
This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of October 9, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/346/6206/256.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2014/10/08/346.6206.256.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/346/6206/256.full.html#related>

This article **cites 43 articles**, 15 of which can be accessed free:

<http://www.sciencemag.org/content/346/6206/256.full.html#ref-list-1>

This article has been **cited by** 1 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/346/6206/256.full.html#related-urls>

This article appears in the following **subject collections**:

Medicine, Diseases

<http://www.sciencemag.org/cgi/collection/medicine>

a single biopsy from any other region would suggest treatment with a BRAF inhibitor, for which the tumor cells from R3 might be resistant because of the *PIK3CA* mutation (32).

Our study also sheds light on the divergent genomic instability processes involved in NSCLC evolution and their dynamics over time. Evidence for spatial diversity in genomic instability processes suggests that opportunities to exploit such mechanisms therapeutically may be limited in this disease (33). In three tumors, we detected genome-doubling events occurring before subclonal diversification but after acquisition of driver mutations, consistent with findings in colorectal cancer that genome doubling may accelerate cancer genome evolution (34). The relation of chromosomal instability with drug resistance and early tumor recurrence (35, 36) suggests that targeting truncal driver events may be compromised by the initiation of chromosomal instability later in tumor evolution. These results, coupled with the observation that NSCLC tumors may have prolonged latency periods, support continued efforts to optimize methods for earlier detection.

Unexpectedly, we found that despite continuous exposure to the mutagens in tobacco smoke, tumors from smokers showed evidence that an additional genomic instability process (APOBEC-associated mutagenesis) likely contributes to tumor progression. A large proportion of subclonal driver mutations were found to occur in an APOBEC context, which suggests that the differences in mutation spectra over time and space may reflect the activity of the process generating the mutations, as well as the selective advantage of the acquired mutations.

The presence of subclonal, regionally separated driver events, coupled with the relentless and dynamic nature of genomic instability processes observed in this study, highlight the therapeutic challenges associated with NSCLC. Engaging an adaptable immune system may present a tractable approach to manage the dynamic complexity in NSCLC (37). Longitudinal studies will be required to decipher drivers of subclonal expansion, identify the origins of subclones contributing to metastatic recurrence, and resolve the evolutionary principles that underpin the dismal outcome associated with this disease.

REFERENCES AND NOTES

- World Health Organization, www.who.int/cancer/en/ (2013).
- Siegel, D. Naishadham, A. Jemal, *CA Cancer J. Clin.* **63**, 11–30 (2013).
- H. Tanaka *et al.*, *Cancer Res.* **67**, 6007–6011 (2007).
- B. A. Weir *et al.*, *Nature* **450**, 893–898 (2007).
- L. Ding *et al.*, *Nature* **455**, 1069–1075 (2008).
- Z. Kan *et al.*, *Nature* **466**, 869–873 (2010).
- S. L. Carter *et al.*, *Nat. Biotechnol.* **30**, 413–421 (2012).
- T. I. Zack *et al.*, *Nat. Genet.* **45**, 1134–1140 (2013).
- R. Govindan *et al.*, *Cell* **150**, 1121–1134 (2012).
- E. D. Pleasance *et al.*, *Nature* **463**, 184–190 (2010).
- W. Lee *et al.*, *Nature* **465**, 473–477 (2010).
- G. P. Pfeifer, P. Hainaut, *Mutat. Res.* **526**, 39–43 (2003).
- S. A. Roberts *et al.*, *Nat. Genet.* **45**, 970–976 (2013).
- M. B. Burns, N. A. Temiz, R. S. Harris, *Nat. Genet.* **45**, 977–983 (2013).
- L. B. Alexandrov *et al.*, *Nature* **500**, 415–421 (2013).
- Detailed information on methods is available on Science online.
- S. V. Sharma, D. W. Bell, J. Settleman, D. A. Haber, *Nat. Rev. Cancer* **7**, 169–181 (2007).
- N. Bolli *et al.*, *Nat. Commun.* **5**, 2997 (2014).
- M. Imielinski *et al.*, *Cell* **150**, 1107–1120 (2012).
- M. S. Lawrence *et al.*, *Nature* **505**, 495–501 (2014).
- P. Van Loo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16910–16915 (2010).
- S. Nik-Zainal *et al.*, *Cell* **149**, 994–1007 (2012).
- A. Malhotra *et al.*, *Genome Res.* **23**, 762–776 (2013).
- S. F. Bunting, A. Nussenzweig, *Nat. Rev. Cancer* **13**, 443–454 (2013).
- C. Z. Zhang, M. L. Leibowitz, D. Pellman, *Genes Dev.* **27**, 2513–2530 (2013).
- S. C. Baca *et al.*, *Cell* **153**, 666–677 (2013).
- M. Gerlinger *et al.*, *Nat. Genet.* **46**, 225–233 (2014).
- M. Gerlinger *et al.*, *N. Engl. J. Med.* **366**, 883–892 (2012).
- M. Reck, D. F. Heigener, T. Mok, J. C. Soria, K. F. Rabe, *Lancet* **382**, 709–719 (2013).
- H. Davies *et al.*, *Nature* **417**, 949–954 (2002).
- S. Kang, A. G. Bader, P. K. Vogt, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 802–807 (2005).
- H. Shi *et al.*, *Cancer Discov.* **4**, 80–93 (2014).
- R. A. Burrell, N. McGranahan, J. Bartek, C. Swanton, *Nature* **501**, 338–345 (2013).
- S. M. Dewhurst *et al.*, *Cancer Discov.* **4**, 175–185 (2014).
- R. Sotillo, J. M. Schwartzman, N. D. Socci, R. Benezra, *Nature* **464**, 436–440 (2010).
- A. J. Lee *et al.*, *Cancer Res.* **71**, 1858–1870 (2011).
- J. C. Soria *et al.*, *Eur. J. Cancer* **49** (suppl. 2), 3408 (2013).

ACKNOWLEDGMENTS

E.B. is a Rosetrees Trust fellow; M.J.H. has a Cancer Research UK fellowship; N.Mu. received funding from the Rosetrees Trust; M.G. is funded by the UK Medical Research Council; I.V. is funded by Spanish Ministerio de Economía y Competitividad subprograma Ramón y Cajal; R.C.R. and D.M.R. are partly funded

by the Cambridge Biomedical Research Centre and Cancer Research UK Cancer Centre; P.V.L. is a postdoctoral researcher of the Research Foundation—Flanders (FWO); S.M.J. is a Wellcome Senior Fellow in Clinical Science; and C.S. is a senior Cancer Research UK clinical research fellow and is funded by Cancer Research UK, the Rosetrees Trust, European Union Framework Programme 7 (projects PREDICT and RESPONSIFY, ID:259303), the Prostate Cancer Foundation, the European Research Council and the Breast Cancer Research Foundation. This research is supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. We would like to thank Servicio Santander Supercomputación for their support. P.C. is a paid consultant for and holds equity in 14M Genomics Ltd. The results published here are in part based upon data generated by the Cancer Genome Atlas pilot project established by the National Cancer Institute and National Human Genome Research Institute, NIH. Information about TCGA and the investigators and institutions who constitute the TCGA research network can be found at <http://cancergenome.nih.gov/>. The data were retrieved through dbGaP authorization (accession no. phs000178.v5.p5). Sequence data have been deposited at the European Genome-Phenome Archive (EGA, www.ebi.ac.uk/ega/), which is hosted by the European Bioinformatics Institute (EBI), under accession numbers EGAS00001000809, EGAS00001000888 and EGAS00001000889.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6206/251/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S4
References (38–64)

17 March 2014; accepted 4 July 2014
10.1126/science.1253462

LUNG CANCER EVOLUTION

Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing

Jianjun Zhang,^{1,2} Junya Fujimoto,³ Jianhua Zhang,⁴ David C. Wedge,⁵ Xingzhi Song,⁴ Jiexin Zhang,⁶ Sahil Seth,⁴ Chi-Wan Chow,³ Yu Cao,¹ Curtis Gumbs,¹ Kathryn A. Gold,² Neda Kalhor,⁷ Latasha Little,¹ Harshad Mahadeshwar,⁴ Cesar Moran,⁷ Alexei Protopopov,⁴ Huandong Sun,⁴ Jiabin Tang,^{4*} Xifeng Wu,⁸ Yuanqing Ye,⁸ William N. William,² J. Jack Lee,⁹ John V. Heymach,^{2,10} Waun Ki Hong,² Stephen Swisher,¹¹ Ignacio I. Wistuba,³ P. Andrew Futreal^{1,12†}

Cancers are composed of populations of cells with distinct molecular and phenotypic features, a phenomenon termed intratumor heterogeneity (ITH). ITH in lung cancers has not been well studied. We applied multiregion whole-exome sequencing (WES) on 11 localized lung adenocarcinomas. All tumors showed clear evidence of ITH. On average, 76% of all mutations and 20 out of 21 known cancer gene mutations were identified in all regions of individual tumors, which suggested that single-region sequencing may be adequate to identify the majority of known cancer gene mutations in localized lung adenocarcinomas. With a median follow-up of 21 months after surgery, three patients have relapsed, and all three patients had significantly larger fractions of subclonal mutations in their primary tumors than patients without relapse. These data indicate that a larger subclonal mutation fraction may be associated with increased likelihood of postsurgical relapse in patients with localized lung adenocarcinomas.

Intratumor heterogeneity (ITH) may have impacts on tumor biopsy strategy, characterization of actionable targets, treatment planning, and drug resistance (1–6). ITH has recently been elucidated in substantial detail in sev-

eral cancer types with the use of next-generation sequencing (NGS) approaches (7–14). Recent evidence supports a model of branched evolution leading to variable ITH in different tumors (9, 13, 15, 16). Studies in clear cell renal carcinoma

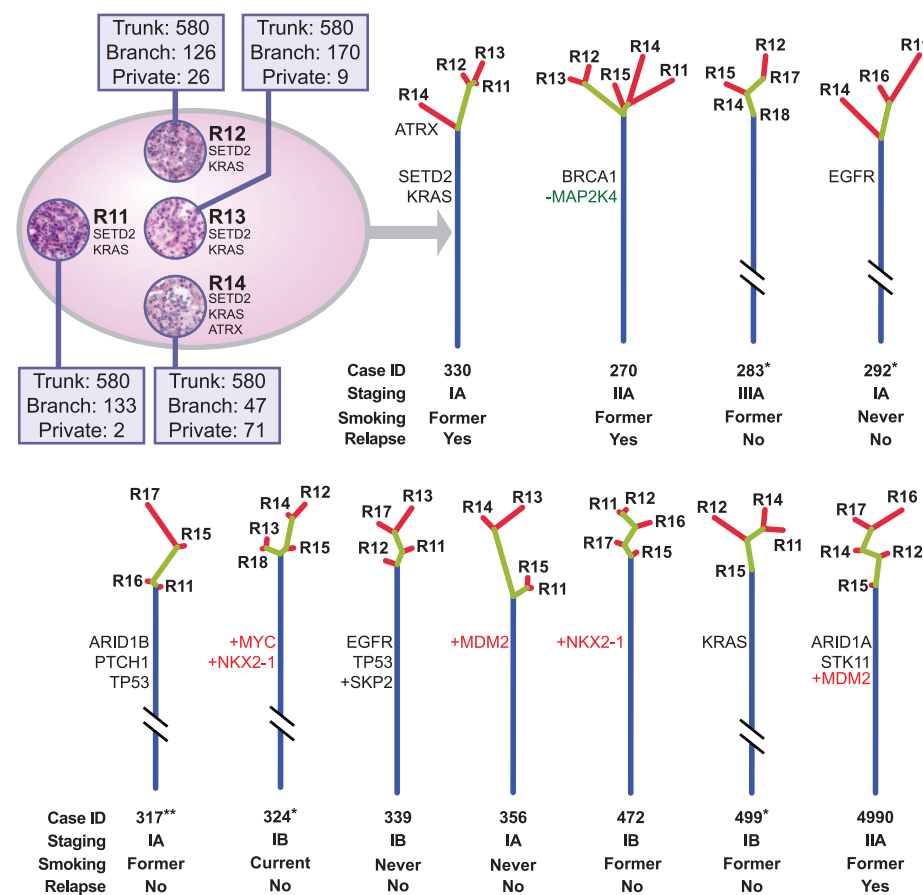
(ccRCC) have demonstrated substantial ITH, with the majority of mutations in known cancer genes confined to spatially separated tumor regions, except that *VHL* loss is a ubiquitous event (16, 17). These data suggest that a single biopsy may be inadequate for identifying all cancer gene mutations from a tumor, as it presents an incomplete view of potential targets for therapy. Critically, the extent to which these observations in ccRCC apply to other solid tumors is currently not clear.

To characterize ITH in localized lung adenocarcinomas, we applied multiregion whole-exome sequencing (WES) on 48 tumor regions from 11 resected lung adenocarcinomas (eight stage I, two stage II, and one stage III tumors, tumor size 2 to 4.6 cm), all from patients who had surgery with curative intent (Fig. 1 and table S1). WES was conducted at a mean depth of 277×. In total, 7269 mutations were identified, and 7026 (97%) somatic mutations were validated by a separate bespoke capture sequencing experiment at mean depth of 863× (table S2). The numbers of mutations varied substantially between tumors (fig. S1), but no significant correlations were identified between mutation burden and age, gender, tumor size, lymph node status, or smoking status.

A useful approach when considering ITH is to depict a given tumor as a tree structure with the trunk representing ubiquitous mutations present in all regions of the tumor; branches representing heterogeneous mutations present in only some regions of the tumor; and private branches representing mutations that are present only in one region of the tumor—analogueous to a phylogenetic tree. Placement of mutations on trunks versus branches reflects relative molecular time of acquisition, with branch mutations occurring, by definition, subsequent to trunk mutations. We applied this approach to multiregion sequencing data from these 11 lung adenocarcinomas. Evi-

dence for ITH was found in each tumor studied. On average, 76% of all mutations were detected in all regions of the same tumors. However, the phylogenetic structure varied considerably between tumors (Fig. 1). We then characterized known cancer gene mutations, defined as nonsynonymous mutations identical to those previously reported in known cancer genes (18–23) or truncating mutations in known tumor suppressor genes, in the context of the derived phylogenetic tree structures. Thirteen of 14 known cancer gene point mutations were mapped to the trunks of the phylogenetic trees (Fig. 1 and table S3), which indicated that these mutations were acquired relatively early during evolution of these 11 tumors. In contrast to ccRCC, these data suggest that single-region sampling may be sufficient to identify the majority of known cancer gene mutations in localized lung adenocarcinomas.

We were also able to evaluate copy number changes relative to ITH. In contrast to ccRCC (16, 17, 24), we did not observe substantial differences in large-scale chromosome aberrations (fig. S2A), and the log2 ratio profiles were similar between different regions within the same tumors (fig. S2B and table S4). Furthermore, amplification or deletion of known cancer genes (22), as well as their relative placement on the phylogenetic trees, were delineated for these 11 lung adenocarcinomas. All of these events were mapped to the trunks of the phylogenetic trees (Fig. 1), which suggested that, like known cancer gene point mutations discussed above, amplification and/or deletion of known cancer genes were also early molecular events for these 11 tumors. Previous work in breast cancer also suggested that known cancer gene mutations were relatively early genetic events shared by all subclones of



© 2014 The University of Texas M.D. Anderson Cancer Center

Fig. 1. Assessment of ITH of 11 lung adenocarcinomas by multiregion sequencing. An example of regional mutation distribution (case 330) of resected tumors (represented by the ellipsis) is shown at the upper left corner. Mutated cancer genes are indicated to the right of representative hematoxylin and eosin-stained pathology image of each sequenced tumor region. Numbers of trunk, branch, and private branch mutations for each region are indicated in associated windows. A phylogenetic tree was generated from all validated mutations by using the Wagner parsimony method in PHYLIP. Blue, yellow, and red lines represent trunk, branch, and private branches, respectively. Trees are anchored at a germline DNA sequence obtained from peripheral blood of the relevant patients. Known cancer gene mutations are mapped to the trunks and branches as indicated. Point mutations, amplifications, and deletions of known cancer genes are presented as black, +red, and –green, respectively. Trunk and branch lengths are proportional to the numbers of mutations acquired on the corresponding trunk or branch. Note: Five tumors have their trunk lengths reduced to 10%* or 2%** of original length for visualization purposes.

¹Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. ²Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA.

³Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. ⁴Applied Cancer Science Institute, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. ⁵Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK. ⁶Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA.

⁷Department of Pathology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. ⁸Department of Epidemiology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. ⁹Department of Biostatistics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. ¹⁰Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA.

¹¹Department of Thoracic and Cardiovascular Surgery, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. ¹²Honorary Faculty, Wellcome Trust Sanger Institute, Hinxton, UK CB10 1SA.

*Present address: Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY, 10065, USA. †Corresponding author. E-mail: afutreal@mdanderson.org

individual breast cancers (13). Taken together, these results indicate that different cancer types may have different relative timing of acquisition of cancer gene mutations. Further, the data would suggest in this subset of lung adenocarcinomas, there are likely mutations in noncanonical cancer genes that drive tumor development and subclonal divergence.

With a median follow-up of 21 months after surgery at the time of this report, three patients have had disease relapse. These three patients had a significantly larger proportion of subclonal nontrunk mutations (branch plus private branch mutations) in their primary tumors than patients without relapse (average 40% in relapsed patients versus 17% in patients without relapse, $P = 0.006$ by t test) (Fig. 1). Although the sample size is small, these findings suggest the possibility that subclonal mutations may be important for cancer progression and that larger subclonal mutation fraction may be associated with an increased likelihood of postsurgical relapse in this subset of lung adenocarcinoma patients.

Analysis of NGS data relies heavily on adequate sequencing depth to make high-accuracy consensus base calls. We compared our WES data (average sequencing depth 277 \times) with deep sequencing data (average sequencing depth 863 \times) to assess the effect of sequencing depth on detecting known cancer gene mutations. In tumor 499, a canonical *KRAS* p.G12C mutation was detected in only one of four tumor regions at exome depth but was detected in all four tumor regions at increased sequencing depth (table S2). Extending this analysis, we then compared deep sequencing data with WES data in defining ITH. The result showed many branch and private branch mutations defined by WES were detectable in all regions of individual tumors with increasing sequencing depth (Fig. 2). Taken together, these results indicate that considerable depth of sequencing will be necessary to detect cancer gene mutations and to accurately characterize ITH of lung adenocarcinomas.

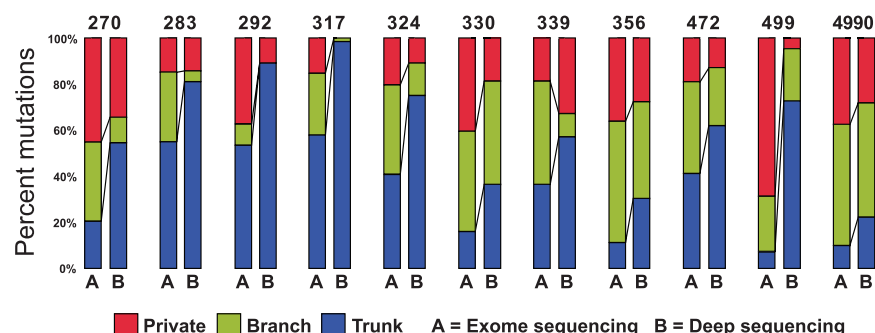
Next, we analyzed the mutational spectrum of these 11 lung adenocarcinomas. Consistent with previous studies (18–20, 25), different mutation spectra were observed in smokers and nonsmokers. Three never-smokers (cases 292, 339, and 356) showed C>T-predominant mutation profiles, as in nonsmokers. Three former smokers who had quit more than 20 years before (cases 270, 472, and 4990) and one former smoker who had a 25 pack-year (pack-year = number of packs per day \times number of years) history of smoking and quit 6 years ago (case 283) also showed C>T-predominant mutation profiles, as in nonsmokers. Two former smokers who had a >50 pack-year history of smoking and had quit 5 years before (cases 317 and 499) and one former smoker, who had a 25 pack-year history of smoking and had quit only 2.5 years before (case 330) showed C>A-predominant mutation profiles consistent with the mutation profile of cigarette smoke exposure. The only current smoker, who had a 20 pack-year history of smoking but had cut down to two cigarettes a day at the time of cancer diagnosis (case 324), showed

an equivalent portion of C>T (26%) versus C>A (21%) substitutions in her tumor (Fig. 3A). These results indicate that tumor mutation spectra in former smokers reflect not only quantity of smoking exposure but also time since smoking cessation.

We next compared the mutational spectrum of trunk versus nontrunk mutations to explore the relative contribution of mutational processes over time. Significant differences in mutational spectrum were observed in six tumors, which indicated that specific mutational processes were

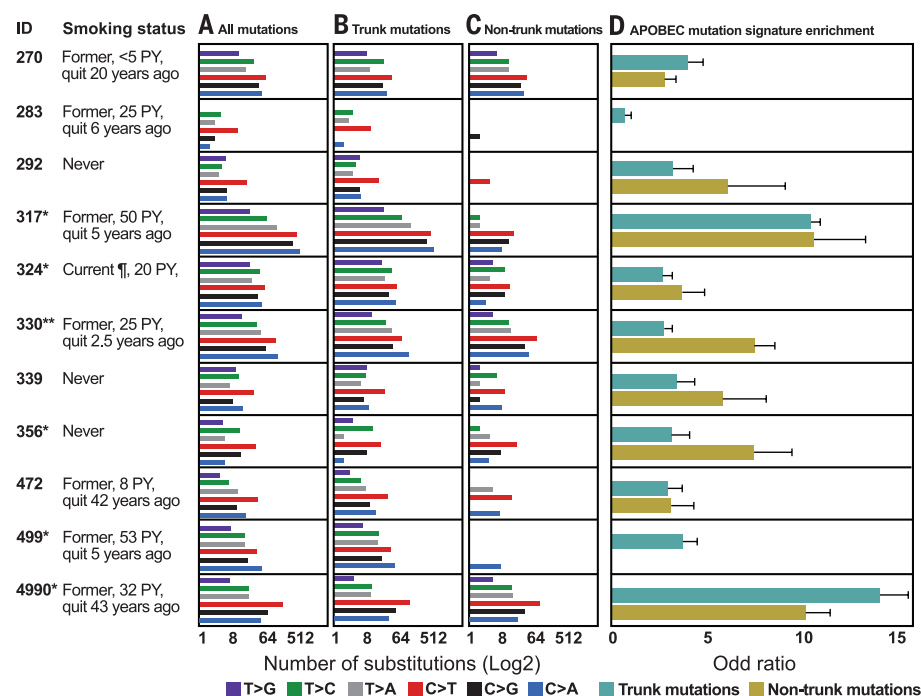
likely operative at different times during development of these tumors (Fig. 3, B and C). Of interest, two former smokers (cases 317 and 330) and the current smoker (case 324) showed significant differences between trunk and nontrunk mutation spectrum with a shift from smoking-associated C>A transversions in trunk mutations to nonsmoker-associated C>T transitions in nontrunk mutations.

Recent evidence has suggested that APOBEC activity is a major source for C>T and C>G mutations



© 2014 The University of Texas M.D. Anderson Cancer Center

Fig. 2. Distribution of trunk, branch, and private branch mutations defined by exome sequencing (average sequencing depth of 277 \times) versus deep sequencing (average sequencing depth of 863 \times). Only validated mutations meeting the following criteria are included: total counts in tumor DNA ≥ 100 ; total counts in germline DNA ≥ 50 ; variant allele frequency (VAF) of $\geq 5\%$ in tumor DNA, and VAF = 0 in germline DNA.



© 2014 The University of Texas M.D. Anderson Cancer Center

Fig. 3. Mutation spectrum of the 11 lung adenocarcinomas. (A) Mutation spectrum of all validated mutations. (B) Mutation spectrum of trunk mutations. (C) Mutation spectrum of nontrunk mutations. The difference of mutation spectrum between trunk and nontrunk mutations in each patient was evaluated with Fisher's exact test, and significant P values are shown as * $P < 0.05$ and ** $P < 0.01$. (D) APOBEC mutation signature enrichment odds ratio for trunk and nontrunk mutations. The 95% confidence intervals for Fisher's exact test are indicated. PY, pack-year. ¶ indicates that the patient had cut down to two cigarettes a day at the time of cancer diagnosis.

(12, 26). We therefore investigated whether there is evidence of an APOBEC mutational process in this subset of lung adenocarcinomas. On average, 28% of all mutations had a specific substitution pattern (C>T/G at TpCpW sites, where W is A or T), consistent with an APOBEC-mediated process (fig. S3). APOBEC mutation signature enrichment was found to be more pronounced for nontrunk mutations compared with trunk mutations in 7 of the 11 patients; however this difference was statistically significant only for case 330 (Fig. 3D). These data suggest that an APOBEC-like process is contributing substantially to the mutations found in this subset of lung adenocarcinomas and that the process tends to be more pronounced in later, subclonal mutations; this further highlights the dynamic nature of mutational processes in play.

Substantial variation in the allele frequency of somatic mutations within each individual tumor region from a given tumor was observed in this set of lung adenocarcinomas (fig. S4A). To more formally characterize subclonal fraction within each tumor region, we used the ABSOLUTE algorithm (27). These analyses demonstrated that at least 29 out of 48 individual tumor regions showed evidence of intraregional subclonal populations. The distribution of clonal and subclonal mutations was different among the sampled regions within the same tumors in some patients (fig. S4B), which suggests that single-biopsy analysis would be inadequate to fully represent ITH in these tumors.

To explore the possible implications of these data on translation to routine ITH assessment for in clinical practice, we repeated the ABSOLUTE analysis on the combined sequencing data from all tumor regions of each patient to assess the global ITH on a per-patient level, defined by the relative proportion of subclonal mutations. Similar to the phylogenetic analyses, all three patients with relapsed disease had larger subclonal fractions in their primary tumors (average 41% in patients with relapse versus 24% in patients without relapse, $P = 0.045$ by t test) (fig. S5A). Use of a complementary Bayesian Dirichlet process (13) on the per-patient combined data revealed the same trend (average subclonal mutations 66% in patients with relapse versus 36% in patients without relapse, $P = 0.035$ by t test)

(fig. S5B). These results suggest that a measure of overall subclonal fraction may be of interest from a prognostic standpoint in this population of patients.

Resectable localized disease accounts for 30 to 50% of all non-small cell lung cancers, with increasing prevalence as screening is more widely implemented (28–30). Given that this subset of patients has tumors surgically resected as standard of care, there is an opportunity to confirm these preliminary observations by deep sequencing multiregion samples obtained from resected tumors. The questions of whether sequencing-targeted cancer gene panels versus WES will yield sufficient mutation data for meaningful analyses and the most appropriate algorithms for analyses will need to be addressed in order to fully test if the clinical correlation suggested in these data are borne out in larger patient cohorts.

Evidence of marked regional ITH in ccRCC suggested substantial challenges to personalized oncology based on single-tumor biopsy to portray the mutational landscape. This study, however, provides evidence that ITH patterns may be different between cancer types. With the caveat of limited sample size fully acknowledged, these data suggest that, although multiregion sampling is needed to fully assess ITH complexity, single-biopsy analysis at appropriate depth might be sufficient to identify the majority of known cancer gene mutations in this subset of lung adenocarcinomas. Studies in much larger cohorts, ideally with comprehensive clinical annotation and repeat biopsy at relapse, are needed to fully understand the clinical impact of ITH and insights afforded by these types of analyses. Furthermore, extension of research to epigenetic and phenotypic assessment through regional DNA methylation, chromatin state, and RNA and/or protein expression studies over time and under treatment is needed to fully understand the impact of ITH on the biology of the cancer itself and its impact on the clinical phenotype of cancer patients.

REFERENCES AND NOTES

1. T. A. Yap, M. Gerlinger, P. A. Futreal, L. Pusztai, C. Swanton, *Sci. Transl. Med.* **4**, 127ps10 (2012).
2. C. Swanton, *Cancer Res.* **72**, 4875–4882 (2012).
3. C. Swanton et al., *Genome Med.* **2**, 53 (2010).

4. S. Horswell, N. Matthews, C. Swanton, *Cancer Lett.* **340**, 220–226 (2013).
5. R. Fisher, L. Pusztai, C. Swanton, *Br. J. Cancer* **108**, 479–485 (2013).
6. N. A. Saunders et al., *EMBO Mol. Med.* **4**, 675–684 (2012).
7. N. Navin et al., *Genome Res.* **20**, 68–80 (2010).
8. N. Navin et al., *Nature* **472**, 90–94 (2011).
9. K. Anderson et al., *Nature* **469**, 356–361 (2011).
10. A. Sottoriva et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4009–4014 (2013).
11. P. J. Campbell et al., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13081–13086 (2008).
12. S. Nik-Zainal et al., *Cell* **149**, 979–993 (2012).
13. S. Nik-Zainal et al., *Cell* **149**, 994–1007 (2012).
14. S. Yachida et al., *Nature* **467**, 1114–1117 (2010).
15. S. P. Shah et al., *Nature* **461**, 809–813 (2009).
16. M. Gerlinger et al., *N. Engl. J. Med.* **366**, 883–892 (2012).
17. M. Gerlinger et al., *Nat. Genet.* **46**, 225–233 (2014).
18. M. Imielinski et al., *Cell* **150**, 1107–1120 (2012).
19. L. Ding et al., *Nature* **455**, 1069–1075 (2008).
20. R. Govindan et al., *Cell* **150**, 1121–1134 (2012).
21. I. R. Watson, K. Takahashi, P. A. Futreal, L. Chin, *Nat. Rev. Genet.* **14**, 703–718 (2013).
22. B. Vogelstein et al., *Science* **339**, 1546–1558 (2013).
23. S. A. Forbes et al., *Nucleic Acids Res.* **39** (Database), D945–D950 (2011).
24. P. Martinez et al., *J. Pathol.* **230**, 356–364 (2013).
25. P. Hainaut, G. P. Pfeifer, *Carcinogenesis* **22**, 367–374 (2001).
26. L. B. Alexandrov et al., *Nature* **500**, 415–421 (2013).
27. S. L. Carter et al., *Nat. Biotechnol.* **30**, 413–421 (2012).
28. H. Bülzebruck et al., *Cancer* **70**, 1102–1110 (1992).
29. P. J. Mahadevia et al., *JAMA* **289**, 313–322 (2003).
30. T. R. Church et al., *N. Engl. J. Med.* **368**, 1980–1991 (2013).

ACKNOWLEDGMENTS

This study was supported by the Cancer Prevention and Research Institute of Texas (R120501), the University of Texas (UT) Systems Stars Award (PS100149), the Welch Foundation Robert A. Welch Distinguished University Chair Award (G-0040), Department of Defense PROSPECT grant (W81XWH-07-1-0306), the UT Lung Specialized Programs of Research Excellence grant (P50CA70907), the MD Anderson Cancer Center Support grant (CA016672), NIH T32 Research Training in Academic Medical Oncology grant (CA-009666), the MD Anderson Institutional Support for the Center for Translational and Public Health Genomics, and the A. Lavoy Moore Endowment Fund. The authors thank L. Chin and R. Verhaak for constructive discussions. Sequence data have been deposited at the European Genome-phenome Archive (EGA, www.ebi.ac.uk/ega/), which is hosted by the European Bioinformatics Institute, under accession number EGAS00001000930.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6206/256/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S4
References (31–44)

3 June 2014; accepted 29 August 2014
10.1126/science.1256930