Development of Expression Based Biomarkers in NSCLC: A Study of Intratumor Heterogeneity Using FFPE Tissue

Timothy S. Davison[1], Robert J. Holt[1], Sian Dibben[1], Janet Taylor[1], Claire Wilson[1], Ian Paul[2], Kieran McManus[3], Paul J. Kelly[3], Donna McIlwaine[1], Kerry Lavery[1], Rachael Wheavil[1], Lisa McCoy[1], Vitali Proutski[1], Paul Harkin[1,2], Peter Kerr[1], Richard Kennedy[1,2], Dean A. Fennell[2] and Jacqueline A. James[2]

[1] Almac Diagnostics, [2] The Centre for Cancer Research and Cell Biology, Queen's University, Belfast, [3] Belfast Health and Social Care Trust

There is considerable interest in the use of gene expression data to generate predictive and prognostic biomarkers from archived non-small cell lung cancer (NSCLC) tissue (Potti *et al*., 2006).

In early stage disease, up to twenty FFPE blocks are available from each patient. Since NSCLC tumors exhibit histological diversity, it is possible that areas of individual tumors represented by different FFPE blocks will exhibit distinct molecular profiles (Sakurada *et al*., 2008). This intratumoral heterogeneity represents a potential problem for the development of biomarkers from NSCLC tissue.

This study examines gene expression data in multiple FFPE blocks taken from individual patients. Variations in gene expression are analysed between:

- FFPE blocks of individual patients
- Whole FFPE sections and macrodissected tumor tissue
- Individual patients
- Disease histologies (squamous and adenocarcinoma)

The study also evaluates the impact of these variations on powering clinical biomarker discovery studies.

Methods

STUDY DESIGN

• FFPE blocks from ten patients were chosen from a single sample cohort (6 adenocarcinoma and 4 squamous, Table 1). Block selection aimed to constrain block age, stage (Goldstraw *et al.,* 2007) and % tumor (at least 50%). Between three and five blocks were available for each sample.

Sample Name	Gender	Age at Surgery	Histology	Number of Blocks	Block Age (years)	AJCC 6 Stage	AJCC 7 Stage
SQ1	Male	63	Squamous	3	7.8	IB	IA
SQ2	Female	70	Squamous	3	8.0	IB	IIA
SQ3	Male	81	Squamous	5	6.9	IIIA	IIB
SQ5	Male	75	Squamous	4	8.7	IB	IA
AD1	Male	59	Adeno	4	6.2	IB	IB
AD2	Female	61	Adeno	4	7.4	IIIA	IIIA
AD3	Male	69	Adeno	3	7.8	IA	IA
AD4	Male	68	Adeno	4	7.9	IV	IV
AD5	Male	71	Adeno	3	8.3	IV	IV
AD6	Male	73	Adeno	3	11.3	IIIB	IIIB

Table 1. Sample Characteristics

SAMPLE PROCESSING

• Block cutting and all subsequent processing steps were randomised to eliminate confounding batch effects. RNA was extracted from whole FFPE sections as well as from macrodissected tumor tissue

• cDNA was amplified and biotin labelled using the Nugen FFPE V2 Amplification and FL-Ovation kit. Samples were then processed onto the Lung Cancer DSA™, a lung cancer specific DNA microarray. 70 arrays were available for data analysis

DATA ANALYSIS

• Array data was pre-processed using RMA background correction, normalisation and summarisation implemented in Partek® 6.4. Affymetrix control probes were filtered out prior to subsequent analysis

• Unsupervised PCA was performed using the correlation matrix for all 60355 non-control probesets

• The independent variables (histology, patient, block, macrodissection, and block age) were fit to a mixed-model ANOVA and evaluated for their absolute and relative contribution to the sources of variation as seen in the dependent variables (normalised gene expression values)

• The impact of gene expression variation within patient block and between patient on powering clinical biomarker discovery and validation studies using either whole or macrodissected FFPE material from either Squamous or Adenocarcinoma was assessed following methods outlined in Pintilie *et al.*, 2009:

• The error in biomarker measurement (e_k) due to within and between patient variability (heterogeneity) is defined as:

$$e_k = \frac{\hat{\sigma}_w^2 / k}{\hat{\sigma}_w^2 / k + \hat{\sigma}_B^2}$$

where k is the number of replicates measured within patient and $\hat{\sigma}_w^2$ and $\hat{\sigma}_B^2$ are the within patient variance and between patient variance respectively

• The resulting power equation becomes $\mathbf{Z}_{1-\beta} = \sqrt{n_d}\,\hat{\sigma}_B \sqrt{1-e_k}\,\log(HR) - z_{1-\alpha/2}$ where HR is the hazard ratio, n_d is the number of events, and α and β are the types I and II error respectively, $Z_{1-\beta}$ and $Z_{1-\alpha/2}$ are the quantiles of the standard normal distribution for 1- β and 1- $\alpha/2$ respectively and e_k and $\hat{\sigma}_B$ are defined above

• Simulated signatures: gene expression data was weighted and summarised into a single 'signature expression' (SX) value for each of the combinations of signature size (n = 4, 8, 10, 50, 100, 500 and 1000 features) and sampling strategy (random, median and sliding scale). Random weights were generated by sampling n/2 weights from a normal distribution with mean 0.5 and standard deviation 0.25, where n is signature length. The sign of these weights was then reversed to total n weights with sum 0

Results

PRINCIPLE COMPONENT ANALYSIS

• Unsupervised principal component analysis of global gene expression data indicates that variations in gene expression are present between FFPE blocks taken from a single patient (Figure 2). However, the variations in gene expression observed between individual patients and disease states are greater

• The intratumor heterogeneity of adenocarcinomas appears to be less than that of squamous samples

• These results are also reflected in the PCA for both macrodissected tissue and whole FFPE sections (Figures 3 & 4)

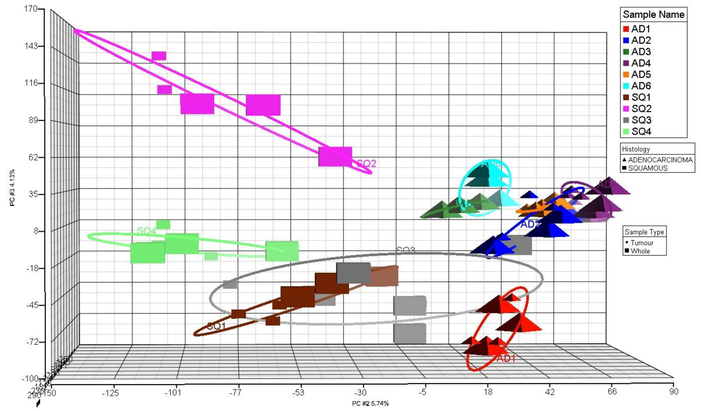

Figure 2. PCA of Global Gene Expression (all samples)

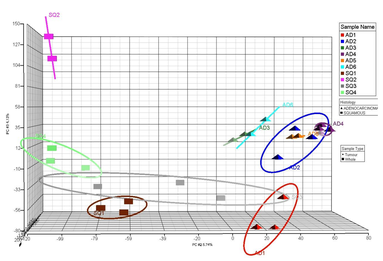

Figure 3. PCA of Macrodissected Tumor Tissue

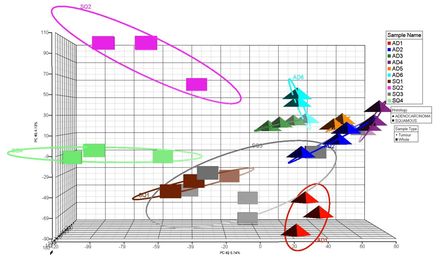

Figure 4. PCA of Whole FFPE Sections

SOURCES OF VARIATION

• A mixed-model ANOVA model was fit to the gene expression data to assess the statistical significance of observed variations within and between samples and histologies (Figure 5).

• This indicates that variations due to histology and patient variation are greater than block to block variation within patient (same tumor)

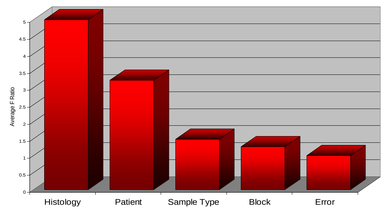

Figure 5. Sources of Variation

POWERING CLINICAL BIOMARKER STUDIES

• The number of events required to power a single biomarker study were calculated for Adenocarcinoma and Squamous FFPE samples profiled as either macrodissected or whole slide sections (Table 2, α=0.05, β=0.20, HR=2)

• Calculations are shown for one to three patient replicate samples with either the median or 25th percentile estimates of measurement error as well as the naïve application of the power equation where measurement error is not considered

• The simulation of different signature sizes results in a reduction of observed heterogeneity, with a corresponding decrease in the number of events required to power an equivalent study (Table 3, α=0.05, β=0.20, HR=2; random sampling of features; 100 000 simulations per sample type/ signature size)

• Failure to consider biomarker variability within tumor tissue results in underestimates of required events sufficient to power the biomarker study by as much as 57% depending on histology and section type

Sample Type		Naïve	1x	2x	3x	1x	2x	3x
Adeno	Tumor	62	85	73	70	75	69	67
	Whole	70	103	86	81	89	79	76
Squamous	Tumor	61	81	71	67	70	65	64
	Whole	75	118	97	90	98	87	83
			Median Error			25th Percentile Error		

Table 2. Estimated Number of Events Required to Power a Biomarker Study

Sample Type /Signature Size		4	8	10	50	100	500	1000
Adeno	Tumor	235	195	186	158	153	149	149
	Whole	264	217	208	170	164	161	160
Squamous	Tumor	212	172	166	137	132	127	126
	Whole	284	228	217	172	162	159	158

Table 3. Estimated Number of Events Required to Power a Signature Based Biomarker Study (Random Sampling)

Conclusions

• Minor variations in gene expression are present between FFPE blocks of an individual patient; however the variations observed between patients and histologies are significantly larger

• The use of samples with high tumor content allows whole FFPE sections to be used for biomarker development without the need for macrodissection

• Failure to consider intratumor heterogeneity may result in underpowered biomarker development studies

• As the contribution of information per feature is reduced, and more features are required to observe the same effect size, simulations suggest studies with fewer events will lead to the development of signatures with more features.

• Almac Diagnostics are currently developing a NSCLC prognostic gene signature using FFPE samples. This work will allow the generation of a robust signature using a single FFPE sample from each patient

References

1. Potti *et al*., 2006. N Engl J Med. 355:570-80
2. Sakurada *et al*., 2008. J Thorac Oncol. 3:527-529
3. Goldstraw *et al*., 2007. J Thorac Oncol. 2:706-714
4. Tanney *et al*., 2008. BMC Medical Genomics. 1:20
5. Mixed model ANOVA was performed using Partek® software, v6.4 Copyright © 2009 Partek Inc., St. Louis, MO, USA
6. Pintilie *et al*., 2009. J. Clin. Oncol. 9:1517-1521

Acknowledgements

The authors would like to thank Ken Arthur at the CCRCB Tissue Core Lab.