Epidermal Growth Factor Receptor Mutation Status in Circulating Free DNA in Serum

From IPASS, a Phase III Study of Gefitinib or Carboplatin/Paclitaxel in Non-small Cell Lung Cancer

Koichi Goto, MD, PhD, Yukito Ichinose, MD,† Yuichiro Ohe, MD, PhD,‡ Nobuyuki Yamamoto, MD, PhD,§ Shunichi Negoro, MD,‖ Kazuto Nishio, MD, PhD,¶ Yohji Itoh, PhD,# Haiyi Jiang, MD,# Emma Duffield, MSc,** Rose McCormack, PhD,** Nagahiro Saijo, MD, PhD,¶ Tony Mok, MD,†† and Masahiro Fukuoka, MD, PhD¶*

Introduction: In IPASS (IRESSA Pan-Asia Study), clinically selected patients with pulmonary adenocarcinoma received first-line gefitinib or carboplatin/paclitaxel. This preplanned, exploratory analysis was conducted to increase understanding of the use of surrogate samples, such as serum, versus tumor biopsy samples for determining *EGFR* mutation status in the Japanese cohort ($n = 233$).

Methods: *EGFR* mutations were assessed using tumor tissue-derived DNA ($n = 91$) and circulating free (cf) DNA from pretreatment serum samples ($n = 194$).

Results: Fewer patients were *EGFR* mutation positive when assessed using pretreatment cfDNA (23.7%) versus tumor tissue-derived DNA (61.5%). cfDNA results identified no false positives but a high rate of false negatives (56.9%). There was a significant interaction between cfDNA *EGFR* mutation status and treatment for progression-free survival (PFS) ($p = 0.045$). PFS was significantly longer and objective response rate (ORR) higher with gefitinib than carboplatin/paclitaxel in the cfDNA *EGFR* mutation-positive subgroup (PFS: hazard ratio [HR], 0.29; 95% confidence interval [CI], 0.14–0.60; $p < 0.001$; ORR: odds ratio [OR], 1.71; 95% CI, 0.48–6.09; 75.0% versus 63.6%; $p = 0.40$). There was a slight numerical advantage in PFS and ORR for gefitinib over carboplatin/paclitaxel in the cfDNA *EGFR* mutation-negative subgroup, likely due to the high rate of false negatives within this subgroup.

Conclusions: These results merit further investigation to determine whether alternative sources of tumor DNA, such as cfDNA in serum, could be used for determining *EGFR* mutation status in future; currently, where a sample is available, analysis of tumor material is recommended.

Key Words: EGFR, Mutation, Gefitinib, NSCLC, Serum.

(*J Thorac Oncol.* 2012;7: 115–121)

*National Cancer Center Hospital East, Chiba; †National Kyushu Cancer Center, Fukuoka; ‡National Cancer Center Hospital, Tokyo; §Shizuoka Cancer Center, Shizuoka; ‖Hyogo Cancer Center, Hyogo; ¶Kinki University School of Medicine, Osaka; #AstraZeneca, Osaka, Japan; **AstraZeneca, Macclesfield, United Kingdom; and ††The Chinese University of Hong Kong, Shatin, Hong Kong.

Disclosure: K. Goto has received honoraria from AstraZeneca, Chugai Pharmaceutical Co., Ltd, and Ono Pharmaceutical Co., Ltd; S. Negoro and K. Nishio have received honoraria from AstraZeneca; Y. Itoh, H. Jiang, E. Duffield, and R. McCormack are employees of AstraZeneca and hold stock in the company; M. Fukuoka has received honoraria from AstraZeneca, Chugai Pharmaceutical Co., Ltd, Daiichi-Sankyo, and Ono Pharmaceutical Co., Ltd; T. Mok has acted as a consultant to AstraZeneca, Roche, Eli Lilly, Taiho Pharmaceutical Co., Ltd, Merck Serono, Boehringer Ingelheim, AVEO Pharmaceuticals Ltd, and Pfizer and has received honoraria from AstraZeneca, Roche, Eli Lilly, Boehringer Ingelheim, and Merck Serono and research funding from AstraZeneca.

Address for correspondence: Koichi Goto, Division of Thoracic Oncology, National Cancer Center Hospital East, Kashiwanoha, 6-5-1, Kashiwa, Chiba 277-8577, Japan. E-mail: kgoto@east.ncc.go.jp

Yuichiro Ohe is currently at National Cancer Center Hospital East, Chiba, Japan.

Presented in part at the 50th Annual Meeting of the Japan Lung Cancer Society, Tokyo, Japan, November 12–13, 2009; the European Organization for Research and Treatment of Cancer, the National Cancer Institute, and the American Society of Clinical Oncology (EORTC-NCI-ASCO) Annual Meeting, Brussels, Belgium, October 15–17, 2009; the 13th World Conference on Lung Cancer, July 31–August 4, 2009, San Francisco, CA, USA; the 49th Annual Meeting of the Japanese Respiratory Society, June 12–14, 2009; and the 45th Annual Meeting of the American Society of Clinical Oncology, May 29–June 2, 2009, Orlando, FL, USA.

ISSN: 1556-0864/12/0701-0115

The epidermal growth factor receptor (EGFR) superfamily has been implicated in the regulation of tumor cell biology and, as such, has emerged as a therapeutic target.[1] In 2004, mutations in the *EGFR* were reported to be associated with sensitivity to EGFR tyrosine kinase inhibitors (EGFR-TKIs).[2–4] The presence of such mutations in tumor tissue is associated with a number of clinical factors including Asian origin, female sex, adenocarcinoma histology, and a never-smoking history, and these factors have additionally been correlated with response to gefitinib (IRESSA, AstraZeneca, Macclesfield, UK), an EGFR-TKI.[5]

The IRESSA Pan-Asia Study (IPASS) compared gefitinib with carboplatin/paclitaxel as first-line treatment in 1217 never-smokers/light ex-smokers with advanced adenocarcinoma of the lung in East Asia.[6] Subgroup analysis of patients with *EGFR* mutations ($n = 261$) detected in DNA derived from tumor tissue samples demonstrated significantly longer progression-free survival (PFS) with gefitinib versus carboplatin/paclitaxel (hazard ratio [HR], 0.48; 95% confidence interval [CI], 0.36–0.64; $p < 0.001$).[6] In the *EGFR* mutation-negative (M−) subgroup ($n = 176$), PFS was significantly longer with carboplatin/paclitaxel versus gefitinib (HR, 2.85; 95% CI, 2.05–3.98; $p < 0.001$). Objective response rates (ORR) were 71.2% versus 47.3% ($p < 0.001$) and 1.1% versus 23.5% ($p = 0.001$) with gefitinib versus carboplatin/paclitaxel in *EGFR* M+ and M− patients, respectively.

The difficulties of collecting sufficient tumor tissue for biomarker analyses have stimulated interest in analyses using surrogate samples, such as serum and plasma samples, which frequently contain circulating free (cf) DNA derived from tumor tissues. Previous studies in relatively few patients had detected *EGFR* mutations in cfDNA in serum or plasma samples and suggested that using such methodology to predict response to gefitinib was worthy of further evaluation.[7–12] However, most of these studies were retrospective.

Herein, we report the evaluation of *EGFR* mutations in cfDNA from serum samples of patients in the IPASS study recruited in Japan. This preplanned, exploratory analysis was conducted to increase the understanding of the use of surrogate samples, such as serum, versus tumor biopsy samples for determining *EGFR* mutation status.

MATERIALS AND METHODS

Study Design and Patients

Full details of the IPASS study design (ClinicalTrials.gov identifier NCT00322452) have been published previously.[6] Planned objectives of this substudy of IPASS were evaluations of efficacy between the gefitinib and carboplatin/paclitaxel treatment groups by cfDNA *EGFR* mutation status from pretreatment serum samples and evaluation of the concordance between *EGFR* mutation status in pretreatment cfDNA versus tumor. Comparison of *EGFR* mutation status in pretreatment versus postprogression serum samples was also performed; however, not all patients with a pretreatment sample had a postprogression sample, which limited the comparison. In addition, comparisons with postprogression serum and pretreatment pleural effusion samples are reported in Supplemental Digital Content 1 (Methods http://links.lww.com/JTO/A152). Preplanned analysis of the Japanese subset of the IPASS population was performed to meet Japanese regulatory requirements.

All patients provided written informed consent. Provision of samples for biomarker research was optional and involved separate consent procedures for tumor and serum sampling. An independent ethics committee at each participating institution approved the study protocol. The study was conducted in accordance with the Declaration of Helsinki, the International Conference on Harmonisation Guidelines for Good Clinical Practice, applicable regulatory requirements, and AstraZeneca's policy on bioethics.

Biomarker Analyses

Sample collection and DNA extraction are described in Supplemental Digital Content 1 (Methods http://links.lww.com/JTO/A152). *EGFR* mutations were detected using the DxS *EGFR* Mutation Test Kit for Research Use Only (DxS, Manchester, UK), which combines Amplification Refractory Mutation System (ARMS) (allele-specific polymerase chain reaction [PCR]) with the Scorpions real-time PCR technology.[13,14] Modified run conditions and cutoffs (delta Ct values [dCt]) used to define M+ samples for cfDNA derived from serum and pleural effusion samples were as follows: 50 cycles of PCR were carried out and the dCt for exon 19 deletions was 12, L858R was 14, and T790M was 8 (for tumor DNA, 40 cycles of PCR were carried out and the dCt cutoffs were 9, 11, and 8, respectively). In analyses of tumor DNA, all 29 mutations detected by the kit were assayed (19 deletions in exon 19, L858R, T790M, L861Q, G719X [S, A, or C], S768I, and 3 insertions in exon 20); whereas for serum and pleural effusion samples, the 21 most common mutations (19 deletions in exon 19, L858R, and T790M) were assayed (to make the best use of limited cfDNA yield). Samples were tested in duplicate, and only if both replicates were positive for at least one of the mutations was the sample defined as M+. Patients without a tumor sample evaluable for mutation analysis and samples which were not successfully analyzed were classified as *EGFR* mutation unknown. Biomarker samples were assayed blinded to clinical outcome and randomized treatment.

Statistical Analyses

Serum samples were collected for patients recruited in Japan and who consented to this optional analysis. Analyses of efficacy end points comparing treatment groups in the Japanese subset (intent-to-treat [ITT] population) were assessed as described previously for the overall IPASS population.[6] However, for the analyses in the cfDNA M+ and M− subgroups, the prespecified covariates of World Health Organization (WHO) performance status (PS), smoking history, and sex could not be included as covariates because of the small number of patients who had a WHO PS 2, were ex-smokers, or were males; therefore, models without covariates were used. Because of the lack of power to detect treatment differences, the result of the Japanese subset should be interpreted with caution, taking into account the associated variability and overlap in plausible range of effects (CIs). Analyses comparing treatment groups were performed for PFS (by Cox proportional hazards model) and ORR (by logistic regression model) in subgroups defined by cfDNA *EGFR* mutation status. A test for interaction between cfDNA *EGFR* mutation status (M+ or M−) and treatment was used to assess whether the PFS treatment effect was statistically different between subgroups.

Comparison of pretreatment cfDNA versus tumor *EGFR* mutations was based on the 21 mutations analyzed for cfDNA using patients with known mutation status (M+ or M−) in both samples. The sensitivity, specificity, positive

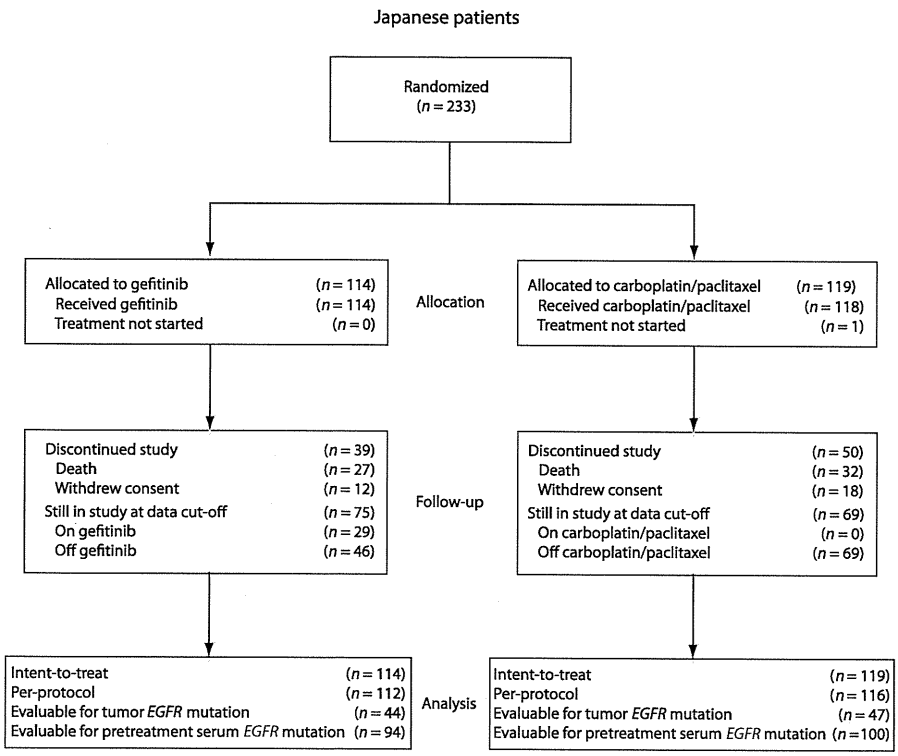


FIGURE 1. CONSORT diagram representing patient disposition (including number of patients with tumor tissue or serum evaluable for *EGFR* mutation status). *EGFR*, epidermal growth factor receptor.

and negative predictive values and their exact 95% CIs, and the kappa coefficient and 95% CI, for *EGFR* mutation status in serum samples, were evaluated assuming that the *EGFR* mutation status in tumor tissue was a true reflection of tumor biology. The proportion of concordance between cfDNA and tumor was calculated on a similar basis by excluding patients judged as unknown using either cfDNA or tumor samples.

RESULTS

Patients

In total, 233 patients from Japan were randomized to study treatment (19.1% of the overall IPASS population). Preplanned evaluations of efficacy, quality of life, and safety for the overall Japanese study population have been previously presented[15,16] and are summarized in Supplemental Digital Content 2 (Results http://links.lww.com/JTO/A153) and 3 (Figure http://links.lww.com/JTO/A154). The patient disposition for the Japanese subset of IPASS is shown in Figure 1.

EGFR Mutation Status

An evaluable DNA sample for *EGFR* mutation status derived from tumor tissue was available for 91 patients; of these, 56 (61.5%) patients were *EGFR* M+, with a lower proportion of *EGFR* M+ patients in the gefitinib group compared with the carboplatin/paclitaxel group (52.3% [23/44] versus 70.2% [33/47]) (Figure 2). A total of 194 patients provided a pretreatment serum sample for mutation analysis; all were evaluable. Of these, 46 (23.7%) patients were cfDNA *EGFR* M+ (25.5% [24/94] and 22.0% [22/100] in the gefitinib and carboplatin/paclitaxel groups, respectively) (Figure 2). Data from pretreatment pleural effusion (9 patients) and postprogression serum analyses (144 patients) are presented in Supplemental Digital Content 2 (Results http://links.lww.com/JTO/A153) and 4 (Table http://links.lww.com/JTO/A155).

Demographic and Baseline Characteristics of Patients with Known *EGFR* Mutation Status

Key demographic and baseline characteristics for patients with known (i.e., evaluable) cfDNA or tumor *EGFR* mutation status were generally consistent with the overall Japanese study population (Table 1).

Pretreatment cfDNA *EGFR* Mutation Status and Clinical Outcomes

The subset of patients with known cfDNA *EGFR* mutation status could be assumed to be representative of the overall Japanese study population (and therefore the overall study population) as shown by similar PFS and ORR results (Table 1).

A significant interaction between cfDNA *EGFR* mutation status and treatment was evident for PFS (interaction test $p = 0.045$). PFS was significantly longer with gefitinib than carboplatin/paclitaxel in the cfDNA *EGFR* M+ subgroup (HR, 0.29; 95% CI, 0.14–0.60; $p < 0.001$) (Figure 3*A*). In the cfDNA *EGFR* M− subgroup, there were no significant differences for PFS with gefitinib compared with carboplatin/paclitaxel (HR, 0.88; 95% CI, 0.61–1.28; $p = 0.50$) (Figure 3*B*). However, the HR was not constant over time. We

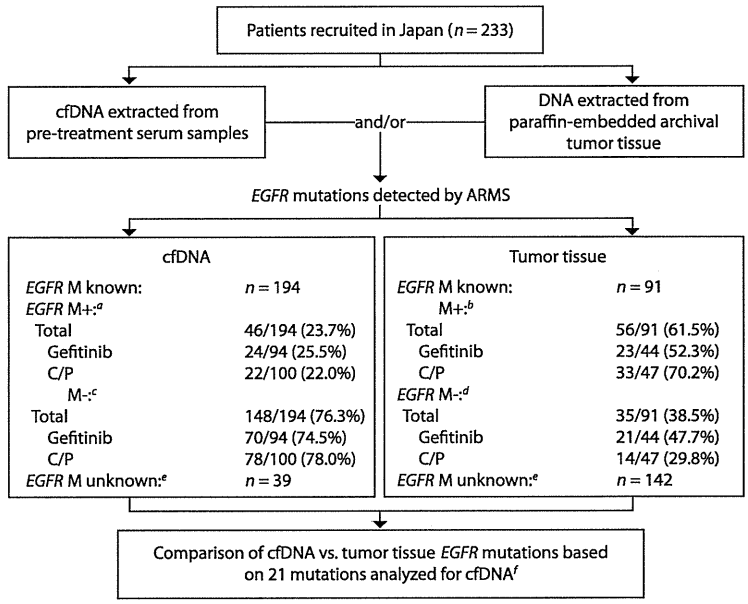


FIGURE 2. Flow and results of *EGFR* mutation analysis. [a]Sample positive for ≥1 of 21 mutations tested; detected 19 deletions in exon 19, L858R, and T790M. [b]Sample positive for ≥1 of 29 mutations tested; detected 19 deletions in exon 19, L858R, T790M, L861Q, G719S, G719A, G719C, S768I; 3 insertions in exon 20. [c]Sample negative for all 21 mutations tested. [d]Sample negative for all 29 mutations tested. [e]Unknown *EGFR* mutations: no sample available or failed analysis. [f]86 patients had known mutation status by both tumor tissue and cfDNA. C/P, carboplatin/paclitaxel; *EGFR*, epidermal growth factor receptor; M, mutation; M+, mutation-positive; M−, mutation-negative.

TABLE 1. Patient Demographics, Baseline Characteristics, and Efficacy (PFS and ORR) for Patients with Samples (cfDNA or Tumor) Evaluable for *EGFR* Mutation Status Compared with the Overall Japanese[a] Study Population (Japanese ITT Population)

	Evaluable for *EGFR* Mutation Status (cfDNA) (n = 194)[b]	Evaluable for *EGFR* Mutation Status (Tumor) (n = 91)[b]	Overall Japanese Study Population (n = 233)
Demography, n (%)			
Female	172 (88.7)	84 (92.3)	204 (87.6)
WHO PS 0/1	185 (95.4)	89 (97.8)	223 (95.7)
Never-smoker	177 (91.2)	83 (91.2)	212 (91.0)
Stage IIIB	66 (34.0)	27 (29.7)	73 (31.3)
Age <65 yr	97 (50.0)	45 (49.5)	121 (51.9)
Efficacy			
PFS HR[c] (95% CI)	0.68 (0.49–0.95)	1.08 (0.68–1.72)	0.69 (0.51–0.94)
ORR OR[d] (95% CI)	1.45 (0.80–2.61)	0.99 (0.41–2.40)[e]	1.34 (0.78–2.30)

[a] Refers to the country of recruitment and not necessarily to racial origin.
[b] Includes both mutation-positive and mutation-negative samples.
[c] HR <1 indicates a difference in favor of gefitinib.
[d] OR >1 indicates a greater chance of response on gefitinib.
[e] These results should be interpreted with caution as the logistic regression model did not converge.
cfDNA, circulating free DNA; CI, confidence interval; *EGFR*, epidermal growth factor receptor; HR, hazard ratio; ITT, intent-to-treat; OR, odds ratio; ORR, objective response rate; PFS, progression-free survival; PS, performance status; WHO, World Health Organization.

believe that this result was due to the high rate of false negative results as described later (i.e., this group included both tumor *EGFR* M+ and M− patients).

In the cfDNA M+ subgroup, ORR was not significantly different in the gefitinib group compared with carboplatin/paclitaxel treatment (75.0% [18/24] and 63.6% [14/22], respectively; odds ratio [OR], 1.71; 95% CI, 0.48–6.09; p = 0.40). In the cfDNA M− subgroup, there were no significant differences in ORR with gefitinib compared with carboplatin/paclitaxel (27.1% [19/70] and 21.8% [17/78], respectively; OR, 1.34; 95% CI, 0.63–2.84; p = 0.45) (Figure 4). Again, this subgroup included both tumor *EGFR* M+ and M− patients as described later.

The results for clinical outcome by *EGFR* mutation status (M+, M−) for the Japanese subset of patients with known tumor *EGFR* mutation status (n = 91) are included in Supplemental Digital Content 2 (Results http://links.lww.com/JTO/A153).

Comparison of *EGFR* Mutation Status in Pretreatment cfDNA and Tumor Tissue

A total of 108 patients had a known mutation result by cfDNA but not by tumor; 5 patients had a known mutation

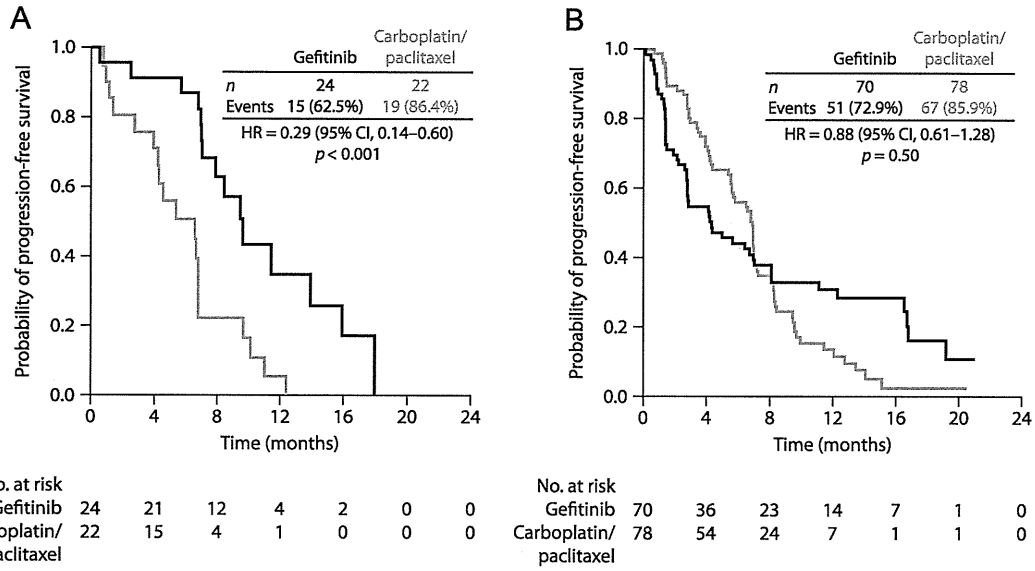


FIGURE 3. Kaplan-Meier curves of progression-free survival in cfDNA *EGFR* mutation-positive (*A*) and cfDNA *EGFR* mutation-negative (*B*) patients in the Japanese subset of IPASS. HR <1 indicates a difference in favor of gefitinib. CI, confidence interval; cfDNA, circulating free DNA; *EGFR*, epidermal growth factor receptor; HR, hazard ratio.

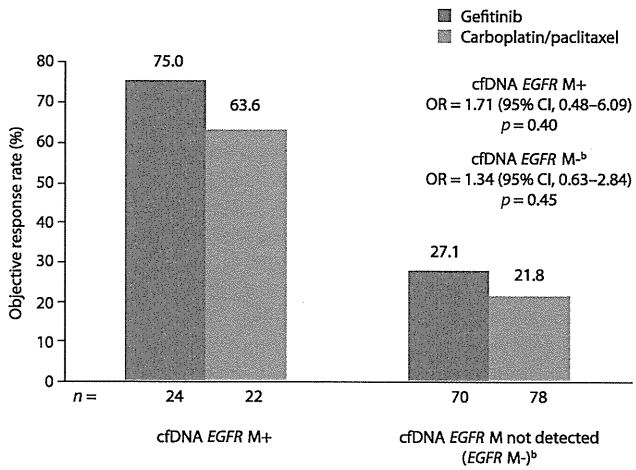


FIGURE 4. Objective response rates by treatment and by cfDNA (serum) *EGFR* mutation status (Japanese ITT population[a]). [a]Refers to the country of recruitment and not necessarily to racial origin. [b]There was a high rate of false-negative results, i.e., this group included both tumor *EGFR* M+ and M− patients. OR >1 implies a greater chance of response on gefitinib. OR, CI, and *p* values from logistic regression. cfDNA, circulating free DNA; CI, confidence interval; *EGFR*, epidermal growth factor receptor; ITT, intent-to-treat; M+, mutation-positive; M−, mutation-negative; OR, odds ratio.

result by tumor but not cfDNA (no serum sample provided); and 86 patients had a known mutation status by both tumor and cfDNA.

Of the 86 patients who had a known tumor and cfDNA mutation status, no false positives were identified (i.e., no samples were tumor M− but cfDNA M+). All 22 patients identified as cfDNA *EGFR* M+ were tumor *EGFR* M+, i.e., the positive predictive value was 100% (all samples that were cfDNA M+ were tumor M+) and the specificity was 100% (all samples that were tumor M− were cfDNA M−) (Table 2). However, the rate of false negatives was high: 29/51 (56.9%) of patients identified as tumor *EGFR* M+ were cfDNA *EGFR* M− (Table 2).

TABLE 2. Comparison of *EGFR* Mutation Status in cfDNA and Tumor Samples in 86 Patients with a Known *EGFR* Mutation Status Using Both Methods (Japanese[a] ITT Population)

	Mutation Status (Tumor Tissue), *n*		
	M+	M−	Total
Mutation status (cfDNA), *n*			
M+	22	0	22
M−	29	35	64
Total	51	35	86

Sensitivity = 43.1% (22 cfDNA M+ out of 51 tumor M+).[b]
Specificity = 100% (all 35 tumor M- were cfDNA M−).[b]
Positive predictive value = 100% (all 22 cfDNA M+ were tumor M+).[b]
Negative predictive value = 54.7% (35 tumor M- out of 64 cfDNA M−).[b]
Concordance = 66.3% (cfDNA and tumor results agreed in 57 of 86 cases).[b,c]

[a] Refers to the country of recruitment and not necessarily to racial origin.
[b] Those with a known *EGFR* mutation status using both methods.
[c] Kappa coefficient 0.38 (95% CI, 0.24–0.53).
cfDNA, circulating free DNA; CI, confidence interval; *EGFR*, epidermal growth factor receptor; ITT, intent-to-treat; M+, mutation positive; M−, mutation negative.

EGFR Mutation Types in Pretreatment cfDNA and Tumor Tissue

Of the patients classified as *EGFR* M+ at pretreatment by both tumor and cfDNA, all had the same mutation type in

TABLE 3. *EGFR* Mutations in Pretreatment cfDNA vs. Tumor Samples (Japanese[a] ITT Population)

	Tumor *EGFR* Mutation[b]						
cfDNA *EGFR* Mutation	Exon 19 Deletions Only	Exon 20 T790M Only	Exon 21 L858R Only	Exon 20 T790M and Exon 21 L858R	Negative	Unknown	Total
Exon 19 deletions only	11	0	0	0	0	15	26
Exon 20 T790M only	0	0	0	1	0	1	2
Exon 21 L858R only	0	0	10	0	0	8	18
Exon 20 T790M and exon 21 L858R	0	0	0	0	0	0	0
Negative	18	0	11	0	35	84	148
Unknown	2	1	0	0	2	34	39
Total	31	1	21	1	37	142	233

The categories are mutually exclusive. The categories "Exon 19 deletions and exon 20 T790M" and "Exon 19 deletions and exon 21 L858R" were 0 for both tumor and cfDNA and have been omitted from the table.

[a] Refers to the country of recruitment and not necessarily to racial origin.

[b] Mutations that were tested in tumor tissue samples but not serum included: exon 20 insertion, exon 21 L861Q, exon 18 G719X, and exon 20 S768I. Two patients with tumor samples had these mutations (1 with exon 20 insertion and 1 with exon 21 L861Q). These patients were excluded from the comparative analysis of mutation detection by sample type.

cfDNA, circulating free DNA; *EGFR*, epidermal growth factor receptor; ITT, intent-to-treat.

tumor and cfDNA except one patient who had exon 20 T790M and exon 21 L858R by tumor but exon 20 T790M only by cfDNA (Table 3).

DISCUSSION

The feasibility of using cfDNA to detect *EGFR* mutations was assessed in the Japanese subset of patients from the IPASS study. The proportion of patients identified as *EGFR* M+ was lower when assessed in cfDNA (23.7%) compared with tumor tissue (61.5%). Although cfDNA results identified no false positives, a high rate of false negatives (56.9%) was observed, with more than half of the tumor M+ patients not detected by cfDNA testing (of patients with evaluable mutation status from both cfDNA and tumor). Further research into appropriate methods and analysis needs to be performed before it could be accepted as an option in the diagnostic or screening setting. If larger patient series confirmed the absence of false-positive results and demonstrated an improvement or lowering of false-negative results, serum testing may prove useful for patients for whom tumor samples are not available.

Testing of biopsied tumor tissue remains the current recommended method for *EGFR* mutation analysis.[8] However, tumor tissue is often difficult to obtain, particularly from patients with advanced non-small cell lung cancer (NSCLC), and a lack of tumor cells in a given sample and subsequently failure on pathological examination can make *EGFR* mutation analysis very difficult. The increased recognition of the relevance of mutation testing to treatment selection may stimulate efforts to better obtain tissue for *EGFR* mutation testing in the future. In the meantime, detection of *EGFR* mutation status in cfDNA derived from serum/plasma may allow patients without diagnostic tumor material the opportunity to benefit from personalized treatment and also has a use in the clinical trial setting where tumor material is not always available.

Although minimally invasive, the use of serum as a nontumor surrogate sample may be limited by the amount of cfDNA available in the sample, meaning that some positive samples are not detected. In addition, some patients may not have cf tumor DNA as their tumors may not be releasing this material into the bloodstream, giving rise to false-negative results. Because of the limited yields of cfDNA obtained from serum, two changes (in addition to duplicate tests) were made to the *EGFR* mutation ARMS kit used to detect *EGFR* mutations in this study: an increase in the number of PCR cycles and an alteration of the cutoffs used to define M+ samples (dCt values). Further analysis is underway to investigate whether these conditions are the most appropriate and whether less stringent settings could result in more true positives (fewer false negatives) while retaining no false positives.

There have been several reports on the detection of cfDNA *EGFR* mutation status using different methods. A significant correlation between cfDNA *EGFR* mutation status and clinical response to gefitinib was found in two previous small studies that assessed cfDNA *EGFR* mutation status using the ARMS method of detection, a highly sensitive (1% sensitive) targeted technique to detect specific known *EGFR* mutations.[9,11] Other screening techniques detect all *EGFR* mutations, known and novel variants, by PCR amplification followed by sequencing, pyrosequencing, or melt analysis (10–30% sensitivity).[8] However, although these methods are widely used for *EGFR* mutation analysis of DNA derived from tumor tissue, not all of these methods have demonstrated utility for *EGFR* mutation analysis of cfDNA. In a small study that used DNA sequencing to detect *EGFR* mutations in serum, mutations were more frequently observed in patients experiencing partial response or stable disease compared with those whose disease progressed, although the difference did not reach statistical significance.[10] No statistically significant association between cfDNA *EGFR* mutation status and PFS by multivariate analysis (HR, 1.48; 95% CI, 0.93–2.36; $p = 0.09$) was found in the study by Rosell et al.[12] which assessed *EGFR* mutations by PCR-based methods in the presence of a protein nucleic acid (PNA) clamp in the cfDNA extracted from serum of 164 patients

treated with erlotinib. In another study that used denaturing high-performance liquid chromatography to analyze for mutations in exons 19 and 21 from matched plasma and tumor samples, patients with plasma *EGFR* mutations had significantly higher ORR and prolonged PFS.[7] The present study using ARMS demonstrated that the treatment effect for the Japanese cfDNA *EGFR* M+ subgroup followed the same pattern as the tumor *EGFR* M+ subgroup of the overall IPASS population (i.e., PFS HR significantly in favor of gefitinib and higher ORR with gefitinib versus carboplatin/paclitaxel).[6] There was a significant interaction between cfDNA *EGFR* mutation status and treatment for PFS.

Any variance in concordance rates for mutation results between pretreatment serum versus tumor tissue (66.3% in our study and between 58 and 93% in previously reported studies)[7,9–11] may be attributed to different methods of extraction, detection, run conditions, the size and yield of the DNA fragments, and the fact that cfDNA may not be present in the circulation of all patients with NSCLC. For example, targeted sequences amplified by ARMS are short, at 100–150 bp, leading to decreased assay failure rates (particularly from formalin-fixed paraffin-embedded material or fragments of cfDNA) compared with sequencing methods, which tend to involve the amplification of longer target sequences of 150–250 bp or above.[8,13,14,17,18]

In patients who were cfDNA *EGFR* M− in this study, no significant difference for PFS was seen with gefitinib compared with carboplatin/paclitaxel; however, the HR was not constant over time (as was observed for the overall Japanese study population). These results should be interpreted with caution as there was a high rate of false negatives, and this subgroup is likely to include tumor *EGFR* M+ and M− patients.

In conclusion, these results merit further investigation to determine whether alternative samples, including serum or plasma, may be considered for determining *EGFR* mutation status in future, particularly in cases where diagnostic tumor material is not available. Currently, analysis of tumor material is the recommended method for determining *EGFR* mutation status.

ACKNOWLEDGMENTS

Supported by AstraZeneca.

The authors thank the patients and investigators for their participation in this study and Annette Smith, PhD, from Complete Medical Communications, who provided medical writing support funded by AstraZeneca.

REFERENCES

1. Modjtahedi H, Essapen S. Epidermal growth factor receptor inhibitors in cancer treatment: advances, challenges and opportunities. *Anticancer Drugs* 2009;20:851–855.
2. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004;350:2129–2139.
3. Paez JG, Jänne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004;304:1497–1500.
4. Pao W, Miller V, Zakowski M, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci USA* 2004;101:13306–13311.
5. Mitsudomi T, Kosaka T, Yatabe Y. Biological and clinical implications of EGFR mutations in lung cancer. *Int J Clin Oncol* 2006;11:190–198.
6. Mok TS, Wu Y-L, Thongprasert S, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 2009;361:947–957.
7. Bai H, Mao L, Wang HS, et al. Epidermal growth factor receptor mutations in plasma DNA samples predict tumor response in Chinese patients with stages IIIB to IV non-small-cell lung cancer. *J Clin Oncol* 2009;27:2653–2659.
8. Eberhard DA, Giaccone G, Johnson BE. Biomarkers of response to epidermal growth factor receptor inhibitors in Non-Small-Cell Lung Cancer Working Group: standardization for use in the clinical trial setting. *J Clin Oncol* 2008;26:983–994.
9. Kimura H, Kasahara K, Kawaishi M, et al. Detection of epidermal growth factor receptor mutations in serum as a predictor of the response to gefitinib in patients with non-small-cell lung cancer. *Clin Cancer Res* 2006;12:3915–3921.
10. Kimura H, Kasahara K, Shibata K, et al. EGFR mutation of tumor and serum in gefitinib-treated patients with chemotherapy-naive non-small cell lung cancer. *J Thorac Oncol* 2006;1:260–267.
11. Kimura H, Suminoe M, Kasahara K, et al. Evaluation of epidermal growth factor receptor mutation status in serum DNA as a predictor of response to gefitinib (IRESSA). *Br J Cancer* 2007;97:778–784.
12. Rosell R, Moran T, Queralt C, et al. Screening for epidermal growth factor receptor mutations in lung cancer. *N Engl J Med* 2009;361:958–967.
13. Newton CR, Graham A, Heptinstall LE, et al. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res* 1989;17:2503–2516.
14. Whitcombe D, Theaker J, Guy SP, et al. Detection of PCR products using self-probing amplicons and fluorescence. *Nat Biotechnol* 1999;17:804–807.
15. Ohe Y, Ichinose Y, Nishiwaki Y, et al. Phase III, randomized, open-label, first-line study of gefitinib vs carboplatin/paclitaxel in selected patients with advanced non-small cell lung cancer (IPASS): evaluation of recruits in Japan. Poster 8044 presented at the ASCO Annual Meeting, Orlando, FL, USA, May 29 to June 2, 2009.
16. Ichinose Y, Nishiwaki Y, Ohe Y, et al. Analyses of Japanese patients recruited in IPASS, a phase III, randomized, open-label, first-line study of gefitinib vs carboplatin/paclitaxel in selected patients with advanced non-small cell lung cancer. Poster PD3.1.3 presented at the 13th World Conference on Lung Cancer, International Association for the Study of Lung Cancer, San Francisco, CA, USA, July 31 to August 4, 2009.
17. Board RE, Thelwell NJ, Ravetto PF, et al. Multiplexed assays for detection of mutations in PIK3CA. *Clin Chem* 2008;54:757–760.
18. Holland PM, Abramson RD, Watson R, et al. Detection of specific polymerase chain reaction product by utilizing the 5′—-3′ exonuclease activity of Thermus aquaticus DNA polymerase. *Proc Natl Acad Sci USA* 1991;88:7276–7280.

Proteomic Biomarkers for Acute Interstitial Lung Disease in Gefitinib-Treated Japanese Lung Cancer Patients

Fredrik Nyberg[1,2*☯], Atsushi Ogiwara[3,4☯], Chris G. Harbron[5☯], Takao Kawakami[3,4], Keiko Nagasaka[3], Sachiko Takami[3], Kazuya Wada[3], Hsiao-Kun Tu[3], Makiko Otsuji[3], Yutaka Kyono[3], Tae Dobashi[3], Yasuhiko Komatsu[3], Makoto Kihara[3], Shingo Akimoto[3,4], Ian S. Peers[5], Marie C. South[5], Tim Higenbottam[6¤a], Masahiro Fukuoka[7], Koichiro Nakata[8], Yuichiro Ohe[9], Shoji Kudoh[10,11], Ib Groth Clausen[12], Toshihide Nishimura[3,4], György Marko-Varga[6,13¤b,c], Harubumi Kato[13,14]

1 Global Epidemiology, AstraZeneca R&D, Mölndal, Sweden, **2** Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden, **3** Research and Development Division, Medical ProteoScope Company, Tokyo, Japan, **4** Clinical Proteome Center, Tokyo Medical University, Tokyo, Japan, **5** Discovery Statistics, AstraZeneca R&D, Macclesfield, United Kingdom, **6** AstraZeneca in Respiratory Biological Sciences, AstraZeneca R&D, Lund, Sweden, **7** Department of Medical Oncology, Kinki University School of Medicine, Osaka, Japan, **8** Nakata Clinic, Tokyo, Japan, **9** Department of Thoracic Oncology, National Cancer Center Hospital East, Chiba, Japan, **10** Japan Anti-Tuberculosis Association Fukujuji Hospital, Tokyo, Japan, **11** Division of Pulmonary Medicine, Infectious Diseases and Oncology, Department of Internal Medicine, Nippon Medical School, Tokyo, Japan, **12** Respiratory Biological Sciences, AstraZeneca R&D, Lund, Sweden, **13** Department of Surgery, Tokyo Medical University, Tokyo, Japan, **14** Niizashiki Central General Hospital, Saitama, Japan

Abstract

Interstitial lung disease (ILD) events have been reported in Japanese non-small-cell lung cancer (NSCLC) patients receiving EGFR tyrosine kinase inhibitors. We investigated proteomic biomarkers for mechanistic insights and improved prediction of ILD. Blood plasma was collected from 43 gefitinib-treated NSCLC patients developing acute ILD (confirmed by blinded diagnostic review) and 123 randomly selected controls in a nested case-control study within a pharmacoepidemiological cohort study in Japan. We generated ~7 million tandem mass spectrometry (MS/MS) measurements with extensive quality control and validation, producing one of the largest proteomic lung cancer datasets to date, incorporating rigorous study design, phenotype definition, and evaluation of sample processing. After alignment, scaling, and measurement batch adjustment, we identified 41 peptide peaks representing 29 proteins best predicting ILD. Multivariate peptide, protein, and pathway modeling achieved ILD prediction comparable to previously identified clinical variables; combining the two provided some improvement. The acute phase response pathway was strongly represented (17 of 29 proteins, $p = 1.0 \times 10^{-25}$), suggesting a key role with potential utility as a marker for increased risk of acute ILD events. Validation by Western blotting showed correlation for identified proteins, confirming that robust results can be generated from an MS/MS platform implementing strict quality control.

Citation: Nyberg F, Ogiwara A, Harbron CG, Kawakami T, Nagasaka K, et al. (2011) Proteomic Biomarkers for Acute Interstitial Lung Disease in Gefitinib-Treated Japanese Lung Cancer Patients. PLoS ONE 6(7): e22062. doi:10.1371/journal.pone.0022062

Editor: Scott A. Coonrod, Cornell University, United States of America

Received March 28, 2011; **Accepted** June 14, 2011; **Published** July 20, 2011

Funding: This study was funded by AstraZeneca and conducted in collaboration by researchers from Academia, AstraZeneca and Medical Proteoscope Company. All collaborating parties participated in the study design, data collection and analysis, decision to publish, and preparation of the manuscript.

Competing Interests: FN, CGH, ISP, MCS and IGC are employees of AstraZeneca and own shares in the company. AO, TK, K. Nagasaka, ST, KW, HKT, MO, Y. Kyono, TD, Y. Komatsu, MK, SA, and TN are/were employees of Medical Proteoscope Co., Ltd. TH is an employee of Chiesi Farmaceutici. TH and GMV are previous employees of AstraZeneca. MF has received honoraria from AstraZeneca. K. Nakata, YO, SK, and HK have declared that no conflicts of interest exist.

* E-mail: Fredrik.Nyberg@astrazeneca.com

¤a Current address: Chiesi Farmaceutici, Parma, Italy
¤b Current address: Division of Clinical Protein Science and Imaging, Biomedical Center, Department of Measurement Technology and Industrial Electrical Engineering, Lund University, Lund, Sweden
¤c Current address: Department of Surgery, Tokyo Medical University, Tokyo, Japan

☯ These authors contributed equally to this work.

Introduction

Interstitial lung disease (ILD) affects the pulmonary parenchyma or alveolar region [1]. When associated with drug treatment, it can present precipitously as acute diffuse alveolar damage (DAD), sometimes with a fatal outcome [2]. Patients often have severe breathlessness and chest radiology shows 'ground glass' appearance. No specific treatment is available, but supportive therapy includes oxygen, corticosteroids, or assisted ventilation. Acute ILD events may develop *de novo*, but an existing chronic ILD condition increases the risk considerably [3], as observed in recent studies of patients with idiopathic pulmonary fibrosis (IPF), the most common chronic form [4].

ILD, especially IPF, is a known co-morbidity in patients with non-small-cell lung cancer (NSCLC) [5]. Acute ILD events have been reported with many lung cancer therapies at rates up to ~10% [6–11]. ILD is recognized as more common in Japan than elsewhere, both in the population and among patients with NSCLC [5,6,12,13], although it is unclear why.

EGFR tyrosine kinase inhibitors (TKIs) are an established treatment for advanced NSCLC. Unlike much chemotherapy, they are typically well tolerated and without cytotoxic side effects.

The EGFR TKI gefitinib (IRESSA) was approved in 2002 in Japan for treatment of advanced NSCLC. Although some ILD-type events were observed in clinical trials and compassionate clinical use, only after approval did an increasing number of spontaneous reports for ILD appear in Japan as the drug became more widely available.

At that point, better understanding of ILD was urgently needed: baseline incidence on different treatments, risk factors, and the potential association of gefitinib with ILD risk. An independent academic team together with AstraZeneca scientists therefore designed and conducted a cohort and nested case-control pharmacoepidemiological study of ILD in Japanese NSCLC patients treated with either gefitinib or chemotherapy, with clinical results reported previously [3]. As one exploratory sub-study component, patients receiving gefitinib (both subsequent ILD cases and control patients) were sampled for plasma proteomics, with two main objectives: 1) to identify proteomic predictors of ILD that might ultimately be developed into a personalized medicine diagnostic to identify patients at greater risk of ILD; 2) to increase understanding of the mechanisms underlying the development of acute ILD events.

Using a multiple biomarker approach such as proteomics (the simultaneous study of large parts of the human proteome to give a global view of differential expression of proteins in blood or tissue), rather than simply a conventional single biomarker, potentially increases predictive power both through increased robustness deriving from multiple measurements and the opportunity to combine information from multiple biological processes. To support high-quality generation of such information, we combined in a novel way several key study components: robust study design, well-defined phenotypic definitions, careful sample collection procedures, stable advanced liquid chromatography (LC)-tandem mass spectrometry (MS/MS)-based peptide separation and detection methods, statistical analysis incorporating proteomic and clinical information, stringent methods for database protein annotation of detected peptide peaks, and biological interpretation using literature mining software, plus extensive quality control and validation, reported below.

Results

Characteristics of the study population

The non-randomized cohort included 3,166 Japanese patients with advanced/recurrent NSCLC who were followed for 12 weeks after initiating gefitinib (n = 1,872 treatment periods) or chemotherapy (n = 2,551). From the gefitinib-treated sub-cohort, 103 suspected ILD cases (79 subsequently confirmed and 24 rejected by the Case Review Board [CRB]), as well as 252 controls, were registered into the case-control study. Proteomics samples for this sub-study were available from 43 confirmed ILD cases, 123 control subjects, and 15 CRB-rejected initially diagnosed ILD cases (Table 1). Clinical characteristics of the cases and controls are described in Table S1.

Exploratory analysis of LC-MS/MS data generated under quality controlled conditions reveals large batch variation that needs to be controlled in subsequent statistical analyses

Quality assessment of sample processing and data generation. After immunoaffinity depletion, remaining serum albumin was <8% for all 181 baseline samples (Table S2). The subsequent tryptic hydrolysis resulted in a remaining undigested protein portion ranging from 3.0% to 32.3% (mean 15.3%) (Table S2). The variation in these processing steps was independent of case/control status (data not shown).

LC-MS/MS measurements for the 181 individual baseline samples were performed in 11 batches, with 19 and 20 samples from batches 1 and 3 repeated in batches 10 and 11, respectively (Table 1), resulting in 220 discrete proteomics measurements. Four of the 11 batches initially failed the quality control criteria (coefficient of variation [CoV] >20% for any one of the six control

Table 1. Composition of the LC-MS/MS measurement batches for 181 blood plasma samples from Japanese patients with NSCLC.

Batch number	Number of study samples not analyzed in previous batches	Number of analyzed samples[a]	Type of study subject		
			ILD case	Control	Rejected case[b]
1	20	20	3	15	2
2	20	20	5	13	2
3	20	20	6	12	2
4	20	20	3	15	2
5	20	20	6	12	2
6	20	20	5	14	1
7	20	20	6	13	1
8	20	20	3	16	1
9	20	20	6	12	2
10[c]	1[c]	20[c]	3	15	2
11[c]	0[d]	20[d]	6	12	2
Total	181	220	52 (43[e])	149 (123[e])	19 (15[e])

[a]Each batch also contained 3 experimental control samples in positions 1, 12, and 23 of total batch size of 23.
[b]Case Review Board (see Materials and Methods) did not confirm clinical ILD diagnosis after blinded diagnostic review.
[c]19 samples from batch 1 repeated, 1 new control analyzed.
[d]All 20 samples from batch 3 repeated.
[e]Number of unique study subjects.
doi:10.1371/journal.pone.0022062.t001

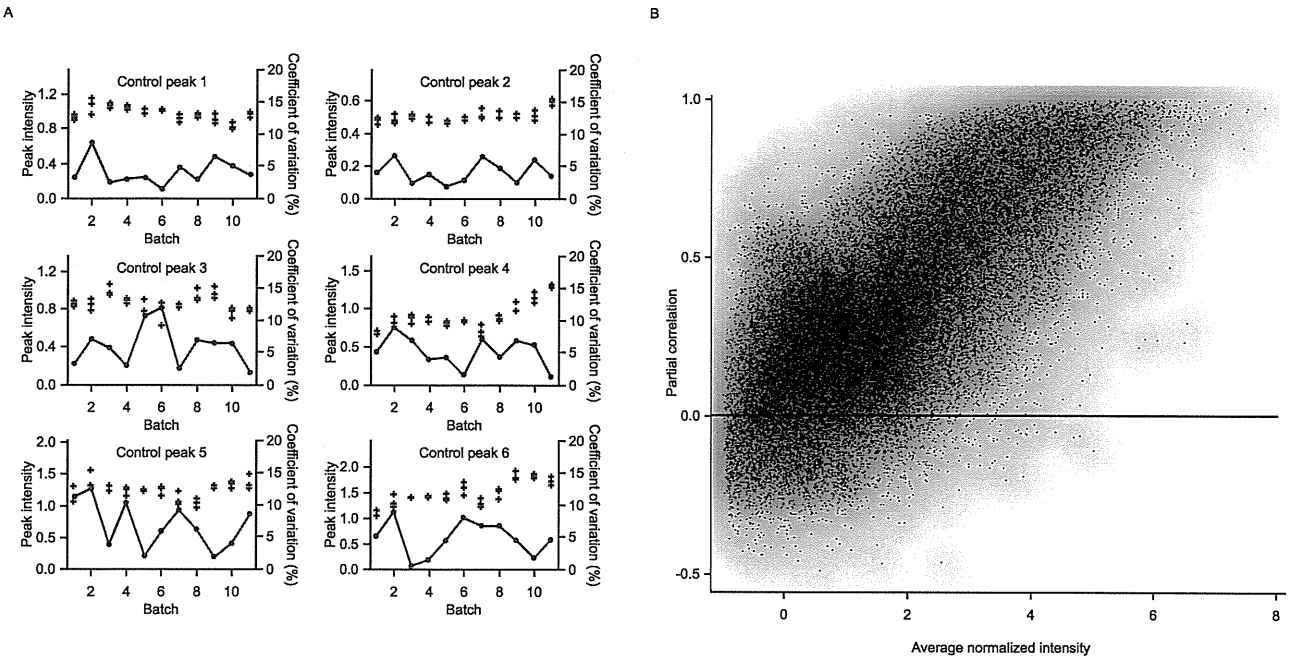


Figure 1. Quality control: reproducibility of control samples and sample duplicates. (A) Reproducibility of 6 control peaks for the 3 standard quality control samples, plotted as '+', in each analysis batch (peak intensity, left axis). The coefficients of variation (%, right axis) between the 3 control samples in each batch are plotted as points joined by a line. (B) Reproducibility of peptide intensities for 39 samples with duplicate analyses in different analysis batches. Partial correlation, after removing between batch differences, plotted against the average normalized intensity for each peptide. Higher intensity peptides show high reproducibility in their intensities between repeated batches.
doi:10.1371/journal.pone.0022062.g001

peptides among the three within-batch control samples) on the first measurement run, but passed the criteria on repeated measurement. A quality control summary of acceptable batch runs is given in Figure 1A.

Figure 1B shows a plot of partial correlations between the duplicate samples in batches 1 and 10, 3 and 11, after allowing for any batch effect, against the average normalized intensity over the complete sample set for each signal. Peptides with higher average intensities show higher reproducibility between batches as evidenced by generally high partial correlations.

Exploratory data analysis of MS signal intensities. We then used a principal component analysis (PCA) to explore the data in order to identify the largest sources of variation. Figure 2A shows a plot of this analysis, with each sample colored according to batch. Measurements from the same batch tend to cluster together, separate from other batches, implying that the largest differences between samples arise from the batch-wise processing.

Figure 2B shows the results of PCA on the pairs of repeated batches (1 and 10, 3 and 11), with duplicate samples joined by a line, plotted against the first two principal components. The lines are generally horizontal and parallel, again suggesting that the largest source of variability or the greatest overall differences in profiles between samples (first principal component) relates to inter-batch variability, and that the ordering of samples on the second principal component, i.e. in the next largest source of variability or overall differences between samples, is in strong agreement between the repeated batches. After allowing for consistent differences between batches, these results thus confirm that inter-sample differences are reproducible with the method used.

Whereas Figure 2B compares results summarized over all measured peptides, Figure 2C shows the repeated run results for an example peptide. Although there are large between-batch differences, within each batch there is high correlation between the intensities for the same subject on replicate runs. Of the 41 differentially expressed peptide peaks used to identify the proteins listed in Table 2, 25 (61%) show a partial correlation after removing the batch effect greater than 0.8 and all show a partial correlation in excess of 0.35.

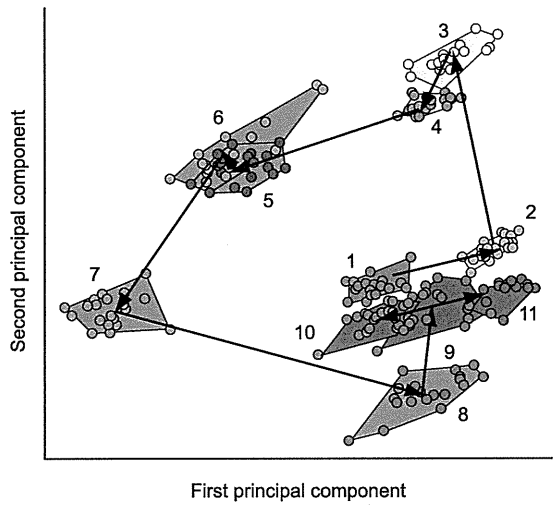
Clear differences in peptide and protein patterns between ILD cases and controls

The subsequent analyses aimed to identify peptides and proteins that effectively discriminated between cases and controls, so rejected cases were now excluded. Repeated samples in batches 10 and 11 were excluded, and given the large between-batch differences identified in the exploratory analyses, the control subject measured in batch 10 was also excluded, leaving 43 confirmed ILD cases and 122 controls with one sample measurement each.

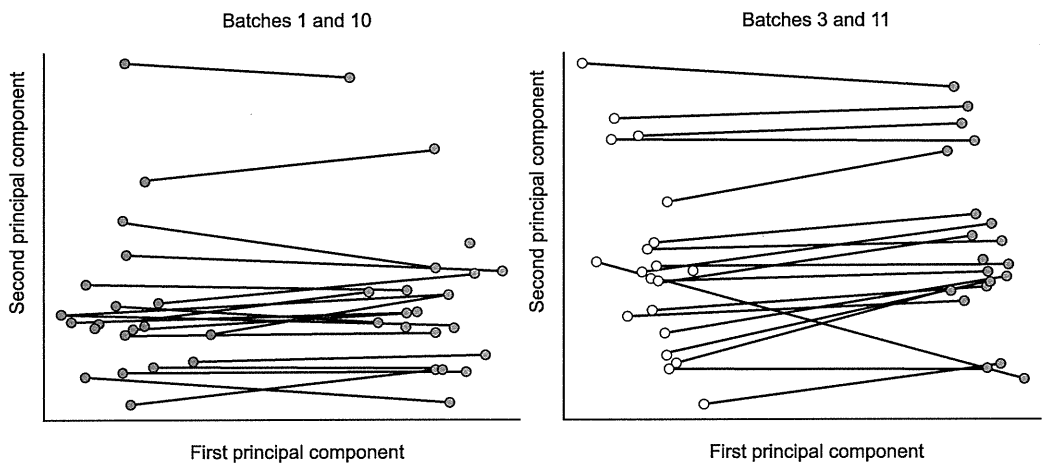
Identification of discriminating peptides and proteins. Figure 3 shows the results of the univariate (individual peptide) analyses using analysis of covariance (ANCOVA), displayed as histograms of the p-values for the comparison between cases and controls. Allowing for batch as an analysis covariate, to remove inter-batch variation, substantially increases the power of the analysis, identifying approximately twice as many peptides showing statistically significant differential expression at the 5% level. Figures S1 and S2 explore and explain this relationship in more detail. Further accounting for the within-batch order only slightly decreases the number of significant peptides, suggesting that any within-batch order effect is marginal and attempts to model it will not increase power.

On the other hand, allowing for key clinical variables (WHO performance status [PS], smoking history, extent of normal lung

A


Second principal component
First principal component
1
2
3
4
5
6
7
8
9
10
11

B


Batches 1 and 10
Batches 3 and 11
Second principal component
First principal component
Second principal component
First principal component

C

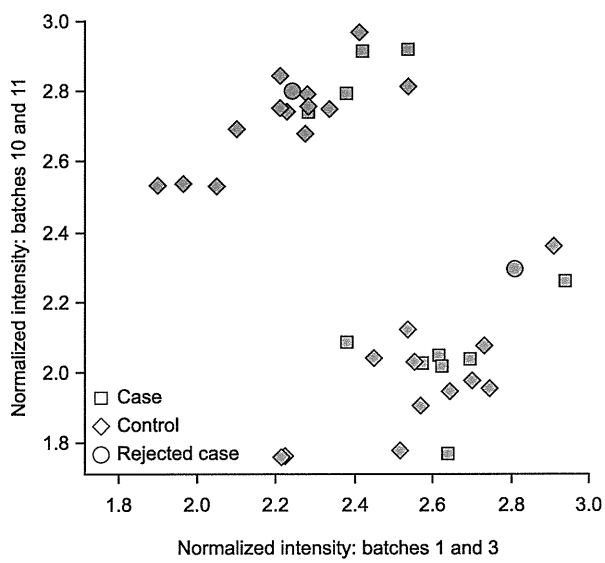

Normalized intensity: batches 10 and 11
Normalized intensity: batches 1 and 3
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Case
Control
Rejected case

Figure 2. Exploratory data analysis of MS signal intensities using PCA. (A) Plot of first two principal components from PCA analysis of the full proteomic data from all 11 analysis batches (numbered 1–11 in time sequence). Each sample is represented by a single point, with the range of points within each batch being shown by a polygon joining the extreme points in that batch. (B) Plots of the first two principal components for the repeated batches of samples (1 and 10, 3 and 11). Individual samples are represented by a line, connecting the two replicates in different batches. (C) Reproducibility of an example differentially expressed peptide between two duplicate batch runs of proteomic analysis. The intensities of the first and second runs for each replicated sample are plotted against each other. Samples colored by batch (batch 1 repeated as batch 10 – blue; batch 3, repeated as batch 11 – red). Allowing for between-sequence differences there is a good correlation between replicate runs.
doi:10.1371/journal.pone.0022062.g002

coverage on CT scan, and severity of pre-existing ILD) in the analysis consistently reduced the number of peptides being detected as differentially expressed, reflecting that much of the information carried in the most significant peptides is duplicating information carried by the clinical variables (Figure 3). Figure S3 shows an example of this, where higher levels of the peptide, higher performance status score and case status are all associated with each other, and much of the increased peptide intensity of the cases compared with the controls can be explained by their association with higher performance status score, so the peptide intensity is adding much less information when considered in combination with this clinical variable.

Table 2. List of 29 proteins representing 27 protein identifications from the 41 selected peaks, with pathway assignments according to ingenuity analysis and the validation of ms/ms results using western blots on 12 subjects (6 ILD cases and 6 controls).

Protein name	Acute Phase Response pathway	Correlation between expression levels using MS/MS and Western blot
alpha-1-acid glycoprotein 1	YES	0.717
alpha-1-antitrypsin	YES	0.512
alpha-1B-glycoprotein		
Leucine-rich alpha-2-glycoprotein		
alpha-1-antichymotrypsin	YES	0.744
Antithrombin-III		
Apolipoprotein A-I	YES	0.468
Apolipoprotein B-100		
Apolipoprotein C-III		
Armadillo repeat-containing protein 2		
Complement C3	YES	0.242
Complement C4-A, Complement C4-B[a,b]	YES	0.768
Complement component C9	YES	
Plasma kallikrein	YES	
alpha-2-HS-glycoprotein	YES	0.808
Gelsolin		0.873
Hemoglobin alpha		
Hemoglobin beta, Hemoglobin delta[b]		
Haptoglobin	YES	0.859
Haptoglobin-related protein		
Histidine-rich glycoprotein	YES	
Inter-alpha-trypsin inhibitor heavy chain H4	YES	
Retinol binding protein 4	YES	
Serum amyloid P-component	YES	
Serotransferrin	YES	
Transthyretin	YES	
Ig kappa chain V-III region Ti		

[a]C4 beta chain (common to C4A and C4B).
[b]Dual identification of 2 closely related proteins from the same protein family.
doi:10.1371/journal.pone.0022062.t002

Based on the p-value, the top 100 peaks from both the analyses including and excluding clinical variables were identified. These peaks were subsequently restricted to 41 according to the following criteria: 1) normalized LC retention time between 5 and 75 min, and 2) full scan m/z value of the precursor ion between 450 and 1,500. Next, peptide identifications included in the 41 peptide peaks were selected using the following criteria: 1) a Mascot ion score more than the identity threshold value given to the individual amino acid sequence of the peptide; and 2) >3 samples with the corresponding peptide identification. This resulted in 45 valid peptide identifications from 28 of the 41 peaks, including two peptides from the spiked lysozyme (Table S3). The plasma-derived 43 peptides represented 27 distinct identifications with 2 dual identifications of closely related proteins, for a total of 29 proteins. These are listed in Table 2, with more detail concerning their identification given in Table S3.

Acute phase response identified as an important pathway likely to be involved in acute ILD events

This set of proteins was then used in the biological interpretation analyses, using the Ingenuity Pathway Analysis (IPA) system. The most significant pathway found when overlaying the proteins onto Ingenuity-curated canonical pathways was the acute phase response signaling pathway, with which 17 of the 29 proteins could be associated ($p = 1.0 \times 10^{-25}$). Other pathways showing a high overlap with the list of proteins included the complement and coagulation pathways, but p-values were less significant due to the smaller number of proteins involved (Figure 4).

Entering the 29 proteins into IPA, 5 networks were formed. The most significant network contains 24 of the 29 proteins (Figure 5). Proteins added to the network by the tool to connect the marker proteins include IL1 and NF-κB, suggesting that these proteins could also be involved in generating the observed pattern. Combining the two networks with the highest scores further adds IL1-beta, HNF1A, HNF4A, HNF6 (ONECUT1), and CEBPB as central components (Figure S4).

Validation of the MS/MS data shows good reproducibility and reasonable agreement with Western blot

Within the MS/MS platform there was strong agreement between replicate runs of the same samples after allowing for batch effects, as described above (Figure 2C).

Validating with another method, Table 2 shows the correlation in intensities derived from the MS/MS and Western blotting (WB) for a selection of 9 proteins. Considering that WB targets the intact protein, whereas the present MS/MS can detect peptides derived from the intact proteins, these 9 proteins show quite a strong level of agreement between the technologies, with 6 of the 9 proteins

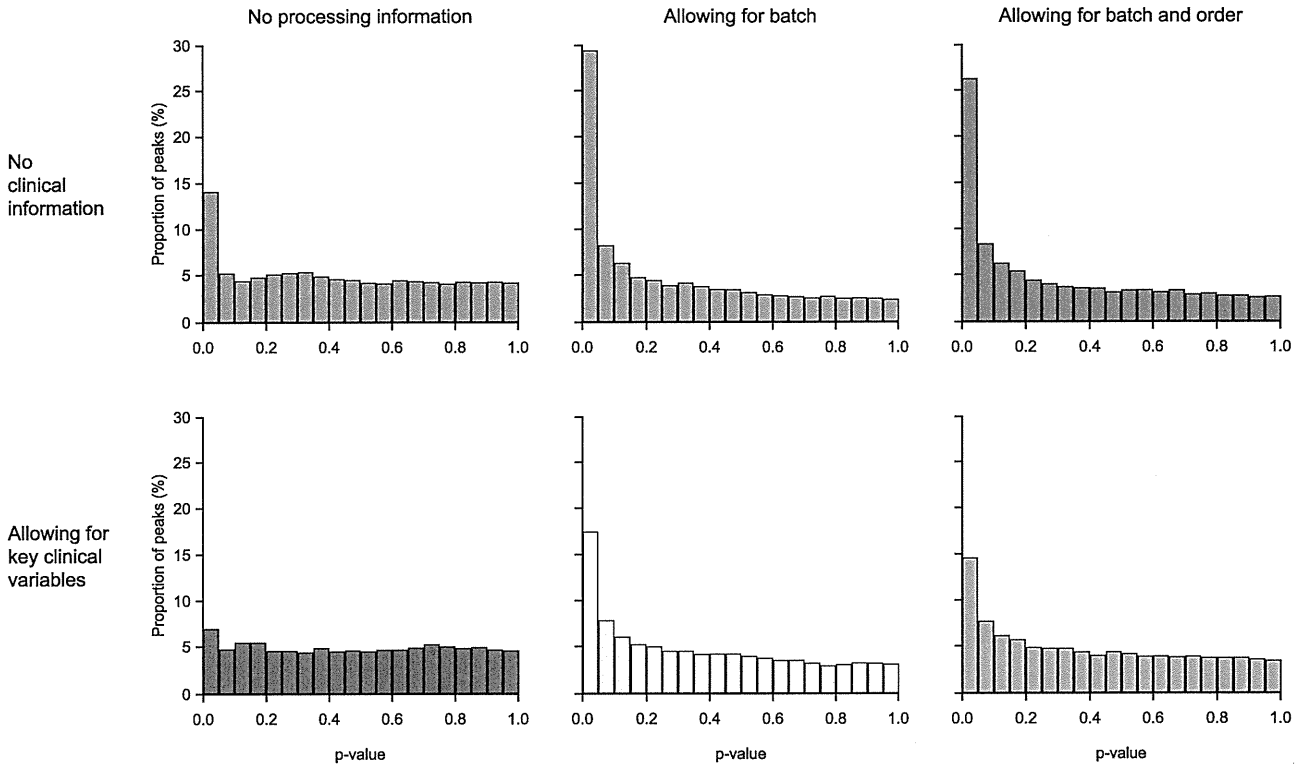


Figure 3. Distribution of significance tests of differential expression between cases and controls for individual peptides. The figure shows the effect on the distribution of p-values for differential expression of including analysis processing information and clinical variables.
doi:10.1371/journal.pone.0022062.g003

exhibiting a correlation in excess of 0.7. Scatterplots comparing the MS/MS and WB protein intensities are shown in Figure S5 and the WB images in Figure S6.

Prediction of ILD using proteins and clinical data

Modeling phenotype based on multiple peptide markers. Figure 6A shows the predictive power based on leave-one-out cross-validation for models built using a range of different numbers of peptides in combination. Substantial improvements on random prediction were obtained from just a few peptides, and increasing the number of peptides further did not substantially improve the model predictions. The predictive power of the model even decreased when using very many peptides.

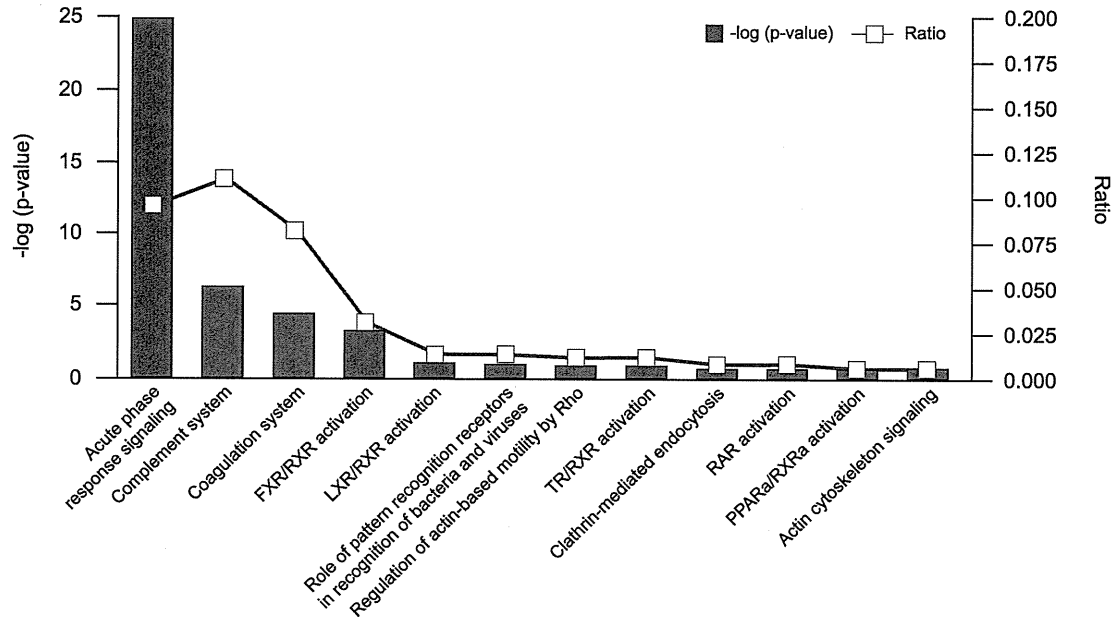
For robustness, alternative multivariate modeling approaches were compared. Using random forests instead of partial least squares discriminant analysis (PLS-DA) and logistic regression within the modeling framework yielded approximately the same predictive level as evidenced by the area under the curve (details not shown).

A subgroup analysis restricting the set of cases to those with the DAD acute ILD pattern (20 of the 43 cases) was hampered by small sample size and was unable to improve the overall predictive power.

Modeling phenotype based on multiple peptide markers and clinical data. Figure 6B compares the predictive power, based on leave-one-out cross validation, for models built on clinical/radiological data alone using a logistic regression, peptide data alone, and a combination. Both data types alone provided similar prediction. Some improvement was obtained by combining the two data types, but it was far less than additive. This is consistent with the results from the analysis of individual peptides, suggesting that the discriminating peptides partly carry information also available from the clinical/radiological variables.

Modeling phenotypes based on proteins and clinical data. Figure 7 shows the p-values for distinguishing cases and controls obtained from the proteins (i.e. combined constituent peptide score), all their constituent peptides, and the combined acute phase pathway intensity measure (i.e. the combined score of the 16 included constituent proteins). For most proteins, the estimated protein intensity is more significant than most of the measured peptide intensities associated with that protein, but only improves on the significance of the best peptide for a few proteins. As these results were obtained within the same dataset that was used to identify and select the constituent peptides, some over-fitting may be occurring, and the protein expression intensity incorporating information from many peptides may be a more robust measure to apply in a wider context. The combined acute phase pathway intensity measure shows a more significant response than any of the constituent proteins. A similar picture is obtained when we consider the additional information provided by the peptide, protein and pathway measures on top of the known clinical variables in predicting ILD status (Figure S7).

Figure 8A shows the acute phase response pathway intensity plotted against a combined clinical variable score measuring the likelihood of a subject being a case calculated from a logistic regression of case-control status against the clinical variables WHO PS, smoking history, extent of normal lung coverage on CT scan, and severity of pre-existing ILD. This shows both sources of information contributing to predicting ILD outcome, although these two measures are fairly strongly correlated, so that much of the information is duplicated. Figure 8B considers the implications



Name	p-value	Ratio
Acute phase response signaling	1×10^{-25}	17/178 (0.096)
Complement system	7.78×10^{-7}	4/36 (0.111)
Coagulation system	5.66×10^{-5}	3/37 (0.081)
FXR/RXR activation	1.1×10^{-3}	3/100 (0.03)
LXR/RXR activation	1.46×10^{-1}	1/85 (0.012)

Figure 4. Significant associated pathways with ILD status. The most significant pathways from an analysis linking the identified 29 proteins from the study to curated pathways in the Ingenuity Pathway Analysis system are shown, ordered according to the ratio between the number of protein markers that can be associated with the pathway and the number of proteins in the pathway.
doi:10.1371/journal.pone.0022062.g004

for predicting ILD by showing the receiver operating characteristics (ROC) curves for the clinical variables, the acute phase pathway intensity, and the combination of the two sources of information. This shows comparable levels of predictive power from the clinical variables and acute phase pathway intensities, and some potential benefit from combining them together which, however, is limited, reflecting their correlation.

Discussion

We present here results from a proteomic analysis applied to a large-scale pharmacoepidemiological study and demonstrate that with considerable attention to study design and experimental procedures throughout the entire process required to generate high-quality data, it is possible to derive valuable knowledge from both a scientific and a diagnostic perspective. However, there are numerous potential sources of data variation and bias in this process. The integrity of all of the steps of the process is critical to generating useful data and failure to ensure high quality in any one of them may compromise the validity and value of the entire study.

Methodological aspects

Study design and sample preparation. We applied careful phenotyping with blinded diagnostic review to ensure an accurate ILD diagnosis, and incorporated measures to ensure that all cases of ILD occurring in the source cohort would be captured. Controls were selected from the actual population generating the cases, ensuring comparability between cases and controls so that contrasts seen may be attributed to case status. Participation rates were high (>90%) and similar for cases and controls [3], making selection bias unlikely. Proteomics samples were obtained after separate informed consent from approximately half of all gefitinib-treated cases and controls. Steps to ensure high-quality proteomic data for our large-scale epidemiological investigation included randomization of the processed samples, careful quality assessment of sample preparations and optimized preparation protocols to ensure stability in all procedures for a large number of samples. We have previously described the general strategy that we decided to use in the study based on a number of experimental pilot phase rounds [14].

Experimental measurement batch effects. Two-dimensional polyacrylamide gel electrophoresis is one common conventional technology used for protein analysis in serum/plasma [15,16]. Currently, LC-MS/MS has become widely accepted for high-resolution proteome-wide profiling from a complex peptide mixture [17]. Recent advances in this methodology including improved stability of peptide separation and detection has enabled comparison of ion intensity between LC-MS/MS profiles [18]. Our proteomics analysis system applied LC-MS/MS after immunoaffinity depletion of the most abundant constituent proteins in blood plasma, and proteolytic enzyme treatment of the depleted plasma sample.

We identified that the LC-MS/MS measuring process has systematic measurement errors, as one might expect, which we took measures to eliminate by introducing batch processing with quality control, designing the order of sample processing to

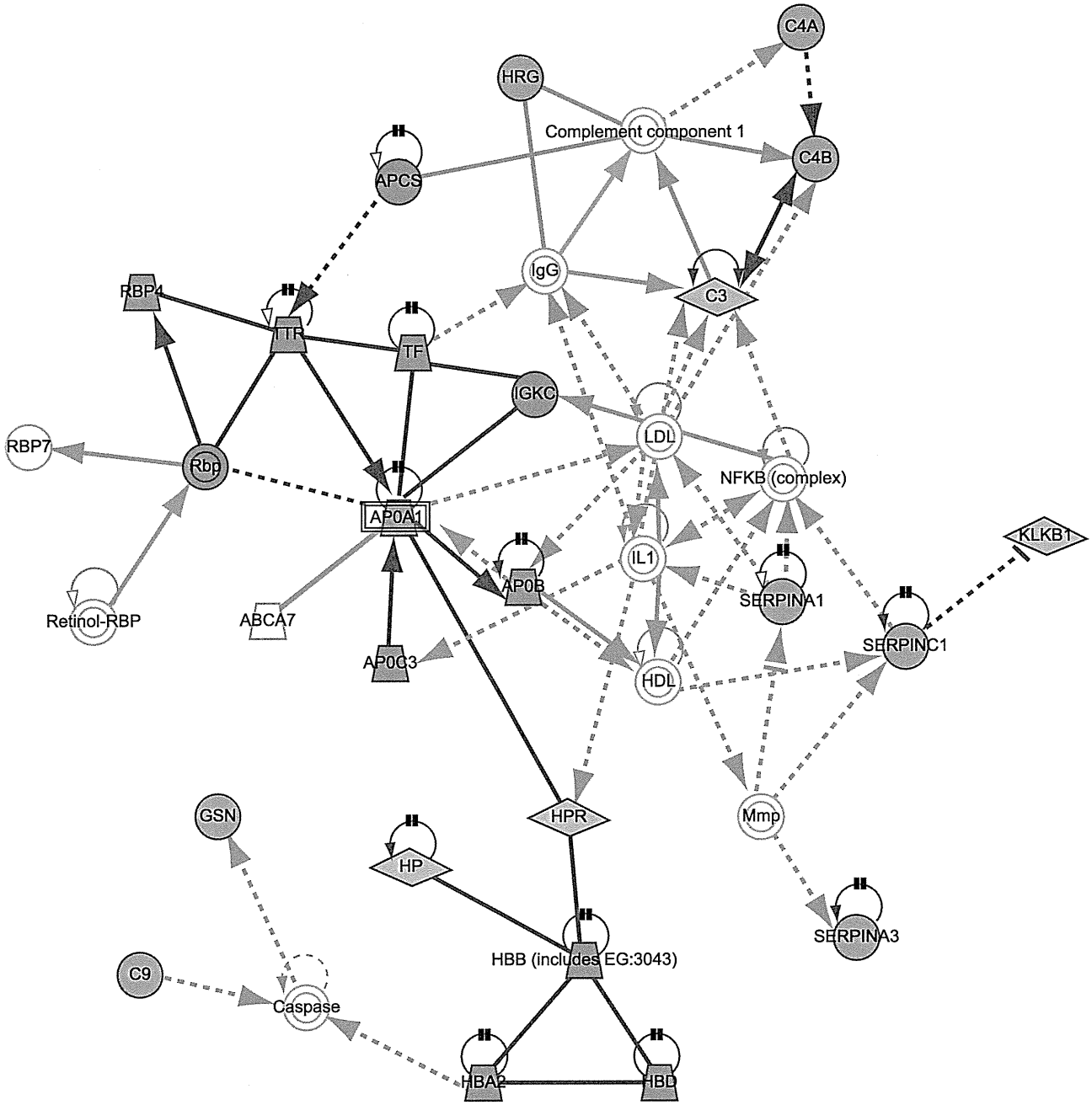


Figure 5. Highest scoring Ingenuity Pathway Analysis network. Highest scoring network generated from entering the identified 29 proteins into the Ingenuity Pathway Analysis system, with proteins identified in the study shaded grey and connecting proteins identified by Ingenuity Pathway Analysis non-shaded. Dark blue shapes and lines = proteins identified as predictors in this study and interactions between them. Grey shapes and lines = proteins identified by Ingenuity to generate the network and interactions between them. Light blue lines = interactions between proteins identified by Ingenuity to generate the network and the proteins identified in the study. Figure S4A shows this figure with the interaction relationships labeled. Proteins identified in the study and included in the network: SERPINA1 = alpha-1-antitrypsin; SERPINA3 = alpha-1-antichymotrypsin; SERPINC1 = antithrombin-III; APOA1 = apolipoprotein A-I; APOB = apolipoprotein B-100; APOC3 = apolipoprotein C-III; C3 = complement C3; C4A, C4B = complement C4-A; complement C4-B; C9 = complement component C9; GSN = gelsolin; HBA2 = hemoglobin alpha; HBB, HBD = hemoglobin beta/delta; HP = haptoglobin; HPR = haptoglobin-related protein; HRG = histidine-rich glycoprotein; KLKB1 = plasma kallikrein; IGKC = Ig kappa chain V-III region Ti; RBP4, Rbp = retinol binding protein 4; APCS = serum amyloid P-component; TF = serotransferrin; TTR = transthyretin. Proteins identified in the study and not included in the network: ORM1 = alpha-1-acid glycoprotein 1; A1BG = alpha-1B-glycoprotein; LRG1 = leucine-rich alpha-2-glycoprotein; ARMC2 = armadillo repeat-containing protein 2; AHSG = alpha-2-HS-glycoprotein; ITIH4 = inter-alpha-trypsin inhibitor heavy chain H4.
doi:10.1371/journal.pone.0022062.g005

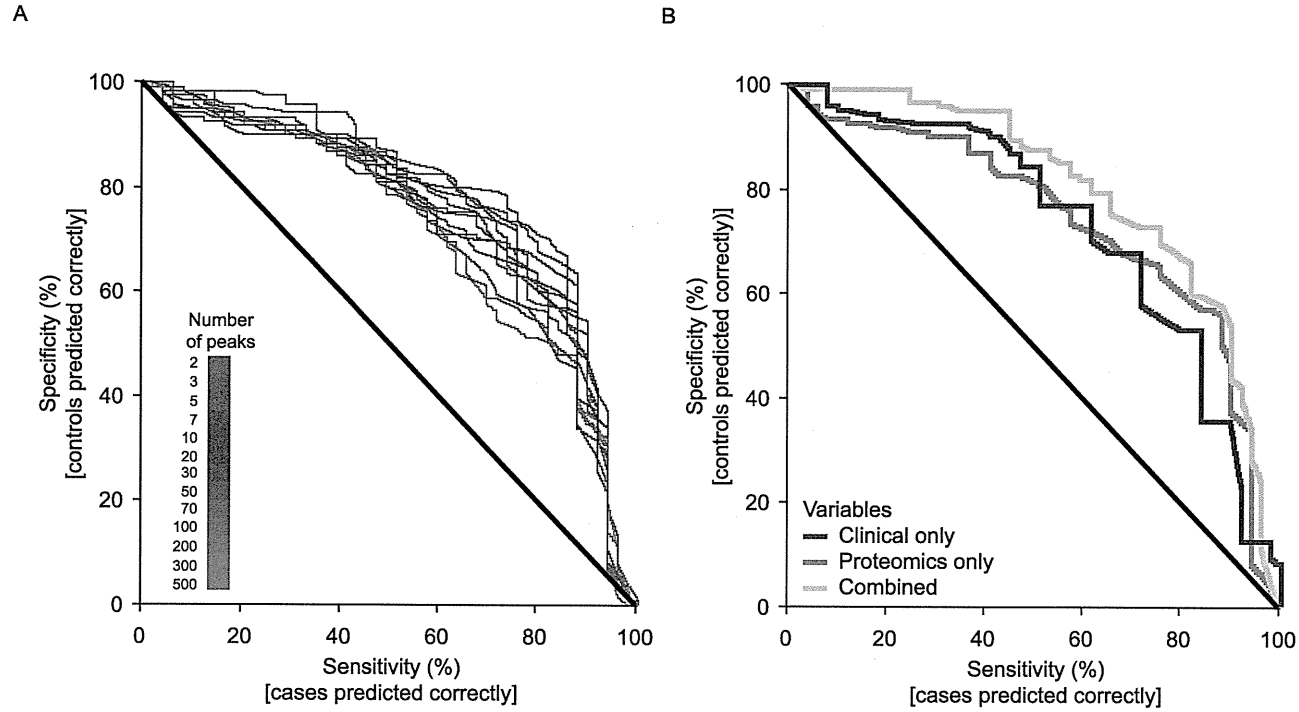


Figure 6. Receiver operating characteristics curve of cross-validated predictions. (A) from peptides, for different number of peptides included in the proteomic prediction model, and (B) from clinical data, proteomic data, and a combination of both clinical and proteomic data.
doi:10.1371/journal.pone.0022062.g006

minimize the effects of any systematic errors, and then eliminating batch effects in the statistical analysis. The consideration of batch effects in the statistical analysis appeared to improve the process of detecting discriminating peaks, and we therefore based our final protein identification step on this analysis approach.

Alignment algorithms – internal standard-guided Optimal Profile ALignment (i-OPAL). In order to allow comparative quantification between samples, the sample measurements have to be aligned – i.e. the correspondence of ion signals must be identified. Various methods have been proposed for this, e.g. based on stable isotope labeling [19–21], utilizing comparative identification [22,23], or with direct comparison of the respective peptide ion signals [18]. The i-OPAL method belongs to the last category. In spite of the relatively low accuracy and mass resolution of m/z measurement, the MS instrument of ion-trap type allows a long-term stable measurement without any calibration operation [24,25]. Consequently, the m/z values are directly comparable in a large set of samples without further transformations.

Biological findings and implications

Potential biological mechanisms underlying acute ILD events. In our IPA mapping to canonical biological pathways, acute phase proteins came out as the strongest signal, followed by the complement and coagulation pathways. Activation of acute phase response with connection to the complement and coagulation systems have been suggested as key processes in acute ILD events following blood transfusions (transfusion related acute lung injury; TRALI) [26] and in patients with idiopathic thrombocytopenic purpura (ITP) [27]. Acute phase responses can be induced by bronchoscopy with bronchoalveolar lavage [28]. Available evidence strongly suggests that balance between injury and repair is fundamental for regulating injury repair and protecting the lung [29]. In our study, clinical findings suggested that patients had increased risk of acute ILD events early after lung cancer diagnosis, if they had pre-existing chronic ILD, and if their remaining normal lung coverage as assessed on CT scan was low [3], suggesting that factors associated with active or extensive disease processes and/or cancer diagnostic procedures were important. Interestingly, another key protein signal (outside the acute phase pathway) for ILD risk in our study was gelsolin. Gelsolin was recently highlighted by comparative expression profiling and animal experiments as necessary for the development of modeled pulmonary inflammation and fibrosis, and caspase-3 mediated gelsolin fragmentation was shown to be an apoptotic effector mechanism and a marker of lung injury, again emphasizing the balance between injury and repair [30]. Interesting protein connections were also revealed by the IPA-generated networks. For example, CEBPB (NFIL6) is a principal effector of cyclin D1 activity in human cancer and an enhancer of e.g. IL-6 transcription, which plays an important role in the acute phase response [31,32]. It is important to clarify that IPA-generated networks are not the same as canonical biological pathways, but rather connect different proteins and genes based upon a wide range of interactions reported in the scientific literature.

Biomarker validation. Within this study we have validated our conclusions on several different levels. Technical validation by repeating the analysis of the same samples has confirmed the ability of the technology to reproducibly measure the levels of the peptides within the samples. This has been strengthened by the use of alternative technologies to confirm the intensities of key proteins. Together, these two sets of validation data show that the protein intensities derived from this MS/MS analysis are both reproducible and in agreement with those found from other technologies. In combination, they provide strong evidence of the

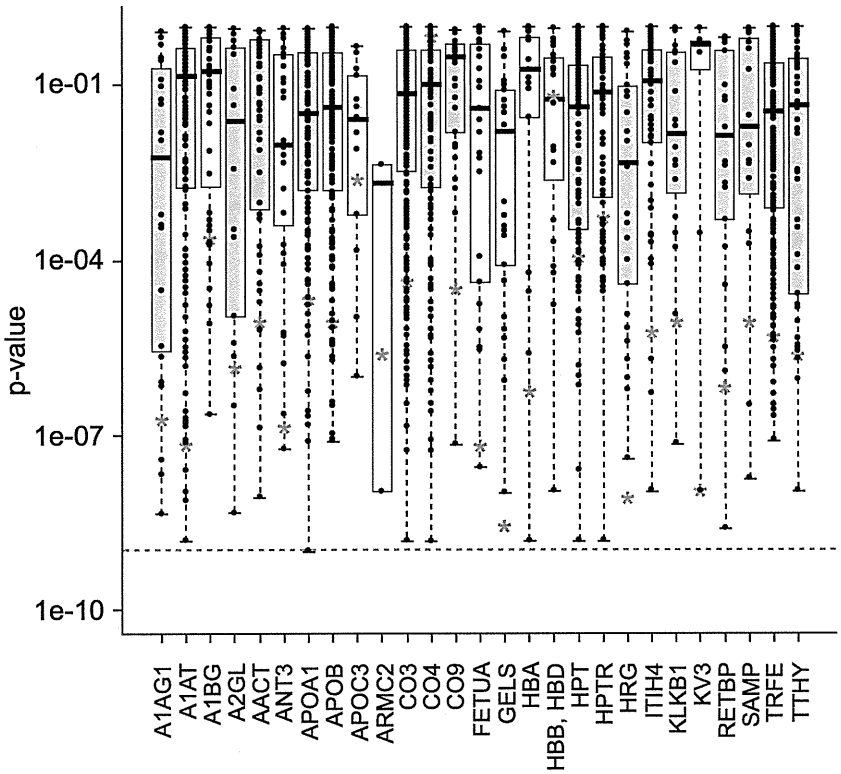


Figure 7. Significance levels from the proteins, constituent peptides, and acute phase pathway intensities. p-values for the proteins are shown by red stars, p-values for individual peptides are shown by points, and the distribution of these for each protein is shown by a boxplot. In each boxplot, the upper and lower sides of the box represent the higher and lower quartile values (Q3 and Q1), respectively. The black bar in each box represents the median value. The p-value for the acute phase pathway is represented by the dashed line; boxplots for proteins in the acute phase response pathway are shaded. A1AG1 = alpha-1-acid glycoprotein 1; A1AT = alpha-1-antitrypsin; A1BG = alpha-1B-glycoprotein; A2GL = leucine-rich alpha-2-glycoprotein; AACT = alpha-1-antichymotrypsin; ANT3 = antithrombin-III; APOA1 = apolipoprotein A-I; APOB = apolipoprotein B-100; APOC3 = apolipoprotein C-III; ARMC2 = armadillo repeat-containing protein 2; CO3 = complement C3; CO4 = complement C4-A; complement C4-B; CO9 = complement component C9; FETUA = alpha-2-HS-glycoprotein; GELS = gelsolin; HBA = hemoglobin alpha; HBB, HBD = hemoglobin beta/delta; HPT = haptoglobin; HPTR = haptoglobin-related protein; HRG = histidine-rich glycoprotein; ITIH4 = inter-alpha-trypsin inhibitor heavy chain H4; KLKB1 = plasma kallikrein; KV3 = Ig kappa chain V-III region Ti; RETBP = retinol binding protein 4; SAMP = serum amyloid P-component; TRFE = serotransferrin; TTHY = transthyretin.
doi:10.1371/journal.pone.0022062.g007

reliability of these data, giving confidence in further interpretation of the results. Biological validation has examined peptides from the same proteins and identified strong levels of correlation, strengthening and helping to confirm our hypotheses. Internally, cross-validation was used as an efficient method for avoiding over-fitting of a multivariate dataset and estimating an error rate of the modeling process whilst still maximizing the use of a limited number of samples. A final level of validation, which has yet to be addressed in this work, would be to validate our conclusions in a completely independently collected set of samples. Given the relative rarity of ILD and the difficult diagnosis, such a dataset remains to be assembled.

Biomarker properties. The true practical utility of any molecular diagnostic is not only its ability to make a prediction of outcome, but also its ability to add additional, alternative, and more timely information to assist the physician in treating the patient, at a reasonable cost and effort. With ILD, using clinical and radiological information which may often be easily available, a physician is able to make some assessment of the risk of a patient developing ILD, although this evaluation at present is inexact and difficult to apply consistently. While a patient's proteomic profile appears to provide similar prediction using an alternative method, and may even improve the accuracy of risk assessment when added to clinical/radiological information, that improvement is limited. However, the more objective and possibly more reproducible character of a proteomic measurement might provide advantages over a purely clinical assessment. It may also be noted that the additional value of any component of a risk score as assessed by ROC change is often very small, as has been demonstrated for well-known clinical lab tests such as HDL, HbA1C, and hsCRP in the context of clinically validated Reynolds Risk Score for cardiovascular disease [33], but that the individual components contributing to a score or data compilation used for a clinical decision may nevertheless all contribute to elevating the combined information above the threshold of clinical utility. Nevertheless, from both a medical and a commercial perspective and considering that further validation in independent sets of patients is still required, as well as development of a practical, cost-efficient, timely, and clinically available assay, it is not obvious at this point whether the possible added value justifies further development of the technology as a potential diagnostic.

Conclusion

This study has identified proteomic markers (peptides) that show a reasonable predictive power for ILD. However, as might be expected, the information they carry appears to overlap partly

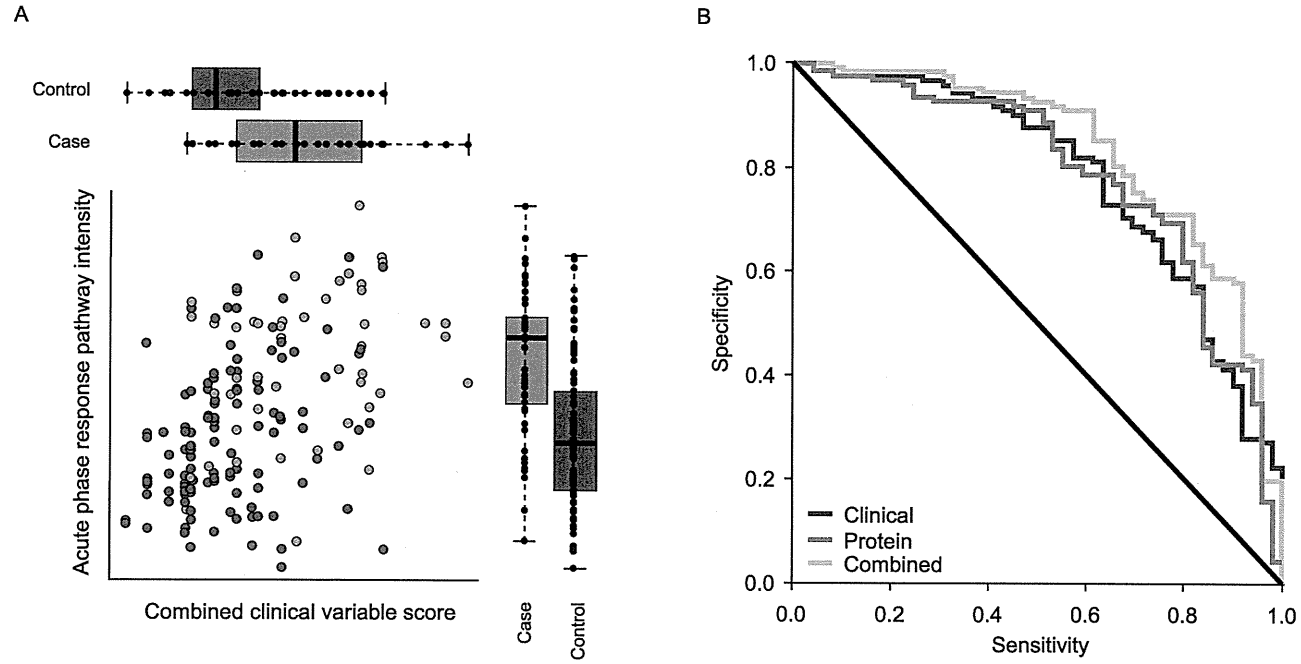


Figure 8. Plots illustrating the relationship between acute phase response pathway intensity score and clinical variable score. (A) Plot of the acute phase response intensity against the combined clinical variable score measuring the likelihood of a subject being a case calculated from a model predicting case-control status based only on the clinical variables WHO PS, smoking history, extent of normal lung coverage on CT scan, and severity of pre-existing ILD, with boxplots comparing the distribution of these measures in cases and controls. In each boxplot, the upper/right and lower/left sides of the box represent the higher and lower quartile values (Q3 and Q1), respectively. The black bar in each box represents the median value. (B) Receiver operating characteristics curve of cross-validated predictions from clinical data, the acute phase response intensity and a combination of the clinical data and acute phase response intensity.
doi:10.1371/journal.pone.0022062.g008

with that from clinical/radiological data previously shown to predict ILD, including WHO performance status score, smoking history, lung coverage on CT scan, and severity of pre-existing ILD. When combining the two, some gain in predictive power is obtained, although this is limited based on evaluation by ROC curves. Basing the predictive model upon the proteins/pathways identified as interesting shows potential for greater predictive power. In particular, the data suggest that activated acute phase response could be a marker for increased risk of acute ILD events following drug exposure. Whether this is a general mechanism that may be true also for acute ILD exacerbations in other settings, as some recent reports would suggest, is an important question for further research. In any case, our findings again highlight the potential importance of the balance between destruction and repair mechanisms for maintaining a functioning lung. If the results regarding the acute phase pathway can be confirmed, this may lead to a better mechanistic understanding of the basis for ILD events occurring, which would have potentially great future clinical utility as ILD events are an important consideration in the development of many potent new drugs, particularly in the areas of respiratory disease and oncology. In addition, such understanding would allow a more targeted approach to identifying and defining proteomic biomarkers with higher predictive value and clinical utility than was possible in this exploratory study.

Materials and Methods

Study design – patients and data collection

This non-randomized cohort study with a nested case-control component was conducted in November 2003 to February 2006 in 50 centers across Japan. Patients with advanced or recurring NSCLC with at least one previous chemotherapy regimen were eligible. Patients and their physicians selected a treatment (gefitinib 250 mg or chemotherapy), and follow-up was extended for up to 12 weeks after treatment initiation. Study design has been described in more detail previously [3]. This proteomics sub-study was performed only with gefitinib-treated patients from the case-control study component.

Patients who developed acute ILD during follow-up were registered to the case-control study nested within the cohort, as clinically diagnosed potential cases. For each potential case, four controls were randomly selected from patients then registered to the cohort who had not developed ILD.

To ensure valid and sensitive ILD diagnosis, the study included 1) an information card describing the symptoms of ILD distributed to all cohort patients; 2) internationally agreed criteria for ILD diagnosis and a diagnostic algorithm developed from an international consensus statement [1]; and 3) an independent CRB of radiologists and clinicians for blinded diagnostic review of all clinically diagnosed potential ILD.

For cases and controls, detailed patient data were collected covering NSCLC treatment, demography, cancer histology, clinical staging, WHO PS, smoking, previous cancer treatments, past and current medical history, surgical history, and concomitant medication and therapy. For gefitinib-treated cohort members consenting to the proteomics sub-study, one baseline 6-mL sodium-heparinized blood sample was drawn immediately (1–3 h) after the first gefitinib dose at registration into the cohort. Samples were spun for 10 min at 3,000 rpm and the plasma was stored at –80°C.

Subjects were recruited from 50 centers in Japan, and the Institutional Review Board of each site approved the study and informed consent forms. Written informed consent was required for registration to the study cohort, and separately for registration to the case-control study and for participation in the proteomics study.

Sample and data processing

The study was performed with quality control procedures at each sample processing, data generation, and data processing step, as described below.

Immunoaffinity depletion of serum albumin and IgG from the blood plasma and tryptic hydrolysis of plasma proteins. Depletion of the blood plasma samples was carried out using a dual Albumin and IgG Removal Kit (GE Healthcare UK Ltd, Amersham Place, Buckinghamshire, UK) according to the manufacturer's instructions, with some modifications. An affinity resin of a single production lot was used for the depletion process throughout this investigation. The depletion procedure for the 181 baseline samples was carried out in 18 batches (Table S1). As an experimental control, pooled human plasma (Sigma-Aldrich Inc., St. Louis, MO, USA) was simultaneously subjected to the same experimental procedures. Prior to the depletion the small debris part of the biofluid was removed by filtration. Aliquots (70 μL) of the plasma solution were diluted with 4.0 mL of the suspended gel slurry containing slurry beads with immobilized protein G binding polyclonal antibodies against both human serum albumin and IgG [34]. The sample was incubated on a rotator (5 rpm, 30 min, room temperature) to keep the gel slurry in suspension. Subsequently, the serum albumin/IgG-binding slurry beads were removed from the sample solution by centrifugation (7,000× *g*, 5 min) using a Vivaspin 6 column with polyethersulfone membrane (pore size 0.2 μm; Sartorius AG, Goettingen, Germany). The recovered liquid fraction (approximately 3.3 mL), containing unbound plasma proteins, was subjected to a buffer-exchanging process as follows: the fraction was condensed on a 3,000 molecular weight cutoff membrane of polyethersulfone in a diafiltration vessel (Vivaspin 2 column, Sartorius AG, Goettingen), followed by dilution with excess volume (2 mL) of 50 mM ammonium bicarbonate. This cycle was repeated three times. Finally the resulting solution was condensed to less than 100 μL on the same membrane and adjusted to 200 μL with 50 mM ammonium bicarbonate. The total protein concentration of the depleted solution was measured according to Lowry et al [35], using a DC protein assay kit (Bio-Rad Laboratories Inc., Hercules, CA, USA) with bovine serum albumin as the calibration standard. To confirm the depletion treatment, the concentration of the human serum albumin remaining in the depleted protein solution was measured as follows: An aliquot of the depleted solution was subjected to sodium dodecylsulfate (SDS) polyacrylamide gel electrophoresis (PAGE) [36]. Protein bands on the polyacrylamide gel were stained with a SYPRO Ruby fluorescence dye (Invitrogen Co., Carlsbad, CA, USA), followed by scanning of the gel slab on a LAS-3000 imaging system (FUJIFILM Co., Tokyo, Japan). Finally the fluorescence intensity ratio of the serum albumin band to the all protein bands detected on the gel lane was calculated using a Multi Gauge image analyzing software (FUJIFILM). The samples were then stored at −80°C until use.

Tryptic hydrolysis of the plasma proteins. The hydrolysis procedures were carried out in a single batch (Table S1). An aliquot containing 200 μg of the depleted plasma sample was spiked with 250 pmol of egg white lysozyme (Sigma-Aldrich Inc., St Louis, MO, USA) used as a source of exogenous internal standard peptides. Next the samples were denatured by incubating by gentle agitation in 200 μL of 100 mM ammonium bicarbonate containing 25% (v/v) acetonitrile at 37°C for 60 min. The resulting solution was immediately subjected to reductive *S*-carboxyamidomethylation of the sulfhydryl groups of the cysteine residues: incubation by gentle agitation at 37°C for 45 min with addition of 1 μmol of Tris(2-carboxyethyl)phosphine (TCEP) (20 μL of 50 mM solution), followed by the addition of 5 μmol of iodoacetamide (20 μL of 250 mM solution) at room temperature for 60 min in the dark. The Cys-modified proteins were further subjected to tryptic hydrolysis by the addition of 4 μg of porcine trypsin (20 μL of 0.2 mg/mL solution) (Promega Co., Madison, WI, USA) and incubation at 37°C for 16 h. The resulting peptide mixture was stored at −80°C until use.

To measure the degree of hydrolysis, aliquots of the sample solutions before and after the hydrolysis treatment were subjected to SDS PAGE [36], followed by obtaining a fluorescence gel image as described in the previous section. The total fluorescence intensities of the protein bands were compared between both solutions to calculate the protein portion remaining due to incompleteness of the hydrolysis.

LC-MS/MS measurement procedures. The peptide mixture was then dissolved in 1.0% v/v trifluoroacetic acid (TFA) aqueous solution with the final peptide concentration of 0.1 mg/mL, and analyzed using an LC-MS/MS system with a Finnigan LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) [25] in a fully automated manner. Briefly, peptide separation was performed with a Paradigm MS4 LC instrument (Michrom BioResources Inc., Auburn, CA, USA) containing a MAGIC C_{18} capillary LC column (0.2 mm id, 50 mm length, 3 μm particle size, and 200 Å pore size; Michrom BioResources). The mobile phase consisted of formic acid, acetonitrile, and water at a volume ratio of 0.1:2:98 for mobile phase A, and 0.1:90:10 for mobile phase B. The initial flow of 100 μL/min was reduced by a flow splitter to approximately 1 μL/min. 10 μL of the peptide solution, containing 1 μg peptide, was applied using an HTS PAL autoinjector (CTC Analytics AG, Zwingen, Switzerland) onto a Peptide CapTrap column (0.5 mm id, 2.0 mm length, bed volume 0.5 μL; Michrom BioResources) equilibrated with a solution of TFA, acetonitrile, and water at a volume ratio of 0.1:2:98. The peptides concentrated and purified on the trap column were injected onto the C_{18} capillary LC column by valve switching. The peptides were continuously eluted at a rate of 1 μL/min on a gradient mode: The initial ratio of 5% of mobile phase B was increased linearly to 40% B during 70 min, followed by the increase to 95% B during the next 5 min. After washing with a non-gradient flow at 95% B, the column was equilibrated again with the solvent of 5% B for the next separation. The total analysis time was 90 min. For gasification of the protonated peptides, the LC effluent was interfaced with an electrospray ionization (ESI) source in a positive ion mode, on a Finnigan LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific) [25]. The ESI used a FortisTip spray emitter (20 μm id, top Teflon-coated; AMR Inc., Tokyo, Japan) directly connected with the outlet of the LC column. The set parameters included a spray voltage of 2.0 kV and a capillary temperature of 200°C. No sheath gas was supplied during the ESI. The other parameters on the ion separation and detection were optimized according to an Autotune function in the mass spectrometer instrument. For MS/MS, protonated peptides in a gas phase were sequentially analyzed by data-dependent scanning mode of a full scan at an m/z range of 450 to 2,000 and subsequent product ion scans of the three most intense precursor ions. The data acquisitions were made in a Centroid mode for

both scans. The product ion scan was performed under conditions including an intensity threshold of 1×10^3, 30% normalized collision energy, 2.0 Da isolation *m/z* width, and dynamic exclusion for 30 sec. The ESI-MS/MS operation and data acquisition were carried out on an Xcalibur Revision 1.4 SR1 system controller (Thermo Fisher Scientific).

The LC-MS/MS measurement was continuously performed with alternate injection of the sample solutions and blank solutions for system washing. A single continuous LC-MS/MS measurement batch comprised 23 samples; 20 peptide solutions from patients and 3 experimental control samples of pooled plasma at the initial, middle, and final positions. The analysis order of patient samples was randomized across and within batches, whilst ensuring reasonable balance between batches in case/control status (Table 1).

Measurement variability of the ion intensity within each analysis batch was assessed as the CoV between the three control samples for the relative ion signal area of each of 6 selected common peptides, i.e. the ratio of the total signal area of the differently charged ions shown in Table 3, to the signal area of the most stably detected internal standard peptide (IS1) derived from the spiked lysozyme (Table 3).

Signal normalization, signal alignment, and peak detection in the LC-MS data. Two-dimensional profile data consisting of LC retention time and full MS scan (LC-MS data) were extracted from all 220 LC-MS/MS measurements (Table 1) and processed using the i-OPAL algorithm (Patent # WO/2004/09526 AI).

i-OPAL is a dynamic programming algorithm that maximizes the alignment between LC-MS profiles through shrinking or extending the retention time axis to maximize the similarity in peak shapes within the chromatograms, and the similarity of the mass spectra, using a wider range of criteria than alternative data processing methods including dynamic time warping (DTW) [37] or correlation optimized warping (COW) [37,38].

An important feature of the i-OPAL algorithm is its utilization of internal standards, which are forced to be aligned. This reduces the linear programming problem from the whole range of retention time to a series of small time sections, increasing the accuracy of alignment and reducing the computational time.

The i-OPAL program thus consists of 3 parts: 1) signal intensity normalization using one or more internal standards; 2) alignment of LC-MS data using internal standard signal sets and a dynamic programming algorithm; and 3) peak detection. First, the intensity of the whole signal was normalized across all samples using IS1. Second, alignment of the three internal standard signal sets (Table 3) was forced across the LC-MS data. Alignment of the remaining regions was carried out based on the dynamic programming algorithm. Following signal alignment, peak detection was performed using an iterative process.

A clear benefit of i-OPAL relative to other signal alignment approaches [39] is that peak detection is carried out after alignment rather than before. Most peak detection algorithms utilize the shape of peaks, making the integration of weak signals difficult. Aligning prior to peak detection increases the likelihood of detection of a peak across the range of samples, which can increase the numbers of peaks confidently detected.

Statistical analysis and modeling of peptide data

First, Variance Scaling Normalization [40] was applied to the peaks to scale the signal intensities from each sample to a common level and also remove any dependency between the mean and variance of the intensities. PCA [41] of the scaled data was used for exploratory data analysis to identify the main sources of variation within the proteomic data.

The analysis to identify single peptide markers associated with case status then proceeded using ANCOVA, testing significance using type III sums of squares for each peptide separately, with normalized peptide intensity as the response and case/control status as the explanatory variable, with adjustment for some or all

Table 3. Peptides used as internal standards and for assessment of the ion intensity variations across the LC-MS/MS measurement batches.

	Amino acid sequence (From – To)[a]	Protein name	Retention time, min[b]	Ion *m/z* value[c] (charge)	Swiss-Prot[d] accession number
Peptides for assessment of the ion intensity variations					
	EGTC[e]PEAPTDEC[e]KPVK (347–362)	Transferrin	10.8–16.6	910.0 (2+), 607.0 (3+)	P02787
	LRTEGDGVYTLNNEK (117–131)	Haptoglobin	19.8–25.2	1,709.9 (1+), 855.4 (2+), 570.6 (3+)	P00738
	AVGDKLPEC[e]EADDGC[e]PKPPEIAHGYVEHSVR (78–108)	Haptoglobin	27.7–32.5	1,717.9 (2+), 1145.6 (3+), 859.5 (4+)	P00738
	DYVSQFEGSALGK (52–64)	Apolipoprotein A–I	38.0–45.0	1,401.5 (1+), 701.3 (2+)	P02647
	HSTIFENLANKADRDQYELLC[e]LDNTR (226–251)	Transferrin	47.8–53.7	1,569.7 (2+), 1046.8 (3+), 785.4 (4+)	P02787
	TSESGELHGLTTEEEFVEGIYKVEIDTK (69–96)	Transthyretin	58.7–66.0	1,571.7 (2+), 1048.1 (3+), 786.4 (4+)	P02766
Internal standard (IS) peptides, two exogenous and one endogenous					
IS1	FESNFNTQATNR (52–63)	Lysozyme	19.5±2.0	714.8±2.4 (2+)	P00698
IS2	NTDGSTDYGILQINSR (64–79)	Lysozyme	36.2±2.0	877.4±2.4 (2+)	P00698
IS3	ITPNLAEFAFSLYR (50–63)	alpha-1-Antitrypsin	67.9±2.0	821.5±2.4 (2+)	P01009

[a]Residue numbers in the unprocessed precursor.
[b]Maximum range for all the analysis batches.
[c]Tolerance of ±0.5 *m/z* unit for the peak area calculation.
[d](http://expasy.org/sprot/).
[e]S-Carboxyamidomethyl cysteine residue.
doi:10.1371/journal.pone.0022062.t003