



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<sup>1,2,§</sup>, Shinichi Yachida<sup>3,§</sup>, Laura J Mudie<sup>1</sup>, Philip J Stephens<sup>1</sup>, Erin D Pleasance<sup>1</sup>, Lucy A Stebbings<sup>1</sup>, Laura A Morsberger<sup>3</sup>, Calli Latimer<sup>1</sup>, Stuart McLaren<sup>1</sup>, Meng-Lay Lin<sup>1</sup>, David J McBride<sup>1</sup>, Ignacio Varela<sup>1</sup>, Serena A Nik-Zainal<sup>1</sup>, Catherine Leroy<sup>1</sup>, Mingming Jia<sup>1</sup>, Andrew Menzies<sup>1</sup>, Adam P Butler<sup>1</sup>, Jon W Teague<sup>1</sup>, Constance A Griffin<sup>3</sup>, John Burton<sup>1</sup>, Harold Swerdlow<sup>1</sup>, Michael A Quail<sup>1</sup>, Michael R Stratton<sup>1,4,\*</sup>, Christine Iacobuzio-Donahue<sup>3,\*</sup>, and P Andrew Futreal<sup>1,\*</sup>

<sup>1</sup>Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton, UK <sup>2</sup>Department of Haematology, University of Cambridge, Cambridge, UK <sup>3</sup>Departments of Pathology and Oncology, Johns Hopkins Medical Institutions, Baltimore, Maryland <sup>4</sup>Institute for Cancer Research, Sutton, Surrey, UK

### SUMMARY

Pancreatic cancer is an aggressive malignancy with 5-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<sup>1–5</sup>, but genomic rearrangements have not been characterised in detail. Despite the clinical importance of metastasis, there remain fundamental questions about the clonal structures of metastatic tumours<sup>6,7</sup>, including phylogenetic relationships among metastases, the scale of on-going parallel evolution in metastatic and primary sites<sup>7</sup>, and how the tumour disseminates. Here, we harness advances in DNA sequencing<sup>8–12</sup> 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

Users may view, print, copy, download and text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

\*Joint corresponding authors, **Address for correspondence:** Dr Andy Futreal & Prof Mike Stratton, Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire CB10 1SA, UK. Phone: +44 1223 494951, Fax: +44 1223 494809, [paf@sanger.ac.uk](mailto:paf@sanger.ac.uk); [mrs@sanger.ac.uk](mailto:mrs@sanger.ac.uk), Dr Christine Iacobuzio-Donahue, Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins Medical Institutions, 1550 Orleans St (CRB2, Rm 343), Baltimore, MD 21231. Phone: +1 410 955 3511, Fax: +1 410 614 0671, [ciacobu@jhmi.edu](mailto:ciacobu@jhmi.edu).

§Joint first authors

### AUTHOR CONTRIBUTIONS

PJC undertook the analysis of the sequencing data, assisted by PJS, EDP, LAS, MLL, DJM, IV, SNZ, CL, MJ, AM, APB and JWT. Sample collection, processing, establishment of cell lines, DNA extraction and cytogenetic studies were performed by SY, LAM, CAG and CID. PCR genotyping, capillary sequencing and downstream validation studies were performed by LJM, with assistance from CL and SM. JB, HS and MAQ were responsible for generating libraries and running sequencers. PJC, SY, MRS, CID and PAF directed the research and wrote the manuscript, which all authors have approved.

### 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 EBI, under accession number EGAS00000000064. Reprints and permissions information is available at [npg.nature.com/reprintsandpermissions](http://npg.nature.com/reprintsandpermissions). The authors have no conflicts of interest to report. Correspondence and requests for materials should be addressed to [pc8@sanger.ac.uk](mailto:pc8@sanger.ac.uk).

control, namely dysregulated G1-S phase transition with intact G2-M checkpoint. These initiate amplification of cancer genes and occur predominantly in early cancer development rather than later stages of disease. Genomic instability frequently persists after cancer dissemination, resulting in on-going, parallel and even convergent evolution among different metastases. We find evidence that there is genetic heterogeneity among metastasis-initiating cells; seeding metastasis may require driver mutations beyond those required for primary tumours; and phylogenetic trees across metastases show organ-specific branches. These data attest to the richness of genetic variation in cancer, hewn by the tandem forces of genomic instability and evolutionary selection.

## MAIN TEXT

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 37bp from 400–500bp fragments of genomic DNA (supplementary figures 1–2). Putative rearrangements were screened by PCR and capillary sequencing across the breakpoint, allowing annotation to base-pair resolution and distinction between germline and somatic rearrangements<sup>13,14</sup>. For three patients (PD3644–PD3646), samples were early passage cell lines from resected primary pancreatic tumours. For the other 10 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 10 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 (figure 1A, supplementary tables 2–3), classified into 7 categories (supplementary table 4). The consequences of these rearrangements for protein-coding genes are discussed in supplementary results (also supplementary figures 3–4, supplementary tables 5–6). There was considerable inter-patient heterogeneity in patterns of genomic instability, with differences in numbers (3–65/patient) and types of rearrangement ( $p < 0.0001$ ; figure 1A). Genomic landscapes showed striking disparity within the cohort (figure 1B, supplementary figure 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 intercrossing 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’ (figure 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 likely mechanism to be breakage-fusion-bridge cycles<sup>15,16</sup> (supplementary results, supplementary figure 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 striking similarities to data from mouse models<sup>17</sup>, frequently underpins and initiates amplification of cancer genes (supplementary results, supplementary figures 7–8).

The distribution of rearrangements in pancreatic cancer is different to that observed in breast cancer<sup>14</sup> ( $p < 0.0001$ ; figure 1D, supplementary figure 9). In particular, deletions (22% vs 13%) and fold-back inversions (16% vs 2%) were more frequent in pancreatic cancer, whereas tandem duplications (8% vs 31%) and amplicon-related rearrangements (17% vs 28%) were less frequent.

Taken together, these data suggest that pancreatic cancer has a distinctive flavour of genomic instability. Breakage-fusion-bridge cycles predicate specific abnormalities of cell-cycle control, namely dysregulation of G1-to-S transition and an intact G2-M checkpoint. Duplication of DNA breaks in S-phase implies 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 dsDNA break initiating breakage-fusion-bridge repair results from telomere loss<sup>5,17–20</sup>.

To understand clonal relationships among metastases in pancreatic cancer, we genotyped 206 rearrangements across multiple lesions from 10 patients (figures 2–4, supplementary figures 10–11, 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 (figure 2A), with considerable inter-individual heterogeneity (figure 2B).

In comparison with other classes of rearrangement, fold-back inversions were significantly more likely to be found in all metastases from that patient ( $p = 0.003$ ; figure 2C), implying 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<sup>5,18</sup>, whereas telomere attrition is not implicated in the pathogenesis of, for example, interstitial deletions or tandem duplications<sup>21</sup>. Telomerase, the gene that maintains telomere length, shows low expression during early pancreatic carcinogenesis before markedly increasing expression in the invasive tumour<sup>5,18,22</sup>. The genome-stabilising 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 suggest 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 relationships<sup>23</sup>. Many rearrangements occur in the primary tumour before metastasis commences, and are therefore present in all metastases (figure 2B). However, in several patients, there is evidence for on-going 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 (figure 2D–E), with a similar pattern in PD3642 (figure 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 cancer<sup>24</sup>.

We also find evidence of clonal evolution within metastases. Rearrangements private to the index lesion were found in 7/10 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* passage<sup>8</sup> could give similar findings. Additionally, we found five rearrangements in PD3640 present in the index lesion and another metastasis but *not* the primary tumour (figure 2A,D), with similar patterns in PD3637 (figure 3A) and PD3641 (figure 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 controversial<sup>25</sup>. In PD3637, 8 rearrangements were not found in the primary pancreatic tumour despite being present in all metastases (figure 3A–B, supplementary figure 10). That all metastases are so phylogenetically distant from the primary tumour suggests that one or more driver mutations, which might either be among the 8 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 identified<sup>8</sup>. 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 figure 7A). Since 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 (figure 3G–H). Within the index lesion, the rearrangements cause marked copy number changes, suggesting 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 expansion<sup>26</sup>.

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 (figure 4). In PD3827, we identified two overlapping, out-of-frame deletions of exon 6 of *PARK2* (figure 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 (figure 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 in the previous section, several of the lung-specific rearrangements might have conferred further selective advantage on that clone. In PD3828, 8 rearrangements were restricted to lung metastases: these clustered around *MYC* and resulted in amplification not seen in abdominal metastases (supplementary figure 11). Similarly, in PD3827, 4/12 rearrangements restricted to lung metastases further amplified the *CCNE1* cancer gene (supplementary figure 8B).

There are two explanations for organ-specific branches of phylogenetic trees. Firstly, particular genotypes might drive metastasis to particular organs. That lung metastases in these two patients were associated with additional driver mutations (amplification of *MYC* or *CCNE1*) suggests that tumour cells from subclones carrying these rearrangements were more likely to survive in the lung. Secondly, 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 colonising a given organ might depend upon 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 discovery of drug targets.

## METHODS

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 hours of death, as described<sup>27</sup>. 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-micron 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<sup>13,14</sup>. Genomic DNA from the tumour samples was randomly fragmented, and fragments 400–500bp in size selected by gel purification. Libraries were synthesised

following our standard protocol, as described<sup>28</sup>, and sequenced on a Genome Analyzer II (Illumina Inc) to give 37bp reads from both ends of 50–150 million DNA fragments. In our experience, this identifies ~50–60% of rearrangements in a sample<sup>13,14</sup>. This level of genome coverage is insufficient to allow accurate identification of point mutations<sup>11</sup>, but allows patterns of genomic rearrangement to be studied across multiple cancer samples without bias in size or type of rearrangement.

Sequencing data were aligned to the human reference genome (NCBI build 36) using the MAQ algorithm<sup>29</sup>. Clusters of anomalously mapping reads spanning putative rearrangements were identified informatically<sup>13</sup>. 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 10 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 1/1000 cells<sup>30</sup>, considerably better than can be achieved for point mutations.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

This work was supported by the Wellcome Trust (grant reference 077012/Z/05/Z). PJC is funded through a Wellcome Trust Senior Clinical Research Fellowship (grant reference WT088340MA). SY 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). IV is supported by a fellowship from The International Human Frontier Science Program Organization. We would like to thank Ultan McDermott for helpful discussions and a critical reading of the manuscript.

## REFERENCES

1. Jones S, Zhang X, Parsons DW, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science*. 2008; 321:1801–1806. [PubMed: 18772397]
2. Harada T, Chelala C, Bhakta V, 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, Hezel AF, Aguirre AJ, et al. Genomic alterations link Rho family of GTPases to the highly invasive phenotype of pancreas cancer. *Proc Natl Acad Sci U S A*. 2008; 105:19372–19377. [PubMed: 19050074]
5. Gisselsson D, Jonson T, Petersen A, et al. Telomere dysfunction triggers extensive DNA fragmentation and evolution of complex chromosome abnormalities in human malignant tumors. *Proc Natl Acad Sci U S A*. 2001; 98:12683–12688. [PubMed: 11675499]
6. Klein CA. Parallel progression of primary tumours and metastases. *Nat Rev Cancer*. 2009; 9:302–312. [PubMed: 19308069]
7. Kuukasjarvi T, Karhu R, Tanner M, 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, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature*. 2010; 464:999–1005. [PubMed: 20393555]
9. Mardis ER, Ding L, Dooling DJ, 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, Cheetham RK, Stephens PJ, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature*. 2010; 463:191–196. [PubMed: 20016485]
11. Pleasance ED, Stephens PJ, O'Meara S, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature*. 2010; 463:184–190. [PubMed: 20016488]
12. Shah SP, Morin RD, Khattra J, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature*. 2009; 461:809–813. [PubMed: 19812674]
13. Campbell PJ, Stephens PJ, Pleasance ED, et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat Genet*. 2008; 40:722–729. [PubMed: 18438408]
14. Stephens PJ, McBride DJ, Lin ML, 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, Santarius T, Pole JC, 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, Chang S, Maser RS, 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. *Nat Rev Cancer*. 2002; 2:897–909. [PubMed: 12459728]
19. Maser RS, Choudhury B, Campbell PJ, 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? *Nat 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; 12:17–28. discussion 28–19. [PubMed: 17960465]
23. Campbell PJ, Pleasance ED, Stephens PJ, et al. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc Natl Acad Sci U S A*. 2008; 105:13081–13086. [PubMed: 18723673]
24. Liu W, Laitinen S, Khan S, et al. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat Med*. 2009; 15:559–565. [PubMed: 19363497]
25. Nguyen DX, Massague J. Genetic determinants of cancer metastasis. *Nat Rev Genet*. 2007; 8:341–352. [PubMed: 17440531]
26. Klein CA, Blankenstein TJ, Schmidt-Kittler O, et al. Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. *Lancet*. 2002; 360:683–689. [PubMed: 12241875]
27. Embuscado EE, Laheru D, Ricci F, 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, Kozarewa I, Smith F, et al. A large genome center's improvements to the Illumina sequencing system. *Nat 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
30. Flohr T, Schrauder A, Cazzaniga G, 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]

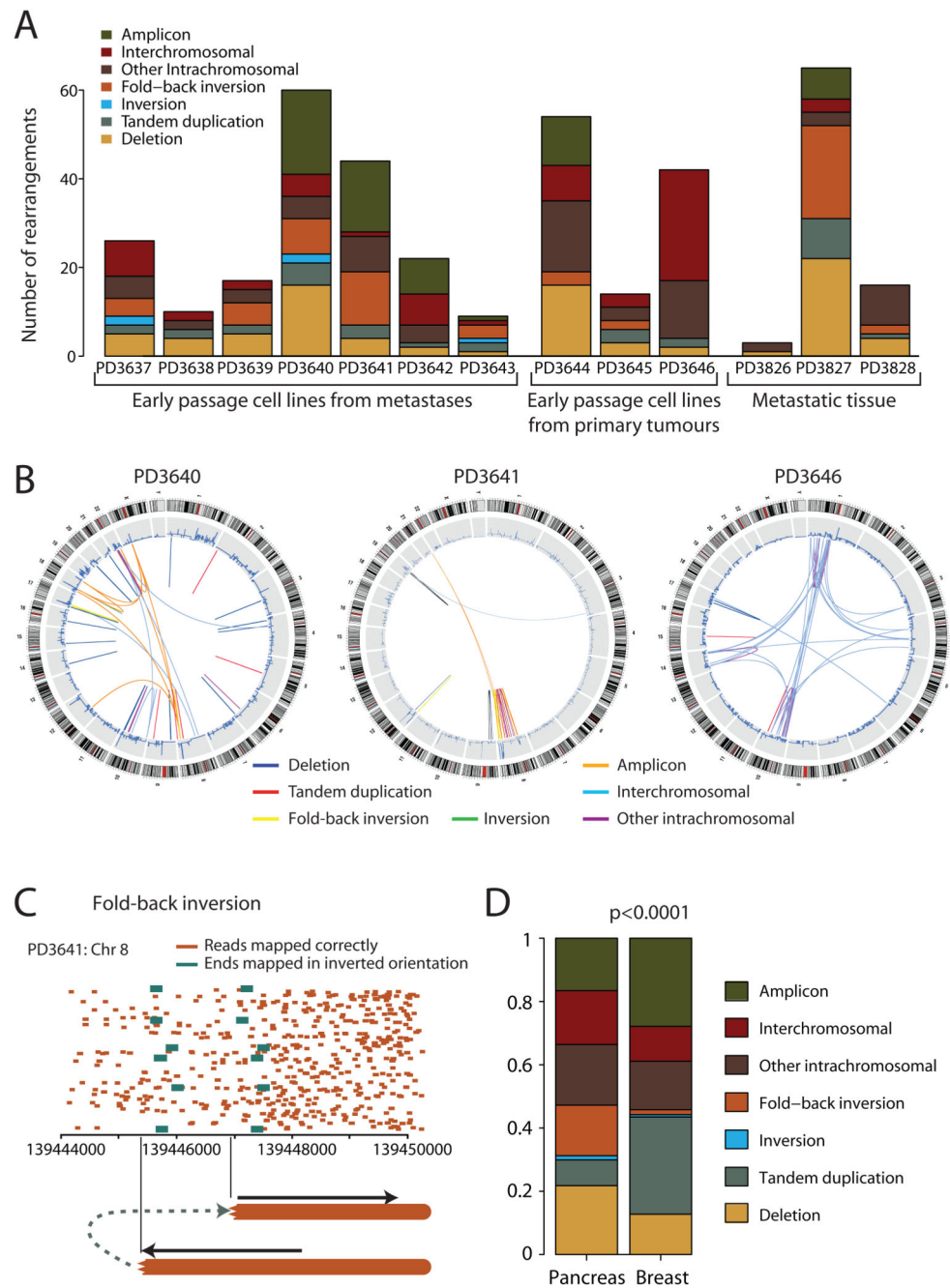
Author Manuscript

Author Manuscript

Author Manuscript

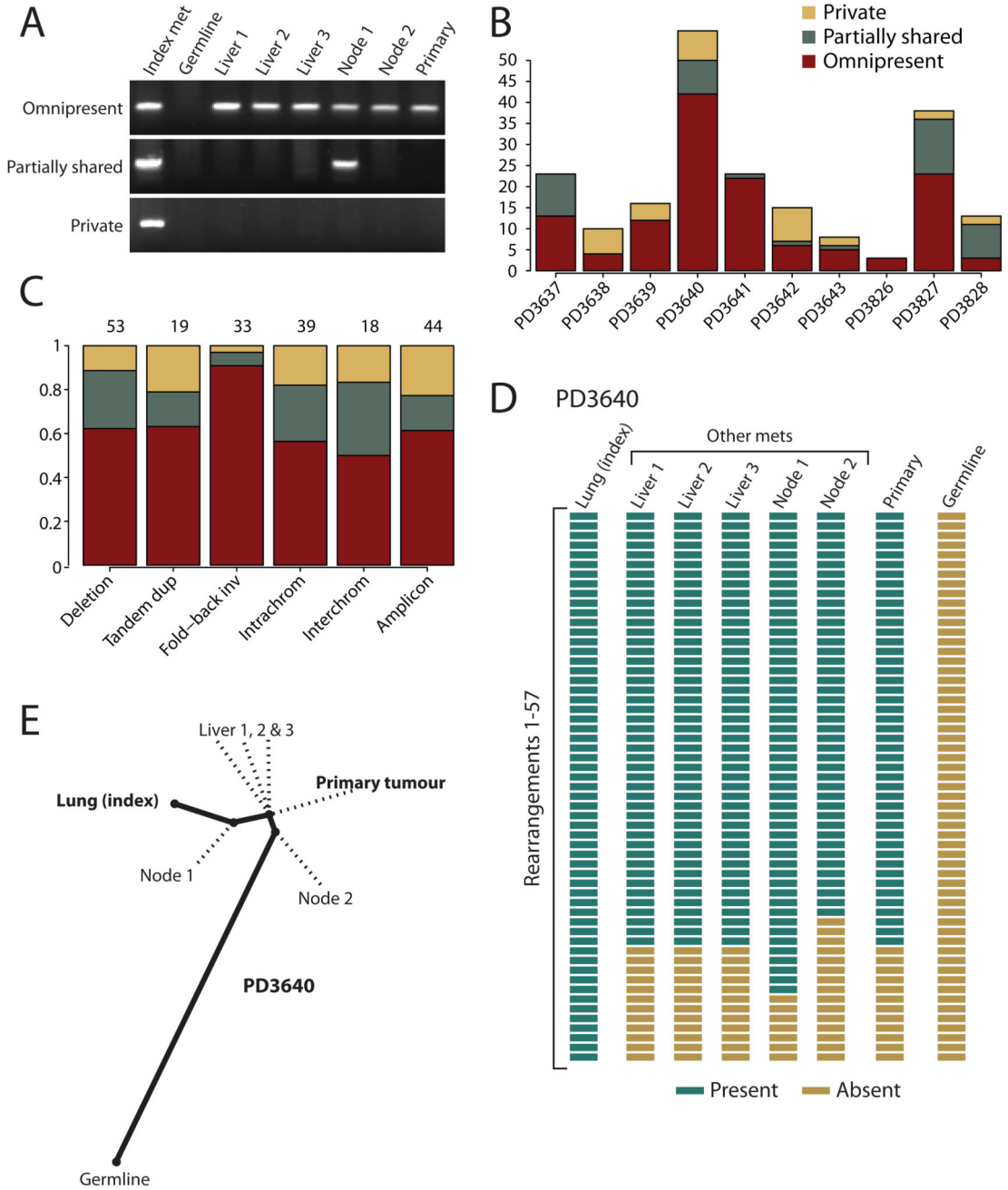
Author Manuscript





**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, suggesting that the copy number is higher here. The change in copy number is demarcated by anomalously mapping paired reads (green), aligning ~2kb apart on the genome and in inverted orientation. The only genomic structure which 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. (D) The distribution of types of rearrangement was significantly different between breast cancer and pancreatic cancer ( $p < 0.0001$ ).



**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

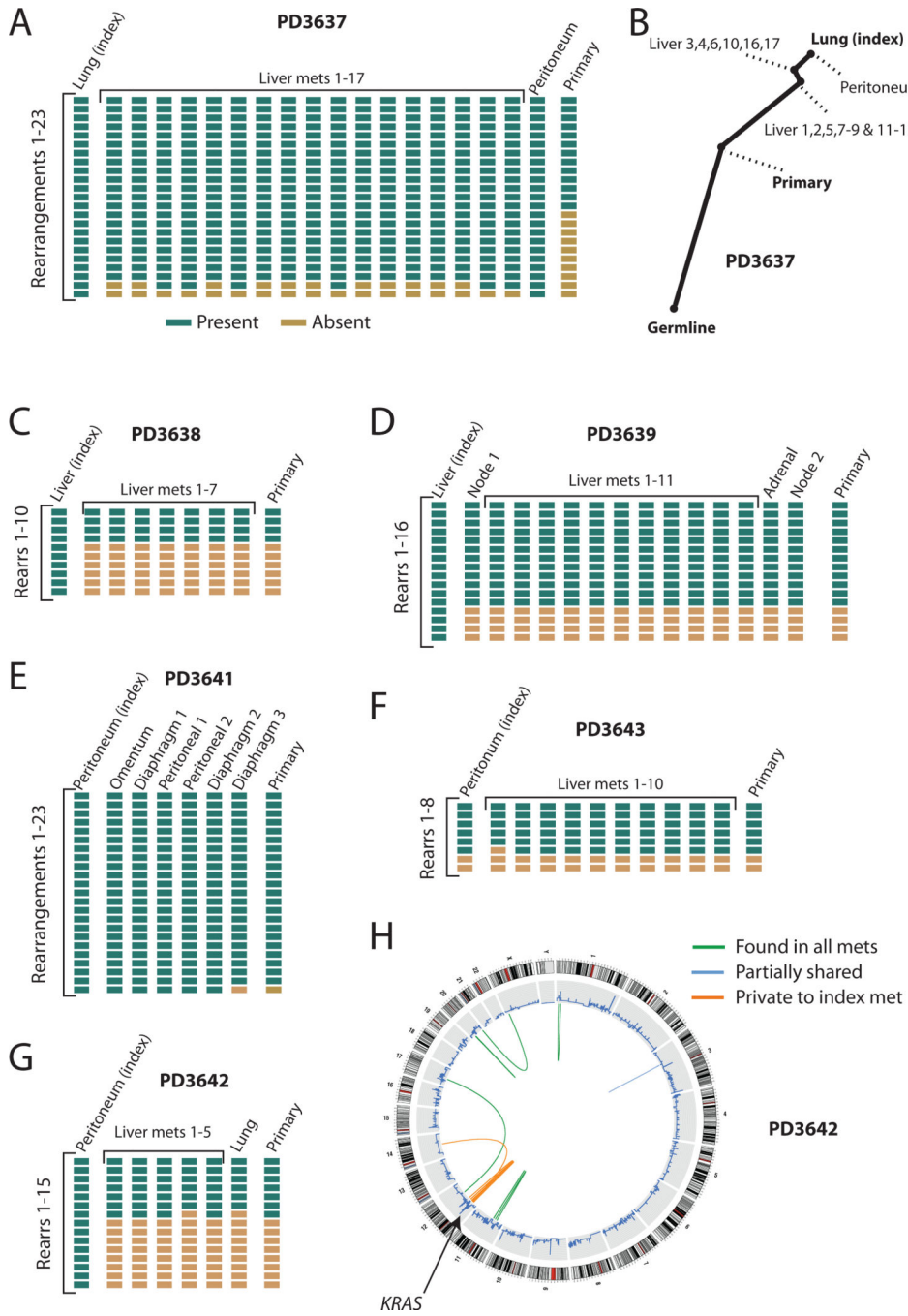
metastasis sequenced. (C) Patterns across six broad categories of rearrangement in the proportions of variants that are omnipresent across metastases, partially shared by some but not all lesions or are private to the index metastasis sequenced. The numbers of rearrangements in each category are shown 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.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**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) Genotyping results for PD3638, as well as (D) PD3639, (E) PD3641, (F) PD3643 and (G) PD3642. (H) Circle plot showing that the rearrangements generating the

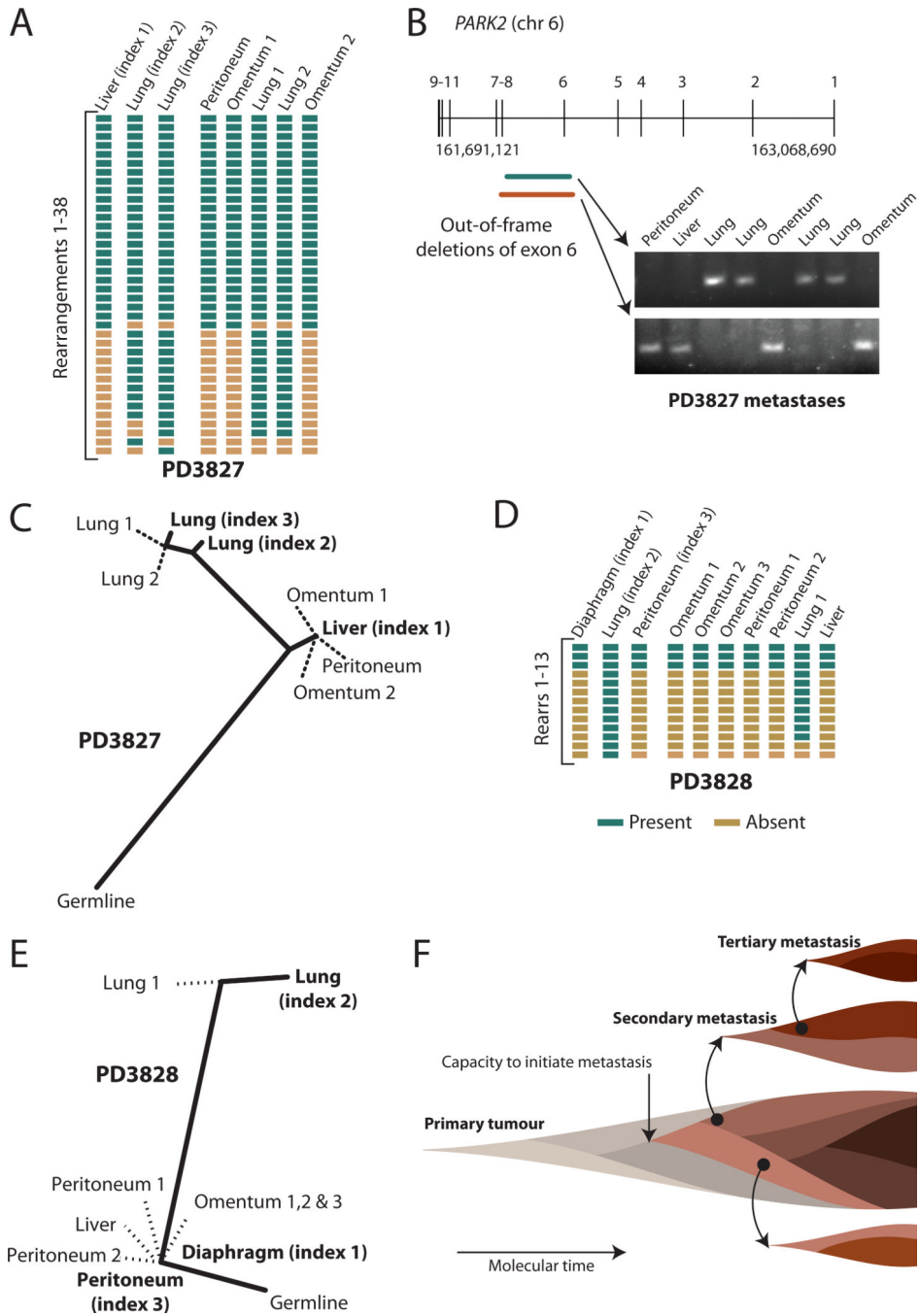
amplicon of *KRAS* on chromosome 12 in PD3642 were only found in the index metastasis sequenced, and none of the other metastases or the primary tumour.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 4.** Organ-specific signatures of metastasis. (A) Results of PCR genotyping for 38 rearrangements across the three index metastases and five 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 (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 metastasise (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.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript