Epidermal Growth Factor Receptor (EGFR) Pathway Genes and Interstitial Lung Disease: An Association Study

Chong Li1,2, Rongrong Wei1, Yava L. Jones-Hall3, Ragini Vittal4, Min Zhang5 & Wanqing Liu1,6

1Department of Medicinal Chemistry and Molecular Pharmacology, College of Pharmacy, Purdue University, West Lafayette, Indiana, 2Department of Respiratory Medicine, the Third Affiliated Hospital of Suzhou University, Changzhou, Jiangsu, China, 3Department of Comparative Pathobiology, College of Veterinary Medicine; Purdue University, West Lafayette, Indiana, 4Center for Immunobiology, Indiana University School of Medicine, Indianapolis, Indiana, 5Department of Statistics, Purdue University, West Lafayette, Indiana, 6Purdue University Center for Cancer Research, Purdue University, West Lafayette, Indiana.

The etiology and pathogenesis of idiopathic interstitial lung disease (ILD) remain incompletely understood. Genetic susceptibility to ILD has been demonstrated in previous studies. It is well known that EGFR inhibitors can induce ILD in human lung cancer patient with ethnic differences, which prompted us to hypothesize that genetic variation in EGFR pathway genes confer susceptibility to ILD. We aimed in this study to investigate whether functional polymorphisms of EGFR and its ligands genes (EGF and TGFA) were associated with ILD. Three EGFR [216G/T (rs712830), 191A/C (rs712829), 497R.K(A/G) (rs2227983)], one EGF [61A/G, (rs4444903)] and one TGFA (rs3821262C/T) polymorphisms previously demonstrated to alter gene functions were genotyped in 229 sporadic idiopathic ILD patients and 693 normal healthy individuals. Allelic and genotypic association tests between these polymorphisms and ILD were performed. The EGF 61A/G polymorphism was significantly associated with elevated risk of ILD, with the frequency of G allele significantly increased in the ILD patient population (OR = 1.33, 95%CI = 1.07–1.66, P = 0.0099). None of the other polymorphisms were associated with risk of ILD. Our study suggested that the EGF 61A/G polymorphism may be associated with sporadic ILD. While a false positive finding cannot be excluded, independent studies are warranted to further validate this result.

Interstitial lung disease (ILD) refers to a broad range of chronic lung disorders with diverse pathogenesis and complex histopathology, together accounting for 15% of respiratory care practice. Most entities are manifested as epithelial injury, followed by fibroblastic proliferation and development of fibroblastic foci with exuberant deposition of matrix - typical hallmarks of pulmonary fibrosis. Over two thirds of ILD cases do not have a known cause and are thus named idiopathic interstitial pneumonia (IIP). Although the incidence of ILD in the US is low (approximately 30 cases per 100,000 persons per year), the disease can be progressive and fatal. The mean survival time of ILD patients is only about 3 years. The etiology and pathogenesis of most ILD entities remain unknown, thus greatly hampering progress in the development of therapeutics for the disease. To date, no proven drug therapy for most entities has been recognized.

It is now widely accepted that the development of ILD has a strong genetic basis. Substantial evidence demonstrates that ILD is a heritable complex disease determined by genetic factors with involvement of environmental stimuli, such as tobacco smoke. Family-based studies have been conducted in an attempt to identify genes predisposing to ILD, and causal mutations have been identified in several genes, e.g. telomerase reverse transcriptase gene (TERT), the telomerase RNA component gene (TERC), surfactant proteins A2 (SPA2) and C (SPC) genes. More recently, a few genome-wide association studies have identified a number of single nucleotide polymorphisms (SNP) located at or close to TERT, TERC, MUC5B, FAM13A, DSP, OBFC1, ATP11A, DPP9, TOLLIP and SPPL2C genes significantly (P < 5 x 10^-8) associated with IPF and/or IIP as an overall phenotype. However, these polymorphisms together were estimated to account for about only one third of the risk of IIP, suggesting additional genetic component yet to be identified.

The epidermal growth factor receptor (EGFR) is a tyrosine kinase receptor for various growth factors including EGF (epidermal growth factor), TGF-α (transforming growth factor-α) and other EGF-like ligands. The EGFR pathway plays an important role in pulmonary physiology especially the function of epithelial cells via signaling transduction that regulates key cellular processes such as self-renew, wound-healing, proliferation, survival,
adhesion, migration and differentiation. EGFR inhibitors have been widely used in treatment of non-small cell lung cancer (NSCLC). However, ILD has been consistently reported as one of the uncommon but severe adverse reactions of EGFR inhibitors. A strong association between the incidence of ILD and anti-EGFR treatments has been reported in a large case-cohort study that included over 4,000 subjects. The study showed a 3.23-fold increase in risk of ILD in patients who received gefitinib when compared with those who underwent conventional chemotherapy. Furthermore, significant inter-ethnic differences in the incidence of ILD in patients treated with EGFR inhibitors has been consistently observed. According to the U.S. Food and Drug Administration (FDA), an overall ILD incidence of 1% was demonstrated in 50,005 patients receiving gefitinib, including 18,960 patients from Japan and 23,000 from the U.S. Interestingly, the incidence of ILD was higher in Japanese patients (1.7%) compared to patients from the US (0.3%). There was also a significant difference in the median time to onset (TTO) of ILD between Japanese and U.S. patients. The TTO was about 24 days in the former but around 42 days in the latter. These findings have been confirmed in other independent studies. Taken together, these observations suggest that certain genetic factors related to the EGFR pathway may confer susceptibility to ILD in general. In order to corroborate this hypothesis, we set out in this study to test the genetic association between functional polymorphisms in EGF, EGF and TGFA genes and ILD. These polymorphisms have been previously demonstrated to alter gene expression, function or other related phenotypes in our and other studies.

Methods

Ethics statement. Research conducted in this study was performed on anonymous adult individuals without intervening with patients, and is therefore not considered to involve ‘human subjects’. Samples were collected with written informed consent obtained from participants with approval of institutional review boards (IRBs) at the Lung Tissue Research Consortium (LTRC, http://www.ltrcpublic.com) and The University of Chicago. The study was carried out in accordance with the approved guidelines by the Purdue University IRB (approval number 1307013815) and was in compliance with the Helsinki Declaration.

Study population. DNA extracted from peripheral blood of ILD patients (n = 227) were obtained from the Lung Tissue Research Consortium (LTRC, http://www.ltrcpublic.com) and The University of Chicago. These samples came from patients regularly visiting the clinic, excluding individuals with any respiratory symptoms according to the ICD-9 classification. All control patients are self-reported Caucasians. Demographic data for ILD patients and controls were summarized in Table 1.

| Disease status | Gender | Age | Current | Ever | Never | BMI |
|---------------|--------|-----|---------|------|-------|-----|
|               | Total  |     | Mean ± SD |     |       | Mean ± SD |
| UIP/IFD       | 84     | 55  | 64.4 ± 7.7 | 1   | 52    | 31   | 30.3 ± 5.2 |
| NSIP          | 27     | 10  | 59.1 ± 8.4 | 1   | 15    | 11   | 30.5 ± 6.8 |
| DIP           | 9      | 7   | 54.1 ± 7.7 | 3   | 4     | 2    | 31.3 ± 8.7 |
| RB-ILD        | 22     | 13  | 58.1 ± 11.5 | 2  | 17    | 3    | 32.5 ± 7.1 |
| COP           | 10     | 7   | 62.2 ± 11.9 | 2  | 6     | 2    | 29.0 ± 6.7 |
| HP            | 8      | 1   | 52.0 ± 16  | 0   | 1     | 7    | 31.7 ± 5.7 |
| LIP           | 10     | 7   | 49.9 ± 11  | 0   | 0     | 0    | 27.4    |
| AIP           | 9      | 6   | 74         | 0   | 0     | 0    | 26.6    |
| UF            | 67     | 40  | 64.8 ± 10.4 | 1  | 45    | 21   | 29.4 ± 5.6 |
| Control       | 693    | 360 | 55.7 ± 13.2 | -  | -     | -    | -       |

SD, standard deviation; BMI, body mass index.

Results

We used a case-control study design to examine whether polymorphisms of EGF and its ligand genes were associated with susceptibility to ILD. The observed genotype frequencies of these SNPs were all in agreement with the HWE in the control subjects (df = 1, P > 0.05 for all tests, data not shown). The genotype and allele distributions of the five SNPs between the cases and controls are summarized in Table 2.

| Genotype | OR (95% CI) |
|----------|-------------|
| G/A      | 1.07–1.66   |
| G/G      | 1.07–1.66   |
| A/A      | 1.07–1.66   |

Discussion

We observed a statistically significant association between an EGF polymorphism that previously demonstrated to alter EGF gene expression and ILD. While the association was weak, the results...

Table 1 | Demographic and covariates data associated with ILD patients and controls

| Disease status | Total number | Gender | Male [N] | Age | Cigarette Smoking | BMI | SD |
|---------------|--------------|--------|----------|-----|------------------|-----|-----|
|               |              |        |          |     |                  |     |     |
| UIP/IFD       | 84           | 55     |          |     |                  |     |     |
| NSIP          | 27           | 10     |          |     |                  |     |     |
| DIP           | 9            | 7      |          |     |                  |     |     |
| RB-ILD        | 22           | 13     |          |     |                  |     |     |
| COP           | 10           | 7      |          |     |                  |     |     |
| HP            | 8            | 1      |          |     |                  |     |     |
| LIP           | 10           | 7      |          |     |                  |     |     |
| AIP           | 9            | 6      |          |     |                  |     |     |
| UF            | 67           | 40     |          |     |                  |     |     |
| Control       | 693          | 360    |          |     |                  |     |     |
epithelial cells were shown to develop PF phenotypes. On the expression of TGF-
prevented in another bitransgenic mouse model with constitutive
SCIENTIFIC
REPORTS
higher in EGFRi treatment than chemotherapy. Meanwhile, pre-
Although induction of ILD was observed with many drugs, it was
demonstrated to induce ILD in lung cancer patients.
variation in the EGFR pathway may confer risk to general ILD. This result can be further replicated, it would indicate that genetic
result lacked further validation with additional sample sets.
Unfortunately, our
for age and gender. The Cochran-Armitage test for trend also sug-
ing induces ILD. However, this plausibly contradicts the
for EGFR inhibition may at least in part contribute to the natural history of ILD. Interestingly, a recent GWAS observed that two polymorphisms (rs79842896 and rs76795398) close to (~33 kb downstream) the EGFR gene were top-ranked to possess associations with IPF (P < 10^-7), although this association did not reach the genome-wide significance. Continued study on more EGFR polymorphisms is thereby necessary to further clarify the role of EGFR variation in ILD.

The polymorphisms tested in this study have been proven to be functional in determining the expression or activity of EGFR or its ligands in both our and others' studies. We have demonstrated that the two EGFR polymorphisms -216G/T and -191C/A were associated with increased EGFR promoter activity and gene/protein expression. They were also shown to be associated with drug-induced toxicities e.g. skin rash in EGFR inhibitor treatment in our previous study. A SNP at EGFR codon 497 results in an Arg (R) to Lys (K) substitution, which has been associated with decreased EGFR activity. The TGFA intronic polymorphism rs3821262C/T was also found to be associated with TGFA gene expression as well as sensitivity to EGFR inhibitors in cancer cell lines in our previous study. With regard to EGF 61A/G polymorphism, it was initially demonstrated to affect EGF protein expression, with the G allele associated with a higher EGF level relative to the A allele. A recent meta-analysis of 41 case-control studies on various cancers have shown that the G allele was significantly associated with increased cancer risk. It is thus possible that the relatively lower EGF level associated with the A allele might be the reason underlying its association with ILD in our study. This supports the notion that a relatively lower EGFR-axis activity might be a risk factor for ILD, which indicates a possible additive effect of the risk allele.

It should be further noted that previous GWAS did not identify this locus. While most previous GWAS were focused on IPF rather than EGFR inhibition, our study included patients with a variety of ILDs, suggesting that the risk allele may be specific to ILDs.
than ILD as a phenotype, it might also be important due to the population difference. It is commonly observed that many loci identified in GWAS actually exert different effect size in different populations. Nevertheless, without independent validation, a false positive result in the association observed in our study could not be excluded, in particular that our study was limited by the relatively small sample size. Therefore, our findings should be used and interpreted with caution.

Conclusions
Our study provides a new investigation of the relationship between functional EGFR pathway gene polymorphisms and risk of ILD. The findings suggested a possible association between EGF 61A/G polymorphism and ILD. Further validation of this genetic association in independent sample sets is warranted.

Acknowledgments
We thank the Lung Tissue Research Consortium (http://www.ltrcpublic.com) and the National Cancer Institute (R03 HL105838) for providing the ILD patients and control samples, respectively. This study was supported in part by the NIH/NHLBI grant (R03 HL097016) (W.L.) and the University of Chicago for providing the ILD patients and control samples, respectively. We thank the Lung Tissue Research Consortium (http://www.ltrcpublic.com) and the National Cancer Institute (R03 HL105838) for providing the ILD patients and control samples, respectively. This study was supported in part by the NIH/NHLBI grant (R03 HL105838) (W.L.) and the University of Chicago for providing the ILD patients and control samples, respectively.
Author contributions
C.L. and R.W. conducted the experiments and participated in the manuscript writing. Y.L.J.-H. participated in the data interpretation and manuscript writing. R.V. participated in the study design and data interpretation. M.Z. analyzed the data. W.L. conceived the study and wrote the main text. All authors reviewed the manuscript.

Additional information
Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Li, C. et al. Epidermal Growth Factor Receptor (EGFR) Pathway Genes and Interstitial Lung Disease: An Association Study. Sci. Rep. 4, 4893; DOI:10.1038/srep04893 (2014).

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. The images in this article are included in the article’s Creative Commons license, unless indicated otherwise in the image credit; if the image is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the image. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/