

Published in final edited form as:

Nature. 2010 October 28; 467(7319): 1109–1113. doi:10.1038/nature09460.

The patterns and dynamics of genomic instability in metastatic pancreatic cancer

Peter J. Campbell^{1,2,*}, Shinichi Yachida^{3,*}, Laura J. Mudie¹, Philip J. Stephens¹, Erin D. Pleasance¹, Lucy A. Stebbings¹, Laura A. Morsberger³, Calli Latimer¹, Stuart McLaren¹, Meng-Lay Lin¹, David J. McBride¹, Ignacio Varela¹, Serena A. Nik-Zainal¹, Catherine Leroy¹, Mingming Jia¹, Andrew Menzies¹, Adam P. Butler¹, Jon W. Teague¹, Constance A. Griffin³, John Burton¹, Harold Swerdlow¹, Michael A. Quail¹, Michael R. Stratton^{1,4}, Christine Iacobuzio-Donahue³, and P. Andrew Futreal¹

¹Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK

²Department of Haematology, University of Cambridge, Cambridge CB2 2XY, UK

³Departments of Pathology and Oncology, Johns Hopkins Medical Institutions, Baltimore, Maryland 21287, USA

⁴Institute for Cancer Research, Sutton, Surrey SM2 5NG, UK.

Abstract

Pancreatic cancer is an aggressive malignancy with a five-year mortality of 97–98%, usually due to widespread metastatic disease. Previous studies indicate that this disease has a complex genomic landscape, with frequent copy number changes and point mutations^{1–5}, but genomic rearrangements have not been characterized in detail. Despite the clinical importance of metastasis, there remain fundamental questions about the clonal structures of metastatic tumours^{6,7}, including phylogenetic relationships among metastases, the scale of ongoing parallel evolution in metastatic and primary sites⁷, and how the tumour disseminates. Here we harness advances in DNA sequencing^{8–12} to annotate genomic rearrangements in 13 patients with pancreatic cancer and explore clonal relationships among metastases. We find that pancreatic cancer acquires rearrangements indicative of telomere dysfunction and abnormal cell-cycle control, namely dysregulated G1-to-S-phase transition with intact G2–M checkpoint. These initiate amplification of cancer genes and occur predominantly in early cancer development rather than the later stages of the disease. Genomic instability frequently persists after cancer dissemination, resulting in ongoing, parallel and even convergent evolution among different metastases. We find evidence that there is genetic heterogeneity among metastasis-initiating cells,

©2010 Macmillan Publishers Limited. All rights reserved

Correspondence and requests for materials should be addressed to P.A.F. (paf@sanger.ac.uk) or C.I.-D. (ciacobu@jhmi.edu).

*These authors contributed equally to this work

Author Contributions P.J.C. undertook the analysis of the sequencing data assisted by P.J.S., E.D.P., L.A.S., M.-L.L., D.J.M., I.V., S.A.N.-Z., C.L., M.J., A.M., A.P.B. and J.W.T. Sample collection, processing, establishment of cell lines, DNA extraction and cytogenetic studies were performed by S.Y., L.A.M., C.A.G. and C.I.-D. PCR genotyping, capillary sequencing and downstream validation studies were performed by L.J.M. with assistance from C.L. and S.M. J.B., H.S. and M.A.Q. were responsible for generating libraries and running sequencers. P.J.C., S.Y., M.R.S., C.I.-D. and P.A.F. directed the research and wrote the manuscript, which all authors have approved.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author Information Genome sequence data have been deposited at the European Genome-Phenome Archive (EGA, <http://www.ebi.ac.uk/ega/>), which is hosted by the European Bioinformatics Institute (EBI), under accession number EGAS00000000064. Reprints and permissions information is available at www.nature.com/reprints. Readers are welcome to comment on the online version of this article at www.nature.com/nature.

The authors declare no competing financial interests.

that seeding metastasis may require driver mutations beyond those required for primary tumours, and that phylogenetic trees across metastases show organ-specific branches. These data attest to the richness of genetic variation in cancer, brought about by the tandem forces of genomic instability and evolutionary selection.

We performed massively parallel paired-end sequencing to identify somatically acquired genomic rearrangements in 13 patients with pancreatic adenocarcinoma (Supplementary Table 1). For each sample, we generated 50–150-million paired sequences of 37 base pairs (bp) from 400–500-bp fragments of genomic DNA (Supplementary Figs 1 and 2). Putative rearrangements were screened by polymerase chain reaction (PCR) and capillary sequencing across the breakpoint, allowing annotation to base-pair resolution and distinction between germline and somatic rearrangements^{13,14}. For three patients (patient IDs PD3644–PD3646), samples were early-passage cell lines from resected primary pancreatic tumours. For the other ten patients, multiple metastases were collected at autopsy. In seven of these (PD3637–PD3643), we performed paired-end sequencing on an early-passage cell line derived from a single metastasis per patient. In one patient (PD3826), we sequenced DNA from a bulky metastasis and in two patients (PD3827–PD3828), we separately sequenced three metastases per patient. Hereafter, we refer to lesions sequenced as ‘index’ metastases. For the ten patients with samples from multiple metastases, lesions not sequenced, as well as germline DNA, were genotyped by PCR for the presence or absence of each rearrangement.

We identified 381 somatically acquired and 177 germline rearrangements (Fig. 1a, Supplementary Tables 2 and 3), classified into 7 categories (Supplementary Table 4). The consequences of these rearrangements for protein-coding genes are discussed in Supplementary Results (see also Supplementary Figs 3 and 4 and Supplementary Tables 5 and 6). There was considerable inter-patient heterogeneity in patterns of genomic instability, with differences in numbers (3–65 per patient) and types of rearrangement ($P < 0.0001$; Fig. 1a). Genomic landscapes showed marked disparity within the cohort (Fig. 1b and Supplementary Fig. 5). For example, patient PD3640 had rearrangements evenly scattered across the genome, whereas 35/44 (80%) breakpoints from PD3641 involved chromosome 8. Intrachromosomal rearrangements generally predominated over those between chromosomes, but in PD3646, an inter-crossing patchwork of joins among five chromosomes was the major feature in an otherwise quiet genome.

One sixth of rearrangements show a distinctive pattern we have termed ‘fold-back inversions’ (Fig. 1c). A copy number change is demarcated by read-pairs aligning close together but in inverted orientation. Thus, a genomic region is duplicated, but the two copies head away in opposite orientations from the breakpoint. We believe the most probable mechanism to be breakage-fusion-bridge cycles^{15,16} (Supplementary Results and Supplementary Fig. 6). A double-stranded DNA break occurring in G0-1 phase is replicated during S phase, leading to two identical DNA ends. Repair pathways directly join these, leading to a fold-back inversion pattern at the junction and an unstable dicentric chromosome. We find that this form of genomic instability is an early event in the development of pancreatic cancer and, with marked similarities to data from mouse models¹⁷, frequently underpins and initiates amplification of cancer genes (Supplementary Results and Supplementary Figs 7 and 8).

The distribution of rearrangements in pancreatic cancer is different to that observed in breast cancer¹⁴ ($P < 0.0001$; Fig. 1d and Supplementary Fig. 9). In particular, deletions (22% versus 13%) and fold-back inversions (16% versus 2%) were more frequent in pancreatic cancer, whereas tandem duplications (8% versus 31%) and amplicon-related rearrangements (17% versus 28%) were less frequent.

Taken together, these data indicate that pancreatic cancer has a distinctive pattern of genomic instability. Breakage–fusion–bridge cycles predicate specific abnormalities of cell-cycle control, namely dysregulation of the G1-to-S transition and an intact G2–M checkpoint. Duplication of DNA breaks in S phase implies that repair was not required before DNA replication and end-to-end fusion of the duplicated breaks implies active G2–M surveillance. End-to-end chromosome fusions are often seen in association with telomere erosion and it may be that the double-strand DNA break initiating breakage–fusion–bridge repair results from telomere loss^{5,17–20}.

To understand clonal relationships among metastases in pancreatic cancer, we genotyped 206 rearrangements across multiple lesions from ten patients (Figs 2–4, Supplementary Figs 10, 11 and Supplementary Table 7). Rearrangements followed three patterns: omnipresent across all lesions; partially shared by some but not all metastases; or unique to the index metastasis sequenced (Fig. 2a), with considerable inter-individual heterogeneity (Fig. 2b).

In comparison with other classes of rearrangement, fold-back inversions were significantly more likely to be found in all metastases from that patient ($P=0.003$; Fig. 2c), implying that fold-back inversions occur early in cancer development, before tumour cells disseminate. Breakage–fusion–bridge cycles, resulting in fold-back inversions, are often initiated by telomere loss^{5,18}, whereas telomere attrition is not implicated in the pathogenesis of, for example, interstitial deletions or tandem duplications²¹. Telomerase, the gene that maintains telomere length, shows low expression during early pancreatic carcinogenesis before markedly increasing expression in the invasive tumour^{5,18,22}. The genome-stabilizing effects of telomerase re-expression would therefore have more impact on reducing rates of fold-back inversion in advanced disease than other classes of rearrangement. In contrast, our data indicate that other types of rearrangement occur throughout the cancer life cycle, although the biological pathways underlying these forms of genomic instability remain unclear.

Subclonal evolution within tumours allows reconstruction of phylogenetic relationships²³. Many rearrangements occur in the primary tumour before metastasis commences, and are therefore present in all metastases (Fig. 2b). However, in several patients, there is evidence for ongoing clonal evolution in the primary tumour among cells capable of initiating metastases. Three rearrangements in PD3640 are found in the primary tumour and four metastases, but not the fifth (Fig. 2d, e), with a similar pattern in PD3642 (Fig. 3g). The most likely explanation is that two genetically distinct subclones of the primary independently seeded metastases. We cannot disprove that the discrepant metastasis lost the relevant rearrangements during clonal evolution, but the three events in PD3640 were on different chromosomes, making this unlikely. Importantly, these data indicate that metastasis is clonal, with individual deposits seeded by one or a few genetically similar cells, as described for prostate cancer²⁴.

We also find evidence of clonal evolution within metastases. Rearrangements private to the index lesion were found in seven out of ten patients. Most of these probably occurred in the developing metastasis, although rearrangements acquired either in a subclone of the primary beneath the sensitivity of PCR or during *in vitro* passage⁸ could give similar findings. Additionally, we found five rearrangements in PD3640 present in the index lesion and another metastasis but not in the primary tumour (Fig. 2a, d), with similar patterns in PD3637 (Fig. 3a) and PD3641 (Fig. 3e). These rearrangements might have arisen from clonal evolution in either a secondary metastasis that then itself seeded tertiary metastases or in a subclone of the primary that we have not sampled. Either way, there is considerable genetic heterogeneity among cells capable of initiating metastasis.

Whether metastasis requires mutations beyond those required to drive the primary tumour is controversial²⁵. In PD3637, eight rearrangements were not found in the primary pancreatic tumour despite being present in all metastases (Fig. 3a, b and Supplementary Fig. 10). That all metastases are so phylogenetically distant from the primary tumour indicates that one or more driver mutations, which might either be among the eight rearrangements or among point mutations acquired contemporaneously, have conferred a selective advantage for metastatic spread. In published genomes from a matched breast cancer, brain metastasis and xenograft, there was similar enrichment in the metastasis and xenograft for 10–20 mutations at low prevalence in the primary, although driver mutations for metastasis could not be identified⁸. Taken together, these data imply the existence of a metastasis-promoting genomic signature in at least some patients.

We also find evidence for selection and adaptation within developing metastases after dissemination. For example, in a peritoneal metastasis from PD3642, *KRAS* is amplified to ~8–10 copies (Supplementary Fig. 7A). Because relevant sequencing reads all report the G12V mutation, amplification targeted the activating allele of *KRAS*. Remarkably, all rearrangements driving *KRAS* amplification were found only in the index metastasis and not in any other metastases or the primary (Fig. 3g, h). Within the index lesion, the rearrangements cause marked copy number changes, indicating that each is present in all tumour cells from that metastasis. This implies that rearrangements cumulatively amplifying mutant *KRAS* occurred early during establishment of the metastasis, driving successive waves of clonal expansion²⁶.

Little is known about whether metastases from a given organ system are more closely related to one another than to metastases from different organs. We therefore sequenced three metastases from two patients (Fig. 4). In PD3827, we identified two overlapping, out-of-frame deletions of exon 6 of *PARK2* (Fig. 4b). One was present in all four lung metastases but no abdominal deposits, whereas the other was carried by all four abdominal lesions but no lung deposits. Thus, the two deletions probably arose in separate clones, one of which founded the lung metastases and the other seeded the abdominal metastases. Similarly, in PD3828, lung metastases were on a separate branch of the phylogenetic tree from abdominal lesions (Fig. 4d).

In both patients, the lung lesions were further evolved than the abdominal metastases, and indeed, the additional rearrangements targeted cancer genes. Thus, similar to the *KRAS* amplicon in PD3642 described earlier, several of the lung-specific rearrangements might have conferred further selective advantage on that clone. In PD3828, eight rearrangements were restricted to lung metastases: these clustered around *MYC* and resulted in amplification not seen in abdominal metastases (Supplementary Fig. 11). Similarly, in PD3827, four out of twelve rearrangements restricted to lung metastases further amplified the *CCNE1* cancer gene (Supplementary Fig. 8B).

There are two explanations for organ-specific branches of phylogenetic trees. First, particular genotypes might drive metastasis to particular organs. The fact that lung metastases in these two patients were associated with additional driver mutations (amplification of *MYC* or *CCNE1*) indicates that tumour cells from subclones carrying these rearrangements were more likely to survive in the lung. Second, metastatic spread might be a stepwise process, occurring more readily within organ boundaries than between organs. These explanations are not mutually exclusive. Overcoming the barrier to colonizing a given organ might depend on a subclone of cancer cells acquiring particular adaptive changes, which, once established, can then disseminate through the organ with relative ease.

At first glance, the remarkable genetic diversity and adaptability of cancer under different selection pressures glimpsed here has ominous implications for our attempts to find curative therapies for metastatic disease. Nevertheless, for most patients studied here, more than half the rearrangements were found in all metastases and the primary tumour. The ability of studies such as this one to identify and understand these early mutations provides a route to the discovery of drug targets.

METHODS SUMMARY

Thirteen patients with pancreatic cancer were studied, with written informed consent for sample collection and analysis. Ten patients had multiple metastases collected at autopsy performed within 6 h of death, as described²⁷. We also studied primary tumours collected from three patients undergoing resection with curative intent. Representative samples of primary carcinoma or metastases were minced with sterile blades, and the tissues gently pressed through a 45- μ m mesh to disaggregate epithelial and stromal cells. For low-passage cell lines, filtered cells were resuspended into culture media and passaged up to five times to remove contaminating fibroblasts.

Protocols for massively parallel paired-end sequencing have been described in detail elsewhere^{13,14}. Genomic DNA from the tumour samples was randomly fragmented, and fragments 400–500 bp in size selected by gel purification. Libraries were synthesized following our standard protocol, as described²⁸, and sequenced on a Genome Analyser II (Illumina) to give 37-bp reads from both ends of 50–150-million DNA fragments. In our experience, this identifies ~50–60% of rearrangements in a sample^{13,14}. This level of genome coverage is insufficient to allow accurate identification of point mutations¹¹, but allows patterns of genomic rearrangement to be studied across several cancer samples without bias in size or type of rearrangement.

Sequencing data were aligned to the human reference genome (National Center for Biotechnology Information (NCBI) build 36) using the MAQ algorithm²⁹. Clusters of anomalously mapping reads spanning putative rearrangements were identified informatically¹³. PCR across the breakpoint was performed in tumour and normal DNA, allowing rearrangements to be classified as somatically acquired, germline or artefactual. PCR products underwent capillary sequencing to annotate breakpoints to base-pair resolution. In ten patients, primers for somatic rearrangements were used to genotype by PCR all other metastases and, where available, the primary tumour from that patient. The sensitivity of PCR for detection of genomic rearrangements is at least 1/1,000 cells³⁰, considerably better than can be achieved for point mutations.

Acknowledgments

This work was supported by the Wellcome Trust (grant reference 077012/Z/05/Z). P.J.C. is funded through a Wellcome Trust Senior Clinical Research Fellowship (grant reference WT088340MA). S.Y. has support from the Uehara memorial foundation. We would also like to acknowledge the financial support of the Skip Viragh Foundation and the Michael Rolphe Foundation for the autopsy programme, and funding from the National Institutes of Health (grants CA106610 and CA140599). I.V. is supported by a fellowship from The International Human Frontier Science Program Organization. We would like to thank U. McDermott for discussions and a critical reading of the manuscript.

References

1. Jones S, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science*. 2008; 321:1801–1806. [PubMed: 18772397]
2. Harada T, et al. Genome-wide DNA copy number analysis in pancreatic cancer using high-density single nucleotide polymorphism arrays. *Oncogene*. 2008; 27:1951–1960. [PubMed: 17952125]

3. Fu B, Luo M, Lakkur S, Lucito R, Iacobuzio-Donahue CA. Frequent genomic copy number gain and overexpression of GATA-6 in pancreatic carcinoma. *Cancer Biol. Ther.* 2008; 7:1593–1601. [PubMed: 18769116]
4. Kimmelman AC, et al. Genomic alterations link Rho family of GTPases to the highly invasive phenotype of pancreas cancer. *Proc. Natl Acad. Sci. USA.* 2008; 105:19372–19377. [PubMed: 19050074]
5. Gisselsson D, et al. Telomere dysfunction triggers extensive DNA fragmentation and evolution of complex chromosome abnormalities in human malignant tumors. *Proc. Natl Acad. Sci. USA.* 2001; 98:12683–12688. [PubMed: 11675499]
6. Klein CA. Parallel progression of primary tumours and metastases. *Nature Rev. Cancer.* 2009; 9:302–312. [PubMed: 19308069]
7. Kuukasjarvi T, et al. Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res.* 1997; 57:1597–1604. [PubMed: 9108466]
8. Ding L, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature.* 2010; 464:999–1005. [PubMed: 20393555]
9. Mardis ER, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. *N. Engl. J. Med.* 2009; 361:1058–1066. [PubMed: 19657110]
10. Pleasance ED, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature.* 2010; 463:191–196. [PubMed: 20016485]
11. Pleasance ED, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature.* 2010; 463:184–190. [PubMed: 20016488]
12. Shah SP, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature.* 2009; 461:809–813. [PubMed: 19812674]
13. Campbell PJ, et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nature Genet.* 2008; 40:722–729. [PubMed: 18438408]
14. Stephens PJ, et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature.* 2009; 462:1005–1010. [PubMed: 20033038]
15. McLintock B. The stability of broken ends of chromosomes in *Zea mays*. *Genetics.* 1941; 26:234–282. [PubMed: 17247004]
16. Bignell GR, et al. Architectures of somatic genomic rearrangement in human cancer amplicons at sequence-level resolution. *Genome Res.* 2007; 17:1296–1303. [PubMed: 17675364]
17. O'Hagan RC, et al. Telomere dysfunction provokes regional amplification and deletion in cancer genomes. *Cancer Cell.* 2002; 2:149–155. [PubMed: 12204535]
18. Bardeesy N, DePinho RA. Pancreatic cancer biology and genetics. *Nature Rev. Cancer.* 2002; 2:897–909. [PubMed: 12459728]
19. Maser RS, et al. Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. *Nature.* 2007; 447:966–971. [PubMed: 17515920]
20. Sahin E, Depinho RA. Linking functional decline of telomeres, mitochondria and stem cells during ageing. *Nature.* 2010; 464:520–528. [PubMed: 20336134]
21. Blow JJ, Gillespie PJ. Replication licensing and cancer—a fatal entanglement? *Nature Rev. Cancer.* 2008; 8:799–806. [PubMed: 18756287]
22. Hashimoto Y, Murakami Y, Uemura K, et al. Telomere shortening and telomerase expression during multistage carcinogenesis of intraductal papillary mucinous neoplasms of the pancreas. *J. Gastrointest. Surg.* 2008; 12:17–29. [PubMed: 17960465]
23. Campbell PJ, et al. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc. Natl Acad. Sci. USA.* 2008; 105:13081–13086. [PubMed: 18723673]
24. Liu W, et al. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nature Med.* 2009; 15:559–565. [PubMed: 19363497]
25. Nguyen DX, Massague J. Genetic determinants of cancer metastasis. *Nature Rev. Genet.* 2007; 8:341–352. [PubMed: 17440531]

26. Klein CA, et al. Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. *Lancet*. 2002; 360:683–689. [PubMed: 12241875]
27. Embuscado EE, et al. Immortalizing the complexity of cancer metastasis: genetic features of lethal metastatic pancreatic cancer obtained from rapid autopsy. *Cancer Biol. Ther.* 2005; 4:548–554. [PubMed: 15846069]
28. Quail MA, et al. A large genome center's improvements to the Illumina sequencing system. *Nature Methods*. 2008; 5:1005–1010. [PubMed: 19034268]
29. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res*. 2008; 18:1851–1858. [PubMed: 18714091]
30. Flohr T, et al. Minimal residual disease-directed risk stratification using real-time quantitative PCR analysis of immunoglobulin and T-cell receptor gene rearrangements in the international multicenter trial AIEOP-BFM ALL 2000 for childhood acute lymphoblastic leukemia. *Leukemia*. 2008; 22:771–782. [PubMed: 18239620]

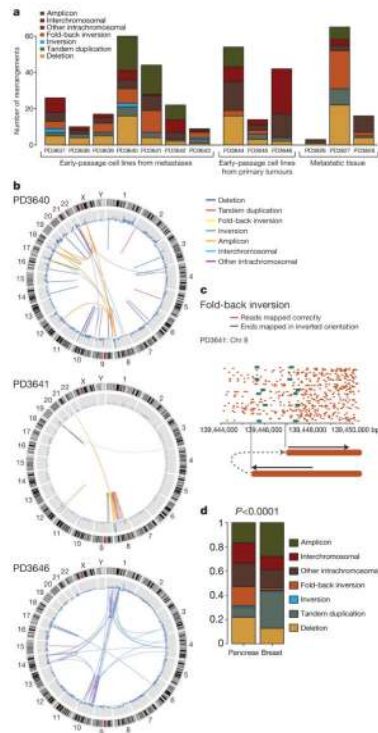


Figure 1. Patterns of somatically acquired genomic rearrangements in pancreatic cancer

a, Histogram showing the distribution of the number and types of rearrangement observed in 13 patients with pancreatic cancer. **b**, Circle plots showing the genomic landscape of rearrangements in three representative samples. Chromosome ideograms are shown around the outer ring with copy number plots on the inner ring. Individual rearrangements are shown as arcs joining the two genomic loci, each coloured according to the type of rearrangement. **c**, Example of a so-called ‘fold-back inversion’. Correctly mapping paired reads (orange) show much greater density on the right half of the figure than the left, indicating that the copy number is higher here. The change in copy number is demarcated by anomalously mapping paired reads (green), aligning ~2 kb apart on the genome and in inverted orientation. The only genomic structure that can explain this pattern is a rearrangement in which the abnormal chromosome is ‘folded back’ on itself leading to duplicated genomic segments in head-to-head (inverted) orientation. *x* axis, genomic position. Chr, chromosome. **d**, The distribution of types of rearrangement was significantly different between breast cancer and pancreatic cancer ($P < 0.0001$). *y* axis, proportion of all rearrangements.

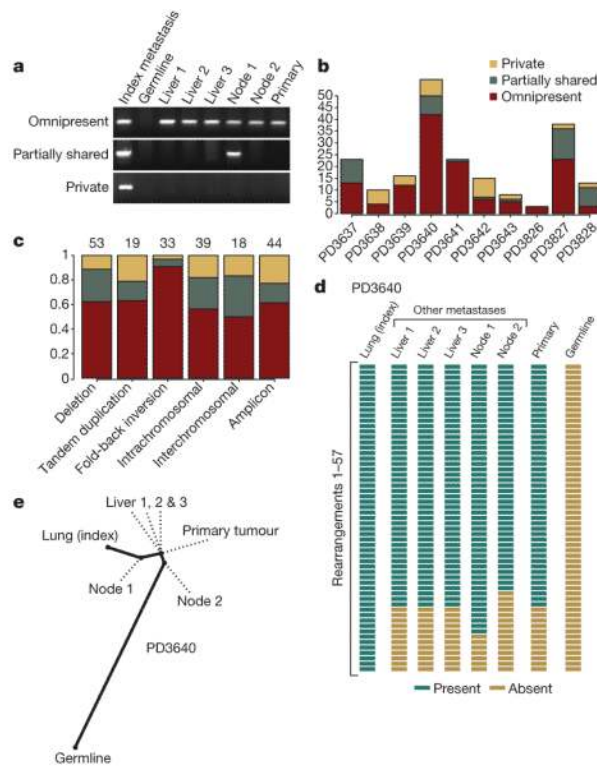


Figure 2. Phylogenetic relationships of different metastases within a patient

a, PCR genotyping of three rearrangements across DNA from the index metastasis sequenced, other metastases from the same patient, the primary tumour and germline tissue. Somatic rearrangements may be present in all cancer samples but not the germline (omnipresent); present in some but not all metastases (partially shared); or present just in the index metastasis sequenced (private). **b**, Inter-individual differences in the proportions of rearrangements that are omnipresent across metastases, partially shared by some but not all lesions or are private to the index metastasis sequenced. **c**, Patterns across six broad categories of rearrangement in the proportions of variants that are omnipresent across metastases, partially shared by some but not all lesions or are private to the index metastasis sequenced. The numbers of rearrangements in each category are shown at the top. The difference in proportions between fold-back inversions and the other categories was statistically significant ($P=0.003$). **d**, Genotyping of 57 rearrangements in PD3640 shows a coherent, nested structure, with 42 found in all metastases and the primary tumour, 7 found uniquely in the index tumour and 8 partially shared by some but not all metastases. **e**, The nested structure of rearrangements defines a phylogenetic tree of relationships among the metastases and primary tumour. The length of heavy black lines is proportional to the genetic distance between nodes. Dotted lines delineate the departure points of other, unsequenced lesions from the lineage between the germline genome and that of the index metastasis.

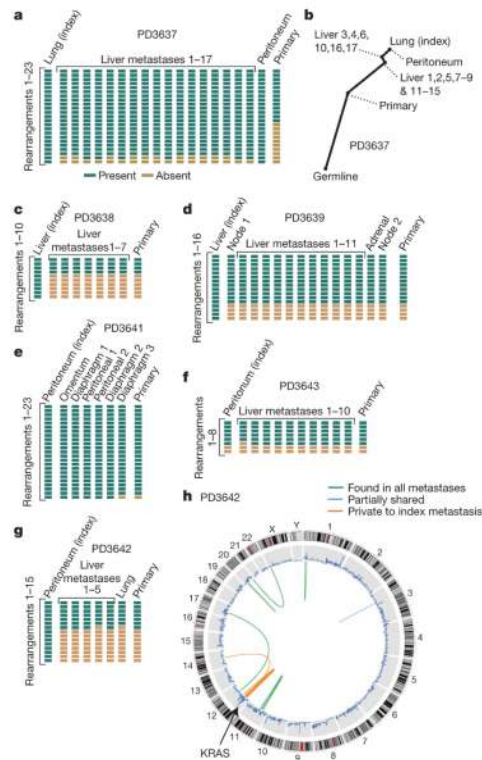


Figure 3. Phylogenetic relationships among different metastases and the primary tumour
a, Results of PCR genotyping for 23 rearrangements across 19 metastases and the primary tumour from patient PD3637. **b**, Phylogenetic tree showing the relatedness of different metastases and the primary tumour. Note the early divergence of the primary tumour from all metastases. **c–g**, Genotyping results for PD3638 (**c**), PD3639 (**d**), PD3641 (**e**), PD3643 (**f**) and PD3642 (**g**). **h**, Circle plot showing that the rearrangements generating the amplicon of *KRAS* on chromosome 12 in PD3642 were only found in the index metastasis sequenced, and none of the other metastases or the primary tumour.

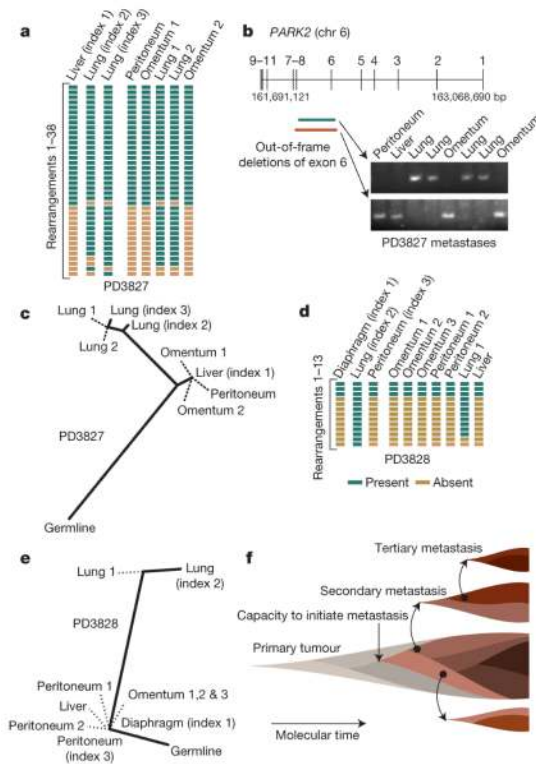


Figure 4. Organ-specific signatures of metastasis

a, Results of PCR genotyping for 38 rearrangements across the 3 index metastases and 5 other metastases from patient PD3827. **b**, Overlapping out-of-frame deletions of exon 6 of *PARK2* were mutually exclusive to either the four lung metastases or the four abdominal metastases. The numbers above the gene refer to axon numbers and those below to the genomic position. **c**, A phylogenetic tree of relationships for metastases from patient PD3827, showing a clade of abdominal metastases and a further evolved clade of lung metastases. The length of heavy black lines is proportional to the genetic distance between nodes. Dotted lines delineate the departure points of other, unsequenced lesions from the lineage between the germline genome and that of the index metastasis. **d**, Results of PCR genotyping for PD3828. **e**, Phylogenetic tree of relationships for metastases from PD3828. **f**, Model for the clonal evolution of metastases derived from the patterns of phylogenetic relationships observed. Molecular time proceeds from left to right, and is associated with subclonal evolution and expansion within the developing primary tumour. Eventually a subclone within the primary tumour acquires the capacity to metastasize (pink), but this subclone continues to acquire genetic lesions (darkening shades of brown) such that different metastases may be founded from different clones. Within the developing metastases, clonal evolution continues, and these newly developed subclones can themselves seed tertiary metastases.