

Selection of metastasis competent subclones in the tumour interior: TRACERx renal

Yue Zhao

The Francis Crick Institute

Xiao Fu

The Francis Crick Institute <https://orcid.org/0000-0002-8611-891X>

Jose Lopez

Cruces University Hospital <https://orcid.org/0000-0003-0842-5348>

Andrew Rowan

The Francis Crick Institute

Lewis Au

The Francis Crick Institute

Annika Fendler

The Francis Crick Institute

Steve Hazell

The Royal Marsden NHS Foundation Trust

Hang Xu

The Francis Crick Institute

Stuart Horswell

The Francis Crick Institute

Scott Shephard

The Francis Crick Institute

Lavinia Spain

The Royal Marsden Hospital NHS Foundation Trust

Fiona Byrne

The Francis Crick Institute <https://orcid.org/0000-0003-3800-0500>

Gordon Stamp

The Francis Crick Institute

Tim O'Brien

Guy's & St Thomas' NHS Foundation Trust

David Nicol

The Royal Marsden NHS Foundation Trust

Marcellus Augustine

The Francis Crick Institute

Ashish Chandra

Guy's & St Thomas' NHS Foundation Trust

Sarah Rudman

Guy's & St. Thomas' NHS Foundation Trust

Antonia Toncheva

Guy's & St Thomas' NHS Foundation Trust

Lisa Pickering

The Royal Marsden Hospital NHS Foundation Trust

James Larkin

The Royal Marsden Hospital NHS Foundation Trust

Erik Sahai

The Francis Crick Institute

Paul Bates (✉ paul.bates@crick.ac.uk)

The Francis Crick Institute <https://orcid.org/0000-0003-0621-0925>

Charles Swanton (✉ Charles.Swanton@crick.ac.uk)

The Francis Crick Institute <https://orcid.org/0000-0002-4299-3018>

Samra Turajlic (✉ Samra.Turajlic@crick.ac.uk)

The Francis Crick Institute

Kevin Litchfield (✉ k.litchfield@ucl.ac.uk)

The Francis Crick Institute <https://orcid.org/0000-0002-3725-0914>

Article

Keywords: tumour evolution, renal cell carcinoma, macroscopic imaging, and spatial growth

Posted Date: August 31st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-61979/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Ecology & Evolution on May 17th, 2021. See the published version at <https://doi.org/10.1038/s41559-021-01456-6>.

Abstract

While the genetic evolutionary features of solid tumour growth are becoming increasingly described, the spatial and physical nature of subclonal growth remains unclear. Here we utilise 102 macroscopic whole tumour images from clear cell renal cell carcinoma (ccRCC) patients, with matched genetic and phenotypic data from 756 biopsies. Utilising a digital image processing pipeline the boundaries between tumour and normal tissue were marked by a renal pathologist, and positions of boundary line and biopsy regions were extracted to X- and Y-coordinates. The coordinates were then integrated with genomic data to map exact spatial subclone locations, revealing how genetically distinct subclones grow and evolve spatially. A phenotype of advanced and more aggressive subclonal growth was present in the tumour centre, characterised by an elevated burden of somatic copy number alterations, higher necrosis, proliferation rate and Fuhrman grade. Moreover, metastasising subclones were found to preferentially originate from the tumour centre. Collectively these observations suggest a model of accelerated evolution in the tumour interior, with harsh hypoxic environmental conditions leading to heightened cellular turnover and greater opportunity for driver SCNA to arise and expand due to selective advantage. Tumour subclone growth was found to be predominantly spatially contiguous in nature, with subclone dispersal a rare event found in two cases, which notably was associated with metastasis. In terms of genetic events, the largest subclones spatially were dominated by driver somatic copy number alterations, suggesting a large selective advantage can be conferred to subclones upon acquisition of these alterations. In conclusion, spatial dynamics is strongly associated with genomic alterations and plays an important role in tumour evolution.

Introduction

Clear cell renal cell carcinoma (ccRCC) has a well described evolutionary basis, with loss of chromosome 3p and inactivation of *VHL* representing pathognomonic features of the disease^{1–4}. These founding clonal events are then classically followed by acquisition of widespread intratumour heterogeneity, which is characterised by the accrual of additional subclonal genetic alteration(s) including *SETD2*, *PBRM1* and *BAP1* mutation, and/or somatic copy number loss of chromosome arms including 14q and 9p^{5,6}. This defined sequence of events, coupled with a general lack of extreme mutagenesis⁷ and strong evidence of selection⁸, renders ccRCC an excellent model for exploring the principles of tumour evolution. While genomic and phylogenetic reconstruction analyses have been extensively applied to study tumour evolution to date⁹, the spatial nature of tumour evolution at a macroscopic (whole tumour) level of resolution remain grossly underexplored. In particular, the modes of spatial growth (spatially contiguous versus dispersed), the origin of metastasising subclones (tumour interior versus tumour margin), and the relationship between genomic & spatial distance remain uncharacterised in large patient cohorts¹⁰. Moreover, in order to achieve a more complete understanding of cancer evolution within its natural ecological context¹¹, spatially and genomic resolved analyses of clinical cohorts are urgently required.

To aid a greater understanding of the genomic and spatial basis of tumour evolution, we implemented a novel image processing pipeline within our multi-centre prospective longitudinal cohort study named TRAcking renal cell Cancer Evolution through Therapy (TRACERx Renal, <https://clinicaltrials.gov/ct2/show/NCT03226886>). As a feature of the study protocol, macroscopic images were taken of whole tumour specimens at the time of surgery, which were then subject to quality control assessment and boundary annotation by a renal pathologist, digital image processing and integration with dense genomic/clinical data. In total, post quality control spatial imaging data, high-depth next generation sequencing results and Ki67/necrosis/pathological grade scores were captured from 756 tumour biopsies taken from 79 ccRCC primary tumour sections (mean 9.6 biopsies per tumour section). In addition, matched metastatic sequencing data was utilised from 79 biopsies, taken from 32 ccRCC metastatic tumours. Here we report findings from our integrated analysis, aiming to identify the spatial origin of metastasising subclones within the primary tumour, as well as determining the dominant mode of spatial growth (contiguous versus dispersed). In addition we assess the relationship between spatial and genetic distance, identify driver genotypes occupying maximal spatial area, and finally we utilise agent based modelling to explore the selective pressure of the local microenvironment.

Results

Macroscopic image workflow and data integration pipeline

Macroscopic primary tumour photos ($n = 102$) were taken before and after multi-region biopsy sampling, for $n=101$ ccRCC patients from the renal TRACERx100 cohort⁵. Photos were taken on the day of surgery from directly above the tumour, capturing the entire tumour specimen and any surrounding normal tissue (**Fig. 1a**). Quality control review of imaging data was conducted by a renal pathologist (J.I.L.), and images were rejected based on three criteria: i) image did not fully capture the entire tumour specimen ($n=9$), ii) the boundary between tumour and normal tissue could not be reliably determined ($n=5$), or iii) exact spatial position of any biopsy regions could not be reliably determined ($n=9$). Post quality control review $n=79$ images were retained for further analysis. The remaining stages in the image processing workflow comprised: tracing of the tumour/normal boundary lines (by pathologist), automated digital capture of (X,Y) coordinates for biopsy and boundary locations, and finally integration of tumour width/length dimensions from the accompanying pathology reports (**methods**) (**Fig. 1a**). This process yielded two-dimensional boundary shape and biopsy coordinates for $n=79$ tumour sections (**Fig. 1b**, **Fig. S1**). Spatial coordinate data was then integrated with previously published⁵ high depth multi-region biopsy ($n=756$) sequencing data (median coverage 613X, driver gene panel, $n=110$ genes, **methods**), together with biopsy level Ki67 immunohistochemistry (IHC) staining to measure proliferation, and pathologist (J.I.L.) scored biopsy level tumour necrosis and pathological Fuhrman grade data (**Fig. 1b**). Across the entire dataset the median tumour size (longest dimension) was 90mm (range 14mm to 180mm) (**Fig. 1c**), median distance from biopsy to boundary was 9mm (range 1mm to 50mm) (**Fig. 1d**) and median distance between biopsies was 34mm (range 2mm to 149mm).

SCNA burden, necrosis and pathological grade are higher at the tumour centre

We began the analysis by assessing phenotypic differences between the tumour centre and tumour margin. Biopsies were classified as being at the tumour margin if they were 10mm or less from the nearest tumour/normal boundary point (**Fig. 2a**), or otherwise they were classed as centrally located. As a control, all analyses were repeated across a range of possible boundary threshold definitions. These produced consistent results, ruling out the possibility of bias due to the threshold used (**Fig. S2**). As a first step we assessed any differences in the somatic copy number alteration (SCNA) burden (as measured using the wGII score¹²) between tumour centre and margin since the prognostic significance of SCNAs has been demonstrated previously⁵. Surprisingly we found SCNA burden to be significantly higher in the tumour centre as compared to tumour margin ($p=0.0016$, **Fig. 2b**). This difference remained significant after correcting for potential confounding factors including tumour size and purity (**Table S1**). Next we assessed for a difference in Ki67, as marker of cellular proliferation, and similarly this was found to be significantly higher in the tumour centre ($p=0.0023$, **Fig. 2c**). Necrosis, a feature associated with tumour aggressiveness¹³, harsh hypoxic environmental conditions¹³ and poor prognosis¹⁴, was additionally found (as expected) to be significantly higher in the tumour centre ($p=0.006$, **Fig. 2d**). Finally, the proportion of Fuhrman grade 4 regions, also a clinical predictor of poor prognosis, was found to be higher in the tumour centre as compared to tumour margin (**Fig. 2e**).

ccRCC metastasising clones are preferentially located in the tumour centre

A key benefit of the renal TRACERx100 cohort is matched primary and metastatic biopsy sequencing data, and accordingly we obtained genetic data from an additional $n=79$ biopsies, taken from $n=32$ sites of metastatic disease. Using previously published phylogenetic data¹⁵, we linked genetically defined tumour subclones present in metastases to their spatial location within each primary tumour. Metastasising subclones were found to be significantly enriched in the tumour centre, as compared to the tumour margin ($p=0.006$, **Fig. 2f**). We note the enrichment of metastasising subclones within the tumour centre is independent of subclone size as a confounding factor, as the test is based on the proportion of metastasising versus non-metastasising subclones at each location (i.e. if the pattern is simply due to metastasising subclones being larger they would be expected to equally occupy a high proportion in both locations) (**Fig. 2f**) Collectively these results suggest a phenotype of advanced and accelerated subclonal growth present in the tumour centre, characterised by elevated wGII, Ki67, necrosis, Fuhrman grade and heightened metastatic potential (**Fig. 2g**). Representative cases were next mapped in detail, to validate that patterns of metastasis from the tumour centre can clearly be observed. For example, in case K153 (**Fig. 3a**) the metastasising subclone shown in purple is centrally located, whereas subclones from the opposing branch are located close to the tumour margin but fail to metastasise. Similar patterns are observed in cases K180, K280 and K446 (**Fig. 3b-d**). Across the four representative cases (**Fig. 3a-d**) a dynamic range of growth patterns are observed, with later (terminal) phylogenetic clones often found in the tumour centre as well as tumour margin, and overall patterns appear to deviate from a simple model of linear growth from the tumour centre outwards. We note the pattern of metastasis from the tumour centre is an enrichment rather than an absolute rule, and metastasising clones are also capable of disseminating from the tumour periphery, albeit at a lower frequency than the tumour centre (**Fig. 2f**).

Modelling associates necrosis with acquisition of SCNAs

To further explore the relationship between necrosis, SCNAs and tumour location we utilised a coarse-grained cellular automata computational model to assess whether necrosis could promote an elevated SCNA burden (**methods**). The model was designed to simulate our ccRCC experimental dataset, with driver mutations and SCNA events introduced into individual model voxels (equivalent to 1 mm³ tumour volume) at random, and then either expanding or declining based on selective advantage (**methods**). To assess the effect of higher necrosis we simulated the experimentally observed scenario by setting an elevated rate of necrosis at the tumour centre (probability of necrosis for each voxel = 0.5), as compared to the tumour margin (necrosis probability = 0.0) (**Fig. 4a-c**). Simulated tumours were then allowed to grow, and an *in silico* biopsy sampling protocol was implemented at the end to observe the results (**methods**). As expected higher necrosis rates were observed at the tumour centre (**Fig. 4d**). Finally, we assessed for a relationship between necrosis and SCNA burden and found that more necrotic biopsies indeed had a higher number of SCNAs ($p=0.001$, **Fig. 4e**). This suggests in a model based framework that heightened necrosis at the tumour centre associates with higher cellular turnover and greater opportunity for driver SCNAs to arise and expand due to selective advantage.

Spatial and genetic distance have a weak to moderate correlation

Next, returning to the experimental data, we compared the relationship between pairwise biopsy spatial distances (in mm, as measured on a two-dimensional plane) and biopsy genetic distances (i.e. the degree of genetic similarity/dissimilarity between two biopsies, as measured using Euclidean genomic distance) (**Fig. 5a**). Across the entire cohort, using all possible pairwise combinations, a weak-to-moderate but detectable correlation was observed ($p<2.2\times 10^{-16}$, $\rho=0.31$, **Fig. 5b**). An interesting question relates to how much spatial area a given subclonal driver genotype occupies, based on the assumption that a larger spatial expansion indicates positive selection due to subclonal expansion. To assess this we ranked all subclonal driver SCNA and mutational events according to the maximum spatial distance between any two biopsies bearing the same event, as a measure of the magnitude of area occupied within a tumour by a single mutation/SCNA (**methods**). We found driver SCNAs to dominate (**Fig. 5c**), with nine of the top ten driver events ranked by spatial distance being SCNAs, and these events included SCNAs previously shown to associate with reduced overall survival including 14q and 9p loss¹⁵ (**Fig. 5c**). By contrast the lower ranked events by spatial distance were enriched for mutational events (**Fig. 5c**), and intriguingly *SETD2* mutations were lowest ranked. Spatial area occupied is also a function of event timing, for example a mutation occurring early will naturally propagate to larger size than those occurring late, assuming balanced fitness conditions. Accordingly, we plotted event timing (as measured by subclonal frequency across the cohort) versus spatial distance, and found driver SCNAs to be enriched in the top right quadrant (i.e. late timing but still large spatial distance) (**Fig. 5d**). Thus, in spite of occurring later during tumour evolution, driver SCNA events including 9p and 14q loss still achieved dominant clonal expansion across the tumour surface area, supporting the notion of these events as potent drivers of malignant growth.

Spatially contiguous growth is a dominant feature of ccRCC evolution

To understand the dynamic modes of spatial growth we examined for evidence of either spatially contiguous (i.e. all subclones are located together) or dispersal driven growth (i.e. subclones are dispersed and non-contiguous). Growth was found to be predominantly spatially contiguous in nature, with n=77 tumour sections matching this pattern and a minority (n=2) showing evidence of dispersal (**Fig. 6a-b**). The two tumours showing evidence of clonal dispersal were assessed based on a strict criteria, whereby the dispersed subclone had to be surrounded by an uninterrupted block of genetically distinct subclone(s), mapping to a different branch in the phylogenetic tree. The subclonal differences had to be supported by distinct genetic mutation(s), rather than copy number alterations, which are at a greater risk of violating the infinite sites model¹⁶. Thus, we assume dispersal only when a subclone cannot visibly have reached a given spatial location without dispersing across a genetically different subclone. We acknowledge however our study is only powered to assess growth in the two-dimensional plane, and we cannot discount an abnormal growth pattern where dispersed clones are connected in three-dimensional space. Accepting this limitation, we mapped the two dispersing clones in detail (**Fig. 6c**) and inferred the directions of growth based on the parent->child relationships in the phylogenetic tree (**methods**). Case K234 displayed dispersal from the left hand branch of the phylogenetic tree, crossing the genetically opposed right hand branch and achieving subsequent subclonal expansion in the tumour centre. Intriguingly, the dispersing subclone in this case was also the metastasising subclone, which seeded the tumour thrombus (**Fig. 6c**). In case K163, again dispersal from a subclone on the left hand branch was observed crossing a clone from the right hand side of the tree (**Fig. 6d**). The distance of dispersal in this case was large in nature, with the subclone crossing an area of 63mm in distance, which presumably would make unsampled spatially contiguous growth in the third dimension unlikely. Case K163 was non-metastatic, so analysis of metastasising clones was not applicable.

Irregularity of tumour boundary associates with SCNA burden

Lastly, we sought to assess the importance of boundary shape, based on the hypothesis that a highly irregular boundary shape may reflect sporadic growth driven by tumour intrinsic factors. For each tumour section the boundary circularity score was calculated, using the formula $4\pi \times \text{area} / \text{perimeter squared}$ (**Fig. S3**). A score of 1 indicates perfect circularity, and lower scores indicate increasing irregularity in shape. The median circularity score across the cohort was 0.66 (range 0.25 to 0.81). We plotted driver genes and SCNAs in order of decreasing circularity and did not observe any notable patterns (**Fig. S4**). Similarly, circularity score was not prognostic for either overall or progression free survival ($p=0.82$ and $p=0.33$ respectively, **Fig. S5**).

Discussion

Here, we have conducted an integrated study of spatial evolution in ccRCC, analysing 756 biopsies across 79 primary tumour sections with spatial, genomic, IHC and clinical data. To our knowledge, this is the largest study of its kind to date, and it reveals novel insight into the spatial characteristics of solid

tumour growth. We contrasted subclonal phenotypes in the tumour centre versus margin, and additionally compared modes of spatial growth and area occupied by individual driver events.

A key finding from this work is that metastatic ccRCC subclones preferentially originate from the interior of the tumour rather than the margin, with the tumour centre characterised by higher levels of proliferation, necrosis, Fuhrman grade and somatic copy number alteration (SCNA) burden. While not functionally validated, our computational simulation of these phenomena supports a model whereby heightened necrosis in the tumour interior facilitates greater cellular turnover and increased SCNA acquisition thus creating a reservoir of subclones adapted to harsh microenvironmental conditions. Given these additional competencies, it is thus unsurprising that subclones from the tumour centre are better adapted to seed secondary metastatic sites. Necrotic primary tumour subclones have been shown to have greater propensity to seed metastases in pre-clinical models¹⁷, and clinically a recent systematic meta-analysis across 34 studies demonstrated tumour necrosis to be associated with reduced cancer-specific survival in RCC¹⁴. Biologically, tumour necrosis is classically associated with chronic hypoxia caused by rapid tumour growth and inadequate blood supply¹⁸. In addition hypoxia has a well-established role as a microenvironmental factor promoting metastasis¹⁹. Of separate note, ccRCC is additionally uniquely characterised by oxygen-independent so called “pseudohypoxia” as a universal feature of the disease due to *VHL* inactivation²⁰. Clinically, necrosis is associated with ischemia, either as a chronic process resulting in gradual tissue damage, or as an acute event such as infarction related to arterial or venous thrombosis. Tumour neovascularisation occurs through ingrowth of blood vessels from peripheral areas. These vessels are physiologically abnormal and thus susceptible to thrombosis rendering tumours or parts of them more prone to acute ischemic events. The chronic or acute nature of tumour necrosis may also in turn shape the nature of tumour evolution, either driving gradual subclonal adaption or a punctuated process of rapid cell death and repopulation by adapted subclones. While outside of the scope of this report, further functional study of necrotic events, hypoxia and processes associated with metastasis (such as epithelial-to-mesenchymal transition) in the tumour centre versus tumour margin would be of significant interest. In addition, comparison of the patterns observed here in ccRCC to other solid tumour types will be of future relevance, to understand if these findings are tumour-type specific (i.e. due to the large size of ccRCCs), or more broadly generalisable.

A unique benefit of this study is the integrated analysis of spatial and phylogenetic data, allowing visualisation of tumour subclones within their spatial location. Using this approach, we can show clear examples of metastasis from the tumour centre, as well as begin to understand the patterns and directions of spatial growth. While only descriptive in nature, the overall the patterns of growth do not appear to conform to a simplistic model of linear growth from the interior outwards towards the margins, instead more advanced evolutionary growth (i.e. later phylogenetic subclones) appear to be able to originate from central tumour regions as well as the margin (Fig. 3a-d, Fig. 6). The evolutionary basis of these patterns remains unclear, however it may be possible that the irregular/non-linear growth is as a result of environmental stimulus within the tumour, whereby subclonal genetic diversification and direction of growth is associated with environmental change (subclonal trophism).

Analysis of genetic versus spatial distance revealed a weak-to-moderate correlation, indicating that tumour subclones tend to be more genetically similar to their neighbouring subclones, as compared to ones further away. This result is congruent with the finding discussed below that spatial growth is predominantly spatially contiguous in nature. The weak-to-moderate strength of this correlation ($\rho = \sim 0.3$) suggests exceptions to this rule are frequent however, with clonal sweeps likely to be a key factor causing this correlation to breakdown and more rarely, clonal dispersal. It should be noted a limitation of this study is the panel based sequencing assay, which while well validated for driver mutation and genome-wide SCNA detection, lacks the dense number of genetic markers that would be available with whole exome or genome sequencing. In terms of individual driver events, we found the spatial area occupied was dominated by SCNAs rather than mutations, with loss of 14q, 9p and other known ccRCC SCNAs²¹ ranking highly. This adds a new dimension of evidence to support SCNAs as the primary driver of ccRCC growth, showing for the first time that they occupy large spatial areas of the primary tumour, despite occurring late within the evolutionary timeframe. Collectively these observations suggest a large selective advantage can be conferred to subclones upon acquisition of driver SCNAs, with growth disproportionately occurring after these events occur. This is consistent with previously reported survival analysis, where 9p loss is highly prognostic even in a setting where all cases have already metastasised¹⁵. The driver event associated with minimal subclonal expansion was *SETD2* mutation, with many subclones bearing this genetic event failing to achieve significant spatial expansion (i.e. event found only in one biopsy). By contrast however *SETD2* is a frequently occurring subclonal event⁵, often found under strong parallel selection, indicating some beneficial growth advantage at a microscopic level which then reduces as clone size gets larger.

Lastly, we identified spatially contiguous growth as the dominant mode of ccRCC spatial growth, found in 77 of 79 tumour sections. This observation provides some insight into ccRCC tumour cell motility, suggesting a model of primary subclones growing predominantly within their localised microenvironment, with limited evidence of dynamic seeding between locations in the primary tumour. We note that $n = 32$ of the tumours in our analysis had already achieved metastasis, so clearly subclones were present within the primary tumour with dispersal/motility competence. These observations would suggest that once motility competence is achieved those subclones are better able to seed and establish within secondary sites rather than other spatial locations in the primary. Previous evidence has indicated reseeding back from metastasis to the primary site in prostate cancer²², as well as frequent rates of dispersal within the primary for glioblastoma²³, so tumour specific patterns are likely to be present. In two cases dispersal was identified and was descriptively found to occur in the same subclone which achieved metastasis. In a second case dispersal was identified to have occurred across a large spatial distance, suggesting possible spread via the tumour vasculature. We note a limitation of this study is a lack of three-dimensional sampling.

In summary, here we present the largest study to date mapping tumour subclones to their spatial locations within the primary tumour site. This study reveals novel insight into the metastatic process, indicating that ccRCC subclones preferentially originate from the tumour interior. This observation is

supported by elevated SCNA burden, higher Fuhrman grade and higher necrosis rates in the tumour centre. Future studies resolved for spatial tumour location and subclonal structure are likely to continue to yield significant insight into the cancer evolutionary process.

Methods

Patient recruitment and sample collection

Patients were from the TRACERx Renal trial (NCT03226886), a multi-centre prospective longitudinal cohort study. The inclusion and exclusion criteria, along with full clinical, histological and follow-up data were described previously^{5,15}. In brief, primary tumours were dissected along the longest axes and spatially separated regions sampled from the “tumour slice” using a 6mm punch biopsy needle. Primary tumour regions were labelled as R1, R2, R3... Rn with locations recorded.

Macroscopic photo acquisition, quality control and pathological review

Macroscopic photos (n = 102) were taken both before and after multi-regional sampling and were taken from directly above the tumour. Photos were then assessed by a renal pathologist (JIL) for quality control to ensure high-quality tumour images including the tumour boundary were visible. Cases in which the whole tumour couldn't be imaged to high quality (n = 9), without a clearly defined boundary between tumour and normal tissues (n = 5), or without exact positions of the biopsy regions (n = 9) were excluded. In total, 79 tumour sections of 66 unique primary tumours were included in this study.

Boundary drawing and digital tumour map processing

Boundaries between tumour and normal tissues were marked independently by a renal pathologist (JIL), after which the boundaries, along with the biopsy regions were extracted into X- and Y-coordinates using WebPlotDigitizer²⁴. 2-dimensional size in millimetres of the tumours was set in accordance with the pathology reports after surgery. For those cases with only maximum length available, the length along the axis perpendicular to that of the maximum length was re-scaled based on the length on the photos. Digital tumour maps were then generated using R (version 3.6.2)²⁵.

Microscopic pathology review, immunohistochemical staining and digital image analysis of Ki67

As previously described, histological sections, along with the immunohistochemical staining and digital image analysis of Ki67 of each region in each case were evaluated by the same pathologist (JIL)^{5,15}. Briefly, tumour type, tumour architecture, tumour grading and the presence of necrosis and microvascular invasion were determined. Immunohistochemical staining of Ki67 was performed on 468 biopsy regions using a fully automated system and ready-to-use optimized reagents according to the manufacturer's recommendations (Ventana Discovery Ultra, Ventana, Arizona, USA). Mouse intestine tissue was used as positive control, and regions containing tumour tissue were identified and marked by a pathologist and

subsequently scanned into digital images, which were then subjected to automated image analysis for Ki67 quantification. Results were depicted as total percentage of Ki67-positive nuclei.

Targeted sequencing and genomic data processing

DNA extraction and multi-regional panel sequencing were as previously described^{5,15}. Driver gene panels (Panel_v3, Panel_v5 and Panel_v6) were used in this study. Panel_v3, Panel_v5 and panel_v6 included 110, 119 and 130 putative driver genes, respectively. Driver genes were selected from genes that are frequently mutated in TCGA and other studies²⁻⁴. Mutation and somatic copy number alteration (SCNA) calling, clonality estimation of genomic alterations and phylogenetic tree reconstruction were done as previously described^{5,15}. Subclonal frequency was defined as the proportion of an event that appeared to be subclonal across all samples. A matrix with presence and absence of nonsynonymous and synonymous point mutations, dinucleotide substitutions (DNVs), INDELs and arm level SCNAs was created for each tumour, and all the events were clustered as previously described^{5,15}.

Definition of tumour centre and tumour margin

Tumour centre was defined as the area within the tumour where the distance between a biopsy region and the tumour boundary was more than or equal to 1cm, while tumour margin was defined as the area within the tumour where the distance between a biopsy region and the tumour boundary was less than 1cm (Fig. 2a).

Spatial and genomic distance calculation

Spatial distance between biopsy regions was measured using the Euclidean metric (Fig. 3a). Each biopsy region was represented as a point on the tumour map, with the X- and Y-coordinates referring to the actual position on the macroscopic photos. Distance between each biopsy region to the tumour boundary was calculated as the Euclidean distance between the biopsy region and its nearest boundary.

Genomic distance was calculated by taking the Euclidean distance of the detected genomic events (non-synonymous mutations and SCNAs). Mutations and SCNAs of all biopsy regions on a slice were represented as a matrix with 1 defined as an existing event and 0 as a non-existing event. Pairwise Euclidean distances were then calculated between all biopsy regions.

Calculation of space occupied by driver events

Driver event status of all biopsy regions was encoded in a matrix with 1 defined as an existing event and 0 as a non-existing event. The space occupied by a driver event was defined as the maximum spatial distance between any two regions containing the same driver event on a tumour slice, measured in millimetres.

Inference of metastasising routes on tumour maps

A matrix for all biopsy regions of all samples showing the relationship of clusters and biopsy regions was generated. Clusters containing nonsynonymous and synonymous point mutations, DNVs, INDELs and arm level SCNAs that appear in all the biopsy regions in a tumour were considered truncal clones, while clusters containing nonsynonymous and synonymous point mutations, DNVs, INDELs and arm level SCNAs that appear only in part of the biopsy regions in a tumour were considered subclones. The parent and child relationship of clusters were based on whether the genetic alterations each cluster contained were detected in other clusters. A secondary subclone was a cluster that contained not only point mutations, DNVs, INDELs and arm level SCNAs detected in the truncal clone, but also point mutations, DNVs, INDELs or arm level SCNAs that were not found in the truncal clone. A third subclone was a cluster that contained not only those genomic alterations detected in a secondary subclone, but also genomic alterations that were not found in the secondary subclone. The parent-and-child relationship went on and on until it reached a terminal clone which harboured genomic alterations that were not shared by any other clusters. Metastasising routes were inferred on tumour maps based on this parent-and-child relationship. Levels of the biopsy regions were set depending on the highest-level subclone it contained. Regions containing a parent clone as the highest-level subclone would point to regions containing a child clone as the highest-level subclone, while regions containing same levels of subclones were lined without arrows.

Definition of clonal expansion patterns

Types of clonal expansion patterns were defined based on the metastasising routes on the tumour maps. Contiguous growth was defined as the growth pattern where nearest subclones from the same branch of the phylogenetic tree were lined one by one, without any subclones from the other branch in between. Clonal dispersal was defined as the growth pattern where nearest subclones from the same branch of the phylogenetic tree crossed the border formed by the subclones from the other branch.

Circularity score calculation

Circularity score was calculated to assess the regularity of tumour boundaries:

$$\text{Circularity score} = \frac{4\pi * \text{Area}}{\text{Perimeter}^2}$$

where area and perimeter of the tumour slice were captured based on the tumour maps using ImageJ version 1.52q. A score of 1 indicates perfect circularity, and lower scores indicate increasing irregularity in shape.

Computational modelling

Simulation of tumour growth and evolution, as determined by the 26 ccRCC drivers, and in the context of necrosis, was achieved by the development of a coarse-grained cellular automata model (12 driver genes and 14 SCNAs). A model tumour comprises units called tumour voxels, each of which reflects a volume

of 1 mm^3 . In the basic model, growth takes place in the form of probabilistic voxel duplication. A tumour voxel can only duplicate if surrounded by at least one empty site. Upon growth, a tumour voxel randomly acquires additional driver events.

The growth probability of a tumour voxel in the model further depends on the drivers harboured within. Two levels of relative driver advantage in growth are set and reflect the ranking of driver advantages by their association with regional Ki67 expression in the TRACERx Renal study. Four SCNAs (1q gain, 7q gain, 4q loss, and 9p loss) endow a tumour voxel with the strongest growth advantage; a tumour voxel harbouring any of these SCNAs grows with a probability of 1.0 every simulation step. Another five SCNAs (8q gain, 12p gain, 20p gain, 1p loss, and 14q loss) endow a tumour voxel with a lesser growth advantage; a tumour voxel harbouring any of these SCNAs, but not none of the above four strongest SCNAs, grows with a probability of 0.5 every simulation step. If a tumour voxel doesn't harbour any of these nine drivers, it grows with a probability of 0.25 every simulation step.

The acquisition probability of driver gene mutations is 1×10^{-4} per simulation step. Mutations in two driver genes, *BAP1* and *PBRM1*, are defined to promote the acquisition of SCNAs. In tumour voxels harbouring these two mutations, the acquisition probability of SCNA is also 1×10^{-4} per simulation step. Otherwise, the acquisition of SCNA is defined to take place with a probability of 1×10^{-7} per simulation step. The loss of 3p and *VHL* in the model are defined to be always truncal and already harboured in the first tumour voxel.

Extending the basic model, we implemented central necrosis to explore the impact of enhanced cell death at the tumour core on tumour growth and, of particular interest, the evolution of SCNA-harbouring subclones. Specifically, we defined that all tumour voxels located farther than 1.5 cm from the tumour surface become necrotic with a probability of 0.5 every simulation step. The chosen distance threshold of 1.5 cm reflects the clinical observations of distances between any regional biopsies having necrosis and the lack of necrosis at the tumour contour (mean: 12.83 mm; range: 3.71 – 27.11 mm). If a tumour voxel is chosen to undergo necrosis, the site occupied by that tumour voxel is labelled as necrotic; consequently, adjacent tumour voxels are allowed to grow and fill in the necrotic site.

Each simulation is performed in three dimensions (3D) until reaching a size of at least 1 million tumour voxels after the last simulation step. Upon completion, a two-dimensional (2D) slice is collected from the middle of the 3D simulated tumour in 3D. Within the tumour slice in 2D, regional biopsies (each with a size reflecting a diameter of 1 cm) are collected uniformly, with a spacing of 1 cm. The necrotic status and driver composition within each regional biopsy, from these simulations, are recorded for further analysis.

The necrotic fraction of a regional biopsy is calculated as the number of tumour voxels that are necrotic divided by the total number of tumour voxels in that biopsy. For the comparison of the number of SCNA within each region, all regional biopsies are ranked according to the necrotic fraction and grouped into “less necrotic” (necrotic fraction between 0 and 0.4) and “more necrotic” (necrotic fraction between 0.4

and 0.8) categories. Due to the small amount of alive tumour voxels, biopsies with a necrotic fraction greater than 0.8 were dropped from the analysis. The number of SCNA s in a regional biopsy is defined as the number of unique SCNA s found in any tumour voxels in that biopsy.

Statistical analyses

Non-parametric Wilcoxon signed-rank test was used to compare medians of groups of continuous variables. Fisher's exact test was used to compare proportions of counts of categorical variables. Non-parametric Spearman's rank correlation coefficient was used to assess the relationship between pairs of variables, as opposed to testing for a difference in medians between two continuous variables. Kaplan-Meier survival curves and log-rank tests were used to analyse patients' survival. Progression-free survival was defined as the time to disease recurrence or relapse, or death without disease recurrence, whichever came first. Overall survival was defined as the time to cancer specific death. Significance level was set 0.05. All statistical analyses were conducted using R version 3.6.2.

Declarations

Code availability statement

Code used for analyses is available at: <https://github.com/kevlitchfield1>.

Funding

K. Litchfield is funded by the UK Medical Research Council (MR/P014712/1), Rosetrees Trust (A2437) and Cancer Research UK (C69256/A30194). J.I.L is funded by MINECO, Spain grant: SAF2016-79847-R. X.F., P.A.B., and E.S. are funded by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001003, FC001144), the UK Medical Research Council (FC001003, FC001144) and the Wellcome Trust (FC001003, FC001144). S.T. is a Cancer Research UK clinician scientist and is funded by Cancer Research UK (grant reference number C50947/A18176) and the National Institute for Health Research (NIHR) Biomedical Research Centre at the Royal Marsden Hospital and Institute of Cancer Research (grant reference number A109). C. S. is a senior Cancer Research UK clinical research fellow and is funded by Cancer Research UK (TRACERx), the Rosetrees Trust, NovoNordisk Foundation (ID 16584), EU FP7 (projects PREDICT and RESPONSIFY, ID: 259303), the Prostate Cancer Foundation, the Breast Cancer Research Foundation, the European Research Council (THESEUS) and National Institute for Health Research University College London Hospitals Biomedical Research Centre.

Author's contributions

Study design: K.L., S.T., C.S., P.A.B.

Data analysis: Y.Z, X.F., J.I.L., M.A., S.H., H. X.

Data contribution and interpretation: All

Manuscript writing: K.L., Y.Z., X.F.

Competing Interests

K.L., S.T., and C.S. have a patent on indel burden and checkpoint inhibitor response pending, and a patent on targeting of frameshift neoantigens for personalised immunotherapy pending. K. L. reports speaker fees from Roche Tissue Diagnostics. ST reports grants from Ventana, outside the submitted work. J.L. reports institutional research support from: BMS, MSD, Novartis, Pfizer, Achilles Therapeutics, Roche, Nektar Therapeutics, Covance, Immunocore, Pharmacyclics, Aveo, and consultancy support from: Achilles, AZ, Boston Biomedical, BMS, Eisai, EUSA Pharma, GSK, Ipsen, Imugene, Incyte, iOnctura, Kymab, Merck Serono, MSD, Nektar, Novartis, Pierre Fabre, Pfizer, Roche / Genentech, Secarna, Vitaccess. C.S. receives grant support from Pfizer, AstraZeneca, BMS, Roche-Ventana, Boehringer-Ingelheim, and Ono. C.S. has consulted for Pfizer, Novartis, GlaxoSmithKline, MSD, BMS, Celgene, AstraZeneca, Illumina, Genentech, Roche-Ventana, GRAIL, Medicxi, and the Sarah Cannon Research Institute. C.S. is a shareholder of Apogen Biotechnologies, Epic Bioscience, GRAIL, and has stock options in and is co-founder of Achilles Therapeutics.

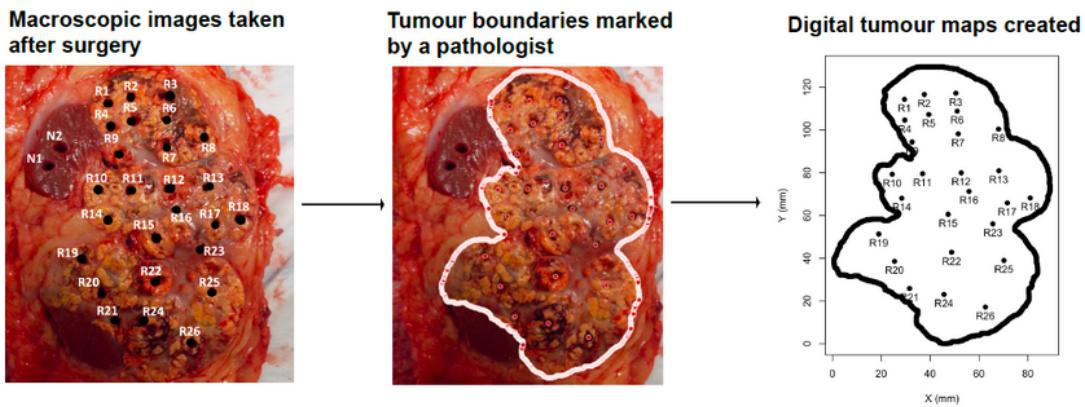
References

1. Cancer Genome Atlas Research, N. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499**, 43 – 9 (2013).
2. Dalgliesh, G.L. *et al.* Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360–3 (2010).
3. Scelo, G. *et al.* Variation in genomic landscape of clear cell renal cell carcinoma across Europe. *Nat Commun* **5**, 5135 (2014).
4. Sato, Y. *et al.* Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat Genet* **45**, 860–7 (2013).
5. Turajlic, S. *et al.* Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. *Cell* **173**, 595–610 e11 (2018).
6. Gerlinger, M. *et al.* Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet* **46**, 225–233 (2014).
7. Priestley, P. *et al.* Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* **575**, 210–216 (2019).
8. Williams, M.J., Werner, B., Barnes, C.P., Graham, T.A. & Sottoriva, A. Identification of neutral tumor evolution across cancer types. *Nat Genet* **48**, 238–244 (2016).
9. Gerstung, M. *et al.* The evolutionary history of 2,658 cancers. *Nature* **578**, 122–128 (2020).
10. Hoefflin, R. *et al.* Spatial niche formation but not malignant progression is a driving force for intratumoural heterogeneity. *Nat Commun* **7**, ncomms11845 (2016).

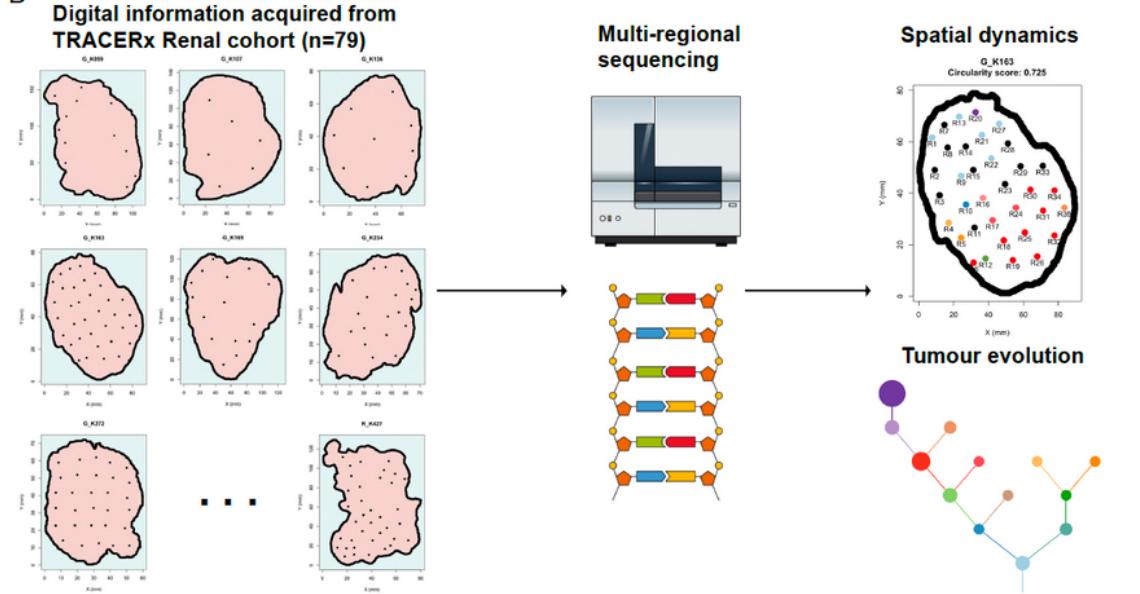
11. Axelrod, R. & Pienta, K.J. Cancer as a Social Dysfunction-Why Cancer Research Needs New Thinking. *Mol Cancer Res* **16**, 1346–1347 (2018).
12. Endesfelder, D. *et al.* Chromosomal instability selects gene copy-number variants encoding core regulators of proliferation in ER + breast cancer. *Cancer Res* **74**, 4853–4863 (2014).
13. Bredholt, G. *et al.* Tumor necrosis is an important hallmark of aggressive endometrial cancer and associates with hypoxia, angiogenesis and inflammation responses. *Oncotarget* **6**, 39676–91 (2015).
14. Zhang, L. *et al.* Tumor necrosis as a prognostic variable for the clinical outcome in patients with renal cell carcinoma: a systematic review and meta-analysis. *BMC Cancer* **18**, 870 (2018).
15. Turajlic, S. *et al.* Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal. *Cell* **173**, 581–594 e12 (2018).
16. Ma, J. *et al.* The infinite sites model of genome evolution. *Proc Natl Acad Sci U S A* **105**, 14254–61 (2008).
17. Bonfil, R.D., Bustuoabad, O.D., Ruggiero, R.A., Meiss, R.P. & Pasqualini, C.D. Tumor necrosis can facilitate the appearance of metastases. *Clin Exp Metastasis* **6**, 121–9 (1988).
18. Kumar, V., Abbas, A.K., Aster, J.C. & Robbins, S.L. Robbins basic pathology. in *Robbins Pathology* 1 online resource (xii, 910 pages) (Elsevier Saunders,, Philadelphia, PA, 2013).
19. Rankin, E.B. & Giaccia, A.J. Hypoxic control of metastasis. *Science* **352**, 175 – 80 (2016).
20. Macklin, P.S. *et al.* Recent advances in the biology of tumour hypoxia with relevance to diagnostic practice and tissue-based research. *J Pathol* **250**, 593–611 (2020).
21. Beroukhim, R. *et al.* Patterns of gene expression and copy-number alterations in von-hippel lindau disease-associated and sporadic clear cell carcinoma of the kidney. *Cancer Res* **69**, 4674–81 (2009).
22. Gundem, G. *et al.* The evolutionary history of lethal metastatic prostate cancer. *Nature* **520**, 353–357 (2015).
23. Foyt, R.A. Tumor cohesion and glioblastoma cell dispersal. *Future Oncol* **9**, 1121–32 (2013).
24. Rohatgi, A. (2020) WebPlotDigitalizer: HTML5 based online tool to extract numerical data from plot images. Version 4.3. <https://automeris.io/WebPlotDigitizer>. (accessed in July 2020).
25. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Figures

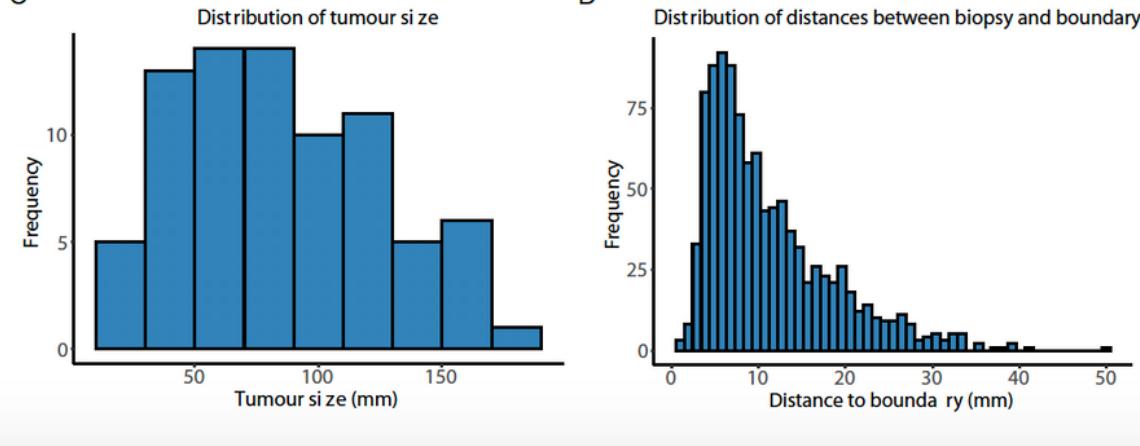
A



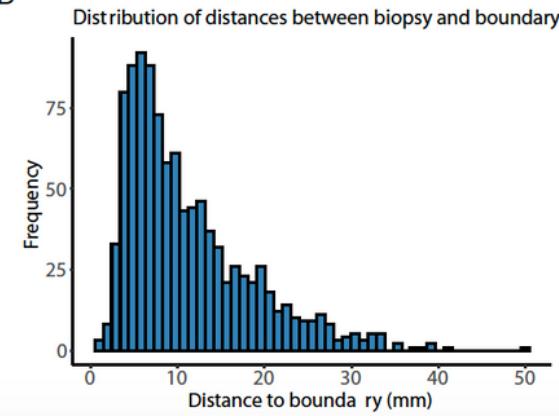
B



C



D

**Figure 1**

Study overview. A) Macroscopic images taken after surgery were reviewed by a pathologist to determine the physical boundary between tumour and normal tissues, and spatial information was subsequently gathered to generate digital tumour maps. B) Spatial data were combined with multi-regional sequencing data to study spatial dynamics and tumour evolution. C) Distribution of tumour size measured in

millimetres. D) Distribution of distances between each biopsy region and its nearest tumour boundary on the tumour maps, measured in millimetres.

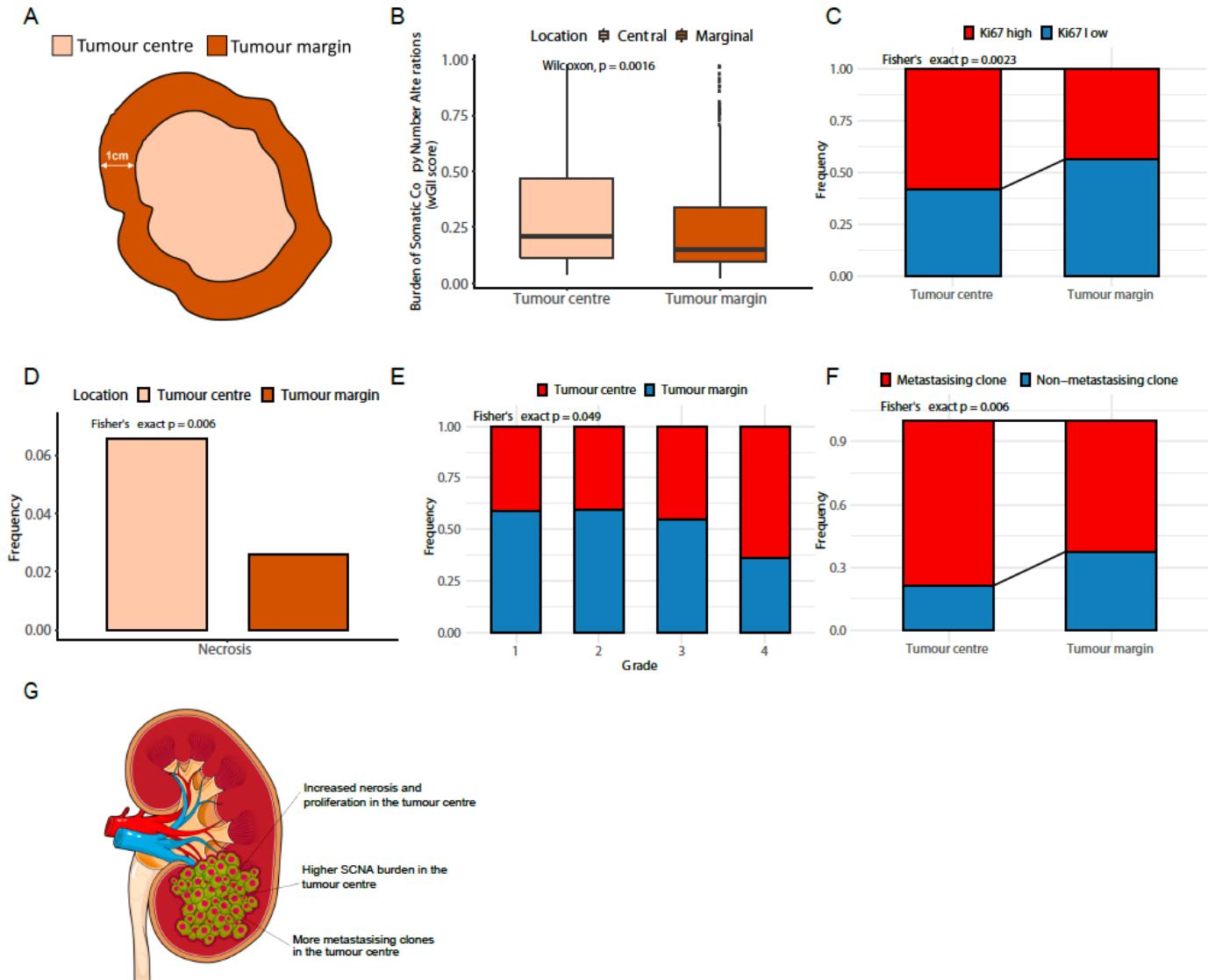


Figure 2

Comparison between regions in the tumour centre and regions in the tumour margin. A) Definition of tumour centre and tumour margin. Biopsies less than 1 centimetre to its boundary were defined as marginal regions, while biopsies more than 1 centimetre to its boundary were defined as central regions. B) Comparison of weighted genome integrity index (wGII) which is a measurement of burden of somatic copy number alterations between tumour centre and tumour margin. C) Comparison of Ki67 between tumour centre and tumour margin. D) Comparison of frequency of necrotic regions between tumour centre and tumour margin. E) Comparison of frequencies of regions of different Fuhrman grades in clear cell renal cell carcinoma between tumour centre and tumour margin. F) Comparison of frequencies of metastasising clones and non-metastasising clones between tumour centre and tumour margin in a subset of metastatic tumours. G) A cartoon summarizing the differences observed between tumour

centre and tumour margin, which might be explained by the differences of blood supply of different parts of the tumour.

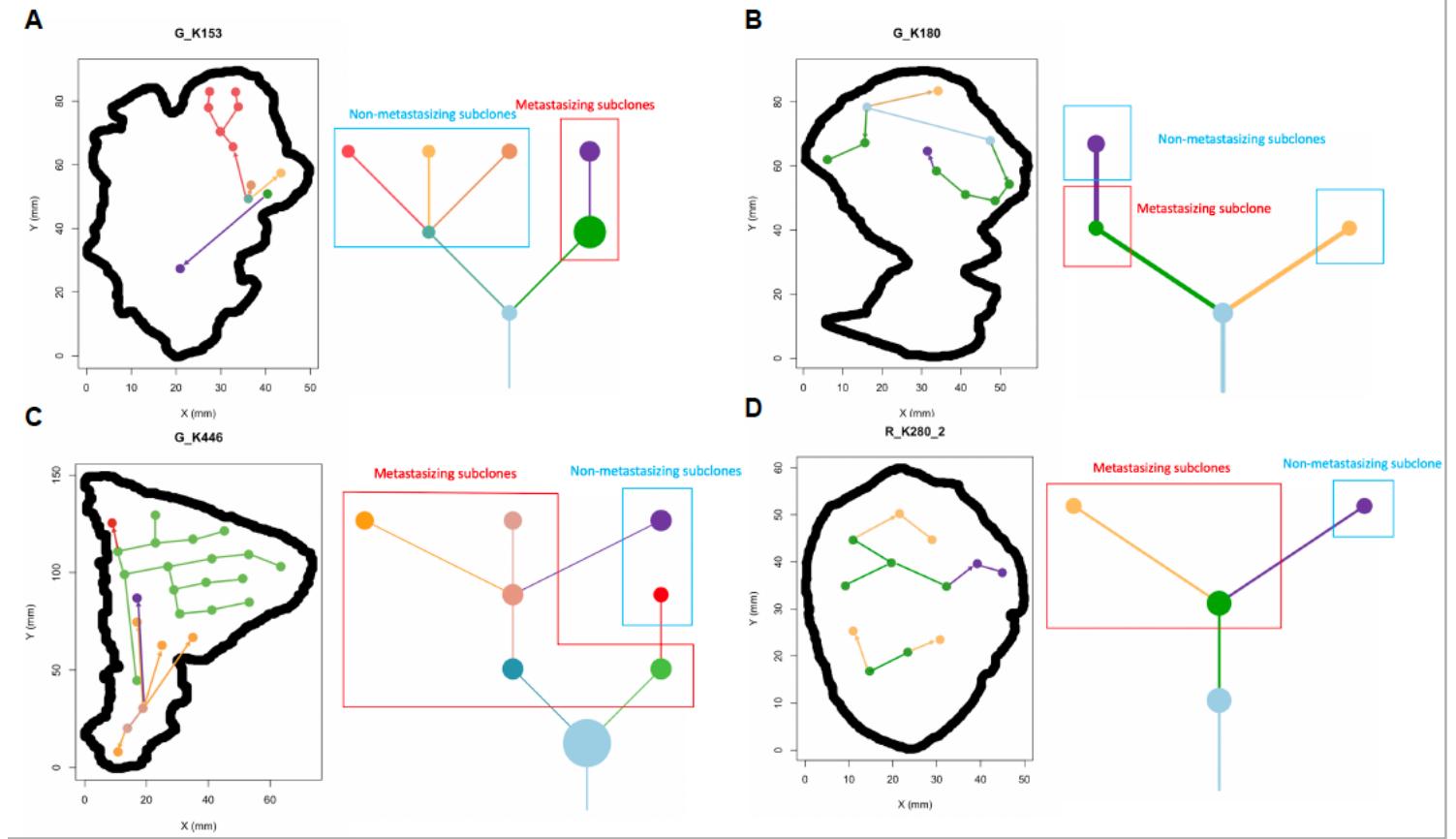


Figure 3

Phylogenetic trees and tumour maps showing the possible metastasising routes of 4 example cases on a 2-dimention level. A) In case K153, metastasising subclones and non-metastasising subclones originate from different branches of the phylogenetic tree. Regions containing metastasising subclones are more enriched in the tumour centre. B) In case K180, the 1st-level metastasising subclone and non-metastasising subclone originate from different branches of the phylogenetic tree, while the 1st-level metastasising subclone also gives rise to the 2nd-level non-metastasising subclone. Regions containing metastasising subclones are more enriched in the tumour centre. C) In case K446, the non-metastasising subclones are all originated from their parent metastasising subclones. Regions containing metastasising subclones are more enriched in the tumour centre and occupy more space. D) In case K280_2, the 1st-level metastasising subclone gives rise to a 2nd-level metastasising subclone and a 2nd-level non-metastasising subclone. Regions containing metastasising subclones are more enriched in the tumour centre and occupy more space.

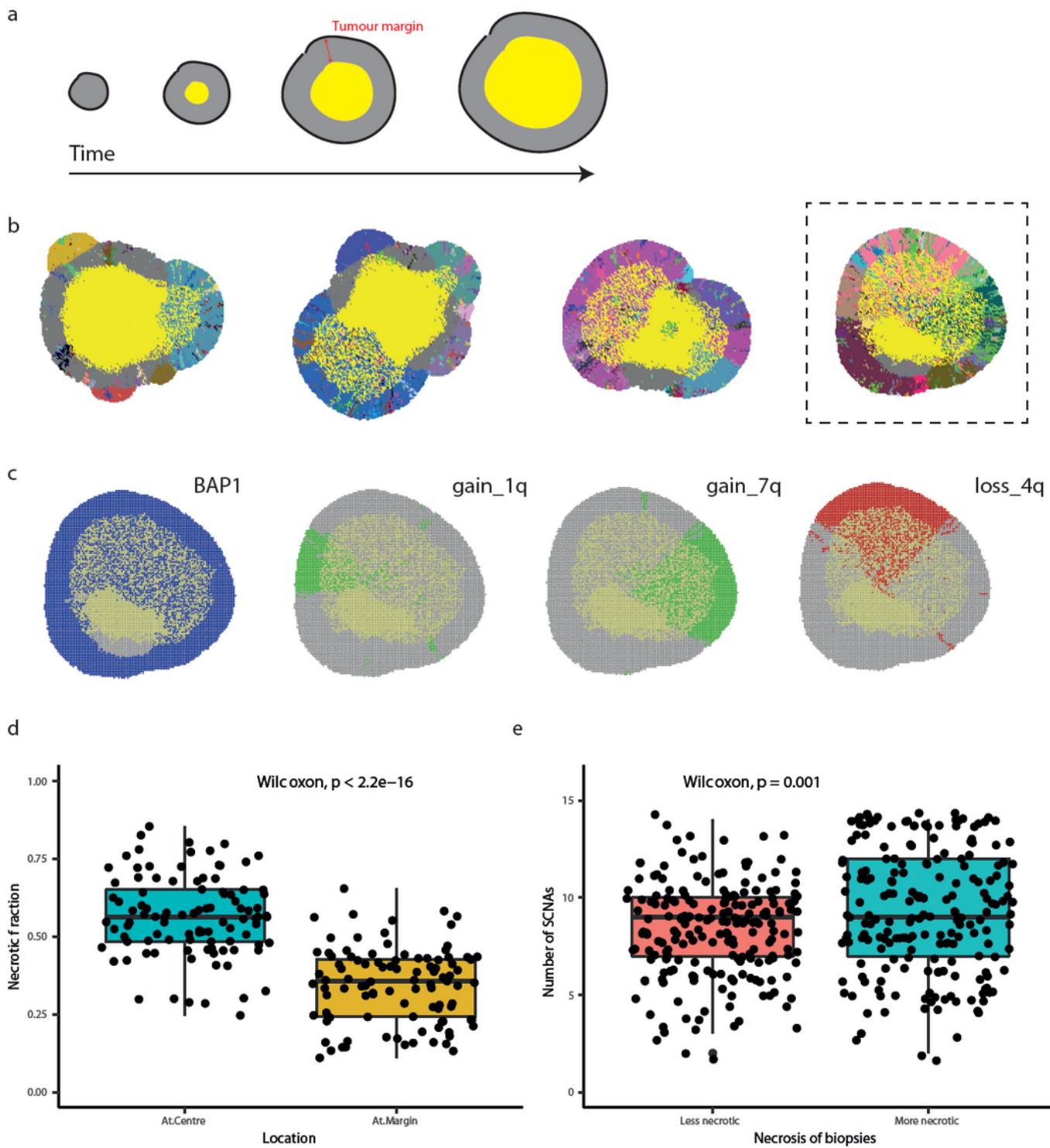


Figure 4

Computational modelling supports the preferential localisation of SCNA subclones in more necrotic regions of a tumour. (a) Schematic illustrating the emergence of necrosis in a growing tumour. Tumour areas in yellow experience a high death rate due to necrosis. (b) Representative simulated tumours and patterns of subclones. Tumour areas in yellow are necrotic; founder clone is in grey and other subclones are in randomly assigned colours. Dashed square highlights the model replica selected to show some

drivers in (c). (c) Spatial patterning of BAP1, 1q gain, 7q gain, and 4q loss in a representative simulated tumour. (d) Necrotic fraction in 20% central-most (“At.Centre”) and marginal-most (“At.Margin”) regional biopsies. (e) Number of SCNA in less and more necrotic regional biopsies. “Less necrotic” biopsies refer to the least necrotic 40% of all regional biopsies while “More necrotic” biopsies refer to the next 40% biopsies in the ranking. The 20% most necrotic biopsies are excluded from analysis.

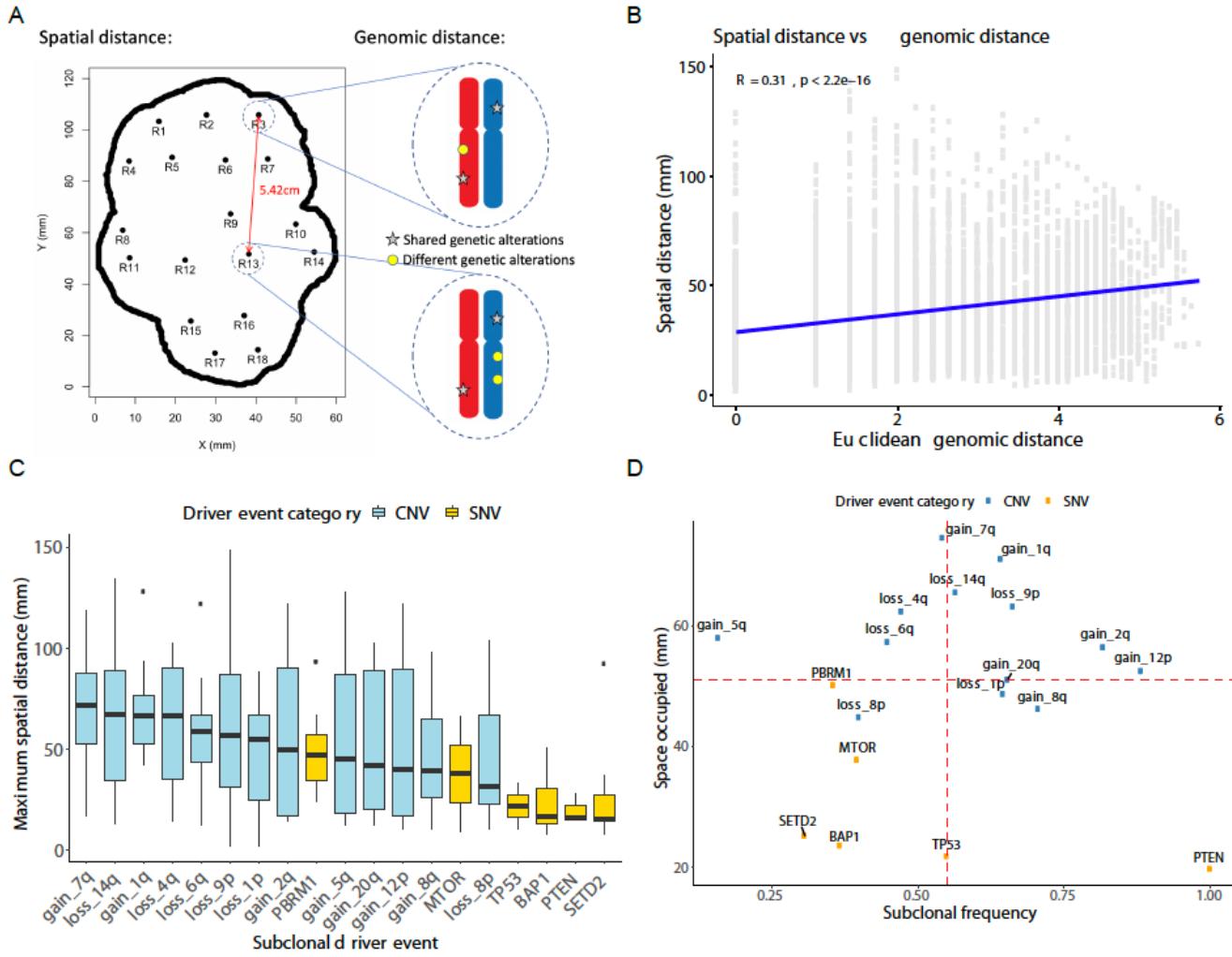


Figure 5

Integrated analysis of genomic and spatial distances. A) A cartoon showing how spatial and genomic distances were measured. Spatial distances between each two biopsy regions were calculated using the spatial data previously retrieved, and genomic distances between each two biopsy regions were calculated based on the genomic alterations they harboured using Euclidean metrics. B) Correlation between spatial distance measured in millimetres and genomic distance across all samples is shown. C) Maximum spatial distance occupied by each subclonal driver event measured in millimetres is shown, ordered from the highest to the lowest. D) Maximum spatial distance occupied in millimetres versus subclonal frequency, showing the relationship between space occupied and timing of the event.

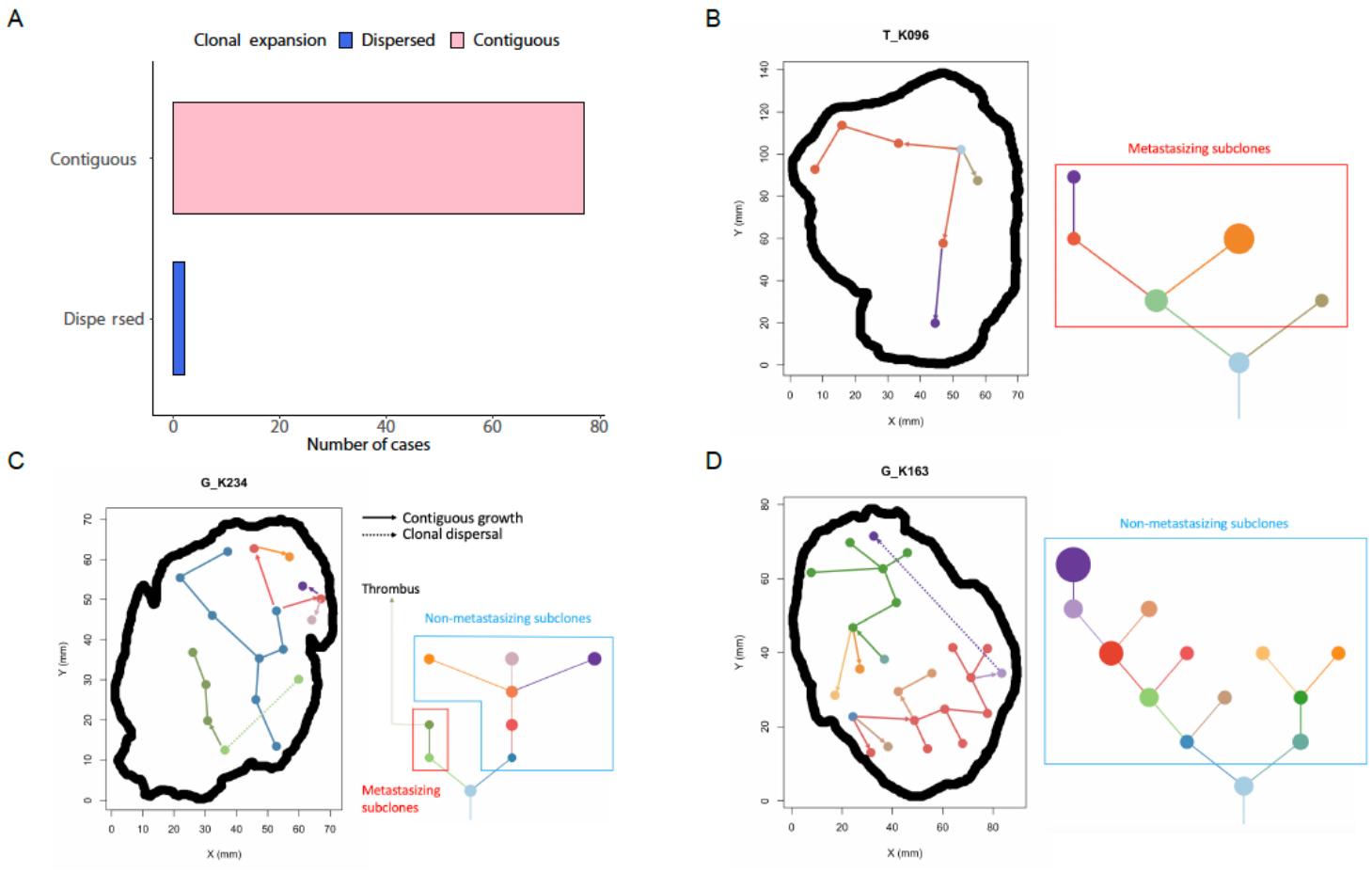


Figure 6

Phylogenetic trees and tumour maps showing the inferred clonal expansion pattern and tumour metastasising routes. A) Most cases had a contiguous clonal expansion, while one case of clonal dispersal was seen in 2 out of 79 samples. B) A tumour map and a phylogenetic tree of case K096 showing a pattern of contiguous clonal expansion. C) Tumour maps and phylogenetic trees of cases G_K234 and D) K163 showing clonal dispersal (dotted line). In case K234, the dispersing clone formed the metastasising branch and gave rise to its descendant clone which was also found in a venous thrombus. In case K163, the dispersing clone wasn't associated with metastasis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementaryfile.pdf](#)