Abbosh, Christopher, Birkbak, Nicolai J., Wilson, Gareth A., Jamal-Hanjani, Mariam, Constantin, Tudor, Salari, Raheleh, Quesne, John Le, Moore, David A, Veeriah, Selvaraju, Rosenthal, Rachel, Marafioti, Teresa, Kirkizlar, Eser, Watkins, Thomas B K, McGranahan, Nicholas, Ward, Sophia, Martinson, Luke, Riley, Joan, Fraioli, Francesco, Bakir, Maise Al, GrÖnroos, Eva, Zambrana, Francisco, Endozo, Raymundo, Bi, Wenya Linda, Fennessy, Fiona M., Sponer, Nicole, Johnson, Diana, Laycock, Joanne, Shafi, Seema, Czyzewska-Khan, Justyna, Rowan, Andrew, Chambers, Tim, Matthews, Nik, Turajlic, Samra, Hiley, Crispin, Lee, Siow Ming, Forster, Martin D., Ahmad, Tanya, Falzon, Mary, Borg, Elaine, Lawrence, David, Hayward, Martin, Kolvekar, Shyam, Panagiotopoulos, Nikolaos, Janes, Sam M, Thakrar, Ricky, Ahmed, Asia, Blackhall, Fiona, Summers, Yvonne, Hafez, Dina, Naik, Ashwini, Ganguly, Apratim, Kareht, Stephanie, Shah, Rajesh, Joseph, Leena, Quinn, Anne Marie, Crosbie, Phil, Naidu, Babu, Middleton, Gary, Langman, Gerald, Trotter, Simon, Nicolson, Marianne, Remmen, Hardy, Kerr, Keith, Chetty, Mahendran, Gomersall, Lesley, Fennell, Dean A., Nakas, Apostolos, Rathinam, Sridhar, Anand, Girija, Khan, Najid, Russell, Peter, Ezhil, Veni, Ismail, Babikir, Irvin-sellers, Melanie, Prakash, Vineet, Lester, Jason, Kornaszewska, Malgorzata, Attanoos, Richard, Adams, Haydn, Davies, Helen, Oukrif, Dahmane, Akarca, Ayse U, Hartley, John A, Lowe, Helen L, Lock, Sara, Iles, Natasha, Bell, Harriet, Ngai, Yenting, Elgar, Greg, Szallasi, Zoltan, Schwar, Roland F, Herrero, Javier, Stewart, Aengus, Quezada, Sergio A, Van Loo, Peter, Dufy, Caroline, Lin, C. Jimmy, Rabinowitz, Matthew, Aerts, Hugo JW L, Hackshaw, Allan, Shaw, Jacqui A, Zimmermann, Bernhard G., Swanton, Charles, Jamal-Hanjani, Mariam, Abbosh, Christopher, Veeriah, Selvaraju, Shafi, Seema, Czyzewska-Khan, Justyna, Johnson, Diana, Laycock, Joanne, Bossard-Carter, Leticia, Goh, Gerald, Rosenthal, Rachel, Gorman, Pat, Murugaesu, Nirupa, Hynds, Robert E, Wilson, Gareth, Birkbak, Nicolai J, Watkins, Thomas B K, McGranahan, Nicholas, Horswell, Stuart, Al Bakir, Maise, GrÖnroos, Eva, Mitter, Richard, Escudero, Mickael, Stewart, Aengus, Van Loo, Peter, Rowan, Andrew, Xu, Hang, Turajlic, Samra, Hiley, Crispin, Goldman, Jacki, Stone, Richard Kevin, Denner, Tamara, Matthews, Nik, Elgar, Greg, Ward, Sophia, Biggs, Jennifer, Costa, Marta, Begum, Sharmind, Phillimore, Ben, Chambers, Tim, Nye, Emma, Graca, Sofia, Al Bakir, Maise, Yoshi, Kroopa, Furness, Andrew, Aissa, Assma Ben, Wong, Yien Ning Sophia, Georgiou, Andy, Quezada, Sergio, Hartley, John A, Lowe, Helen L, Herrero, Javier, Lawrence, David, Hayward, Martin, Panagiotopoulos, Nikolaos, Kolvekar, Shyam, Falzon, Mary, Borg, Elaine, Marafioti, Teresa, Simeon, Celia, Hector, Gemma, Smith, Amy, Aranda, Marie, Novelli, Marcello, Oukrif, Dahmane, Akarca, Ayse U, Janes, Sam M, Thakrar, Ricky, Forster, Martin, Ahmad, Tanya, Lee, Siow Ming, Papadatos-Pastos, Dionysis, Carnell, Dawn, Mendes, Ruheena, George, Jeremy, Navani, Neil, Ahmed, Asia, Taylor, Magali, Choudhary, Junaid, Summers, Yvonne, Califa, Raffaele, Taylor, Paul, Shah, Rajesh, Krysiak, Piour, Rammohan, Kendadai, Fontaine, Eustace, Booton, Richard, Evison, Matthew, Crosbie, Phil, Moss, Stuart, Idries, Faiza, Joseph, Leena, Bishop, Paul, Chaturved, Anshuman, Marie Quinn, Anne, Doran, Helen, Jeek, Angela, Harrison, Phil, Moore, Katrina, Waddington, Rachael, Noviando, Juliette, Blackhall, Fiona, Rogan, Jane, Smith, Elaine, Dive, Caroline, Tugwood, Jonathan, Brady, Ged, Rothwell, Dominic G, Chemi, Francesca, Pierce, Jackie, Gulati, Sakshi, Naidu, Babu, Langman, Gerald, Trotter, Simon, Bellamy, Mary, Pages, Hollie, Kerr, Amy, Kadiri, Salma, Webb, Joanne, Middleton, Gary, Djeamaran, Madava, Fennell, Louise, Bennett, Jonathan, Riley, Joan, Primrose, Lindsay, Martinson, Luke, Anand, Girija, Khan, Najid,
Amadi, Anita, Nicolson, Marianne, Kerr, Keith, Palmer, Shirley, Remmen, Hardy, Miller, Joy, Buchan, Keith, Chetty, Mahendran, Gomersall, Lesley, Lester, Jason, Edwards, Alison, Morgan, Fiona, Adams, Haydn, Davies, Helen, Kornaszewska, Malgorzata, Attanoos, Richard, Lock, Sara, Verjee, Azmina, MacKenzie, Mairead, Wilcox, Maggie, Bell, Harriet, Illes, Natasha, Hackshaw, Allan, Ngai, Yenting, Smith, Sean, Gower, Nicole, Ottensmeier, Christian, Chee, Serena, Johnson, Benjamin, Alzetai, Aiman, Shaw, Emily, Lim, Eric, De Sousa, Paulo, Tavares Barbosa, Monica, Bowman, Alex, Jordan, Simon, Rice, Alexandra, Raubenheimer, Hilgardt, Proli, Chiaara, Elena Cufari, Maria, Ronquillo, John Carlo, Kwayie, Angela, Bhayani, Harshil, Hamilton, Morag, Bakar, Yusra, Mensah, Natalie, Ambrose, Lyn, Devaraj, Anand, Buder, Silviu, Finch, Jonathan, Azcarate, Leire, Chavan, Hema, Green, Sophie, Mashinga, Hillaria, Nicholson, Andrew G, Lau, Kelvin, Sheaff, Michael, Schmid, Peter, Conibear, John, Ezhil, Veni, Ismail, Babikir, Irvin-sellers, Melanie, Prakash, Vineet, Russell, Peter, Light, Teresa, Horey, Tracey, Danson, Sarah, Bury, Jonathan, Edwards, John, Hill, Jennifer, Matthews, Sue, Kitsanta, Yota, Suvarna, Kim, Fisher, Patricia, Keerio, Allah Dino, Shackcloth, Michael, Gosney, John, Postmus, Pieter, Feeney, Sarah, Asante-Siaw, Julius, Constatin, Tudor, Salari, Rahel, Sponer, Nicole, Naik, Ashwini, Zimmermann, Bernhard, Rabinowitz, Matthew, Aerts, Hugo J. W. L., Dentro, Stefan, Dessimoz, Christophe, Swanton, Charles, Jamal-Hanjani, Mariam, Abbosh, Christopher, Shiu, Kai-Keen, Bridgewater, John, Ho chauser, Daniel, Van Loo, Peter, Quezada, Sergio, Beck, Stephan, Parker, Peter, Walczak, Henning, Enver, Tariq, Falzon, Mary, Proctor, Ian, Sinclair, Ron, Lok, Chi-wah, Novelli, Marco, Marafioti, Teresa, Borg, Elaine, Mitchison, Miriam, Trevisan, Giorgia, Lynch, Mark, Brandner, Sebastian, Gishen, Faye, Tookman, Adrian, Stone, Paddy, Sterling, Caroline, Larkin, James, Turajlic, Samra, Attard, Gert, Eeles, Ros, Foster, Chris, Bova, Steve, Sottoriva, Andrea, Chowdhury, Simon, Ashish, Chandra, Spicer, James, Stares, Mark, Lynch, Joanna, Caldas, Carlos, Brenton, James, Fitzgerald, Rebecca, Jimenez-Linan, Merche, Provenzano, Elena, Clurow, Alison, Stewart, Grant, Watts, Colin, Gilbertson, Richard, McDermott, Ulan, Tavare, Simon, Maughan, Tim, Tomllinson, Ian, Campbell, Peter, McNeish, Iain, Biaikin, Andrew, Chambers, Antony, Fraser, Sioban, Oien, Karin, Krebs, Matt, Blackhall, Fiona, Summers, Yvonne, Dive, Caroline, Marais, Richard, Carter, Louise, Nonaka, Daisuke, Quinn, Anne Marie, Dhomen, Nathalie, Fennell, Dean, Le Quesne, John, Moore, David, Shaw, Jacqui, Naidu, Babu, Baijal, Shobhit, Tanchel, Bruce, Langman, Gerald, Collard, Martin, Cockcroft, Peter, Taylor, Joanne, Bancroft, Hollie, Kerr, Amy, Middleton, Gary, Webb, Joanne, Kadir, Salma, Colloby, Peter, Olisemeke, Bernard, Wilson, Rodelaine, Ottensmeier, Christian, Harrison, David, Loda, Massimo, Flanagan, Adreinne, Wilcox, Maggie, McKenzie, Mairead, Hackshaw, Allan, Lederman, Jonathan, Sharp, Abby, Farrelly, Laura and Swanton, Charles 2017. Phylogenetic ctDNA analysis depicts early stage lung cancer evolution. Nature 545 (7655) , pp. 446-451. 10.1038/nature22364

Publishers page: http://dx.doi.org/10.1038/nature22364

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. You are advised to consult the publisher’s version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.
Phylogenetic ctDNA analysis depicts early stage lung cancer evolution

Christopher Abbosh, Nicolai J. Birkbak, Gareth A. Wilson, Mariam Jamal-Hanjani, Tudor Constantin, Raheleh Salari, John Le Quesne, David A Moore, Selvaraju Veeriah, Rachel Rosenthal, Teresa Marafioti, Eser Kirkizlar, Thomas B K Watkins, Nicholas McGranahan, Sophia Ward, Luke Martinson, Joan Riley, Francesco Fraioli, Maise Al Bakir, Eva Grönroos, Francisco Zambrana, Raymundo Endoza, Wenyia Linda Bi, Fiona M. Fennessy, Nicole Sponer, Diana Johnson, Joanne Laycock, Seema Shafi, Justyna Czyzewska–Khan, Andrew Rowan, Tim Chambers, Nik Matthews, Samra Turaajlic, Crispin Hiley, Siow Ming Lee, Martin D. Forster, Tanya Ahmad, Mary Falzon, Elaine Borg, David Lawrence, Martin Hayward, Shyam Kolvekar, Nikolaos Panagiotopoulos, Sam M Janes, Ricky Thakrar, Asia Ahmed, Fiona Blackhall, Yvonne Summers, Dina Hafez, Ashwini Naik, Apratim Ganguly, Stephanie Kareht, Rajesh Shah, Leena Joseph, Anne Marie Quinn, Phil Crosbie, Babu Naidu, Gary Middleton, Gerald Langman, Simon Trotter, Marianne Nicolson, Hardy Remmen, Keith Kerr, Mahendra Chetty, Lesley Comersall, Dean A. Fennell, Apostolos Nakas, Sridhar Rathinam, Girija Anand, Sajid Khan, Peter Russell, Veni Eztil, Babikir Ismail, Melanie Irvin-sellers, Vineet Prakash, Jason F. Lester, Malgorzata Kornaszewska, Richard Attanoos, Haydn Adams, Helen Davies, Dahmane Oukrif, Ayse U Akarca, John A Hartley, Helen L Lowe, Sara Lock, Natasha Iles, Harriet Bell, Yenting Ngai, Greg Elgar, Zoltan Szallasi, Roland F Schwarz, Javier Herrero, Aengus Stewart, Sergio A Quezada, Peter Van Loo, Caroline Dive, C. Jimmy Lin, Matthew Rabinowitz, Hugo JW Aerts, Allan Hackshaw, Jacqui A Shaw, Bernhard G. Zimmermann, the TRACERx consortium, the PEACE consortium & Charles Swanton

This is a PDF file of a peer-reviewed paper that has been accepted for publication. Although unedited, the content has been subjected to preliminary formatting. Nature is providing this early version of the typeset paper as a service to our customers. The text and figures will undergo copyediting and a proof review before the paper is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply.

Cite this article as: Abbosh, C. et al. Phylogenetic ctDNA analysis depicts early stage lung cancer evolution. Nature http://dx.doi.org/10.1038/nature22364 (2017).

Competing financial interests statement: T.C., R.S., E.K., N.S., D.H., A.N., A.G., S.K., J.L., B.G.Z. and M.R. are all employees or former employees of Natera Inc. and own stock and/or options to purchase stock. C.A. and C.S. submitted a patent with UCL business PLC (provisional patent number 1618485.5) based on this work. The title of the patent is “Method of detecting tumor recurrence” and describes a phylogenetic approach to targeted ctDNA profiling, patent is pending. C.A., M.J-H, G.W., C.S. and Natera Inc. employees submitted a patent with Natera Inc. (provisional US application 62/323,589) based on this work. The title of the patent is “Methods for lung cancer detection” and describes a multiplex PCR NGS approach to ctDNA detection in lung cancer, patent is pending. C.S. has had the following financial relationships in the last 36 months - Boehringer Ingelheim - Consulting and speaker fees, Novartis - Consulting and speaker fees, Eli Lilly - Speaker fees, Roche - Consulting and speaker fees, GlaxoSmithKline - Speaker fees, Pfizer - Speaker fees, Celgene - Speaker fees, Servier - Speaker fees, Grail - Scientific Advisory Board; Stock options, APOGEN Biotechnologies - Scientific Advisory Board; Stock options, Epic Biosciences - Scientific Advisory Board; Stock options, Achilles Therapeutics - Founder; Stock Options. In 2015 C.S. was a member of a Natera advisory board and received an honorarium.

Received 27 January; accepted 13 April 2017.
Accelerated Article Preview Published online 26 April 2017.
Phylogenetic ctDNA analysis depicts early stage lung cancer evolution

Christopher Abbosh1,*, Nicolai J. Birkbak1,*, Gareth A. Wilson1,*, Mariam Jamal–Hanjani1,*, Tudor Constantin1,*, Rahelee Salari1,*, John Le Quesne1,*, David A. Moore1,*, Selvaraju Veerarth1,*, Rachel Rosenthal1,*, Teresa Marafioti1,*, Eser Kirkizlar1,*, Thomas B K Watkins1,*, Nicholas McGrath1,*, Sophia Ward1,*, Luke Martinson1,*, Joan Riley1,*, Francesco Fratoli1,*, Maise Al Bakir2,*, Eva Grönroos2,*, Francisco Zambrana3,*, Raymond Endo40,*, Wenya Linda Bi1,*, Fiona M. Fennelly5,*, Nicole Sporer5,*, Diana Johnston5,*, Joanne Laycock6,*, Seema Shaikh2,*, Justyna Czyzewski–Khan7,*, Andrew Rowan4,*, Tim Chambers2,*, Nik Matthews4,*, Samra Turajlic2,*, Crispin Hiley1,*, Siow Ming Lee1,*, Martin D. Forster1,*, Tanya Ahmad1,*, Mary Falcon2,*, Elaine Borg3,*, David Lawrence4,*, Martin Hayward5,*, Shyam Kolvekar6,*, Nikolaos Panagiotopoulos3,*, Sam J. Mares1,*, Ricky Thakrar1,*, Asia Ahmed1,*, Fiona Blackhall1,*, Yvonne Summers6,*, Dina Halef7,*, Ashwin Naik8,*, Aparmit Ganguly9,*, Stephanie Kareth1,*, Rajesh Shah10,*, Leena Joseph10,*, Anne Marie Quinn11,*, Phil Crossbie12,*, Babu Nadula12,*, Gary Middleton13,*, Gerald Langman14,*, Simon Trotter15,*, Marianne Nicolson16,*, Hardy Remmen17,*, Keith Kerr18,*, Mahendran Chetty19,*, Lesley Gomersall20,*, Dean A. Fennell21,*, Apostoalois Nakas22,*, Sridhar Rathinam23,*, Girija Anand23,*, Sajid Khan23,*, Peter Russell24,*, Veni Ezhi25,*, Babirak Ismail26,*, Melanie Irvin – sellers27,*, Vineet Prakash28,*, Jason F. Lester29,*, Malgorzata Kornaszewska30,*, Richard Attanoe31,*, Haydn Adams32,*, Helen Davies33,*, Dahnane Ouikrit34,*, Ayse U Akcar35,*, John A Hartley36,*, Helen L Lowe37,*, Sara Lock38,*, Natasha Isles39,*, Harriet Bell1,*, Yenting Ngai1,*, Greg Elgar1,*, Zoltan Szallas1,*, Roland F Schwarz1,*, Javier Herrero3,*, Aengus Stewart3,*, Sergio A Quezada3,*, Peter Van Loo55,*, Caroline Dive56,*, C. Jimmy Lin1,*, Matthew Rabinowitz3,*, Hugo JW Aerts2,*, Allan Hackshaw4,*, Jacqui A Shaw4,*, Bernhard G. Zimmermann1,*, the TRACERx consortium1,*, the PEACE consortium1,*, & Charles Swanton1,*. 

The early detection of relapse following primary surgery for non–small cell lung cancer and the characterization of emerging subclones seeding metastatic sites might offer new therapeutic approaches to limit tumor recurrence. The potential to non–invasively track tumor evolutionary dynamics in ctDNA of early–stage lung cancer is not established. Here we conduct a tumour–specific phylogenetic approach to ctDNA profiling in the first 100 TRACERx patients.
Lung cancer is the leading cause of cancer death\(^1\,\,2\). Metastatic non-small cell lung cancer (NSCLC) cannot be cured with systemic chemotherapy, yet clinical studies have shown a 5% benefit of post-operative (adjuvant) chemotherapy on overall survival\(^3\). This modest survival benefit may reflect a vulnerability of low volume disease within the context of reduced intra-tumor heterogeneity\(^4\). Circulating tumor DNA (ctDNA) detection in plasma following resection of breast\(^5\,\,6\) and colorectal\(^7\) tumors has been shown to identify patients destined to relapse post-operatively in advance of established clinical parameters.

Identifying, monitoring and genomics characterizing residual disease following primary lung cancer surgery may improve outcomes in the adjuvant setting. This would create a therapeutic setting where only patients destined to recur would receive treatment and where intervention could be directed to the evolving tumor subclone that is seeding metastatic recurrence.

Here, we report a bespoke multiplex-PCR NGS approach to ctDNA profiling within the context of the prospective tumor evolutionary NSCLC TRACERx study. We address determinants of ctDNA detection in early-stage NSCLC and investigate the ability of ctDNA to identify and genomically characterize post-operative NSCLC relapse within a tumor phylogenetic framework.

**Phylogenetic ctDNA profiling**

The TRACERx study monitors the clonal evolution of NSCLC from diagnosis through to death\(^8\,\,9\). Using multi-region exome sequencing (M-Seq) derived tumor phylogenetic trees developed through prospective analysis of a 100 patient TRACERx cohort, we conducted a phylogenetic approach to ctDNA profiling in early-stage NSCLC (Fig. 1). Bespoke multiplex-PCR assay-panels were synthesised for each patient, targeting clonal and subclonal single nucleotide variants (SNVs) selected to track phylogenetic tumor branches in plasma (Fig. 1). SNV detection in plasma was established through a calling algorithm employing negative control samples (see Methods). Analytical validation of the multiplex-PCR NGS platform demonstrated a sensitivity of above 99% for the detection of SNVs at frequencies above 0.1% and the specificity of detecting a single SNV was 99.6% (Extended Data Fig. 1a). At least two SNVs were detected in ctDNA from early-stage NSCLCs analyzed in our published discovery cohort data\(^10\), demonstrating biological sensitivity at a two SNV threshold for ctDNA detection. Therefore, we prospectively selected a threshold of two detected SNVs for calling a sample ctDNA positive for validation within this study; to minimize type I error when testing up to 30 tumour-specific SNVs per time-point in a single patient (see Extended Data Fig. 1b for justification).

**Determinants of ctDNA detection in NSCLC**

We sought to identify clinicopathological determinants of ctDNA detection in early-stage NSCLC by profiling pre-operative plasma samples in 100 TRACERx patients. Samples from four patients could not be analyzed (see cohort design Extended Data Fig. 2a, patient characteristics Extended Table 1a–c, Supplementary Table 1). Individual assay-panels were designed to target a median of 18 SNVs (range 10 to 22) comprising a median of 11 clonal SNVs (range 2 to 20) and a median of 6 subclonal SNVs (range 0 to 16) per patient (Extended Data Fig. 2b,c).

At least two SNVs were detected in ctDNA pre-operatively in 46 of 96 (48%) early-stage NSCLCs and a single SNV was detected in 12 additional cases (Fig. 2a). Centrally reviewed pathological data revealed that ctDNA detection was associated with histological subtype: 97% (30/31) of lung squamous cell carcinomas (LUSCs) and 71% (5 of 7) of other NSCLC subtypes were ctDNA positive, compared with 19% (11/58) of lung adenocarcinomas (LUADs) (Fig. 2a). 94% (16 of 17) of stage I LUSCs were detected compared with 13% (5 of 39) of stage I LUADs (Extended Data Fig. 3a). Passive release of ctDNA into the circulation may be associated with necrosis\(^8\). As expected LUSCs were significantly more necrotic than LUADs\(^12\) and ctDNA positive LUADs formed a sub-group of more necrotic tumors compared with ctDNA negative LUADs (Extended Data Fig. 3b). Necrosis, lymph node involvement, lymphovascular invasion, pathological tumor size, Ki67 labelling indices, non-adenocarcinoma histology and total cell-free DNA input predicted ctDNA detection in univariable analyses (Extended Data Fig. 3c). Multivariable analysis revealed non-adenocarcinoma histology, the presence of lympho-vascular invasion and high Ki67 proliferation index as independent predictors of ctDNA detection (Extended Data Fig. 3c). Since FDG-avidity on positron emission tomography (PET) scans correlates with proliferative indices in early-stage NSCLC\(^13\,\,14\), we investigated tumor PET FDG-avidity and ctDNA detection. PET FDG-avidity predicted ctDNA detection (area under curve = 0.84, P < 0.001, n = 92) (Extended Data Fig. 3d). Within LUADs, driver events in KRAS, EGFR or TP53 were not associated with ctDNA detection (Extended Data Fig. 3e).

We analyzed the distribution of clonal and subclonal SNVs in ctDNA positive patients. Clonal SNVs were detected in all 46 ctDNA positive patients and a median of 94% (range 11% to 100%) of clonal SNVs targeted by assay-panels were detected in the ctDNA of these patients. 40 of 46 ctDNA positive patients had subclonal SNVs targeted by assay-panels and subclonal SNVs were detected in 27 (68%) of these patients. A median of 27% (range 0% to 91%) of subclonal SNVs within individual assay-panels were detected in ctDNA positive patients (Fig. 2b). The mean plasma variant allele frequency (VAF) of clonal SNV's was higher than that of subclonal SNVs (Extended Data Fig. 4a, within patient comparison, Wilcoxon signed-rank test, P < 0.001, n = 27, Supplementary Table 2) supporting the use of clonal alterations as a more sensitive method of ctDNA detection than subclonal alterations\(^10\,\,15\).

In ctDNA positive patients, pathologic tumor size correlated with mean clonal plasma VAF (Spearman’s Rho = 0.405, P = 0.005, n = 46, Extended Data Fig. 4b). CT scan volumetric analyses were available in 38 of 46 ctDNA positive patients (see Extended Data Fig. 4c). Tumor volume correlated with mean clonal plasma VAF (Fig. 3a, Spearman’s Rho = 0.61, P < 0.001, n = 38). A linear relationship between log-transformed volume and log-transformed mean clonal VAF values was observed (Fig. 3a). The line of best fit applied to the data was consistent with the line fitted to NSCLC volumetric data and ctDNA plasma VAFs reported in previously published work\(^16\) (Extended Data Fig. 4d). Linear modelling based on the TRACERx data predicted that a primary tumor burden of 10 cm\(^3\) would result in a mean clonal plasma VAF of 0.1% (95% C.I. 0.05 to 0.17%) (Fig. 3b). Tumor purity was multiplied by tumor volume to control for stromal contamination to determine cancer cell volume corresponding to clonal plasma VAF (Extended Data Fig. 4e). On the assumption that 1 cm\(^3\) of pure tumor contains 9.4 x 10\(^{17}\) cells\(^17\), a plasma VAF of 0.1% would correspond to a primary NSCLC malignant burden of 326 million cells (Fig. 3b, Extended Data Fig. 4f).
To investigate predictors of subclone detection, detected subclonal SNVs were mapped back to M-seq derived tumor phylogenetic trees. 35 of 57 (61%) shared subclones (identified in more than one tumor region through M-Seq analysis) were identified in ctDNA, compared with 26 of 80 (33%) private subclones (detected in a single tumor region only) (Extended Data Fig 4g). This suggested subclone volume influences subclonal ctDNA detection. Subclone volume was estimated based on mean regional subclone cancer cell fraction and cancer cell volume. Detected subclonal SNVs mapped to subclones with higher estimated volumes than subclones containing undetected SNVs (Fig. 3c) and subclone volume correlated with subclonal SNV plasma VAF (Fig. 3d).

Detecting and characterizing NSCLC relapse
The longitudinal phase of the study aimed to determine if ctDNA profiling with patient-specific assay panels could detect and characterize the branched subclone(s) seeding NSCLC relapse. Pre- and post-surgical plasma ctDNA profiling was performed blinded to relapse status in a sub-group of 24 patients (cohort characteristics, Extended Table 1d-e). This included relapse free patients who had been followed-up for a median of 775 days (range 688 to 945 days, n = 10) and confirmed NSCLC relapse cases (n = 14) (cohort design, Extended Data Fig. 2c). Additional PCR assays were added to panels in this phase of the study to attempt to improve ctDNA detection in LUADs, a median of 18.5 SNVs (range 12 to 20) were targeted by LUSC assay-panels and a median of 28 SNVs (range 25 to 30) were targeted by LUAD assay-panels (Extended Data Fig. 2d-e).

Patients were followed up with three to six monthly clinical assessment and chest radiographs. At least 2 SNVs were detected in 13 of 14 (93%) patients with confirmed NSCLC relapse prior to, or at, clinical relapse (Fig. 4a-g, Extended Data Fig. 5). At least two SNVs were detected in 1 of 10 (10%) patients (CRUK0013) with no clinical evidence of NSCLC relapse (Fig. 4h, Extended Data Fig. 6). Excluding a single case where no post-operative plasma was taken prior to clinical relapse (CRUK0041), the median interval between ctDNA detection and NSCLC relapse confirmed on clinically indicated CT imaging (lead-time) was 70 days (range 10 to 346 days). Four of 13 relapse cases exhibited lead-times of more than six months (Fig. 4a-d). In two cases ctDNA detection preceded CT imaging inconclusive for NSCLC relapse by 157 days (CRUK0004, Fig. 4b) and 163 days (CRUK0045, Fig. 4d). ctDNA profiling reflected adjuvant chemotherapy resistance - CRUK0080, CRUK0004 and CRUK0062 had detectable ctDNA in plasma within 30 days of surgery. The number of detectable SNVs increased in all three cases despite adjuvant chemotherapy, with disease recurring within 1 year of surgery (Fig. 4a-c). In contrast, CRUK0013 had 20 SNVs detectable in ctDNA 72 hours after surgery and 13 SNVs detectable prior to adjuvant chemotherapy; 51 days following completion of adjuvant treatment and at post-operative days 457 and 667 no SNVs were detectable and the patient remains relapse free 688 days post-surgery (Fig. 4b). ctDNA profiling detected intracerebral relapse; CRUK0029 had a PET scan performed 50 days prior to surgery demonstrating normal cerebral appearances. ctDNA remained detectable following surgery, 54 days post-operatively the patient was diagnosed with intracranial metastasis, no extracranial disease was evident on CT imaging (Fig. 4e).

In cases where subclonal SNVs were detected in ctDNA post-operatively we predicted subclonal clusters involved in the relapse process by mapping SNVs detected in plasma back to primary M-Seq data (Fig. 4, Extended Data Fig 5b-c). Subclonal SNVs displaying plasma VAFs similar to clonal SNVs from clusters confined to a single phylogenetic branch, were detected post-operatively in the ctDNA of four patients who suffered NSCLC relapse (CRUK0004, CRUK0063, CRUK0065 and CRUK0044) (Fig. 4b-f, Extended Data Fig 5b). This suggested a relapse process dominated by one subclone represented in our assay-panel. The subclone implicated by ctDNA as driving the relapse in the case of CRUK0004 contained an ERRB2 (HER2)

amplification event (>15 copies, triploid background), that may be targetable in NSCLC (Fig. 4b). Relapses involving subclones from more than one phylogenetic branch were evident in patients CRUK0080, CRUK0062 (Fig. 4a,c) and CRUK0041 (Extended Data Fig 5c).

Validation of phylogenetic characterization
To validate subclonal ctDNA analyses, data acquired from sequencing metastatic tissue at recurrence was integrated with M-seq primary tumor data (for biopsy details, Supplementary Table 3). Patient CRUK0063 suffered para-vertebral relapse of their NSCLC. Post-operative ctDNA analysis revealed the detection of the same subclonal SNV (ORS5D18) on four consecutive occasions over a 231-day period (Extended Data Fig 7a). The OR5D18 SNV traced back to a subclonal cluster private to primary tumor region three (Fig. 5a). CT-guided biopsy tissue was acquired from the para-vertebral metastasis at post-operative day 467. Exome sequencing of relapse tissue revealed the subclonal cluster containing the OR5D18 SNV gave rise to the metastatic subclone (Fig. 5a), this supported ctDNA phylogenetic characterization of relapse. The para-vertebral biopsy contained 88 SNVs not called as present in the primary tumor including an ARID1A stop-gain driver SNP. Re-examination of primary tumor region M-Seq data with a lower SNV calling threshold revealed that 16 of 88 SNVs, including ARID1A, were detectable in primary tumor region three, compared to a maximum of 2 of 88 in other tumor regions (Extended Data Fig 7b). These data suggest that ctDNA profiling can resolve the primary tumor region from which a low frequency metastatic subclone derives. CRUK0035 developed two liver and one adrenal metastases (Fig. 5b). Sequencing of the metastatic liver deposit revealed that only 109 of 149 SNVs classed as clonal in the primary tumor were detectable in the metastasis. This was suggestive of an ancestral branching event not resolved through primary M-seq analysis (Fig. 5b). Post-operative ctDNA profiling identified clonal SNVs present in the liver metastasis biopsy but also revealed SNVs representing a subclone from the primary tumor (Extended Data Fig 7c). This subclone was not present in the metastatic liver deposit (Fig. 5b). These data may reflect ctDNA identified from the non-biopsied metastases suggesting multiple metastatic events. CRUK0044 suffered a vertebral and right hilar relapse. Post-operatively the same subclonal SNV (OR10K1), was detected in ctDNA on two occasions 85 days apart (Extended Data Fig 7d). This SNV was represented in a single subclone detected through sequencing hilar lymph-node metastatic tissue, supporting ctDNA findings (Fig. 5c). CRUK0041 suffered an intracerebral, hilar and subcarinal lymph node relapse. Four subclonal SNVs representing both branches of the tumor phylogenetic tree were detectable in ctDNA at relapse. Consistent with these data, sequencing of subcarinal metastatic tissue revealed the presence of subclonal SNVs mapping to both phylogenetic branches (Fig. 5d, Extended Data Fig 7e). Patient CRUK0013 had detectable ctDNA 3 and 38 days post-operatively. Following adjuvant chemoradiotherapy for lymph-node metastases identified in the pathological specimen, ctDNA levels became undetectable (Fig. 4h). Two involved lymph-nodes were sampled for exome analysis together with M-seq of the primary tumor. Four subclonal SNVs detected in ctDNA post-operatively mapped to an ancestral subclone (describing a subclone that existed during the tumor’s evolution) (Extended Data Fig 7f). This ancestral subclone contained a KRAS amplification (>15 copies, triploid background) and was identified as present in primary tumor and sampled lymph-nodes by M-Seq (Fig. 5e). These data provide phylogenetic characterization of post-operative residual disease that responded to adjuvant chemoradiotherapy (Fig. 4h).
structure of this patient's NSCLC (Fig. 6a). All seven metastatic tumor regions arose from a single ancestral subclone represented by phylogenetic cluster 8. Six metastatic regions shared a later phylogenetic origin, cluster 12 (Fig. 6b). The single tumor region not containing phylogenetic cluster 12 was sampled from the para-aortic metastasis at autopsy and contained a private subclone represented by phylogenetic cluster 9 (Fig. 6b).

We designed a bespoke ctDNA assay-panel to retrospectively track metastatic subclonal burden. 20 clonal SNVs and a median of 8 sub-clonal SNVs (range 4 to 15) in each of 9 metastatic subclonal clusters were targeted by the assay-panel (Extended Data Fig. 8). Since 103 variants per time-point were profiled, SNV detection thresholds were increased to maintain platform specificity (see Methods). This resulted in a predicted false positive rate (FPR) of 0.0011 translating to a 10.7% risk of a single false-positive SNV at each time point and a 0.5% risk of 2 false-positive SNVs at each time point when testing 103 SNVs.

Two clonal SNVs were detected by the 103 SNV assay-panel at day 151 post-surgery (Fig. 6c, Extended Data Fig. 8), 189 days prior to the time point ctDNA was detected using the 19 SNV assay-panel in Fig. 4f. At day 242 a single subclonal SNV was detected from phylogenetic cluster 8 (Fig. 6c, Extended Data Fig. 8), within the context of a 10.7% false-positive risk a single SNV call could represent type I error. At day 466, following clinical-relapse at the thoracic para-vertebral site, 18 of 20 SNVs mapping to phylogenetic clusters (8,11 and 12) were detected in ctDNA, these subclonal clusters were shared between six of seven metastatic sites (Fig. 6b-c, Extended Data Fig. 8). Single SNVs from two private subclones (phylogenetic cluster 5 and 9) were also detectable in ctDNA at day 466 (Fig. 6c, Extended Data Fig. 8). These subclones were not identified in the CT guided para-vertebral biopsy taken at day 467 (Fig. 6b). The mean plasma VAF of the SNVs detected in phylogenetic clusters 11, 8, 12, 9 and 5 mirrored their proximity to the clonal cluster (light blue) in the M-Seq derived phylogenetic tree (Fig. 6a,c). This suggested a tiered burden of subclonal disease concordant with M-Seq phylogenetic inferences. Mean clonal VAF fell in response to palliative radiotherapy and chemotherapy, but at day 767 increased (Fig. 6c). Single SNVs mapping to phylogenetic clusters 5 and 9 and two SNVs mapping to phylogenetic cluster 2 were now detectable in ctDNA 90 days before death (Fig. 6a-c, Extended Data Fig. 8). These three phylogenetic clusters represented subclones private to the para-aortic metastases (Fig. 6. a-b). Consistent with these data significant para-aortic progression was observed at post-mortem compared with most recent CT imaging performed 112 days before death - which showed no evidence of para-aortic disease.

**Discussion**

In summary, we find predictors of ctDNA detection in early-stage NSCLC characterized by non-adeno carcinoma histology, necrosis, increased proliferative indices and lymphovascular invasion. Triple negative breast cancers display necrosis19, high proliferative indices20,21 and are associated with increased ctDNA levels compared with other breast cancer subtypes suggesting extension of these observations beyond NSCLC.

Tumor volume correlated with plasma ctDNA VAF (Fig. 3a). A primary NSCLC tumor volume of 10 cm3 predicted a ctDNA plasma VAF of 0.1%. The VAF conferring optimum sensitivity for most ctDNA platforms. Low-dose CT lung screening can identify lung nodules with diameters from 4 mm24,25 and are associated with increased ctDNA levels compared with other breast cancer subtypes suggesting extension of these observations beyond NSCLC.

A limitation to targeted ctDNA profiling is cost, estimated at $1750 per patient for sequencing a single tumor region, synthesis of a patient-specific assay-panel and profiling of five plasma samples. Adjuvant platinum-based chemotherapy in NSCLC improves cure rates following surgery in only 5% of patients and 20% patients receiving chemotherapy experience acute toxicities3. There is a need to increase adjuvant therapy efficacy and better target its use. Bespoke ctDNA profiling can characterize the subclonal dynamics of relapsing NSCLC and identify adjuvant chemotherapy resistance. These findings indicate that drug development guided by ctDNA platforms to identify residual disease, define adjuvant treatment response and target emerging subclones prior to clinical recurrence in NSCLC, with appropriate CLIA validation, are now feasible.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 27 January; accepted 13 April 2017.**

**Published online 26 April 2017.**

---

1. Jamal A, Bray F, Center MM, Ferlay J, Ward E & Forman D. Global cancer statistics. CA: A Cancer Journal for Clinicians 2011;61(2):69–90.

2. Editorial RL, Miller KD & Jemal A. Cancer Statistics. 2017: CA: A Cancer Journal for Clinicians 2017;67(1):7–30.

3. Pignon J-P, Tribodet H, Scagliotti GV, Douillard J-Y, Shepherd FA, Stephens RJ, et al. Lung Adjuvant Cisplatin Evaluation: A Pooled Analysis by the LACE Collaborative Group. Journal of Clinical Oncology 2008;26(21):3592–9.

4. Lahveau D, A Carter SC, Stojanov P, McKenna A, Stevenson C, Lawrence MS, et al. Evolution and Impact of Subclonal Mutations in Chronic Lymphocytic Leukemia. Cell 2015;162(4):714–26.

5. Beaverson, J. et al. Detection of ubiquitous and heterogeneous mutations in cell-free DNA from patients with early-stage non-small-cell lung cancer. Cancer Medicine 2015;7(10):3326–33.

6. Tj, J., Ng, W., Li, C., Li, S., et al. Tracking circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. Science Translational Medicine 2015;7(302):302ra346-ra352.

7. Jamal-Hanjani M, Hackshaw A, Ngai Y, Shaw J, Dive C, Quezada S, et al. Tracking genomic cancer evolution for precision medicine: the lung TRACERx study. PloS Biol 2014;12(7):e1001906.

8. Jamal-Hanjani M, Wilson GA, McGranahan N, Birdkajj NK, Watkins TBK, Veeriss S, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. New England Journal of Medicine (accepted, in press) 2017; DOI: (0.1056/NEJMoa1616288).

9. Jamal-Hanjani M, Wilson GA, Horswell S, Mitter R, Sakarya O, Constantin T, et al. Detection of ubiquitous and heterogeneous mutations in cell-free DNA from patients with early-stage non-small-cell lung cancer. Annals of Oncology 2016;27(5):862–7.

10. Jr L. & Bardelli A. Liquid Biopsies: Genotyping Circulating Tumor DNA. Journal of Clinical Oncology 2014;32(6):579–86.

11. Caruso R, Parisi A, Bonanno A, Paparao D, Quattrocchi E, Branca G, et al. Histologic coagulative tumor necrosis as a prognostic indicator of aggressiveness in renal, lung, thyroid and colorectal carcinomas: A brief review. Oncology Letters 2012;3(1):16–8.

12. Vesselle H, Schmidt RA, Pugsley JM, Li M, Kohlymer SY, Vallières E, et al. Lung Cancer Proliferation Correlates with [F-18]Fluorodeoxyglucose Uptake by Positron Emission Tomography. Clinical Cancer Research 2000;6(10):3837–44.

13. Higashi K, Ueda Y, Yagishita M & Ariasaka Y. FDG PET measurement of the proliferative potential of non-small cell lung cancer. The Journal of Nuclear Medicine 2000;41(1):85.

14. Murata M, Dawson SJ, Pogrebinski K, Rueda OM, Provenzano E, Grant J, et al. Multifocal clonal evolution characterized using circulating tumor DNA in a case of metastatic breast cancer. Nature Communications 2015;6:8760.

15. Newman AM, Bratman SV, To J, Wynne JF, Eclov NCW, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad clinical coverage. Nat Med 2014;20(5):548–54.

16. Del Monte U. Does the cell number 109 still r...

17. Peters S & Zimmermann S. Targeted therapy in NSCLC driven by HER2 insertions. Translational Lung Cancer Research 2014;3(2):84–8.

18. Livas CA, Karaca G, Nanda R, Tretiakova MS, Olopade OI, Moore DT, et al. Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. Mod Pathol 2005;18(2):264–71.

19. Keam B, Im S-A, Kim H-J, Oh D-Y, Kim JH, Lee S-H, et al. Prognostic impact of clinicopathologic parameters in stage II/III breast cancer treated with neoadjuvant docetaxel and doxorubicin chemotherapy: paradoxical features of the triple negative breast cancer. BMC Cancer 2007;7:203.

20. Rhee J, Han SW, Oh DY, Kim JH, Im SA, Han W, et al. The clinicopathologic characteristics and prognostic significance of triple-negative in node-negative breast cancer. BMC Cancer 2008;8:307.
22. Team TNLSTR. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. *New England Journal of Medicine* 2011;365(5):395–402.

23. Newman AM, Lovejoy AF, Klass DM, Kurtz DM, Chabon JJ, Scherer F, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat Biotech* 2016;34(5):547–55.

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We dedicate this manuscript to the memory of Roberto Macina. We thank Samantha Navarro and Antony Tin for facilitating the PEACE ctDNA analysis. We thank the members of the TRACERx and PEACE consortia for participating in this study. C.S. is Royal Society Napier Research Professor. This work is supported by the Francis Crick Institute which receives its core funding from Cancer Research UK (FC001169,FC001202), the UK Medical Research Council (FC001169,FC001202), and the Wellcome Trust (FC001169,FC001202). CS is funded by Cancer Research UK (TRACERx and CRUK Cancer Immunotherapy Catalyst Network), the CRUK Lung Cancer Centre of Excellence, Stand Up 2 Cancer (SU2C), the Rosetrees Trust, Novo Nordisk Foundation (ID 16584), the Prostate Cancer Foundation, the Breast Cancer Research Foundation, the European Research Council (THESEUS) and Support was provided to CS by the National Institute for Health Research, the University College London Hospitals Biomedical Research Centre and the Cancer Research UK University College London Experimental Cancer Medicine Centre. PV.L. is a Winton Group Leader in recognition of the Winton Charitable Foundation’s support towards the establishment of the Francis Crick Institute. The TRACERx (Tracking Non-small Cell Lung Cancer Evolution) study (Clinicaltrials.gov no: NCT01888601) is sponsored by University College London (UCL/12/0279) and has been approved by an independent Research Ethics Committee (13/LO/1546). TRACER is funded by Cancer Research UK (grant number C11496/A17786) and coordinated through the Cancer Research UK & UCL Cancer Trials Centre. The PEACE (Posthumous Evaluation of Advanced Cancer Environment) Study (Clinicaltrials.gov no: NCT03004755), is sponsored by University College London (UCL/13/0165) has been approved by an independent Research Ethics Committee (13/LO/0972). PEACE is funded by Cancer Research UK (C416/A21999) and coordinated through the Cancer Research UK & UCL Cancer Trials Centre.

**Author Contributions** CA., N.J.B., G.A.W., M.J.-H., T.C., R.S., and J.L-Q. contributed equally to this work. C.A. and C.S. co-wrote the manuscript. C.A., M.J.H., and C.S. conceived study design. C.A., N.J.B., G.A.W. and R.R. integrated clinicopathological data, exome data and ctDNA data. M.R. B.G.Z, C.J.L., T.C., R.S., E.K., N.S., D.H., A.N. and A.P., conducted and analysed multiplex-PCR NGS experimental work. N.J.B, G.A.W, T.B.KW, M.A.B, R.R., and N.M. conducted M-Seq analyses of exome data. J.L-Q, T.M. and D.A.M. conducted pathological review. F.F., R.E. and F.Z. conducted radiological review of PET scans. H.W.L.A, W.U.I, F.M. and N.J.B. conducted radiomic analyses. S.V., D.J., J.L., S.S., J.C., K.A.R., T.C., D.O. and A.U.A. conducted TRACERx sample processing. G.E., S.W., N.M. and G.A.W. conducted exome sequencing. L.M., J.R. and J.S. conducted DNA cross-platform validation. M.H., C.D., J.S., and C.S. designed study protocols. C.H., S.L.M., M.D.F., T.A., M.Fa., E.B., D.L., M.H., S.K., N.P., S.M.J., R.T., A.A., F.B., Y.S., R.S., I.J.A., M.Q., P.C., B.N., G.M., G.L., S.T., M.N., H.R., K.K., M.G., L.G., D.A.F., A.N., S.R., S.A., S.K., P.R., V.E., B.I., M.I.-S., V.P., J.F.I., M.K., R.A., H.A., H.D.I., S.L. are clinical members of TRACERx study sites. J.H. and H.L ran the UCL GCLP facility. A.H., H.B., N.J. and Y.N. were involved in study oversight. J.A.S., J.L-Q., Z.S., E.G., S.K., S.T., M.A.B, R.F.S., J.H., A.S., S.Q., P.V.L., C.D. and C.J.L. gave advice and reviewed the manuscript. A.H. gave statistical advice. C.S. provided overall study oversight.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to C.S. (Charles.Swanton@crick.ac.uk).

**Reviewer Information** Nature thanks S. Lippman, R. Rosell and the other anonymous reviewers for their contribution to the peer review of this work.
Figure 1 | Phylogenetic ctDNA tracking. Overview of the study methodology. Multi-region sequencing of NSCLC was performed as part of the TRACERx study. PCR assay-panels were designed based on phylogenetic analysis, targeting clonal and subclonal single nucleotide variants to facilitate non-invasive tracking of the patient-specific tumor phylogeny. Assay-panels were combined into multiplex assay-pools containing primers from up to 10 patients. Cell-free DNA was extracted from pre- and post-operative plasma samples and multiplex-PCR performed, followed by sequencing of amplicons. Findings were integrated with M-Seq exome data to track tumor evolution.
Figure 2 | Clinicopathological predictors of ctDNA detection.
a) Heatmap showing clinicopathological and ctDNA detection data, continuous variables quartiled. Raw data and patient IDs in attached worksheet. b) Detection of clonal and subclonal single nucleotide variants within 46 patients with two or more single nucleotide variants detected in plasma. Histology indicated in panels as LUSC, LUAD and Other.
Figure 3 | Tumor volume predicts plasma variant allele frequency. 

a) Tumor volume (cm$^3$) measured by CT volumetric analysis correlates with mean clonal plasma VAF, $n = 38$, grey vertical lines represent range of clonal VAF; red shading indicates 95% confidence intervals. 
b) Predicted mean clonal VAF at hypothetical volumes ranging from 1 to 100 cm$^3$ based on model in panel a, predicted cancer cell number based on model in extended data 4e. 
c) Estimated effective subclone size, defined as mean CCF of subclone multiplied by tumor volume and purity, influences subclonal SNV detection. For negative calls, median effective subclone size was $1.70 \text{ cm}^3$, range $= 0.21-24.11$, $n = 163$, for positive calls, median effective subclone size $= 4.06 \text{ cm}^3$, range $= 0.31-49.20$, $n = 109$. Wilcoxon rank sum test, $P < 0.001$, data from 34 patients (passed volumetric filters with subclonal SNVs represented in assay-panel). 
d) Estimated effective subclone size correlates with subclonal plasma VAF, $n = 109$ subclonal SNVs, data from 34 patients (passed volumetric filters with detected subclonal SNVs in plasma).
Figure 4 | Post-operative ctDNA detection predicts and characterizes NSCLC relapse. a-h) Longitudinal cell-free DNA profiling. Circulating tumor DNA (ctDNA) detection in plasma was defined as the detection of two tumor-specific SNVs. Detected clonal (circles, light blue) and subclonal (triangles, colors indicates different subclones) SNVs from each patient-specific assay-panel are plotted on graphs colored by M-Seq derived tumor phylogenetic nodes. Mean clonal (blue) and mean subclonal (red) plasma VAF are indicated on graphs as connected lines. Pre-operative and relapse M-Seq derived phylogenetic trees represented by ctDNA are illustrated above each graph.
Figure 5 | Phylogenetic trees incorporating relapse tissue sequencing data. Phylogenetic trees based on mutations found in primary and metastatic tissue (a–d), or primary tumor and lymph node biopsies (e). Colored nodes in phylogenetic trees indicate cancer clones harboring mutations assayed for in ctDNA, grey indicates a clone not assayed. Branch length is proportional to number of mutations unless crossed. Dashed red lines show branches leading to metastatic relapse. Colored bars below show the number of assays per sample detected preoperatively and at relapse (a–d) or in the absence of relapse, post-surgery (e). Thin colored bar shows number of assays in total. Colors match clones on the phylogenetic trees.
**Figure 6 | ctDNA tracking of lethal cancer subclones in CRUK0063.**
Phylogenetic analysis of one relapse biopsy (day 467) and five metastatic biopsies (post mortem) a) To-scale phylogenetic tree of CRUK0063 including M-seq based on metastatic and primary tumor regions. Branch length is proportional to number of mutations in each subclone. b) Phylogenetic trees matching metastatic lesions, colored nodes represent mutation clusters found at each site and assayed for in ctDNA. Open circles represent mutation clusters not detected in ctDNA. c) Tracking plot showing mean VAF of identified mutation clusters in ctDNA. Size of dots indicates number of assays detected. Colors correspond to mutation clusters and match panels a) and b).
METHODS

Patients and samples. The cohort of 100 patients evaluated within this study comprises the first 100 patients prospectively analyzed by the lung TRACERx study (Clinicaltrials.gov no: NCT01888601, approved by an independent Research Ethics Committee, 13/LO/1546) and mirrors the prospective 100 patient cohort by Jamal-Hanjani M et al. 17

Multi-region tumor sampling was performed as described. 18 Relapse tissue samples, excess to diagnostic requirements, were acquired via clinical procedures described in Supplementary Table 3. For patient CRUK0063 post mortem examination was performed through the PEACE study 24 hours following death (Clinicaltrials.gov no: NCT03004755, approved by an independent Research Ethics Committee, 13/LO/0972). Consent was obtained from all subjects for procedures conducted in these studies.

Tissue microarray creation and Ki67 immunohistochemistry. Tissue microarrays (TMAs) were created of 100 NSCLC cases for K67 + immunohistochemistry. Representative primary tumor areas were defined by examination of H&E stained sections from TRACERx cases. Two 2 mm cores were selected from different regions within each specimen and re-embedded in recipient blocks. This resulted in a TMA of 200 cores with four normal lung cores as negative control. 2-5 μm sections from tissue-microarrays containing tumor were cut. Immunohistochemistry with anti-Ki67 monoclonal antibody (Dilution 1:100; clone MB1-1; DAKO Agilent Technologies LDA, UK Limited, Stockport, Cheshire SK8 3GB, UK) was performed using Benchmark Ultra (Ventana/Roche). The percentage of Ki67 positive cells were averaged across two tumor sections for each case. Detection was performed using the peroxidase-based detection reagent conjugate (OptiView DAB IHC Detection Kit; Ventana Medical Systems, Inc).

Central histopathological review. Digital images of diagnostic tumor sections from all cases were reviewed in detail centrally by at least one pathologist, and in cases of uncertainty, by two. Histological images, percentage of necrosis and the presence of lymphovascular invasion were all evaluated on digital images from scanned diagnostic slides blinded to the ctDNA detection status of the patient in question.

Central radiology review & volume estimation. 92 of 96 anonymized diagnostic PET-CT were reviewed by a Nuclear Medicine Physician, blinded to the initial PET-CT reports. Scan images were not available in three cases (CRUK0025, CRUK0039 and CRUK0023) and in one case a pre-operative PET-CT was not performed (CRUK0008). CT and PET images were matched and fused into transaxial, coronal and sagittal images and reviewed on a dedicated PET/CT software platform (AW 4.1/4.2 GE medical systems). The semi-quantitative parameter Standardized Uptake Value (SUVR) was calculated for the primary tumor mass was calculated and recorded along with SUVmax of mediastinal background uptake. Tumor-to-background ratio (TBR) was calculated based on SUVmax of the tumor divided by mediastinal background uptake. 24,25 Tumor volume was determined based on tumor CT scans. CT slices of the primary tumor were measured with 3D Slicer applying the “lung algorithm window” settings. Tumor contours were segmented on each axial CT slice. These steps were performed by an experienced radiologist (W.L.B.), and all contours were confirmed and edited where necessary, by a radiologist with 14 years of experience in cancer imaging (E.M.F.). Effective tumor volume was defined as tumor volume multiplied by the mean purity of the tumor based on M-seq, purity estimates derived from ASCAT analysis as described. 26 Effective subclone size was defined as a mean cancer cell fraction (CCF) across the regions of the mutation cluster multiplied by tumor volume and mean tumor purity.

Tissue exome sequencing and processing. Whole exome sequencing was performed on DNA purified from tumor tissue and normal blood as described, 26 with the exception of CRUK0063 BR_T1-R1. This capture was performed according to the manufacturer’s 200 ng DNA protocol (Agilent). Annotated SNV calls from primary tumors are available in Jamal-Hanjani et al. 17 For this study, metastatic tissue was obtained from four patients (CRUK0044, CRUK0063) and six metastatic samples acquired at post mortem examination of CRUK0063 were obtained. Genomic DNA was purified from all tissue samples, and processed through the TRACERx bioinformatics pipeline as described. 26 Annotated SNV calls are available in Supplementary Table 4.

cDNA library preparation. Forty ul of cfDNA from each plasma sample, which is present as fragments of mononucleosomal and polynucleosomal length, was used as input into Library Prep using the Natera Library Prep kit; in two samples with extremely high cfDNA amounts input was restricted to ~50,000 genome equivalents (165 ng). cfDNA was end repaired and A-tailed. Natera custom adapters were ligated. The libraries were amplified for 15 cycles to plateau and then purified using Ampure beads following the manufacturer’s protocol. The purified libraries were run on the LabChip. Successful libraries had a single peak at ~250 bp.

SNV assay design & optimization. Natera’s standard assay design pipeline was used to generate forward and reverse PCR primers for somatic SNVs detected in tumor samples. For each pair of primers, the probability of forming primer-dimer was calculated and assays were combined into pools such that any primer combination in a pool is not predicted to form primer-dimers. For each patient, assays were prioritized such that, 1) assays covering driver SNVs had highest priority, and 2) there was uniform sampling of the phylogenetic tree. For the baseline cohort, 10 balanced pools were created, each containing the on average 18 assays for 10 patients’ SNVs. For the longitudinal cohort, up to 10 extra assays were generated for adenocarcinoma samples. For patient CRUK0003 post mortem analysis, new assays were designed based on M-seq of metastatic biopsy retrieved at day 467 and of metastatic lesions harvested post mortem. A total of 103 new assays were designed compared to 19 based on the primary tumor alone.

SNV assays were ordered from IDT (Corvallis, OR). Each pool was optimized by running the multiplex PCR and sequencing protocol using one plasma cfDNA library from a healthy subject. For optimization, PCR parameters (primer concentration and annealing temperature) that yielded the best percentage of on-target reads, depth-of-read uniformity (measured as the ratio of the 80th percentile/20th percentile), and number of drop-out assays (defined as assays with <1,000 reads) were determined by sequencing. The PCR conditions that yield the best percentage of on-target reads, depth-of-read uniformity, and the lowest number of drop-outs were determined. For all pools, the optimal conditions were 10 mM primers and 60°C or 62.5°C annealing temperature. Primer pairs contributing to primer formation were removed from each final pool.

Analytical validation. Synthetic spikes representing twenty SNVs randomly selected from Pool 1 were designed and synthesized (IDT, Corvallis, IA) as 160 bp oligos with the respective SNV placed in the middle (position 80). These synthetic spikes were mixed at equimolar ratios and used to prepare a library. This library was titrated into a library prepared from mono-nucleosomal DNA (10,000 copies) from a normal cell line (AG16778 from Coriell, Camden, NJ). The library of 20 synthetic spikes was titrated into the mono-nucleosomal DNA library at 2.5%, 0.5%, 0.25%, 0.1%, 0.05% and 0% (each in triplicate), and 0.01%, 0.005% and 0.001% (each in quadruplicate). Because preparing spiked samples at such low levels is either subject to sampling noise (0.01% spikes into 10,000 genomic copies background is equivalent to one mutant copy), or is not possible (at levels less than 0.01%), samples were mixed as libraries. Following library mixing and sequencing, data was analyzed to detect all the targets in Pool 1 using the same parameters as used for the patient samples.

The measured VAF of each spike for the samples with 2.5% nominal input was used to calculate an input correction factor (measured VAF/2.5%). This correction factor was applied to the other inputs of the corresponding spike titration series. The measured VAF differed from the nominal input most likely because the mononucleosomal fragmentation pattern is not entirely random. Because of this, the actual input levels differ from the nominal input levels. Therefore, analytical sensitivity and specificity were measured based on corrected input intervals (see Extended Data Fig. 1a).

Plasma SNV mPCR-NGS workflow. The library material from each plasma sample was used as input into multiplex PCR using the relevant assay pool and an optimized plasma mPCR protocol. Optimal mPCR conditions were as described. 26,27 Each PCR assay pool was used to amplify the SNV targets from the 10 corresponding samples and 20 negative control samples (plasma libraries prepared from healthy subjects; BioMed IRB #601-01 and E&I West Coast Board IRB00007807, Study #13090-01A and 13090-04A). The mPCR products were barcoded in a separate PCR step. Each amplicon pool was sequenced on one Illumina HiSeq 2500 Rapid Run with 50 cycles paired-end reads using the Illumina Paired End v1 kit with an average target DOR of ~40,000 per assay.

Bioinformatics Pipeline. All the paired-end reads were merged using Pear. 6 Merged reads were mapped to the hg19 reference genome with Novoalign v2.3.4 (http://www.novocraft.com/) and sorted and indexed using SAMtools. 28 Bases that do not have a forward or reverse reads or that have Phred quality score <20 were filtered out to minimize sequencing errors in subsequent steps. Merged reads with mapping quality >30 and at most one mismatch under the sequence of primers were marked as on-target. Targets with <1000 reads were considered failed and were filtered from further analyses. Quality control was performed using
an in-house program checking for a wide list of statistics per sample that included total numbers of reads, mapped reads, on-target reads, number of failed targets, and average error rate.

**Plasma SNV calling algorithm.** For each target SNV a position-specific error-model was built (see Supplementary Methods). Samples with high plasma VAF (>20%) among the putative negatives were considered to have possible germline mutations and were excluded from the error model. A confidence score was calculated for each target SNV based on the error model and a positive plasma SNV call was made if the confidence score passed a threshold of 95% for transitions and 98% for transversions (see Supplementary Methods). There was no difference in depth of read between called and not called SNVs (Extended Data Fig 1c).

Because the post mortem analysis of CRUK0063 involved a larger number of target SNVs per time point being analyzed (103 vs. 19 targets in previous samples) updated calling thresholds were applied to control for false positives. The new updated thresholds were chosen such that the average number of false positives in the 30 negative samples in the run becomes ~1 per sample. All multiplex PCR-NGS ctDNA SNV assays with confidence score data are available in Supplementary Table 5 (Baseline, pre-operative cohort assays), Supplementary Table 6 (Longitudinal Assays), and Supplementary Table 7 (Extended Longitudinal Assays for CRUK0063).

**Cross-platform validation using generic PCR-NGS panel section.** Cross-platform validation was performed in 28 patients with M-Seq confirmed SNV(s) within one or more hotspots targeted by a generic multiplex PCR-NGS panel (Extended Table 2a-b, Supplementary Table 8). 20ng of isolated ctDNA was used for library preparation using the Oncomine™ Lung ctDNA Assay (ThermoFisher Scientific), according to the manufacturer’s instructions. Automated template preparation and chip loading was conducted on the Ion Chef™ instrument using the Ion 520™ & Ion 530™ Kit-Chef (ThermoFisher Scientific). Ultimately, samples were sequenced on Ion 530™ chips using the Ion S5™ System (ThermoFisher Scientific). Sequencing data was accessed on the Torrent suite v5.2.2. Reads were aligned against the human genome (hg19) using Alignment v4.0-r77897, and variants were called using the coverage Analysis v4.0-r77897 plugin. All 18 bespoke-panel ctDNA negative patients had no tumor SNVs detectable in plasma pre-operatively by the generic panel supporting biological specificity of the bespoke targeted approach. 7 negative patients had no tumor SNVs detectable in plasma post-operatively and as a second independent analysis, with tumor purity.

**Phenoma Archive (EGA), which is hosted by the The European Bioinformatics Institute (EBI) and the Centre for Genomic Regulation (CRG), under accession numbers EGAS00001002247 (primary tumor data) and EGAS00001002415 (metastatic tumor data). Further information about EGA can be found on https://ega-archive.org. “The European Genome-Phenome Archive of human data consented for biomedical research.”

24. Hotheinf R, Butof R, Apostolova I, Zohpfe K, Steffen IG, Amthauer H, et al. An investigation of the relationship between tumor-to-liver ratio (TLR) and tumor-to-blood standard uptake ratio (SUR) in oncological FDG PET. EJNMMI Res 2016; 6(1):19.
25. Butof R, Hotheinf R, Zohpfe K, Stadelmann T, Schmolack J, Jentsch C, et al. Prognostic Value of Pretherapeutic Tumor-to-Blood Standardized Uptake Ratio in Patients with Esophageal Carcinoma. J Nucl Med 2015;56(8):1150–6.
26. Zhang J, Klobert K, Foun A, Stambakas A, PEAR: a fast and accurate Illumina Paired-End read merge. Bioinformatics 2014;30(5):614–20.
27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Hardman N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009; 25(16):2078–9.
28. Matichuk S, McPherson AW, Donmez N & Sahinalp CS. Clonality inference in multiple tumor samples using phylogeny. Bioinformatics 2015; 31(9): 1349–56.
Extended Data Figure 1 | Multiplex-PCR next-generation sequencing platform analytical validation. a) Analytical validation of the multiplex-PCR NGS platform was performed by spiking synthetic single nucleotide variants into control cell-free DNA. Sensitivity and specificity of the platform at different spike concentrations was ascertained, 95% binomial confidence interval displayed as error bars. b) Specificity of ctDNA detection based on a 1 SNV and 2 SNV call threshold taking into account parallel testing of multiple SNVs. c) The median depth of read across a position did not vary depending on whether an SNV position was called or not called using the platform error-model. Wilcoxon Test, $P = 0.786$, median depth of read at uncalled positions = 45,777 ($n = 3,745$), range: 0 to 146774, median depth of read at called positions = 45,478, range = 1,354 to 152,974 ($n = 1,124$). Whiskers represent 1.5 times the interquartile range, 2-sided test.
a) The pre-operative study phase cohort consisted of 100 TRACERx patients present in the first 100 patient TRACERx cohort in April 2016. Pre-operative plasma samples were profiled in 96 patients for reasons listed. bii) Contents of patient-specific assay-panels designed in the pre-operative study phase. c) The longitudinal study phase cohort consisted of patients with confirmed NSCLC relapse and patients without relapse. d) Contents of patient-specific assay-panels designed in the longitudinal phases of this study. e) Single nucleotide variant type targeted.
Extended Data Figure 3 | Clinicopathological predictors of ctDNA detection. a) 96 patients in pre-operative cohort stratified by pathological TNM stage. b) LUSCs and ctDNA positive LUADs are significantly more necrotic than ctDNA negative LUADs. Significant differences in necrosis between groups: LUSCs (median necrosis 40%) ( n = 31), ctDNA positive LUADs (median necrosis 15%) ( n = 11) and ctDNA negative LUADs (median necrosis 2%) ( n = 47), Kruskal-Wallis test, P < 0.001, 2-sided pairwise comparisons were performed using Dunn’s procedure with Bonferroni correction. c) Univariate (left) and multivariate analyses (right) were performed, by logistic regression to determine significant predictors of ctDNA detection in early-stage NSCLC. ctDNA detection was defined as detection of two or more SNVs in pre-operative plasma samples. Details regarding multivariable analysis methodology are in methods. d) Receiver operating characteristic curve (ROC) analysis of pre-operative PET scan FDG-avidity (normalized as tumor background ratio (TBR), see methods), as a predictor of ctDNA detection (92/96 PET scans were available for central review). Median PET TBR of detected tumors = 9.01, n = 45. Median PET TBR of undetected tumors = 3.64, n = 47. P-value based on Wilcoxon Rank Sum Test. e) LUAD subtype analyses based on ctDNA detection and the presence of an EGFR, KRAS or TP53 driver mutation.
Extended Data Figure 4 | Predictors of plasma variant allele frequency. a) Plasma variant allele frequencies of SNVs detected in plasma in 46 patients who were ctDNA positive (two or more SNVs detected). Clonal (blue) and subclonal (red) variant allele frequencies indicated, mean shown as horizontal line. Driver variants shown as triangles. b) Mean clonal VAF correlated with maximum tumor size measured in post-surgical specimen (pathological size, n = 46) grey vertical bars represent range of clonal variant allele frequency. Shaded red background indicates 95% confidence interval. c) Filtering steps taken to define a group of ctDNA positive patients with volumetric data considered adequate to model tumor volume and plasma variant allele frequency. d) Scatter plot showing mean clonal VAF relative to tumor volume for TRACERx (blue dots and fitted blue line, n = 38) and VAF relative to volume for previously published data based on CAPP-seq analysis of ctDNA (orange dots and orange fitted line, n = 9). Orange shaded background indicates 95% confidence interval based on CAPP-seq data. e) Mean clonal VAF correlated with tumor volume × tumor purity (cancer cell volume), n = 38. Shaded red background indicates 95% confidence interval. f) Association between number of cancer cells and VAF of clonal SNVs in plasma based on linear modelling of Extended Data Fig 4f. g) Detected subclonal SNVs were mapped back to M-Seq derived tumor phylogenetic trees (process illustrated in graphic). Detected private subclones (subclones identified within only a single tumor region) are coloured red. Shared subclones (subclones detected in more than one tumor regions) are light blue. Shared subclones were mapped back to M-Seq derived tumor phylogenetic trees (process illustrated in graphic).
Extended Data Figure 5 | Longitudinal ctDNA profiling, remaining relapse cases. a) Kaplan-Meier curve demonstrate relapse free survival for patients in whom ctDNA was detected versus patients in whom ctDNA was not detected. b-h) Longitudinal cell-free DNA profiling. Circulating tumor DNA (ctDNA) detection in plasma was defined as the detection of two tumor-specific SNVs. Relapse was based on imaging-confirmed NSCLC relapse, imaging performed as clinically indicated. Detected clonal (circles, light blue) and subclonal (triangles, colors indicates different subclones) SNVs from each patient-specific assay-panel are plotted on graphs colored by M-Seq derived tumor phylogenetic nodes. Mean clonal (blue) and mean subclonal (red) VAF are indicated on graphs. Pre-operative and relapse M-Seq derived phylogenetic trees represented by ctDNA are illustrated above each graph in cases where subclonal SNVs were detected.
Extended Data Figure 6 | Longitudinal ctDNA profiling, non-relapse cases. a–j) Detected clonal (circles, light blue) and subclonal (red triangles) SNVs from each patient-specific assay panel are plotted on graphs. Mean clonal (blue) and mean subclonal (red) VAF are indicated on graphs.
Extended Data Figure 7 | Heatmaps illustrating detection of SNVs in bespoke panel at each sampled time point. a-c.f Bespoke assay panels for CRUK0063, CRUK0035, CRUK0044, CRUK0041 and CRUK0013. Colors indicate originating subclonal cluster based on the phylogenetic trees above the heatmap. Light blue indicates clonal mutation cluster. Full panel with cluster color shown below each heatmap. Filled squares indicates detection of a given variant in plasma ctDNA. Y-axis shows day of sampling. y-axis labels appended with [R] indicates day of clinical relapse. b) Re-examination of primary tumor regions from CRUK0063 with lowered threshold to potentially identify SNVs private to the sequenced relapse biopsy. 16/88 variants were found at very low VAF in region 3, indicating this region from the primary likely gave rise to the metastasis.
Extended Data Figure 8 | Heatmap illustrating detection of SNVs in bespoke panel based on M-seq of metastatic tumor regions for patient CRUK0063 for all sampled time points. Colors indicate originating subclonal cluster based on the phylogenetic trees above the heatmap. Light blue indicates clonal mutation cluster. Full panel with cluster color shown below each heatmap. Filled squares indicates detection of a given variant in plasma ctDNA. Y-axis shows day of sampling.
Extended Data Table 1 | Patient characteristics

a) Clinical characteristics
96 patient pre-operative cohort

| Characteristic | Total |
|----------------|-------|
| Age | <60 | 19 |
|  ≥60 | 77 |
| Sex | Male | 60 |
| Female | 36 |
| ECOG PS | 0 | 49 |
| ≥1 | 1 |
| Histology | Adenocarcinoma | 58 |
| Squamous cell carcinoma | 31 |
| Carcinosarcoma | 2 |
| Large cell carcinoma | 1 |
| Adenosquamous carcinoma | 3 |
| Large cell neuroendocrine carcinoma | 1 |
| TNM stage | Ia | 24 |
| Ib | 35 |
| Ila | 12 |
| Iib | 11 |
| IIa | 13 |
| IIb | 1 |
| Lymph node metastasis | Yes | 24 |
| No | 72 |
| Pleural involvement | Yes | 27 |
| No | 69 |
| Vascular invasion | Yes | 41 |
| No | 55 |
| Resection margin | R0 | 91 |
| R1 | 5 |
| Smoking status | Never smoked | 31 |
| Recent ex-smoker | 30 |
| Ex-smoker | 48 |
| Current smoker | 7 |
| Ethnicity | White British | 85 |
| White-other | 4 |
| White-Irish | 4 |
| Caribbean | 3 |

b) No adjuvant therapy Adjuvant therapy

| TNM Stage | 1a | 24 | 0 |
|           | Ib | 31 | 4 |
|           | Ila | 3 | 9 |
|           | Iib | 4 | 7 |
|           | IIa | 6 | 7 |
|           | IIb | 0 | 1 |

c) Details regarding timing of pre-operative blood sample

| Days pre-surgery | Number | Details |
|------------------|--------|---------|
| Within 24 hours  | 91     | CRUK0051, 0096 |
| 24-72 hours      | 2      | CRUK0073, 0096 |
| 8 days           | 2      | CRUK0089 |
| 31 days          | 1      | CRUK0089 |

d) Clinical characteristics
24 patient longitudinal sub-cohort

| Characteristic | Total |
|----------------|-------|
| Age | <60 | 5 |
|  ≥60 | 19 |
| Sex | Male | 16 |
| Female | 8 |
| ECOG PS | 0 | 12 |
| ≥1 | 12 |
| Histology | Adenocarcinoma | 16 |
| Squamous cell carcinoma | 8 |
| TNM stage | Ia | 3 |
| Ib | 7 |
| Ila | 3 |
| Iib | 7 |
| IIa | 3 |
| IIb | 1 |
| Lymph node metastasis | Yes | 9 |
| No | 15 |
| Pleural involvement | Yes | 7 |
| No | 17 |
| Vascular invasion | Yes | 12 |
| No | 12 |
| Resection margin | R0 | 23 |
| R1 | 1 |
| Smoking status | Never smoked | 1 |
| Recent ex-smoker | 5 |
| Ex-smoker | 16 |
| Current smoker | 2 |
| Ethnicity | White British | 21 |
| White-other | 2 |
| Caribbean | 1 |

e) No adjuvant therapy Adjuvant therapy

| TNM Stage | 1a | 3 | 0 |
|           | Ib | 6 | 1 |
|           | Ila | 0 | 3 |
|           | Iib | 2 | 5 |
|           | IIa | 1 | 2 |
|           | IIb | 0 | 1 |

a) Table of clinical characteristics describing the 96 patient pre-operative cohort and b) demonstrating distribution of stage and whether the patient received adjuvant chemotherapy, c) demonstrating the time-points at which pre-operative plasma was acquired for patients within the cohort, d) table of clinical characteristics describing 24 patient longitudinal cohort and e) demonstrating distribution of stage in the longitudinal cohort and whether the patient received adjuvant chemotherapy.
90% sensitivity at 0.1% VAF and above

Extended Data Table 2 | Cross platform validation using a generic approach to ctDNA profiling

Table 2b - Bespoke panel non-detected NSCLCs - cross platform validation

a) 7/10 (70%) of bespoke-panel ctDNA positive patients had tumor SNVs detectable in plasma preoperatively by a generic hotspot PCR-NGS lung panel (Lung Oncomine, Thermofisher). The three bespoke-panel ctDNA positive patients undetected by the generic panel had mean clonal plasma variant allele frequencies lower than the 0.1% plasma variant allele frequency (VAF) limit of detection reported for the generic panel (shaded yellow). b) Based on CT volumetric assessment of each patient’s primary tumor we predicted plasma VAF corresponding to a tumor of that size (see Fig. 3 and Methods for details of approach). This allowed us to infer platform sensitivities for each patient within the bespoke-panel non-detected cohort. Five LUADs (shaded green, CRUK0037, CRUK0051, CRUK0004, CRUK0039 and CRUK0025) had tumor volumes approximating to a plasma VAF of more than 0.1%. This suggested that these tumors resided within the top platform sensitivity bracket of both the generic and bespoke-panel ctDNA platforms. No ctDNA was detected by either platform in these cases, suggesting biological specificity of the bespoke-panel. c) Hotspot SNVs not identified by either platform in these cases, suggesting biological specificity of the bespoke-panel.

ND - non detected
DOR - depth of read
Combined exome VAF (unfiltered) - Variant allele frequency across all tumor regions analysed (without call filters).

Table 2c - Variants detected by generic PCR-NGS hotspot panel not detected in M-Seq analysis of tumor

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

https://www.thermofisher.com/order/catalog/product/A31149

Oncomine lung panel sensitivity data reported at https://www.thermofisher.com/order/catalog/product/A31149

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.