JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i2.74

Hemoglobin subunit beta (HBB) is a potential biomarker for predicting response to Gefitinib in NSCLC patients

Poh Kuan Chong[1], Huiyin Lee[2], Marie Chiew Shia Loh[2], Wan-Teck Lim[3], Siew Pang Chan[4,5], Ross Andrew Soo[2,6], Yoon Pin Lim*[1,7,8].

[1]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore; [2]Cancer Science Institute of Singapore, National University of Singapore; [3]Department of Medical Oncology, National Cancer Centre of Singapore; [4]Division of Health Research, Faculty of Health Sciences, La Trobe University, Australia; [5] Stata Users Group (Singapore); [6]Department of Hematology/Oncology, National University Hospital, Singapore; [7]Department of Biological Sciences, Faculty of Science, National University of Singapore; [8]Bioinformatics Institute, Agency for Science, Technology and Research, Singapore.

Received: 20 October 2011 **Accepted:** 13 December 2011 **Available Online**: 30 December 2011

Abstract

EGFR mutation status has been reported to correlate well with the response of NSCLC patients to Gefitinib. However, EGFR mutation analysis is invasive in nature and recent studies supported the notion that EGFR mutation was unable to predict response to Gefitinib in some patients. We therefore conducted plasma proteomics to identify potential biomarkers that are less invasive and whose expressions correlate more significantly to response to Gefitinib. To identify protein candidates that correlate with response to Gefitinib, we profiled the relative expression levels of plasma proteins between responders and non-responders prior to Gefitinib treatment. Relative quantification of plasma proteins were analysed using Isobaric Tags for Relative and Absolute Quantification (iTRAQ) and liquid chromatography-electrospray ionization (ESI) tandem mass spectrometry. Proteins that were commonly upregulated or downregulated amongst responders but not the non-responders were selected for validation via immunoblotting. HBB protein was found to be significantly under-expressed in the plasma samples from 6 out of 7 gefitinib-responsive patients but over-expressed in a majority of the non-responders. Our finding showed that HBB is a potential biomarker for predicting response to Gefitinib that may be subject to a larger study to examine its role as a companion biomarker for Gefitinib therapy.

Keywords: NSCLC; Plasma; Proteomics; Gefitinib; iTRAQ; HBB

1. Introduction

The majority of lung cancers are non-small cell lung cancer (NSCLC) and most are diagnosed in late stage resulting in poor prognosis [1,2]. Tyrosine kinase inhibitors show great promises for cancer therapy [3]. Gefitinib (Iressa) is a highly selective tyrosine kinase inhibitor of the epidermal growth factor receptor (EGFR) that is used in the treatment of patients with advanced stage NSCLC. Initial studies have shown that gefitinib improved the quality of life in some NSCLC patients but the response rate was low [4, 5].

Subsequently, sensitizing somatic *EGFR* mutation, predominantly deletions in exon 19 and L858R point mutations in exon 21, was shown to be associated with increased sensitivity to gefitinib [6, 7]. Furthermore, in randomised studies, the response rates of patients with sensitizing *EGFR* mutations treated with gefitinib was 62-74%, while the response rate was 1% for those without mutations [8-10]. This therefore argues for detection of *EGFR* mutations as a means to select patients for gefitinib. However, there are several challenges. First, access to tumour samples for analysis is limited. In many instances, diagnostic samples (e.g. fine needle aspirates) provide poor quality or insufficient amount after diagnostic pathology not to mention its invasive nature. Second, patients without *EGFR* mutation could still respond to gefitinib albeit at low lower rate [9]. This implies that factors other than *EGFR* mutation may predispose patients to gefitinib.

***Corresponding author:** Yoon Pin Lim. Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore. Email Address: bchlyp@nus.edu.sg

Conceivably, markers other than EGFR mutation exist that could predict response to gefitinib. Similarly, it is possible that surrogate marker(s) for EGFR mutation exists in biological fluids that are more easily accessible than tumor tissues. Blood sampling is minimally invasive and does not require repeated biopsies. The notion and feasibility of identifying markers for predicting drug response through blood-related proteomics has been demonstrated [11, 12]. Hence, we aim to profile the prior-to-treatment baseline expression level of proteins in the plasma of patients who responded to gefitinib versus those who did not.

We adopted the Isobaric Tags for Relative and Absolute Quantification (iTRAQ) approach to compare the plasma protein expression profiles of 7 responding and 6 non-responding individual patients. The objective of this pilot study is to identify candidate predictive biomarker that can be subject to larger-scale clinical trials to validate its utility in predicting response to gefitinib.

2. Material and Methods

Blood collection

Patients with histologically or cytologically confirmed advanced stage NSCLC who were planned for gefitinib therapy were enrolled into the study. Blood collection was obtained at baseline (prior to Gefitinib treatment) and at the end of every treatment cycle. Basic patient demographics were collected. Tumour response was evaluated after every two cycles according to the RECIST criteria [13]. Disease control was defined as patients who had stable disease or better. The National University of Singapore Institutional Review Board approved the study and all patients gave written informed consent.

The protocol for blood collection and plasma preparation is based on findings from the Plasma Proteome Project (PPP) [14]. To ensure the consistency in plasma preparation, a quality control measure was implemented where all the blood samples must be processed into plasma in the laboratory within an hour after blood collection in the clinic. The integrity of each plasma sample is further verified by running 1D-SDS-PAGE and stained with SyproRuby fluorescent dye to ensure check for massive protein degradation. Samples that did not satisfy the specified time frame or integrity check would be stored away and will not be used for the study.

Samples from 13 patients were used amongst which 7 responded to gefitinib whereas the remaining 6 had disease progression following gefitinib treatment.

Plasma depletion and iTRAQ labeling

One mL of plasma sample was first subject to delipidation by centrifugation at 130,000 x g at 4°C for 2 hr. The bottom layer of lipids-free plasma was collected and the total protein was estimated using BCA assay. Plasma sample from each patient were then depleted using MARS Hu-7 affinity column (Agilent Technologies, USA) according to the manufacturer's protocol. Total protein after depletion was estimated using BCA assay. Protein samples were then reduced, alkylated, digested and labelled with iTRAQ reagents according to the recommended protocol (Applied Biosystems, Framingham, MA, USA).

LC-MS/MS analysis

For each iTRAQ experiment, the labelled peptides were fractionated into 30 fractions using strong cation exchange using a PolySULFOETHYL™ A Column (PolyLC, Columbia, MD, USA) 5µm of 200mm length × 4.6mm ID, 200Å pore size. These fractions were cleaned-up using a C18 Discovery® DSC-18 SPE column (100mg capacity, Supelco, Sigma-Aldrich). The dried and cleaned fractions were then analyzed using Agilent 1100 nLC system (Agilent) coupled online to a quadruple time of flight mass spectrometer (QStar XL, Applied Biosystems), as described in reference [15]. Eluent from the reverse phase nLC was directly subjected to positive ion nanoflow electrospray analysis in an information dependant acquisition mode (IDA). A ToF MS survey scan was acquired (m/z 370-1600, 0.5 sec), with the 3 most intense multiple charged ions (counts >70) sequentially subjected to MS/MS analysis. The time of summation of MS/MS events was set to be 2 seconds in the mass range of m/z 100-1600.

Similarly to the previous study, protein identification and quantification were carried out using ProteinPilot™ software (version 2.0; Applied Biosystems, MDS-Sciex), searching against IPI human database (version 3.41) [15]. The search was performed using Paragon Algorithm™, which is discussed in detail elsewhere [16]. Only those proteins identified and quantified with at least 95% confidence were taken into account. All results were then exported into Excel for manual data interpretation.

Immunoblotting

For validation of iTRAQ result, plasma samples used for iTRAQ analysis were subjected to immunoblotting for HBB and SAA using mouse monoclonal antibody (Santa Cruz), as described in previous studies [17, 18]. Triplicates blots were carried out for each sample to ensure that the generated data re reliable. For densitometry, images from x-ray film were first captured using Imager Scanner and its corresponding software LabScan version 5.0 (GE Healthcare). Image was then analyzed using ImageQuantTL software v2003.03 (GE Healthcare).

EGFR mutation analysis

Formalin-fixed, paraffin-embedded tumour samples of the cases were obtained from the Pathology departments of the two participating centres i.e. National University Hospital and National Cancer Centre, Republic of Singapore. DNA

was extracted from 5 μm sections of each sample as described previously [19]. Mutations in exons 18–21 of *EGFR* were detected by direct sequencing as described previously [20].

Statistical analysis

To ascertain if there was a significant difference in the levels of HBB protein expressions (obtained from iTRAQ) between the responsive and non-responsive patients, the Mann -Whitney U test was applied. The Fisher exact tests were also performed to determine if there was significant association between EGFR mutation (0: Non wild-type, 1: Wild-type) and the clinical and demographic variables (e.g., gender, ethnicity, smoking status, histology, disease status and response towards treatment). The nonparametric tests were proposed in anticipation of the violation of normality assumption and cell sparsity owing to the small sample size. Using SPSS 16.0 (SPSS Inc, Illinois, USA) and Stata 10.0 (Stata Corp, Texas, USA), all statistical tests were conducted at 5% level of significance.

3. Results and discussion

Proteome profiling of plasma samples from responders and non-responders before gefitinib treatment

The characteristics of the 13 patients recruited for this study is shown in Table 1. Briefly, 7 responded to gefitinib treatment and 6 did not. All the patients, of Asian origin, were diagnosed to suffer from stage IV diseases. They were mostly never smokers (84.6%) and had adenocarinoma (69.2%). Six patients (46%) received gefitinib as first-line, 4 (31%) as second-line and 3 (23%) as third-line treatment. Mutation analysis for *EGFR* mutations was performed in ten patients in whom tumour tissues were available. EGFR mutations were detected in three of these patients.

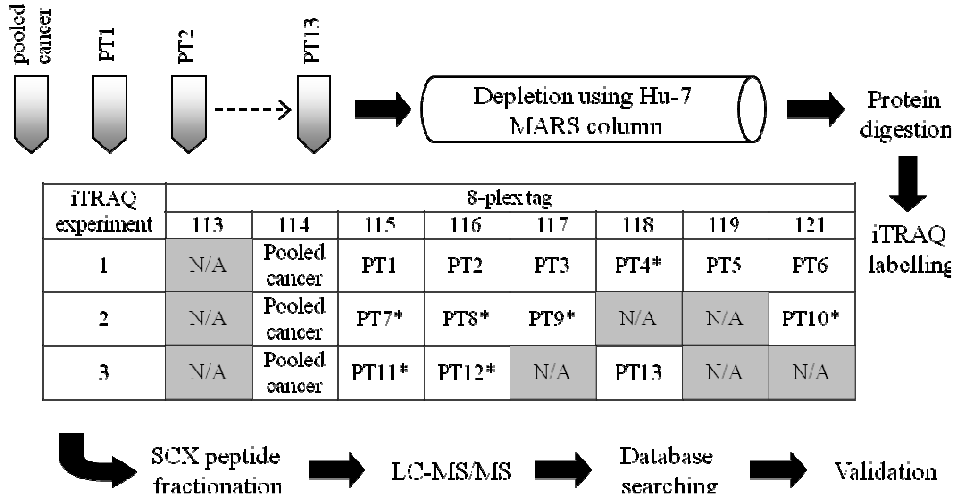
The plasma protein expression profile of individual patients before gefitinib treatment was analyzed in triplicates using quantitative iTRAQ approach. The experimental design showing how all 13 independent samples were analyzed using 8-plex iTRAQ is shown in Figure 1. Baseline plasma samples from all 13 patients were pooled and used as an internal control so that data could be normalized and compared across all the samples. Hence the relative quantification for each individual samples (7 responding and 6 non-responding individual patients) was relative to this pooled sample.

Only common proteins identified across the 3 iTRAQ datasets generated were selected for further consideration. A total of 115 common proteins were identified and quantified. The complete list of the proteins identified in these 3 iTRAQ datasets and their common proteins are tabulated in Supplementary Material 1. The 115 common proteins were then characterised based on their molecular function using Panther Classification (http://www.pantherdb.org) [21]. There were 177 protein assignments were obtained and sorted into 23 molecular classifications, since some of these proteins

Table 1. Sample Characteristics (including clinical data) of NSCLC patients

Patient ID	Age	Gender	Race	Smoking status	Histology	Disease Stage	Line of gefitinib treatment	Response	EGFR mutation status ✝
PT1	59	M	Malay	Never	Adenocarcinoma	IV	1st	No	NA
PT2	55	F	Chinese	Never	Poorly differentiated	IV	1st	No	NA
PT3	62	F	Chinese	Never	Adenocarcinoma	IV	1st	No	exon 21, 858 point MT
PT4	69	M	Chinese	unknown	Large cell carcinoma	IV	2nd	Yes	NA
PT5	75	M	Chinese	Former	Squamous cell	IV	3rd	No	WT
PT6	51	F	Chinese	Never	Adenocarcinoma	IV	3rd	No	exon 19 del 746 mt
PT7	64	F	Chinese	Never	Adenocarcinoma	IV	1st	Yes	WT
PT8	55	M	Chinese	Never	Adenocarcinoma	IV	1st	Yes	WT
PT9	61	F	Chinese	Never	Adenocarcinoma	IV	2nd	Yes	Inconclusive
PT10	62	F	Chinese	Never	Adenocarcinoma	IV	2nd	Yes	WT
PT11	57	F	Chinese	Never	Adenocarcinoma	IV	2nd	Yes	WT
PT12	64	F	Chinese	Never	Poorly differentiated	IV	3rd	Yes	exon 21, 858 point MT
PT13	60	F	Chinese	Never	Adenocarcinoma	IV	1st	No	WT

✝ NA: No available tissue for EGFR analysis; WT: Wild type; Inconclusive : The analysed result was not conclusive.



iTRAQ experiment	8-plex tag							
	113	114	115	116	117	118	119	121
1	N/A	Pooled cancer	PT1	PT2	PT3	PT4*	PT5	PT6
2	N/A	Pooled cancer	PT7*	PT8*	PT9*	N/A	N/A	PT10*
3	N/A	Pooled cancer	PT11*	PT12*	N/A	PT13	N/A	N/A

Figure 1. The overview of experimental workflow employed in the study. A total of 3 8-plex iTRAQ datasets containing the baseline (before gefitinib treatment) plasma proteome profiles of the responders versus non-responders patients towards gefitinib treatment were generated. Patients who responded positively towards gefitinib treatment were denoted with * in the figure.

have multiple classifications assigned (Supplementary Material 2). The majority of the common proteins were grouped under defence/immunity protein followed by protease and transport/carrier protein – typical of proteins found in the plasma. The details protein classification for each group is shown in Supplementary Material 2.

Selection of candidate proteins that may segregate responders from non-responders

From the list of common proteins identified, proteins were considered up or down-regulated when their ratios were >1.3 or <0.77, with their p-value <0.05. This cut-off value was determined from our preliminary analysis, which showed that the technical variation from duplicate sets of iTRAQ experiments was less than 30% (Supplementary Material 3). This technical variation has been consistently obtained in published and unpublished studies from our lab [22, 23]. The biological variations of protein expressions in the plasma of different subjects vary from one protein to another. Some are highly variable while others are not so. As such, we could only implement a cut-off based on technical variation, something that we can determine.

To help short-list proteins that might segregate responders

Table 2. Proteins that exhibited the same expression trend (over or under-expression) in at least 60% of the responder or non-responder group of patients. Proteins were considered differentially expressed (**bold**) when the protein ratio was found to be either >1.3 or <0.77, with p -value <0.05. Other details including p-value and error factor for these proteins can be tabulated in Supplementary Material 1.

Gene symbol	Protein name	Responsive patients						Non-responsive patients					
		PT4 : Pooled cancer	PT7 : Pooled cancer	PT8 : Pooled cancer	PT9 : Pooled cancer	PT10 : Pooled cancer	PT11 : Pooled cancer	PT12 : Pooled cancer	PT1 : Pooled cancer	PT2 : Pooled cancer	PT3 : Pooled cancer	PT5 : Pooled cancer	PT6 : Pooled cancer	PT13 : Pooled cancer
SAA2	Serum amyloid A2 isoform a	**0.44**	**0.56**	**0.27**	**1.81**	**0.20**	**0.75**	**0.63**	0.80	**0.44**	1.14	1.17	0.94	**0.77**
HBB	Hemoglobin subunit beta	**0.58**	**0.34**	0.97	**0.76**	**0.33**	**0.25**	**0.35**	**1.80**	**1.38**	**0.59**	**1.31**	0.91	**3.11**
LOC400682	Similar to hCG1996858	**1.31**	**0.42**	**0.57**	0.78	**0.31**	**0.56**	0.89	0.87	1.15	1.21	**1.86**	**1.30**	1.25
FGB	Fibrinogen beta chain	**1.35**	**1.30**	**0.72**	**0.58**	**1.67**	**1.34**	1.07	0.96	1.28	**0.70**	1.12	0.84	**0.50**
F12	Coagulation factor XII	**1.46**	**1.33**	1.05	1.22	**1.39**	**1.32**	0.80	0.97	1.04	**0.77**	1.02	0.92	1.00
SERPINA1	Isoform 1 of Alpha-1-antitrypsin	**1.52**	**0.31**	**0.47**	**0.69**	**0.72**	**1.63**	1.06	**0.72**	**1.34**	**1.72**	0.84	1.12	**0.24**
LUM	Lumican	**1.60**	**1.69**	0.80	0.96	**1.37**	**1.54**	1.27	0.82	**0.70**	**0.69**	**0.69**	0.91	**0.56**

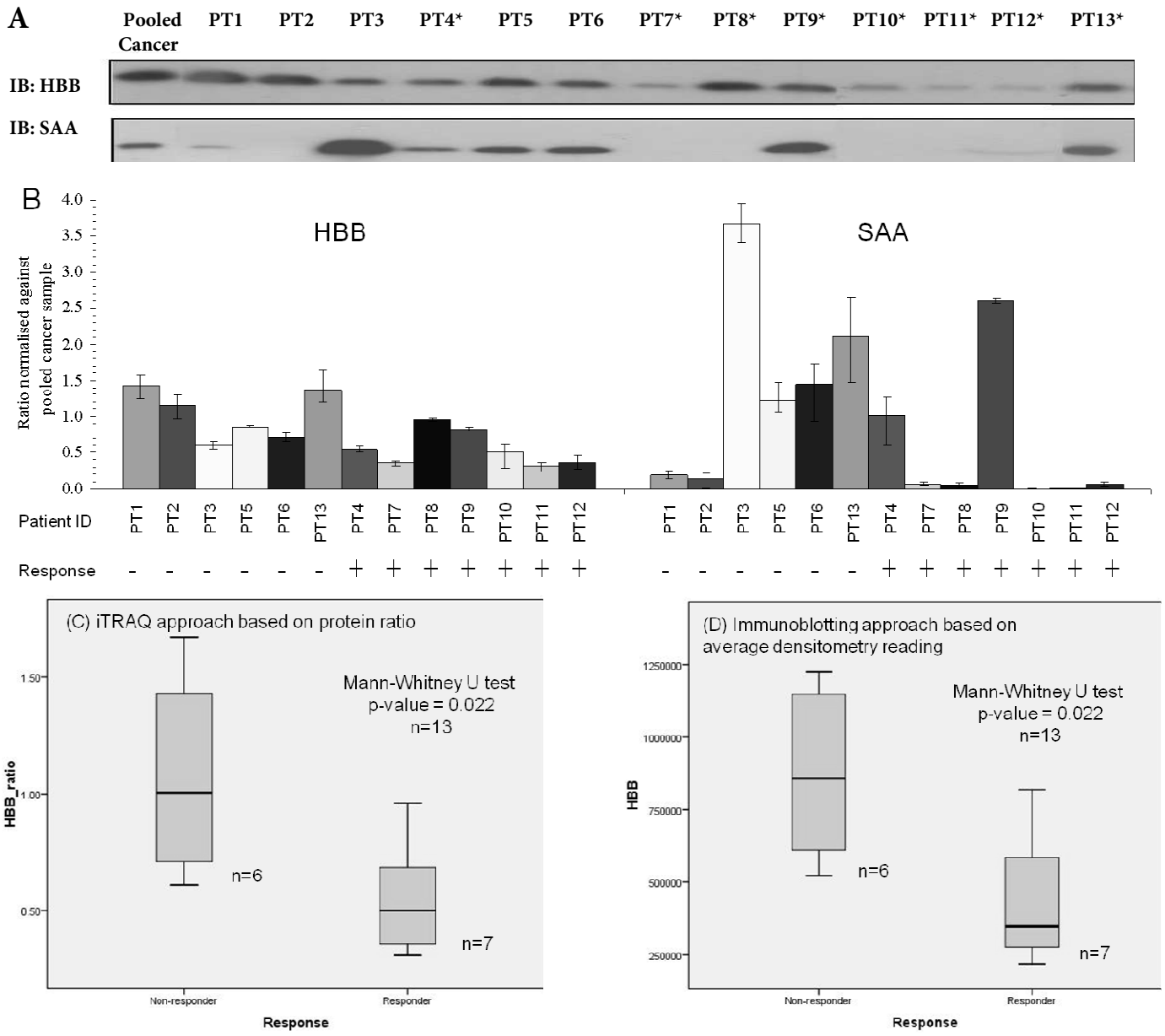


Figure 2. (A) Representative immunoblots for both HBB and SAA protein validation. Patients who responded positively towards gefitinib treatment were denoted with * in the figure. **(B)** Average densitometry readings of triplicate immunoblots for individual NSCLC patients in the study. Box plots showing the distribution of HBB protein expression obtained through **(C)** iTRAQ and **(D)** immunoblotting approaches among the responsive and non-responsive patients. Mann-Whitney U tests revealed that statistically significant difference (p-value<0.001) in the HBB expresion level was observed for both plots.

from non responders, we first grouped the patients into responders or non responders. Next, we selected proteins that displayed the same expression trend (over- or under-expression) in at least 60% of the sample size within the responder or non responder group. This resulted in a total of 7 candidate proteins and they are presented in Table 2. Other proteins showed random expression trend across the responder and non-responders group. Hence they were not considered further in this study.

Verification of HBB and SAA expression levels in the plasma samples from responding and non-responding patients

From Table 2, 2 proteins namely SAA and HBB were found to be under-expressed in 6 out of 7 patients (86%) who responded to gefitinib treatment. This is a higher frequency compared to the rest like LOC400682 and LUM, which were significantly under-expressed and over-expressed in 4 out of 7 cases, respectively. The striking differences in the expression levels of SAA and HBB between the responders and non-responder led us to investigate these two candidates more closely as they might represent bi-

omarkers that predict response to gefitinib. The rest of the candidates like LOC400682 and LUM could not be pursued further in part because of the lack of commercially available antibodies or because the antibodies were of poor quality that did not allow us to interpret the data accurately and with high confidence.

We first set out to examine the protein expression levels of SAA and HBB in the plasma using immunoblotting to determine whether the iTRAQ data could be verified. Figure 2A shows a representative blot each from immunoblotting of SAA and HBB. The densitometry readings for these 2 candidate proteins obtained through immunoblotting were shown in Figure 2B. Overall, the results obtained for both proteins using immunoblotting and iTRAQ approaches were congruent.

Next, statistical analysis was performed to examine whether the divergence in protein expressions of both SAA and HBB among the responsive and non-responsive patients was statistically significant. Using Mann-Whitney U test, the p-values for HBB protein expression level obtained through iTRAQ and immunoblotting were both 0.022 respectively. Figure 2C and 2D show the box plots of HBB protein expression obtained from iTRAQ and immunoblotting approach, respectively. The data indicates that the difference in HBB expression levels between responsive and non-responsive group of patients was significant. In contrast, no statistical significance was obtained for SAA protein (p-value > 0.05) (refer to Supplementary Material 4). Although the remaining 5 candidate proteins were not studied further in this study for reasons discussed above, the results of the statistical analyses of their expression levels between the responsive and non-responsive group of patients were nevertheless included in Supplementary Material 4 as a reference.

EGFR mutation status and response to Gefitinib in the cohort studied

Of the ten patients with EGFR mutation analysis performed, six were *EGFR* wild-type. Four out of 6 patients (67%) with *EGFR WT* responded to gefitinib. This was unexpectedly high given the response rate in patients *EGFR WT* to gefitinib can be as low as 1-3% [8, 24] and even higher range reported is about 38% [25]. Gefitinib has been shown to be more effective in never smokers amongst Asian NSCLC patients[26]. The fact that 5 out of 6 patients with no EGFR mutation in this study were never smokers could have contributed to the higher response rate to gefitinib. On the other hand, we observed only 1 out of 3 patients (33%) with EGFR mutation who responded to gefitinib. This is lower than the reported range of up to 75%. The discrepancy may result from the small sample size. It may also be due to the fact that 1 and 2 patients with EGFR mutation received gefitinib as a first line and third line treatment, respectively. So far, most of the studies on the correlation of EGFR mutation with response to gefitinib were conducted involving the use of gefitinib as 3-rd line treatment. These confounding factors should be considered when designing future larger scale studies. It may also signal the need to discover population-specific biomarkers for predicting response of NSCLC patients to gefitinib therapy. Thus far, we are not aware that such a study has been conducted in Singapore.

To determine whether there exist a candidate surrogate marker for EGFR mutation, statistical correlation analysis between EGFR mutations and relative protein expression data was performed. The results of further assessments (Fisher exact test Mann-Whitney U test) concerning EGFR (0: Non wild-type including an inconclusive case, 1: Wild-type) are shown in Table 3. EGFR mutations status was not significantly associated with gender, age, smoking status, histology and response towards gefitinib treatment. No statistical test could be performed for disease status (all stage IV) and ethnicity (all Chinese). We caution that it is difficult to generalize the reported results of this pilot study owing to the small sample size. However, such a pilot study and the results, although preliminary in nature, are useful for planning a complete study involving more observations.

4. Discussion

There is sufficient evidence showing that not all NSCLC patients respond to gefitinib. As such, there is a need for

Table 3. Statistical analysis of EGFR mutation status.

Variable	MGFR Mutation Status		p-value
	Wild-type	Non wild-type	
Response towards gefitinib treatment			
No	2	2	0.999
Yes	4	2	
Histology			
Adenocarcinoma	5	3	0.667
Large cell carcinoma	1	0	
Poorly differentiated	0	1	
Disease stage			
IV	6	4	N.A.
Ethnicity			
Chinese	6	4	N.A.
Gender			
Female	4	4	0.467
Male	2	0	
Smoking status			
Ex-smoker	1	0	0.999
Non-smoker	5	4	
Age (years)	Median: 61	Median: 61.5	0.830
	Range: 55-75	Range: 51-64	

† Based on Mann-Whitney U test.

companion biomarkers to stratify patients to achieve cost effectiveness in cancer management. To this end, proteomics analysis of lung adenocarcinoma tissues from patients who showed different response to gefitinib was reported. Using a support vector machine algorithm, 9 proteins were selected that could distinguish responders from non responders. Differential expression of one of the nine proteins, heart-type fatty acid-binding protein was successfully validated [27]. While useful, the potential biomarkers identified via analysis of tissues means that invasive surgical methods have to be employed if these biomarkers were to be exploited. Biomarkers are best tested in body fluids like blood that are minimally invasive.

With respect to this, one study conducted serum protein profiling and obtained an algorithm based on 8 distinct mass peaks that could predict the outcomes following gefitinib treatment. In one of their datasets, the median survival in the predicted "good" and "poor" outcome groups was 207 and 92 days, respectively [28]. The identities of the 8 mass peaks are not known. A similar study also generated an algorithm of MALDI-mass peaks from serum profiling that could distinguish patients with different response to treatment with a combination of gefitinib and rofecoxib in platinum-pretreated NSCLC patients [29]. While potentially useful, the use of such "biomarkers" represented by mass peaks with unknown identifies is risky since it is not clear whether they are specific to the sample preparation or analytical method used.

It is likely that an effective "biomarker" would consist of a panel of easy-to-access biomarkers with increased sensitivity and specificity than individual biomarkers. By combining various biomarkers identified from various studies, one may eventually be able to test such a combination of biomarkers for predicting drug response. To contribute to this cause, our study attempted to identify potential biomarkers that could distinguish between patients who respond versus those who don't respond to gefitinib. This was achieved by profiling the plasma proteins from 7 responders and 6 non-responders to gefitinib. There are limitations in this study. First, our study has a higher proportion of females (9 out 13) than males (4 out of 13). However, the adjustment of gender imbalance is not feasible given the small sample size of this pilot study. A multivariate statistical technique which allows such demographic imbalances to be adjusted should be considered in future studies involving more observations. Second, the sample size of the study is small. Since this was meant to be a pilot study, we did not perform sample size calculation prior to data collection. The generated results, based on exploratory statistical techniques, would be useful for calculating the appropriate sample size for a complete study later.

Nonetheless, we have identified two proteins that were differentially expressed in gefitinib-responding and non-responding NSCLC patients were discovered. One of them is Hemoglobin subunit beta (HBB), one of the two polypeptide chains in adult haemoglobin. It plays an important role in oxygen transportation from lung to various peripheral tissues. Various studies focused on the mutation of HBB since its defect can lead to numerous blood disorder diseases such as beta-thalassemia and sickle cell anemia [30]. On the other hand, HBB gene expression level was reported to be down-regulated in breast tumour tissue [31] and anaplastic tyroid cancer cell lines [32] compared with normal tissue/cells. In a glycoproteomic study on human lung adenocarcinoma tissue, HBB was also reported to be down-regulated compared to the normal tissue [33]. In this study, we showed that HBB protein expression was low in the majority of the plasma samples of NSCLC patients who were responsive to gefitinib treatment compared with the non-responders. This implies that HBB may be used to predict patient's response to gefitinib. It is not entirely clear how a lower HBB expression may influence drug response. It is conceivable that a lower HBB expression results in a hypoxic condition and that hypoxic tumours, which are metabolically stressed, may be more susceptible to gefitinib. While it is possible that HBB may be cancer cell-specific, we do not rule out the possibility that lower amount of HBB might be related to physiological processes such the oxygen carrying capacity of the erythrocytes.

5. Conclusion

In conclusion, our pilot study shows that HBB is a candidate discriminatory biomarker that is able to segregate between gefitinib-responding and non-responding NSCLC patients. The key finding from this study justifies the design of a future study with a larger sample size to validate the potential of HBB as a predictive biomarker for gefitinib therapy. Due to the expensive nature of Gefitinib, accrual of patients on this drug for future study will be a challenge. It is envisaged that such a large-scale validation study is best undertaken by the industry or a major cancer program/ consortium.

6. Supplementary Material

Supplementary data and information is available at: http://www.jiomics.com/index.php/jio/rt/suppFiles/74/0

Supplementary Material 1-List of common proteins identified from 3 iTRAQ dataset generated from 13 NSCLC baseline plasma samples (before Iressa treatment); *Supplementary Material 2* - Molecular classification using Panther; *Supplementary Material 3* - Technical variation estimation in iTRAQ labeling; *Supplementary Material 4* - Statistical analysis of protein expression between responsive and non-responsive group of patients.

Acknowledgements

This work was supported by funding from the Singapore Cancer Syndicate, Agency for Science, Technology and Research (A*STAR). Dr Chong PK was supported by the Lee

Kuan Yew Endowment Funds.

References

1. O.S. Breathnach, B. Freidlin, B. Conley, M.R. Green, D.H. Johnson, D.R. Gandara, M. O'Connell, F.A. Shepherd, B.E. Johnson, J Clin Oncol 19 (2001) 1734-1742.
2. N. van Zandwijk, G. Giaccone, Curr Opin Oncol 8 (1996) 120-125.
3. L.K. Shawver, D. Slamon, A. Ullrich, Cancer Cell 1 (2002) 117-123.
4. M. Fukuoka, S. Yano, G. Giaccone, T. Tamura, K. Nakagawa, J.Y. Douillard, Y. Nishiwaki, J. Vansteen kiste, S. Kudoh, D. Rischin, et al, J Clin Oncol 21 (2003) 2237-2246.
5. M.G. Kris, R.B. Natale, R.S. Herbst, T.J. Lynch, Jr., D. Prager, C.P. Belani, J.H. Schiller, K. Kelly, H. Spiri donidis, A. Sandler, et al, Jama 290 (2003) 2149-2158.
6. T.J. Lynch, D.W. Bell, R. Sordella, S. Gurubhagavatu la, R.A. Okimoto, B.W. Brannigan, P.L. Harris, S.M. Haserlat, J.G. Supko, F.G. Haluska, et al, N Engl J Med 350 (2004) 2129-2139.
7. J.G. Paez, P.A. Janne, J.C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F.J. Kaye, N. Lindeman, T.J. Boggon, et al, Science 304 (2004) 1497-1500.
8. T.S. Mok, Y.L. Wu, S. Thongprasert, C.H. Yang, D.T. Chu, N. Saijo, P. Sunpaweravong, B. Han, B. Mar gono, Y. Ichinose, et al, N Engl J Med 361 (2009) 947-957.
9. T. Mitsudomi, S. Morita, Y. Yatabe, S. Negoro, I. Okamoto, J. Tsurutani, T. Seto, M. Satouchi, H. Tada, T. Hirashima, et al, Lancet Oncol 11 (2010) 121-128.
10. M. Maemondo, A. Inoue, K. Kobayashi, S. Sugawara, S. Oizumi, H. Isobe, A. Gemma, M. Harada, H. Yo shizawa, I. Kinoshita, et al, N Engl J Med 362 (2010) 2380-2388.
11. L. Pusztai, B.W. Gregory, K.A. Baggerly, B. Peng, J. Koomen, H.M. Kuerer, F.J. Esteva, W.F. Symmans, P. Wagner, G.N. Hortobagyi, et al, Cancer 100 (2004) 1814 -1822.
12. Z. Xiao, B.T. Luke, G. Izmirlian, A. Umar, P.M. Lynch, R.K. Phillips, S. Patterson, T.P. Conrads, T.D. Veenstra, P. Greenwald, et al, Cancer Res 64 (2004) 2904-2909.
13. P. Therasse, S.G. Arbuck, E.A. Eisenhauer, J. Wan ders, R.S. Kaplan, L. Rubinstein, J. Verweij, M. Van Glabbeke, A.T. van Oosterom, M.C. Christian, S.G. Gwyther, J Natl Cancer Inst 92 (2000) 205-216.
14. G.S. Omenn, D.J. States, M. Adamski, T.W. Black well, R. Menon, H. Hermjakob, R. Apweiler, B.B. Haab, R.J. Simpson, J.S. Eddes, et al, Proteomics 5 (2005) 3226-3245.
15. S. Lim, L.Y. Choong, C.P. Kuan, C. Yunhao, Y.P. Lim, J Proteome Res 8 (2009) 4062-4076.
16. I.V. Shilov, S.L. Seymour, A.A. Patel, A. Loboda, W.H. Tang, S.P. Keating, C.L. Hunter, L.M. Nuway sir, D.A. Schaeffer, Mol Cell Proteomics 6 (2007) 1638 -1655.
17. Y.P. Lim, L.S. Diong, R. Qi, B.J. Druker, R.J. Epstein, Mol Cancer Ther 2 (2003) 1369-1377.
18. Y.P. Lim, B.C. Low, J. Lim, E.S. Wong, G.R. Guy, J Biol Chem 274 (1999) 19025-19034.
19. R. Soong, B.J. Iacopetta, Mod Pathol 10 (1997) 252-258.
20. T.M. Chin, D. Anuar, R. Soo, M. Salto-Tellez, W.Q. Li, B. Ahmad, S.C. Lee, B.C. Goh, K. Kawakami, A. Segal, et al, Clin Chem 53 (2007) 62-70.
21. P.D. Thomas, M.J. Campbell, A. Kejariwal, H. Mi, B. Karlak, R. Daverman, K. Diemer, A. Muruganujan, A. Narechania, Genome Res 13 (2003) 2129-2141.
22. Y. Chen, L.Y. Choong, Q. Lin, R. Philp, C.H. Wong, B.K. Ang, Y.L. Tan, M.C. Loh, C.L. Hew, N. Shah, et al, Mol Cell Proteomics 6 (2007) 2072-2087.
23. P.K. Chong, H. Lee, J. Zhou, S.C. Liu, M.C. Loh, T.T. Wang, S.P. Chan, D.T. Smoot, H. Ashktorab, J.B. So, et al, J Proteome Res (in press) (2010).
24. F.R. Hirsch, M. Varella-Garcia, P.A. Bunn, Jr., W.A. Franklin, R. Dziadziuszko, N. Thatcher, A. Chang, P. Parikh, J.R. Pereira, T. Ciuleanu, et al, J Clin Oncol 24 (2006) 5034-5042.
25. G.Y. Ku, B.A. Haaland, G. de Lima Lopes, Jr., Lung Cancer 74 (2011) 469-473.
26. S.T. Lim, E.H. Wong, K.L. Chuah, S.S. Leong, W.T. Lim, M.H. Tay, C.K. Toh, E.H. Tan, Br J Cancer 93 (2005) 23-28.
27. T. Okano, T. Kondo, K. Fujii, T. Nishimura, T. Taka no, Y. Ohe, K. Tsuta, Y. Matsuno, A. Gemma, H. Kato, et al, Clin Cancer Res 13 (2007) 799-805.
28. F. Taguchi, B. Solomon, V. Gregorc, H. Roder, R. Gray, K. Kasahara, M. Nishio, J. Brahmer, A. Spreafi co, V. Ludovini, et al, J Natl Cancer Inst 99 (2007) 838-846.
29. K.J. O'Byrne, S. Danson, D. Dunlop, N. Botwood, F. Taguchi, D. Carbone, M. Ranson, J Clin Oncol 25 (2007) 3266-3273.
30. G. Lettre, V.G. Sankaran, M.A. Bezerra, A.S. Araujo, M. Uda, S. Sanna, A. Cao, D. Schlessinger, F.F. Costa, J.N. Hirschhorn, S.H. Orkin, Proc Natl Acad Sci U S A 105 (2008) 11869-11874.
31. I.B. Pau Ni, Z. Zakaria, R. Muhammad, N. Abdullah, N. Ibrahim, N. Aina Emran, N. Hisham Abdullah, S.N. Syed Hussain, Pathol Res Pract 206 (2010) 223-228.
32. M. Onda, J. Akaishi, S. Asaka, J. Okamoto, S. Miya moto, K. Mizutani, A. Yoshida, K. Ito, M. Emi, Br J Cancer 92 (2005) 2216-2224.
33. J.H. Rho, M.H. Roehrl, J.Y. Wang, Protein J 28 (2009) 148-160.
34. D.H. Liu, X.M. Wang, L.J. Zhang, S.W. Dai, L.Y. Liu, J.F. Liu, S.S. Wu, S.Y. Yang, S. Fu, X.Y. Xiao, D.C. He, Biomed Environ Sci 20 (2007) 33-40.
35. S. Dai, X. Wang, L. Liu, J. Liu, S. Wu, L. Huang, X. Xiao, D. He, Sci China C Life Sci 50 (2007) 305-311.
36. N. Khan, C.J. Cromer, M. Campa, E.F. Patz, Jr., Can cer 101 (2004) 379-384.
37. W.M. Gao, R. Kuick, R.P. Orchekowski, D.E. Misek, J. Qiu, A.K. Greenberg, W.N. Rom, D.E. Brenner, G.S. Omenn, B.B. Haab, S.M. Hanash, BMC Cancer 5 (2005) 110.
38. W.C. Cho, T.T. Yip, C. Yip, V. Yip, V. Thulasiraman, R.K. Ngan, T.T. Yip, W.H. Lau, J.S. Au, S.C. Law, et al, Clin Cancer Res 10 (2004) 43-52.