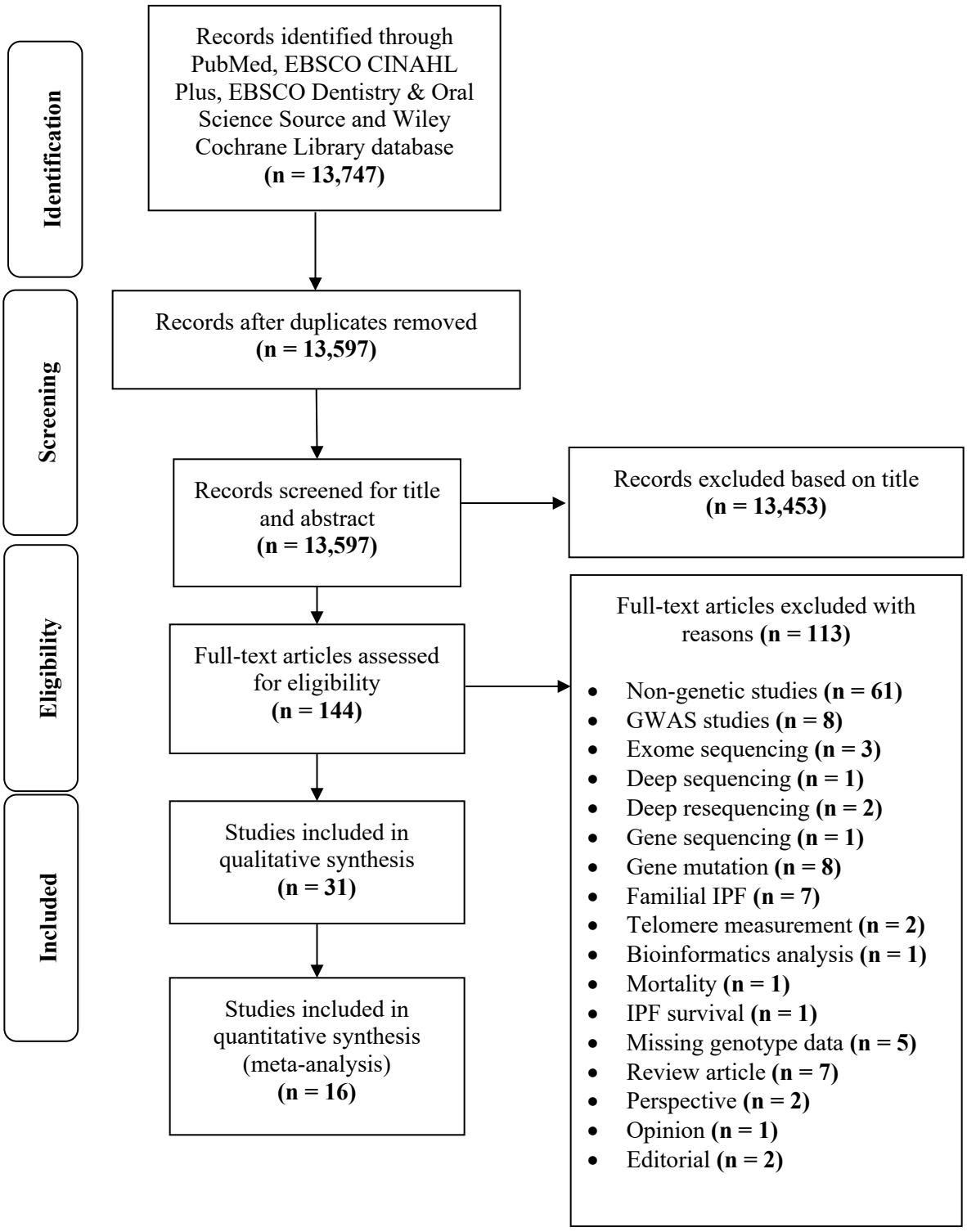
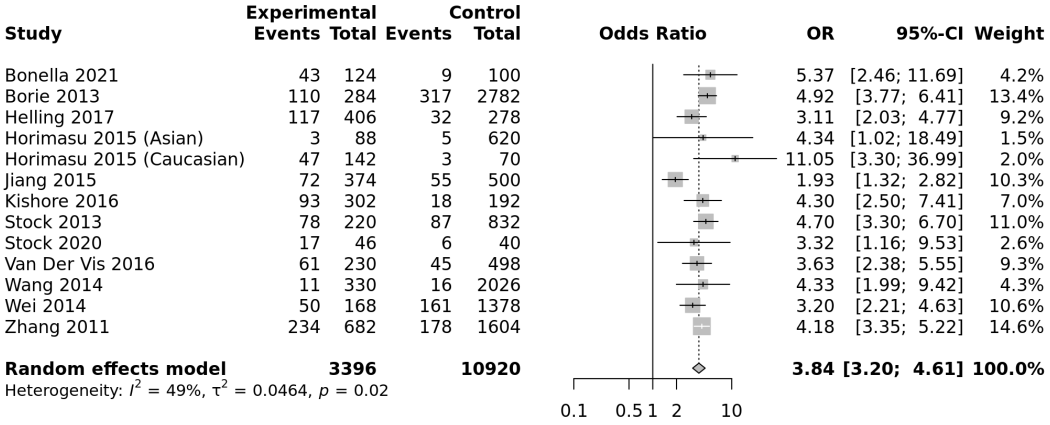


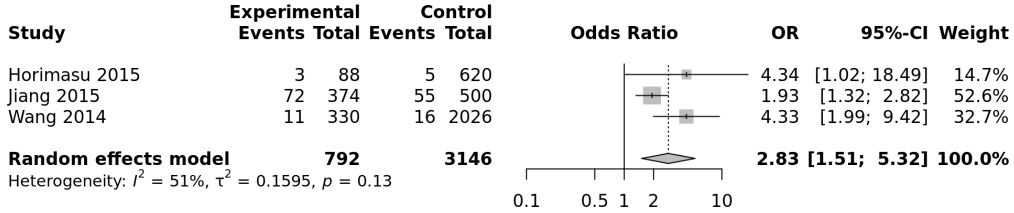
Figure 1. PRISMA flow diagram of study identification, screening, eligibility, and inclusion.

a)

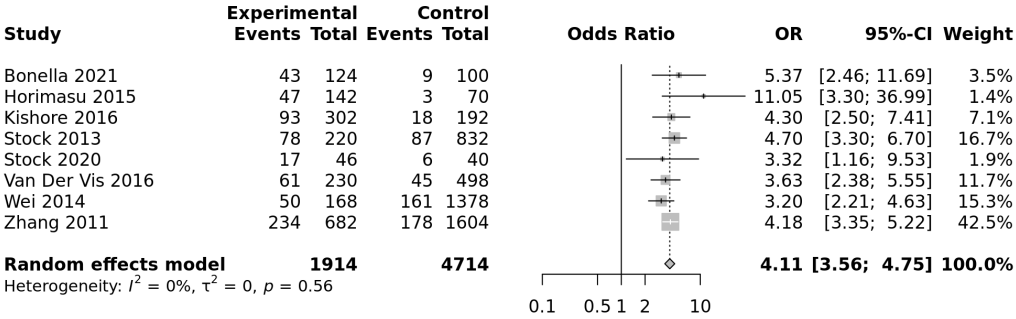
Allele Contrast (T vs. G)



Allele Contrast (T vs. G), Asian

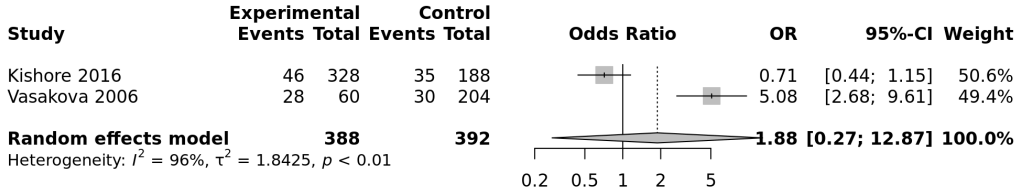


Allele Contrast (T vs. G), Caucasian



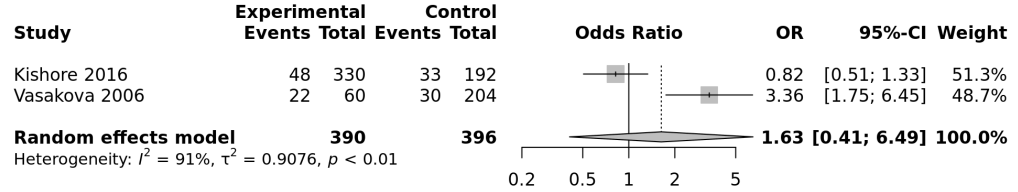
b)

Allele contrast (T vs. C)



c)

Allele contrast (T vs. C)



d)

Allele contrast (A vs. G)

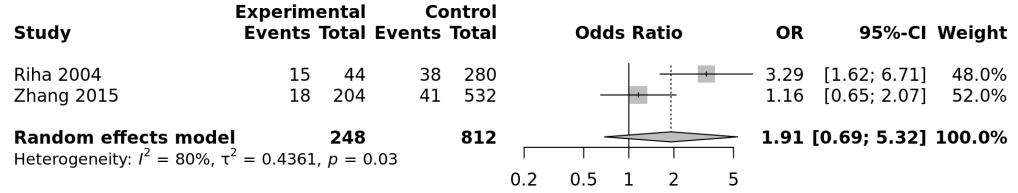


Figure 2. Forest plot demonstrating association between risk of IPF and allelic models of a) *MUC5B rs35705950*, b) *IL-4 rs2243250*, c) *IL-4 rs2070874*, and d) *TNFa -308*. a) *MUC5B rs35705950*; b) *IL-4 rs2243250;* c) *IL-4 rs2070874;* d) *TNFa -308*.

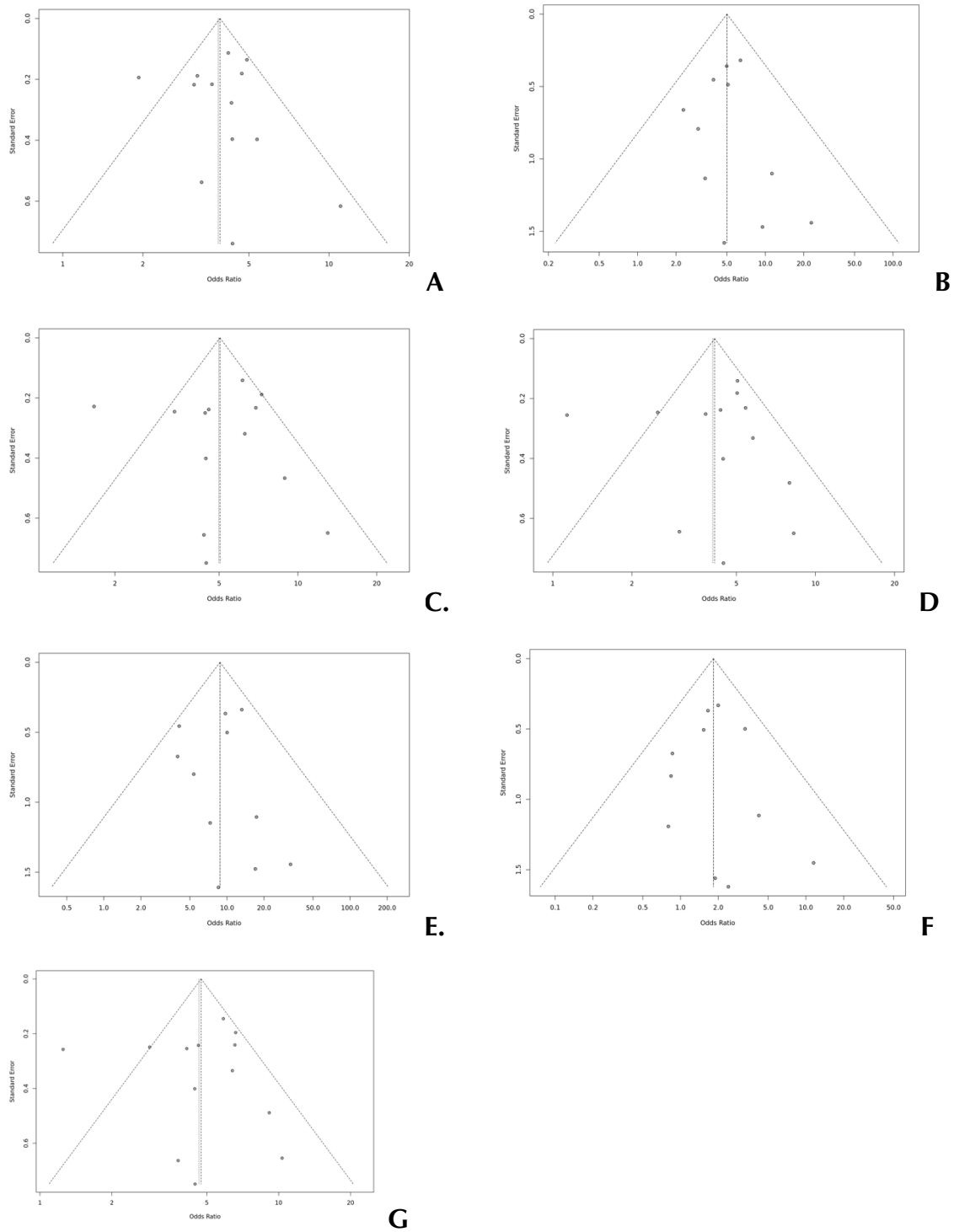


Figure 3. Inverted funnel plots for *MUC5B rs35705950*. A) Allele Contrast (T vs. G), p=0.8858; B) Recessive model (TT vs. TG+GG), p=0.7227; C) Dominant model (TT+TG vs. GG), p=0.9878; D) Overdominant model (TG vs. TT+GG), p=0.9236; E) Homozygote codominant model (TT vs. GG), p=0.9548; F) Heterozygote codominant model (TT vs. TG), p=0.9347; G) Heterozygote codominant model (TG vs. GG), p=0.8867