



Gene expression of lung squamous cell carcinoma reflects mode of lymph node involvement

J.E. Larsen^{*,#}, S.J. Pavay^{*,†}, R. Bowman[#], I.A. Yang^{*,#}, B.E. Clarke⁺, M.L. Colosimo^{*}, N.K. Hayward^{*,†} and K.M. Fong^{*,#}

ABSTRACT: Tumour, node, metastasis staging is essential for lung cancer management. However, similarly staged cancers may have markedly different prognoses, indicating that stage cannot completely explain tumour behaviour. While ipsilateral hilar node involvement is designated N1, the current authors hypothesised that primary tumours involving nodes by direct extension are biologically distinct from those involving nodes through lymphatic metastasis.

Microarrays were used to investigate the gene expression profiles of 59 primary lung squamous cell carcinomas, comparing N0 tumours (n=35), N1 tumours by direct extension (N1d; n=8), and N1/N2 tumours by lymphatic metastasis (N1/N2m; n=16).

Hierarchical clustering using 125 genes differentially expressed between N0 and N1/N2m tumours found N1d tumours clustered with N0 tumours. Class prediction modelling found the expression profiles of all eight N1d tumours were more similar to N0 than to N1/N2m tumours.

The present study demonstrates for the first time that N1 tumours directly invading hilar nodes are genomically different to those that metastasise via lymphatics. Independent reports suggest that tumours with direct, rather than metastatic node involvement have better outcomes. Consequently, the data suggest that there is a need to re-evaluate the N1 staging definition in lung cancer. This is relevant for prognosis prediction and also for clinical management, particularly in selecting those patients most likely to benefit from adjuvant chemotherapy.

KEYWORDS: Expression profiling, nodal disease, nonsmall cell lung carcinoma, squamous cell

Recent randomised trials have demonstrated survival benefit (4–15% at 5 yrs) for adjuvant chemotherapy after resection of nonsmall cell lung cancer (NSCLC) [1–4]. A meta-analysis of 4,584 patients with NSCLC concluded that the benefit of adjuvant cisplatin-based chemotherapy depends largely on disease stage; chemotherapy appeared detrimental in stage IA, had no effect in stage IB, and improved survival in stage II and stage III [5]. Given the cost and toxicity of adjuvant chemotherapy, intense efforts are being made to identify subsets of patients who will respond.

Recently, gene expression profiling has become a major tool to identify sources of biological heterogeneity [6–11]. This heterogeneity is likely to be a major factor underlying the variable

behaviour of lung cancers, including their response to adjuvant chemotherapy.

The present authors hypothesised that tumours currently staged as N1 (involving lymph nodes distal to mediastinal pleural reflection and within visceral pleura: hilar, interlobar, lobar, segmental and subsegmental) [12] comprise two biologically distinct subgroups: 1) N1 by direct extension (involving nodes immediately adjacent to the primary tumour); and 2) N1 by lymphatic metastases (involving nodes nonadjacent to the primary tumour). This has direct clinical relevance, as these two subsets of N1 disease are reported to have significantly different prognoses, with the former having a better outcome [13, 14].

AFFILIATIONS

^{*}Depts of Thoracic Medicine and,
[†]Pathology, The Prince Charles Hospital, Brisbane,
[#]School of Medicine, University of Queensland, and
⁺Human Genetics Laboratory, Queensland Institute of Medical Research, Herston, Australia.

CORRESPONDENCE

J.E. Larsen
Dept of Thoracic Medicine
The Prince Charles Hospital
Brisbane
4032
Australia
Fax: 61 731394957
E-mail: Jill_E_Larsen@health.qld.gov.au.

Received:

December 11 2006

Accepted after revision:

March 14 2007

SUPPORT STATEMENT

J.E. Larsen is the recipient of an Edwin Tooth PhD Scholarship of the University of Queensland. K.M. Fong is the recipient of a Queensland Smart State Fellowship. N.K. Hayward is a recipient of a National Health and Medical Research Council (NHMRC; Canberra, Australia) Fellowship. This work was supported in part by the NHMRC (338200) and The Prince Charles Hospital Research Foundation (Brisbane, Australia).

STATEMENT OF INTEREST

None declared.

For editorial comments see page 5.

[†]This article has supplementary material accessible from www.erj.ersjournals.com

Primary lung squamous cell carcinoma (SCC) has a propensity to arise proximally, resulting in both modes of nodal involvement being observed. To test the hypothesis that SCCs invading regional lymph nodes by direct extension are molecularly distinct from those involving nodes through lymphatic metastases, the present authors analysed microarray gene expression data in resected primary SCC with and without locoregional lymph node metastases.

METHODS

Detailed methods are provided as online supplementary material.

Study subjects

A total of 59 patients undergoing curative-intent surgical resection of primary lung SCC without neoadjuvant chemotherapy or radiotherapy were recruited for the study (table 1). Ethics approval was granted from the Prince Charles Hospital Human Research Ethics Committee (Brisbane, Australia) and patients gave written informed consent. Lobectomy was performed in 51 patients and pneumonectomy in eight patients. Histological examination of the resected tumour, dissected lymph nodes and associated

lung was used to determine pathological stage of the cancer according to the criteria of MOUNTAIN [12]. All patients had clinical follow-up, with a median (range) duration of 58 (5–178) months.

Primary lung SCCs were stratified into three groups after pathological examination: 1) N0 (n=35); 2) hilar N1 nodes by direct extension (N1d; n=8); and 3) N1/N2 nodes by lymphatic metastases (N1/2m; n=16). Direct invasion of N1 lymph nodes (those nodes distal to the mediastinal pleural reflection and within the visceral pleura, including hilar, interlobar, lobar, segmental and subsegmental) [12] was defined as an infiltrative extension of viable tumour cells without loss of continuity from the primary tumour mass to the lymphatic tissue of an adjacent node. Nodes containing tumour cells not in continuity with the primary tumour were assumed to be involved by lymphatic metastasis [15].

Gene expression microarrays and statistical analysis

Gene expression profiles from tumours were obtained using a 22K Human V2.0 Oligo Microarray (Operon Biotechnologies, Cologne, Germany) according to Minimum Information About a Microarray guidelines as described previously [9]. Two statistical analyses were performed to compare the gene expression profiles of SCCs. The first was a supervised analysis, using a Wilcoxon Mann–Whitney U-statistic, to identify genes/transcripts with statistically significant gene expression (p<0.01; 1,000 permutations) between N0 and N1/2m tumours. Significant genes or transcripts were then used to cluster all tumours hierarchically (N0, N1d and N1/2m). Secondly, class prediction models were developed to predict whether N1d tumours were biologically more similar to N0 tumours or to N1/2m tumours. This approach did not represent an attempt to classify N1d tumours as N0 or N1/2m tumours, but rather to identify to which class (N0 or N1/2m) N1d tumours were most similar. Given the reported variability of prediction modelling [16–19], five independent models were used. For each, models were built from the training set of 35 N0 and 16 N1/2m samples (with the prediction error of each model estimated by leave-one-out cross-validation (LOOCV)) and used to predict classification of the test set of eight N1d tumours.

RESULTS

A supervised analysis of N0 and N1/2m tumours identified 125 transcripts differentially expressed (supplementary table 1). A two-dimensional hierarchical cluster of all 59 SCCs, including N1d tumours, was performed using the 125 genes (fig. 1a) and showed that N1d tumours clustered with N0 rather than N1/2m tumours. These genes were chosen because their expression most discriminates (in this cohort) between two distinct categories of lung SCC: 1) tumours with no detectable lymph node metastasis; and 2) tumours with lymphatic metastasis. Figure 1a indicates that in N1d tumours, the expression pattern of these 125 genes share most similarity with that observed in N0 tumours.

The genomic profile of N1d tumours was analysed further by developing class prediction models, with the aim of identifying whether N1d tumours were most similar to the N0 or the N1/2m classes, rather than reclassifying N1d tumours as N0 or N1/2m. Five independent prediction models were used; for each,

TABLE 1 Characteristics of patients and tumours	
Characteristic	
Subjects n	59
Age yrs	68 (62–72)
Sex M/F	73/27
Smoking status	
Never	3
Former	27
Current	70
Smoking history pack-yrs	50 (33–80)
Differentiation	
Well	20
Moderate	31
Poor	42
Stage	
IA T1N0M0	8
IB T2N0M0	22
IIA T1N1M0	2
IIB T2N1M0	16
IIB T3N0M0	3
IIIA T3N1M0	4
IIIA T2N2M0	2
IIIB T4N0M0	2
T stage	
1	17
2	68
3	12
4	3
Nodal status	
Negative	60
Positive	40
Data are presented as mean (interquartile range) or %, unless otherwise stated.	
M: male; F: female.	

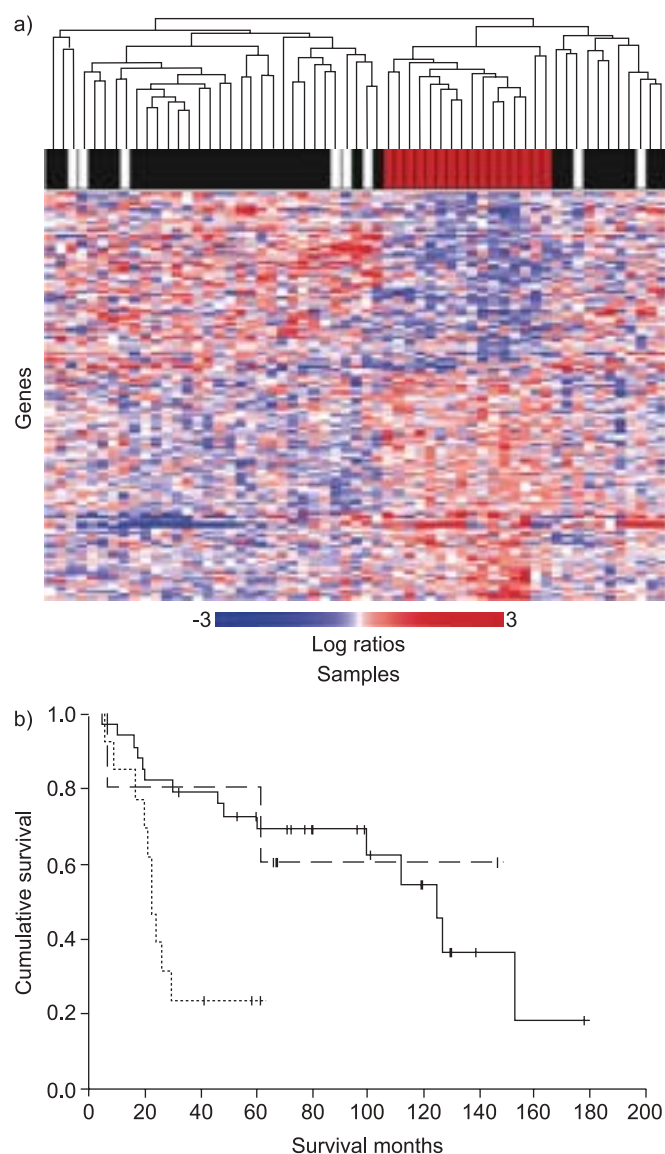


FIGURE 1. Clinical and genomic clustering of N1 tumours spread by direct extension (N1d) with N0 tumours. a) Hierarchical clustering of 59 squamous cell carcinomas (SCCs) based upon expression of 125 genes significantly differentially expressed between N0 tumours and N1/N2 tumours spread by lymphatic metastases (N1/2m) found that N1d tumours clustered within N0 rather than N1/2m samples. Each column represents a sample and each row a gene. Samples are coloured according to nodal stage. Black: N0; red: N1/N2m; white: N1d. The heat map indicates level of gene expression. Red: high expression; Blue: low expression. b) Survival estimates using Kaplan–Meier analysis of 59 SCCs. Tick marks indicate patients whose data were censored at last follow-up. —: N0; ---: N1d;: N1/N2m. $p=0.0031$.

models were built from the training set of 35 N0 and 16 N1/2m samples. Each model selected genes with significantly different ($p<0.01$) expression between the N0 and N1/2m classes, and the prediction error of each model was estimated by LOOCV. Each of the five models was then used to predict the classification (N0 or N1/2m) of the eight N1d tumours. With few exceptions, these models consistently classified N1d tumours as most similar to N0 (table 2).

To ensure that the selected cases were representative in terms of clinical outcomes, Kaplan–Meier survival curves were plotted for the SCC subsets. As expected, log-rank tests showed a significant survival difference between N0 and N1/2m cases ($p=0.0031$) which remained significant following exclusion of N2 samples ($p=0.030$; data not shown). The N1d cases appeared to have an intermediate outcome, although small sample size precluded formal statistical testing (fig. 1b).

DISCUSSION

Expression microarrays, a global measure of cellular function, were used to test the hypothesis that tumours that directly extend to hilar nodes are biologically different from those that involve locoregional hilar/mediastinal nodes by lymphatic metastasis. Using this genomic approach, the present study demonstrates, for the first time to the current authors' knowledge, that lung SCCs which metastasise to lymph nodes are clearly distinct at the mRNA transcript level from those involving N1 nodes by direct extension. Tumours which involve nodes directly by virtue of their proximity, have an expression profile similar to that of N0 tumours.

The main objective of the present study was not to develop a gene expression signature with the potential to become a clinical tool capable of classifying primary tumours with lymphatic metastases as the result of either direct or lymphatic invasion, but to determine whether tumours with N1 direct extension are indeed biologically similar to tumours with N1 lymphatic metastases or, alternatively, more similar to tumours with no lymphatic metastases (N0). While a larger sample size would give more confidence in the finding that N1 tumours by direct extension share more similarity with N0 tumours, it is more important that this is now confirmed in completely independent cohorts. The present results offer a possible explanation for the clinical observation of superior survival of NSCLC patients staged as N1 with nodal involvement by direct extension [13, 14].

A potential limitation of the present study was the fact that involved N2 tumours were grouped with N1 tumours, resulting in a less homogeneous group with lymphatic metastasis. However, the current authors believe this to be justified by the well-recognised occurrence of "skip" metastases [20]. Furthermore, the impact of any bias is unlikely to be large in view of the fact that there were only two N2 tumours in the N1/2m group.

Recent large, randomised clinical trials have shown a benefit for adjuvant chemotherapy after resection of NSCLCs but the benefits appear heterogeneous and stage-dependent. In particular, meta-analyses found that stage II (T1N1M0, T2N1M0, T3N0M0) and III (T3N1M0, T2N2M0, T1N2M0, T3N2M0) NSCLCs appear to benefit more than earlier stages (T1N0M0, T2N0M0) [5]. Therefore, as a group, patients with N1 or N2 nodal involvement appear to benefit from adjuvant chemotherapy, but a significant number of individual patients within these groups derive no benefit while being exposed to measurable risk.

Observational data suggest that direct tumour extension to N1 nodes portends a better prognosis than metastatic involvement of hilar nodes [13, 14], yet both are pathologically classified as N1 [12]. The present study shows the first evidence that the

TABLE 2 Class prediction of N1 squamous cell carcinoma samples by direct invasion (N1d) based upon gene expression of N0 tumours and N1/N2 tumours by lymphatic metastasis (N1/2m)

N1d tumour sample	Compound covariate predictor	Linear discrimination analysis	1-Nearest neighbour	Nearest centroid	Support vector machine
1	N0	N0	N0	N1/2m	N0
2	N0	N0	N0	N0	N0
3	N0	N0	N0	N1/2m	N0
4	N0	N0	N0	N0	N0
5	N0	N0	N0	N0	N0
6	N0	N0	N0	N0	N0
7	N0	N0	N0	N0	N0
8	N0	N0	N0	N0	N1/2m

N1d: n=8; N0: n=35; N1/2m: n=16.

molecular biology of squamous cell carcinoma lung cancer involving locoregional lymph nodes by direct extension differs from that of squamous cell carcinoma involving nodes by lymphatic metastases. These findings have implications for implementation of adjuvant chemotherapy aimed at improving outcome in stage II and III resected non-small cell lung cancer. The present authors are corresponding with the investigators of these adjuvant studies to ascertain the feasibility of testing whether patients with N1 tumours by direct extension benefit to the same extent as patients with metastatic N1 disease. If not, refinement of tumour, node, metastasis-node staging may be needed to adjust for the influence of tumour biology on prognosis and to allow optimal selection of patients who will truly benefit from adjuvant systemic therapy.

ACKNOWLEDGEMENTS

The authors wish to thank M. Windsor, K. Matar, P. Zimmerman, E. Duhig and D. Godbolt at The Prince Charles Hospital (Brisbane, Australia).

REFERENCES

- Arriagada R, Bergman B, Dunant A, Le Chevalier T, Pignon JP, Vansteenkiste J. Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. *N Engl J Med* 2004; 350: 351–360.
- Strauss GM, Herndon J, Maddaus MA, *et al.* Randomized clinical trial of adjuvant chemotherapy with paclitaxel and carboplatin following resection in stage IB non-small cell lung cancer (NSCLC): report of Cancer and Leukemia Group B (CALGB) Protocol 9633. *J Clin Oncol (Meeting Abstracts)* 2004; 22: 7019.
- Winton T, Livingston R, Johnson D, *et al.* Vinorelbine plus cisplatin *vs.* observation in resected non-small-cell lung cancer. *N Engl J Med* 2005; 352: 2589–2597.
- Douillard JY, Rosell R, De Lena M, *et al.* Adjuvant vinorelbine plus cisplatin *versus* observation in patients with completely resected stage IB–IIIA non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): a randomised controlled trial. *Lancet Oncol* 2006; 7: 719–727.
- Pignon JP, Tribodet H, Scagliotti GV, *et al.* Lung Adjuvant Cisplatin Evaluation (LACE): a pooled analysis of five randomized clinical trials including 4,584 patients. *J Clin Oncol (Meeting Abstracts)* 2006; 24: 7008.
- Bhattacharjee A, Richards WG, Staunton J, *et al.* Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci USA* 2001; 98: 13790–13795.
- Garber ME, Troyanskaya OG, Schluens K, *et al.* Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci USA* 2001; 98: 13784–13789.
- Beer DG, Kardia SL, Huang CC, *et al.* Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002; 8: 816–824.
- Larsen JE, Pavey SJ, Passmore LH, *et al.* Expression profiling defines a recurrence signature in lung squamous cell carcinoma. *Carcinogenesis* 2007; 28: 760–766.
- Potti A, Mukherjee S, Petersen R, *et al.* A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *N Engl J Med* 2006; 355: 570–580.
- Raponi M, Zhang Y, Yu J, *et al.* Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. *Cancer Res* 2006; 66: 7466–7472.
- Mountain CF. Revisions in the international system for staging lung cancer. *Chest* 1997; 111: 1710–1717.
- van Velzen E, Snijder RJ, Brutel de la Riviere A, Elbers HJ, van den Bosch JM. Type of lymph node involvement influences survival rates in T1N1M0 non-small cell lung carcinoma. Lymph node involvement by direct extension compared with lobar and hilar node metastases. *Chest* 1996; 110: 1469–1473.
- Marra A, Hillejan L, Zaboura G, Fujimoto T, Greschuchna D, Stamatis G. Pathologic N1 non-small cell lung cancer: correlation between pattern of lymphatic spread and prognosis. *J Thorac Cardiovasc Surg* 2003; 125: 543–553.
- Mountain CF, Dresler CM. Regional lymph node classification for lung cancer staging. *Chest* 1997; 111: 1718–1723.
- Ben-Dor A, Bruhn L, Friedman N, Nachman I, Schummer M, Yakhini Z. Tissue classification with gene expression profiles. *J Comput Biol* 2000; 7: 559–583.

- 17 Dudoit S, Fridlyand J, Speed TP. Comparison of discrimination methods for the classification of tumors using gene expression data. *J Am Stat Assoc* 2002; 97: 77–87.
- 18 Simon R. Diagnostic and prognostic prediction using gene expression profiles in high-dimensional microarray data. *Br J Cancer* 2003; 89: 1599–1604.
- 19 Lee JW, Lee JB, Park M, Song SH. An extensive comparison of recent classification tools applied to microarray data. *Comput Stat Data Anal* 2005; 48: 869–885.
- 20 Cerfolio RJ, Bryant AS. Distribution and likelihood of lymph node metastasis based on the lobar location of nonsmall-cell lung cancer. *Ann Thorac Surg* 2006; 81: 1969–1973.