The Role of Whole Exome Sequencing in Distinguishing Primary and Secondary Lung Cancers

Natalie I Vokes, Jianjun Zhang

1Department of Thoracic and Head and Neck Oncology, University of Texas MD Anderson Cancer Center, Houston, TX, USA; 2Department of Genomic Medicine, University of Texas MD Anderson Cancer Center, Houston, TX, USA

Abstract: Non-small cell lung cancer (NSCLC) that presents with multiple lung tumors (MLTs) poses a challenge to accurate staging and prognosis. MLTs that arise as clonally related secondary metastases from a common primary are higher stage and often require adjuvant chemotherapy or may in fact be incurable stage IV lesions. Conversely, MLTs that represent distinct primaries have a better prognosis and may be overtreated if inappropriately classified as related secondaries. Historically, pathologic and radiographic criteria were used to distinguish between primary and secondary MLTs; however, the advent of genomic profiling has demonstrated limitations to these historic classification systems. In this review, we discuss the use of molecular profiling to distinguish between primary and secondary lung cancers, with a focus on the insights gleaned from whole exome sequencing (WES) analyses. While WES is not yet feasible in routine clinical practice, WES studies have helped elucidate the clonal relationship between primary and secondary lung cancers and provide important context for the application of targeted sequencing panel-based analyses.

Keywords: genomics, non-small cell lung cancer, metastasis

Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related deaths both globally and in the United States. In 2021, there were an estimated 235,760 new cases and 131,880 deaths from NSCLC in the U.S., accounting for 25% of all cancer deaths. Mortality is highly related to stage, with historic five-year survival rates of 10% or less for stage IV disease, compared to 70% or higher for stage I disease. The management of NSCLC is guided by initial staging; whereas stage IV NSCLC is treated palliatively with systemic therapy, early-stage NSCLCs are treated with definitive local therapy. Adjuvant systemic therapy after surgery plays an important role in improving disease-free and overall survival in patients with stage II or III tumors, and the roster of approved adjuvant therapies has recently expanded from platinum-based chemotherapies to include adjuvant Osimertinib in EGFR mutated patients and atezolizumab in patients with PD-L1 > 1%. While adjuvant systemic therapy is considered the standard of care for resected stage II and III NSCLC, systemic therapy does not play a role for earlier stage NSCLC due to historic evidence demonstrating no benefit and in fact a trend toward harm in stage IA patients treated with chemotherapy. Accurately determining cancer staging is therefore crucial for the appropriate management of NSCLC.
One challenge to the accurate staging and management of NSCLC is the occurrence of multiple lung tumors (MLTs) in the same patient. MLTs may arise from local or hematogenous spread of a single primary cancer, or as multiple individual primary cancers. While historically rare, the incidence of MLTs has been rising due to increased use of chest CT scans and longer overall survival in lung cancer patients. MLTs may present at the same time (termed synchronous) or may develop sequentially, with new tumors occurring after prior definitive therapy (termed metachronous). While the exact rates vary, studies estimate that 15–20% of lung cancer patients present with multiple nodules, though the number can be as high as 50%, and the risk of developing a second primary lung cancer is 1–2% per year.

Determining whether a secondary nodule represents a distinct primary or an intrapulmonary metastasis (IPM) therefore has important staging and prognostic implications. The 8th American Joint Committee on Cancer (AJCC) staging edition classifies synchronous nodules related to the primary as T3 (at least stage II) or T4 (stage IIIA) if the tumor is in the same or a different ipsilateral lobe, respectively, and stage M1a (Stage IVa) if the nodule is in a contralateral lobe. Conversely, the AJCC recommends staging multiple primaries individually and managing with definitive therapy as permitted by anatomy and comorbidities. Consistent with these staging guidelines, the prognostic implications of multiple lung nodules depend on whether they are determined to be related or distinct entities. The presence of a histologically presumed intrapulmonary metastasis carries a 10–15% decrease in 5-year survival, whereas multiple primaries are not necessarily associated with worse outcomes. Similarly, as discussed, the appropriate classification of MLTs as distinct primaries or IPMs carries important management implications. Misclassification of distinct early-stage lung primaries as T3 or T4 IPMs can lead to exposure to the toxicities of adjuvant platinum-based chemotherapy with no clinical benefit and potential harm, upstaging contralateral nodules as IPMs can also prevent patients from receiving appropriate definitive local therapies.

Historically, pathologic criteria have been used to distinguish between distinct primaries vs IPMs. However, recent studies with molecular data have demonstrated significant limitations in these criteria, and there is increasing interest in using advanced sequencing to characterize MLTs as distinct primaries vs evolutionarily related secondary malignancies. In this review, we discuss the use of sequencing to distinguish between primary and secondary lung cancers and its implications for clinical practice.

**Pathologic Criteria for Distinguishing Between Primary and Secondary Malignancies**

Until recently, clinicopathologic criteria were the standard in distinguishing between primary and secondary lung cancers. Martini and Melamed published the most widely applied diagnostic standard in 1975, which specified that synchronous tumors were considered distinct primaries if they were physically separated and had different histology; or, if the same histology, originated from carcinoma in situ; had no lymphatic spread; and occurred without extrapulmonary metastases at the time of diagnosis. Similar rules were applied to temporally distinct tumors (metachronous tumors), with the stipulation that historically identical metachronous tumors had to be separated by two years.

Over the years, several revisions to these criteria have been proposed, including the incorporation of DNA ploidy and the extension of the disease-free interval to 4 years. The 2007 edition of the American College of Chest Physicians guidelines attempted to clarify the definitions by classifying synchronous lesions into 1) satellite nodules (same lobe); 2) multiple primaries (different lobes, no N2-N3 lymph nodes or metastases), or 3) hematogenously spread pulmonary metastases (lesion present in different lobes with N2-N3 nodal involvement). However, studies incorporating molecular features including chromosomal abnormalities and mutation data demonstrated that pathologic criteria have limited sensitivity and specificity. The most rigorous histomorphological criteria, put forth in 2009 by Girard et al, were shown to have higher accuracy (91% concordance with molecular assessment compared to 32% using Martini and Melamed criteria), but require skilled pathologic assessment of features including grade, cytology, stromal features, and variant histologic subtypes, making it subject to inter-user variability.

**Limitations of Pathologic Approaches and Incorporation of Molecular Analyses**

The advent of sequencing methodologies led to interest in studying whether molecular assessment of MLTs could
prove more accurate in distinguishing relatedness than pathologic criteria. Even before the advent of high-throughput next-generation sequencing (NGS) of multiple genes, researchers demonstrated that more limited profiling could determine whether spatially and/or temporally distinct cancers shared genetic similarities. One early strategy focused on sequencing TP53, as this gene is frequently mutated in lung cancers, and mutations occur over a large genomic region, making it unlikely that two independent lung cancers will have identical mutations.29 These early studies demonstrated that, in patients with a TP53 mutation in at least one sample, shared or distinct TP53 mutations could identify intrapulmonary metastases or multiple primaries, respectively.29–35 Other orthogonal strategies, including loss of heterozygosity (LOH) analysis,36 LOH integrated with driver gene sequencing;34 oligo gene panel sequencing;37–41 array comparative genomic hybridization studies;10,42 and DNA rearrangement sequencing,43 all similarly demonstrated the utility of molecular profiling in establishing genetic relatedness. In one study, for example, Girard et al performed comparative genomic hybridization, mutational profiling of select genes, and clinicopathologic review on 42 tumors from 20 patients (leading to 22 evaluable tumor-pair combinations).10 While the majority of cases were concordant, classification based on genomic profiling contradicted clinicopathologic diagnosis in 4 (18%) comparisons (independent primaries in 1 case diagnosed as metastasis, metastases in 3 cases diagnosed as independent primaries).

Notably, one limitation of focused sequencing is that it is only helpful in patients in whom at least one abnormality is detected. TP53 sequencing, for example, was helpful in the 40–50% of patients with a TP53 alteration, but added no value in those patients whose tumors were all TP53 wild type.29,33,35 Integration of multiple technologies improved the likelihood that an important difference would be detected,10 but the advent of next-generation sequencing technologies proved even more useful by allowing for the incorporation of more comprehensive sequencing to profile a broader array of genes. In one early study, Murphy et al analyzed DNA rearrangements by NGS in 2014 and demonstrated that paired lung primaries and distant metastases had shared rearrangements identified in all tumor pairs.43 When applied to synchronous pairs, they confirmed that four cases predicted by pathology to be independent shared no somatic rearrangements; concluded that two pathologically indeterminate tumors were not related; and changed the designation in 1 out of 5 putative metastatic cases. Another more recent study using the large MSK-IMPACT panel compared tumors from 60 patients who had undergone sequencing of multiple tumors.44 In 76 tumor pairs, NGS classified 51 into definite distinct primaries and 25 intrapulmonary metastases. Prospective histologic prediction was discordant with NGS in 17 cases (22%); pathologic error in identifying IPMs was most commonly due to histologic progression (7/11 cases), whereas distinct primaries were missed pathologically due to overlap in architectural patterns or cytologic architecture. Other studies using different gene panels or clinical cohorts demonstrated similar utility to this approach,45–51 and confirmed worse outcomes in patients with genomically similar tumors.21 These analyses are summarized in Table 1.

### Whole Exome Sequencing Studies

These studies demonstrated that molecular analyses can be used to determine whether pathologically similar lesions were related or genetically distinct. While not all analyses directly compared the performance of molecular profiling to pathologic classification, those that did demonstrated variable accuracy, with molecular characterization overturning the pathologic determination in 20–40% of cases.10,43,52–54 One challenge, however, in interpreting molecular studies is determining whether tumors may be different or similar genetically for reasons other than relatedness. Specifically, while we would expect most regions of the same tumor to be genomically similar, there may be differences due to subclonal evolution and intratumoral heterogeneity. Conversely, it is also known that certain driver alterations or copy number events are highly recurrent and may be present in two independent tumors by chance or due to germline/environmental predisposition. Whole exome sequencing (WES)-based studies have helped address this knowledge gap by more comprehensively elucidating the clonal relationship between tumors, helping clarify the evolutionary processes that give rise to metastases vs distinct primaries.

The first relevant application of WES in this context were multi-region WES studies that helped delineate the scope of intratumoral heterogeneity within related NSCLC tumors. In one study from our group, Zhang et al performed multi-region WES on 48 regions from 11 resected lung adenocarcinomas. They detected significant intratumoral heterogeneity, with an average shared mutation rate of 76% between all regions in one patient. Putative driver alterations were often clonal and ubiquitous.55 Similarly,
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Set</th>
<th>Sequencing Analysis</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noguchi et al, 1993</td>
<td>30 Nine patients with MLTs</td>
<td>TP53 protein expression</td>
<td>Nine tumors from 4 cases had altered TP53 and were diagnosed genetically as distinct primaries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TP53 mutations (PCR-SSCP)</td>
<td></td>
</tr>
<tr>
<td>Sozzi et al, 1995</td>
<td>11 synchronous lesions from 5 patients</td>
<td>Cytogenetics/LOH TP53 and KRAS sequencing Chromosome 3p deletion</td>
<td>Different genetic lesions observed in all cases regardless of physical proximity, suggestive of multiple primaries.</td>
</tr>
<tr>
<td>Mitsudomi et al, 1997</td>
<td>16 patients with a second, metachronous lung tumors</td>
<td>TP53 sequencing</td>
<td>9/16 patients had at least one TP53 mutation. TP53 mutational status was discordant in all 9 cases, consistent with distinct tumors</td>
</tr>
<tr>
<td>Hiroshima et al, 1998</td>
<td>19 patients with synchronous MLTs and 11 with metachronous MLTs</td>
<td>TP53 protein overexpression using immunohistochemistry TP53 gene abnormalities</td>
<td>11/19 patients had altered TP53 activity. Different genetic changes were detected in 11/19 synchronous (57.9%) and 5/11 metachronous (45.4%) MLTs</td>
</tr>
<tr>
<td>Matsuzoe et al, 1999</td>
<td>20 patients with MLTs</td>
<td>TP53 mutational analysis</td>
<td>Lesions in 7 patients showed TP53 gene abnormalities. One patient had the same mutational status in both tumors.</td>
</tr>
<tr>
<td>Shimizu et al, 2000</td>
<td>14 patients with distinct MLTs</td>
<td>LOH TP53 mutational status</td>
<td>Genetics were concordant with clinical criteria in 10 of 12 cases (83%), discordance in 2 (17%)</td>
</tr>
<tr>
<td>van Rens et al, 2002</td>
<td>64 tumors from 31 patients</td>
<td>TP53 mutational analysis</td>
<td>21 patients had different TP53 mutations, distinct primaries 2 patients had identical TP53 mutations, indicating IPMs. 8 patients were indeterminate.</td>
</tr>
<tr>
<td>Chang et al, 2007</td>
<td>58 patients with synchronous MLTs</td>
<td>TP53 and EGFR expression and mutational analysis</td>
<td>50/58 tumors were distinguishable based on TP53 and EGFR (22 cases related; 28 were distinct).</td>
</tr>
<tr>
<td>Takamochi et al, 2012</td>
<td>82 multifocal lung adenocarcinomas from 36 patients, 31 synchronous and 5 metachronous.</td>
<td>EGFR and KRAS sequencing</td>
<td>EGFR mutations were identified in 36 tumors (44%), KRAS mutations in 19 (23%) Concordance of molecular studies with clinical criteria was observed in 21 (70%) of the 30 cases in which tumor clonality was determinable</td>
</tr>
<tr>
<td>Wu et al, 2014</td>
<td>97 metachronous MLTs</td>
<td>TP53 and EGFR sequencing</td>
<td>75/97 tumors could be distinguished based on the presence of TP53 and/or EGFR. 25/97 (33.3%) of cases had the same clonality, 50 (67%) had different clonality. Distinct clonality occurred even in tumors of the same histologic type (37/61).</td>
</tr>
<tr>
<td>Murphy et al, 2014</td>
<td>22 tumors from 11 patients</td>
<td>Mate-pair NGS sequencing to identify break points</td>
<td>Validated that histologically distinct tumors were genomically distinct without shared genomic rearrangements. Histologically similar tumors had shared rearrangements. 7/11 pairs classified as multiple primaries, 4/11 as related. Largely concordant with histologic classification.</td>
</tr>
</tbody>
</table>

(Continued)
Table I (Continued).

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Set</th>
<th>Sequencing Analysis</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patel et al, 2017</td>
<td>8 primary-metastasis pairs and 11 patients with distinct primaries</td>
<td>NGS using 50-gene AmpliSeq Cancer Hotspot Panel</td>
<td>7/8 primary-metastasis pairs had shared mutations; the last pair had no detectable mutations in either sample. Driver mutations (KRAS, EGFR, and BRAF) were always concordant. 3/11 MLTs were concordant and predicted to be IPMs, 8/11 were discordant. Patients with predicted independent primaries had better survival (p=0.016).</td>
</tr>
<tr>
<td>Saab et al, 2017</td>
<td>52 lung adenocarcinomas from 18 patients.</td>
<td>NGS and ALK gene rearrangement testing</td>
<td>Genomic signatures were helpful in 72% of patients. Genomic and morphologic data together achieved a conclusive diagnosis in 94% of patients.</td>
</tr>
<tr>
<td>Roepman et al, 2018</td>
<td>111 tumors from 50 patients with MLT</td>
<td>TP53 mutation analysis, 50-gene panel NGS</td>
<td>Sequencing was conclusive in all but 2 pairs, and contradicted histopathologic assessment in 19 (39%) of cases.</td>
</tr>
<tr>
<td>Takahashi et al, 2018</td>
<td>82 tumors from 37 patients with surgically resected MLTs</td>
<td>Targeted 20-gene panel NGS</td>
<td>20 (54%) of patients had matching mutations c/w IPMs. In 7 equivocal histopathologic cases, 6 (86%) confirmed as IPMs. Among 17 cases classified as multiple primaries by histology, mutational analysis was discordant in 8 (47%) cases.</td>
</tr>
<tr>
<td>Mansuet-Lupo et al,</td>
<td>240 samples from 120 patients with multiple tumors</td>
<td>Targeted NGS sequencing (22 genes)</td>
<td>109/120 (91%) of pairs had an identifiable mutation. Discordant cases (30/109) were reclassified using combined histomolecular algorithm.</td>
</tr>
<tr>
<td>Murphy et al, 2019</td>
<td>Tumor pairs from 17 patients</td>
<td>Targeted NGS sequencing (8 gene)</td>
<td>53% of tumors had no detected mutations; 4 cases had mutations in one tumor and one had two different KRAS mutations. Two cases shared KRAS mutations despite independence on junction analysis.</td>
</tr>
<tr>
<td>Change et al, 2019</td>
<td>76 tumor pairs from 60 patients</td>
<td>Targets NGS sequencing (MSK-IMPACT panel)</td>
<td>51/76 pairs classified as distinct primaries, 25/76 into IPMs. 11/76 cases were misclassified pathologically</td>
</tr>
<tr>
<td>Liu et al, 2020</td>
<td>40 tumors from 16 pts with multiple lung tumors</td>
<td>Targeted deep sequencing (2000x) using HaploX (464-gene) panel</td>
<td>12/14 patients with histological IPMs did not have shared mutations. Two patients with inferred metastases had shared mutations; two patients had shared driver alterations.</td>
</tr>
<tr>
<td>Goodwin et al, 2021</td>
<td>40 cases of pts treated curatively for two putative primaries with similar histopathology.</td>
<td>Targeted NGS panel (Ion AmpliSeq Colon and Lung Cancer Panel V2 and RNA-fusion panel).</td>
<td>Mutational profiling was concordant with clinicopathologic diagnosis in 33/40 cases (82.5%). 7 cases with mutational overlap suggestive of metastatic disease had reduced overall survival.</td>
</tr>
</tbody>
</table>

Abbreviations: IPM, intrapulmonary metastases; LOH, loss of heterozygosity; MLT, multiple lung tumors; NGS, next-generation sequencing; PCR-SSCP, polymerase chain reaction-single-strand conformation polymorphism.
de Bruin et al performed multi-region WES on 25 regions from 7 NSCLC patients and found that an average of 70% of mutations were shared across regions, with similar enrichment for driver events in the truncal mutations. Subsequent analysis in a larger cohort confirmed that on average 70% of mutations are shared between related lesions, though there was a range of 7–99.5% overlap, demonstrating high variability. These studies corroborated previous findings from matched primary and metastatic tumors in 15 patients assessed by a 182 targeted NGS panel; in this study, likely driver alterations had 94% concordance, and likely passenger alterations had 63% concordance. Notably, de Bruin et al included two patients with synchronous contralateral lung lesions. In one pair, only an EGFR L858R mutation was shared, suggestive of separate primaries with an overlapping driver. In the other pair, 74% of mutations were shared, indicative of intrapulmonary metastasis.

The genomics of multiple primaries was subsequently interrogated more directly by a study by Futreal and colleagues. In this study, researchers performed whole genome sequencing (WGS) or WES with microarray-based comparative genomic hybridization on 16 tumors (15 lung tumors and one lymph node) from a cohort of six Asian patients with multiple synchronous lung cancers. By ACCP 2007 criteria, 5 patients were classified as having satellite nodules, and one as having IPM; therefore, clinically all were defined as having multiple related tumors. When tumor genomics were compared, however, all appeared distinct. In three patients, there were no overlapping mutations between tumors. In patient 1, the tumor and lymph node sample shared 26% of mutations but did not overlap with the other lung sample. Finally, of the other two patients, one had an overlapping EGFR L858R mutation in all three samples but no other shared mutations. The last patient had an overlapping EGFR L858R in tumors 1 and 2, and an overlapping ARHGAP35 E25K mutation shared by tumors 1 and 3, but no other overlap. To determine the likelihood that non-related tumors would share a point mutation, the researchers benchmarked the rate of shared mutations in The Cancer Genome Atlas (TCGA) dataset, and found that the likelihood of overlap in their cohort was no different from TCGA. This allowed the authors to conclude that all tumors in this cohort were distinct primaries, highlighting that single driver alterations (such as EGFR L858R) can recur in otherwise genetically distinct tumors in the same patient. Notably, while clinical staging was consistent with related primaries, more comprehensive pathologic assessment was more congruent with the genomics; 5 out of 6 cases were correctly classified by histologic assessment.

In addition to the small sample size, this study was limited by its focus on Asian patients with minimal smoke exposure, as these features likely selected for a patient population with distinct tumor genomics compared to Caucasian or smoking patients. These findings were therefore corroborated in a follow-up study from our group that performed WES on 10 tumors from 4 Caucasian patients with multiple synchronous lung tumors. Focusing on putatively functional mutations, one patient had no shared mutations across three tumors specimens; the other three patients’ tumors shared a range of mutations between 19% and 91%. Benchmarking against the TCGA confirmed higher rates of shared mutations in this cohort than among unrelated TCGA specimens. In contrast to the previous study, in which unrelated tumors had the same driver alteration, tumors from one patient in this cohort had distinct TP53 mutations despite other overlapping mutations. These two studies, therefore, demonstrate that focused sequencing of single driver events is unreliable due to both convergent and divergent evolution.

Mutational signature analysis in both studies confirmed that gene-level mutational differences reflect different mutational processes. In the Liu et al cohort, which had largely distinct primaries, they observed distinct mutational spectra between different tumors within the same patient. Apolipoprotein B mRNA editing enzyme (APOBEC)-mediated mutations, in particular, may be driven largely by APOBEC-mediated changes in subclonal mutations. In the second study, despite overlapping mutational profiles in three of four patients’ tumors, only one patient had similar mutational signature profiles across all specimens, driven largely by APOBEC. The other three patients, including the two patients with IPMs, had distinct mutational profiles, suggesting that different mutagenic processes can contribute to subclonal evolution in distinct metastatic sites.

A final study by Ma et al analyzed WES on 16 tumor samples (from 11 independent lesions) in four patients. Similar to Liu et al, they observed no overlapping variant sets between any tumor pair; conversely, multi-region sequencing of two larger lesions revealed many shared somatic mutations, similar to prior heterogeneity studies. In this study, most patients had similar
mutational spectra across different tumors; however, one patient had notable discordance across three tumors. Notably, while the specific mutations within each tumor were heterogeneous, the authors observed convergence in pathway dysregulation within each patient. Patient 1 and patient 4, for example, had tumors with altered EGFR signaling, while patients 2 and 3 had tumors with dysregulated MAPK pathway.

**Whole Exome Sequencing to Distinguish Secondary Lung Cancers from Other Primaries**

An additional application of WES is in distinguishing between secondary lung cancers that occur as metastases from other histologic primaries. Within lung cancers, whole exome sequencing demonstrates distinct mutational profiles between adenocarcinomas and squamous cell carcinomas. Vanni et al, for example, sequenced multiple lung primaries in a patient who developed two lung cancers (one adenocarcinoma, one squamous cell carcinoma), and a malignant peritoneal mesothelioma. They found distinct mutational profiles in each primary, and identified germline polymorphisms in DNA repair genes that may confer higher cancer susceptibility. Similarly, Bai et al performed WES of two lung nodules, one adenocarcinoma and one lung squamous cell carcinoma, and identified no overlapping alterations.

Sequencing also confirms distinct genomic profiles in lung cancer vs other cancer types. Sequencing of one patient with a synchronous esophageal squamous cell carcinoma, lung adenocarcinoma, and hepatocellular carcinoma demonstrated distinct somatic genomic profiles; the patient had two germline alterations in cancer susceptibility genes (SPINK1 and JAK3) that might have contributed to his risk. Xue et al analysed WES from 12 patients found to have squamous cell carcinomas of the esophagus and lung. Similar to WES from multiple lung cancers, the authors identified one set of tumor pairs with no overlapping mutations, consistent with distinct lung and esophageal primaries, whereas a second group shared 12–70% of mutations, consistent with likely secondary metastases. These results conflicted with pathologic assessment in 41.7% of cases (5/12 patients), all of whom were pathologically diagnosed with local metastases but genomically consistent with distinct primaries. All these patients did well after surgical lung resection, validating the genomic rather than histologic designation. Another smaller study using more limited profiling demonstrated similar utility to sequencing in this context.

**Clinical Applications of Sequencing to Distinguish Between Primary and Secondary NSCLCs**

Taken together, these studies help to systematically delineate the relationship between multiple lung tumors in individual patients. They highlight the diversity in how lung cancers behave over time and space, with highly variable rates of multiple primaries vs intrapulmonary metastases across studies; even in surgically resected cohorts, which definitionally include small tumors with some clinical probability of being distinct primaries, the rate of IPM can be as high as 33%–77%. These findings suggest that intrapulmonary metastases and synchronous primaries are both common and that clinical features alone are not sufficient to distinguish between these possibilities.

Conversely, comprehensive genomic profiling is highly useful in determining genetic relatedness, and related cancers almost always have genetic features in common. However, the percent overlap can vary dramatically; while focused analyses of primary/metastasis pairs suggest an average of 70% overlap, the WES studies cited above showed a range between 20% and 90%. Mutational signature analysis reveals that distinct primaries arise through distinct mutagenic processes, but even genetically related subclones may have different dominant signatures. Despite this heterogeneity, there appears to be convergence within and across patients into dysregulation of canonical oncogenic pathways; consequently, highly recurrent driver events such as EGFR L858R or KRAS G12C alterations can occur in otherwise unrelated tumors. Conversely, genetically related tumors may have distinct alterations in the same oncogenic drivers or tumor suppressors due to convergent subclonal evolution.

These findings highlight the potential utility of WES in identifying genetically related tumors and the pitfalls that can arise from trying to use more limited sequencing panels for this task. These studies are limited primarily by their small sample size, and by their enrichment for patients with multiple surgically resected tumors, which is not always standard clinical practice. The genomics of tumors not resected surgically due to (small or large) size or due to higher inferred clinical stage may be different and are incompletely assessed by the above studies. Fundamentally, the studies above also demonstrate that
tumor evolution, whether from a primary to a metastasis or the evolution of multiple primaries, is an idiosyncratic process that will likely vary from person to person. Larger studies can help capture and describe different evolutionary routes and can form the basis for models that weight the probability of distinct primaries vs related secondaries on the basis of more widely characterized behavior.

Additionally, while WES provides the most comprehensive means of assessing genetic relatedness, at present it is not practical for routine clinical use. While this may change as sequencing turn-around times and costs continue to decrease, the current role of these WES studies is to provide important baseline metrics to help guide more clinically feasible approaches. One such approach involves utilization of targeted sequencing panels, where the question then becomes how many genes must be sequenced to obtain high sensitivity and specificity. Single driver gene assessment or even assessment of only targetable oncogenes is the cheapest but lacking in sensitivity and specificity; as the studies above demonstrate, many tumors will not have a detectable mutation, and even those that do can have variable rates of discordance (some reports range from 0% to 86% depending on the sequencing technique\textsuperscript{36,57,60,68–75}). Conversely, otherwise genetically distinct tumors may have shared hotspot alterations; Liu et al\textsuperscript{48} and Chang et al,\textsuperscript{44} for example, both observed shared KRAS and EGFR mutations in tumors with otherwise distinct genomic profiles, occurring in 1 in out of 6 patients assessed by WES,\textsuperscript{58} and in 5 out of 51 separate primaries assessed by the large MSK-IMPACT targeted panel.\textsuperscript{44} Indeed, given the high frequency of specific hotspot mutations such as EGFR L858R or KRAS G12C, the odds of co-occurrence by chance can be as high as 1 in 17, and Chang et al found that shared KRAS mutations were almost as likely to occur coincidentally in distinct primaries as in IPMs.\textsuperscript{44}

Different studies have attempted to model the accuracy of various gene panel sizes. Mansuet-Lupo et al used a 22-gene panel and found that 9% of patients were inconclusive, compared to 28% percent who would have been inconclusive using a 5-gene panel.\textsuperscript{50} In another analysis, gene panels of 50 and 182 genes were noninformative in 28% and 14% of cases, respectively.\textsuperscript{28} Chang et al were able to make a definitive assignment in 75 out of 76 tumor pairs using the MSK-IMPACT gene panel (341 or more genes), and, by down-sampling the MSK-IMPACT panel, determined that 4-gene driver panel (EGFR, KRAS, ALK and ROS1) would distinguish between primaries vs IPMs in 60% of cases, and a 50-gene panel in 72% of cases. In general, they found that distinct primaries could be identified with smaller panels, but it would be harder to definitively confirm that tumors were clonally related with smaller compared to larger panels. Computational estimation suggested that panels would need to contain at least 100 frequently mutated genes to allow for confirmation of clonal relatedness in 95% of lung adenocarcinomas.

In contrast, Liu et al compared the accuracy of WES to more comprehensive histologic assessment,\textsuperscript{27,59,60} and found that 5 out of 6 cases could be accurately assessed histologically. However, this was a small study with samples that ended up all being distinct primaries; when Chang et al compared genomic sequencing to histologic prediction, they found that 78% of samples were accurately classified by histologic criteria, and that the rate of discordance was significantly higher for IPMs (44%) than synchronous primaries (12%).\textsuperscript{44}

The role of WES in characterizing the strengths and limitations of pathologic approaches is also an important contribution. While genetic sequencing provides a more direct assessment of clonal relatedness, one practical limitation to sequencing-based approaches is tissue availability. Many of the studies described above are based on surgically resected specimens, which typically have abundant tissue for both histologic and sequencing analyses. In routine clinical practice, however, many synchronous nodules are too small for biopsy, and even larger nodules may produce tissue yields that are inadequate for sequencing.\textsuperscript{76–78} Historic studies have shown inadequate tissue for sequencing in 10–20% of samples,\textsuperscript{79} and the analysis of cytology specimens from fine needle aspirates or pleural fluid studies was found to be particularly problematic (inadequate specimens in 26% of biopsy specimens and 35% of cytology specimens compared to 5% of surgical specimens).\textsuperscript{80} In such cases, histologic assessment may be more clinically feasible, and using genomics as above to benchmark and optimize pathology guidelines may help improve the accuracy of histology-based approaches.

Ultimately, pathologic assessment and genomic sequencing may prove complementary strategies, as pathologic assessment is performed routinely and at present is more cost-effective. However, as biomarker-directed therapies are increasingly becoming the standard-of-care even in early-stage NSCLC,\textsuperscript{7,8} genomic assessment of MLTs will also become more common. At present, the
IASLC guidelines include both clinical, histologic, and biomarker patterns in their criteria for identifying distinct primaries, further studies in larger cohorts using standardized sequencing approaches will help further define how best to systematically incorporate clinical genomics into lung cancer staging.

Disclosure
Dr Natalie I Vokes reports personal fees from Sanoﬁ and Oncocyte, outside the submitted work. Dr. Jianjun Zhang reports personal fees from AstraZeneca, Novartis, Johnson and Johnson, Geneplus; speaker’s fees from BMS, OrigMed, Innoven; and grants from Merck, Novartis, Johnson and Johnson outside the submitted work. The authors report no other conﬂicts of interest in this work.

References


