Title:
Carbon dating cancer: defining the chronology of metastatic progression in a colorectal cancer case

Authors:

These authors contributed equally as first authors to this work.

These authors contributed equally as senior authors to this work.

Affiliations:
1 Division of Molecular Pathology, The Institute of Cancer Research, Sutton, UK
2 Gastrointestinal cancers and Lymphoma Unit, The Royal Marsden NHS Trust, Sutton, UK
3 Centre for Evolution and Cancer, The Institute of Cancer Research, London, UK
4 Department of Oncology, Crosshouse Hospital, Crosshouse, Kilmarnock, UK.
5 Beatson West of Scotland Cancer Centre, Glasgow, UK.
6 Department of Pathology, Southern General Hospital, Glasgow, UK
7 Department of Pathology, Crosshouse Hospital, Crosshouse, Kilmarnock, UK
8 ARC-NET Research Centre University of Verona, Verona, IT
9 Center for Evolution and Cancer, University of California San Francisco, San Francisco, USA
10 Department of Psychology, Arizona State University, Tempe, USA
11 Bodesign Institute, Arizona State University, Tempe, USA
12 Division of Cancer Therapeutics, The Institute of Cancer Research, Sutton, UK
13 Department of Medicine, Surgical Pathology & Cytopathology Unit, University of Padua, Padua, IT

Corresponding authors:
Dr Andrea Sottoriva
Centre for Evolution and Cancer, The Institute of Cancer Research
Sutton, Surrey, SM2 5NG UK
Email: andrea.sottoriva@icr.ac.uk
Dr Nicola Valeri
Centre for Molecular Pathology, The Institute of Cancer Research
Sutton, Surrey, SM2 5PT, UK
Email: nicola.valeri@icr.ac.uk

Author Contributions:
Clinical Data Acquisition: Hazel Lote, Aileen Roy, Alec McDonald, Nabanita Bose, Jeff White, Caroline Findlay and Eddie Kalkman
Experimental Data Acquisition: Inma Spiteri, Michele Simbolo, Andrea Mafficini, Andrea Lampis, Jens Hahne, Francesco Trevisani Zakaria Eltahir, Noori Maka, Margaret Balsitis and Matteo Fassan
Data analysis: Luca Ermini, Alexandra Vatsiou, Marco Punta, Benjamin Werner, Andrea Sottoriva, and Nicola Valeri
Manuscript Drafting: Hazel Lote, Inma Spiteri, Luca Ermini, Stefano Lise, Chiara Braconi, Mel Greaves, Athena Aktipis, Carlo Maley, Andrea Sottoriva and Nicola Valeri
Manuscript approval: all the authors
Study design and supervision: Andrea Sottoriva and Nicola Valeri

Key words: metastatic colorectal carcinoma, cancer evolution, synchronous metastases, phylogenetic tree, tumour age, mutational analysis, whole genome sequencing.

Key message:
We performed whole genome sequencing of metastatic sites and primary colorectal cancer (CRC) in a patient with synchronous and metachronous metastases and exploited the time of needle tract seeding as a stopwatch to determine the timeline of progression. We provide the first experimental evidence documenting the timing of synchronous metastases in CRC and show that genomic instability might be more important than the metastatic potential of the primary cancer in dictating CRC progression.
Abstract

Background
Patients often ask oncologists how long a cancer has been present before causing symptoms or spreading to other organs. The evolutionary trajectory of cancers can be defined using phylogenetic approaches but lack of chronological references makes dating the exact onset of tumours very challenging.

Patients and Methods
Here we describe the case of a colorectal cancer patient presenting with synchronous lung metastasis and metachronous thyroid, chest wall and urinary tract metastases over the course of five years. The chest wall metastasis was caused by needle tract seeding, implying a known time of onset. Using whole genome sequencing data from primary and metastatic sites we inferred the complete chronology of the cancer by exploiting the time of needle tract seeding as an in vivo 'stopwatch'. This approach allowed us to follow the progression of the disease back in time, dating each ancestral node of the phylogenetic tree in the past history of the tumour. We used a Bayesian phylogenomic approach, which accounts for possible dynamic changes in mutational rate, to reconstruct the phylogenetic tree and effectively “carbon date” the malignant progression.

Results
The primary colon cancer emerged between five and eight years prior to the clinical diagnosis. The primary tumour metastasised to the lung and the thyroid within a year from its onset. The thyroid lesion presented as a tumour-to-tumour deposit within a benign Hurthle adenoma. Despite rapid metastatic progression from the primary tumour, the patient showed an indolent disease course. Primary cancer and metastases were microsatellite stable and displayed low chromosomal instability.

Conclusion
Our data provide the first in vivo experimental evidence documenting the timing of metastatic progression in colorectal cancer, and suggest that genomic instability might be more important than the metastatic potential of the primary cancer in dictating colorectal cancer fate.
Introduction

Colorectal cancer (CRC) is the second most common cancer in Europe and the USA [1]. Synchronous and metachronous metastases contribute to the high mortality of CRC [2] and, despite improvements in the management of metastatic CRC, 5-year mortality remains up to 50% [1, 3, 4] with a worse outcome for synchronous versus metachronous metastasis (5-year relative survival 7.2% versus 17.6%) [2]. Discussion around the optimal management of synchronous metastasis from CRC is still open but it is reasonable to hypothesize that the difference in outcome between synchronous and metachronous metastases may not be entirely attributable to clinical variables. International large-scale sequencing efforts will likely identify cancer pathways associated with the molecular makeup of synchronous and metachronous metastatic spread. Critical questions remain open: what is the chronological evolution of a synchronously presenting metastatic CRC? Can we tailor screening and follow-up strategies according to the expected pattern of metastatic progression?

Whole Genome Sequencing (WGS) of distinct metastases allows reconstruction of cancer evolutionary trees but the lack of defined temporal milestones limits the ability to infer the exact timing of cancer progression.

Needle tract seeding is a rare but well-documented phenomenon after diagnostic core needle biopsies and causes metastatic growth of the cancer along the needle tract [5]. Here we performed WGS of primary and multiple metastatic sites of a CRC patient who developed an iatrogenic metastasis due to needle tract seeding. Exploiting the mutational load of this metastasis along with the exact timing of its onset, we calculated the chronology of the primary CRC and metastatic sites.

METHODS

Ethical approval
FFPE material was acquired under the diagnostic samples/pathology archives surplus to clinical requirements protocol at NHS Ayrshire and Arran. The patient provided full written informed consent for the analyses as detailed, and for the case to be published.

The full method list can be found online in the Supplementary Material
Results

Needle tract seeding as stopwatch to calibrate cancer evolution

In September 2008 a 68 year old Caucasian male underwent anterior resection for a pT3pN2(4/17) EMVI negative, poorly differentiated adenocarcinoma of the sigmoid colon. Pre-operative staging revealed a solitary left upper lobe lung lesion of uncertain significance, which was kept under active radiological surveillance for nearly 3 years and remained stable (Fig 1; Supplementary Table 1).

Repeat imaging in July 2011 demonstrated an increase in size of the left upper lobe lung lesion in addition to a new nodule in the left lobe of the thyroid. CT-guided lung biopsy (December 2011) confirmed metastatic colorectal carcinoma with supportive immunohistochemical staining reaction: positive for CK20, CDX2 and CEA but negative for TTF-1 (Fig. 1; Fig. 2A and 2B). The patient proceeded to left upper video-assisted thoracoscopic lobectomy in January 2012. Left thyroid lobectomy (August 2012) demonstrated a follicular Hurthle cell adenoma of the thyroid, within which there was a single tumour-to-tumour deposit of metastatic colorectal cancer (Fig. 1; Fig. 2A and 2C). The patient underwent periodic follow-up with tumour markers and imaging thereafter and remained disease-free.

In September 2013, 5 years from the initial diagnosis, FDG-PET-CT demonstrated a focal unresectable left posterior chest wall mass at the site of previous lung biopsy; histology confirmed metastatic adenocarcinoma from the sigmoid primary in keeping with seeding of malignant cells from the needle biopsy performed two years before. Radical chemo-radiotherapy to the left posterior chest wall mass was completed in February 2014 but repeat CT in May and August 2014 demonstrated residual disease at this site in addition to a new abnormality in the upper pole of the left kidney. Renal and left ureteric biopsy demonstrated metastatic colorectal cancer, with immunohistochemical staining positive for CDX2, CK20, CK7 and BerEP4 [(Fig. 1; Fig. 2A and 2D; Supplementary Fig. S1) a detailed description of the case can be found in the Supplementary Material]. Primary tumour and metastatic deposits showed similar histopathologic characteristics, resembling that of a traditional moderately-differentiated (G2) adenocarcinoma. No MSI-like features (i.e. poor differentiation; presence of mucinous and signet ring cells differentiation; medullary growth pattern; presence of tumour intraepithelial lymphocytes; Crohn’s like lymphocytic reaction) were observed.

Comment [3]: Dates are inconsistent with Figure 1

Comment [4]: Two paragraphs up we say it’s POORLY DIFFERENTIATED!
Genomic profiling

Targeted sequencing of 409 cancer related genes in both the primary CRC and all metastatic sites (lung, thyroid, chest wall and urinary tract) of our case revealed the presence of clonal non-sense mutations in APC (Gln1367*), as well as missense mutations in CTNNB1 (Leu156Gln), KRAS (Gly12Asp) and TP53 (Ser215Ile) (Supplementary Tables 2 and 3). The same mutations were not detected in the normal tissue or in the Hurthle adenoma (Fig. 2E). Two genes included in our panel showed discordance between primary and metastatic cancer: ADAMTS20, a metalloproteinase involved with cancer invasion and migration [6], and AKAP9 an A-kinase anchor protein which binds to the regulatory subunit of protein kinase A [7]. The ADAMTS20 missense mutation Arg1885Thr was observed in the primary cancer but was not detected in either the lung, thyroid or chest wall metastases. Conversely, the AKAP9 missense mutation Ala3077Pro was found in all the metastatic sites but was not detected in the primary cancer. All these mutations were also found using WGS. Furthermore ADAMTS20 and AKAP9 were validated by Sanger Sequencing (Supplementary Fig. S2).

All lesions were microsatellite stable (MSS) [8]. Copy number analysis based on WGS data revealed a relative low level of chromosomal instability, with LOH of chromosome 5, gain of chromosome 7, and a focal amplification on chromosome 13q12.2-12.3, encompassing the VEGFR1 [9] and CDX2 [10] genes (Fig. 2F; Supplementary Fig. S3 and Supplementary Table S4). These aberrations were clonal in all the CRC lesions, whereas the Hurthle adenoma showed a distinct profile characterised only by loss of chr2, chr9 and chr22. Primary tumour and metastatic sites displayed the same dominant age related mutational signature 1 which has previously observed in CRC as well as other cancer types [ref Alexandrov et al. 2013 - Nature] (Supplementary Fig. S4).

Timing Progression

Phylogenetic reconstruction from multiple samples of the same malignancy allows us to turn static snapshots of cancer genomes into the tumour evolutionary history by exploiting genomic intra-tumour heterogeneity (ITH) [11]. However, although phylogenetic trees inform on the relationship between different samples, estimating the exact chronology of the events without a temporal reference remains extremely challenging.

In this exceptional case, the temporal reference is represented by the date of seeding along the needle tract that initiated the chest wall metastasis. This information allowed ‘calibration’ of the timing
for the whole tree and allowed us to follow the progression of the disease back in time, dating each ancestral node in the past history of the tumour.

The phylogenetic tree was inferred using a Maximum Likelihood approach and timed using a Bayesian phylogenomics framework, accounting for possible dynamic changes in mutational rate, to effectively “carbon date” the malignant progression and reveal the complete chronology of the different lesions in this patient (Fig. 3A; see Supplementary Table S5 for the list of mutations used to infer the phylogeny). In the field of evolutionary biology, this type of approach is widely used to time the onset of new species by exploiting temporal information from the fossil record [12].

By considering the time of chest wall metastasis biopsy as time zero (September 2013) and the node between the chest wall biopsy and the lung biopsy (needle-tract seeding event) as a reference, we estimated that metastatic disease originated 9.26 years before, giving rise to the lung and thyroid metastases in parallel. The primary cancer instead originated between 9.96 and 13.17 years before (Fig. 3A and 3B; Supplementary Fig. S5). Hence, the primary adenocarcinoma emerged between five and eight years before clinical diagnosis (2000-2003). Despite the long latency between biological onset of the primary cancer and clinical presentation, within just one year from its inception the tumour began metastasising to the lung and to the thyroid where the metastasis grew within a Hurthle adenoma. Phylogenetic analysis revealed that the lung and thyroid metastasis were more closely related to one another than to the primary tumour. A schematic representation of the timeline of malignant progression is reported in Fig. 3B.

**Neo-antigen Load**

In order to test the contribution of anti-tumor immunity in regulating CRC evolution we applied the pipeline described in Supplementary Fig. S6 (see also supplementary methods) to the samples subjected to multi-region WGS. The analysis resulted in a remarkably low number of strong human leukocyte antigen (HLA) binders [(Colon= 42; Thyroid Met=4 and Lung Metastasis=6) Supplementary Table S6].

**Discussion**

Several studies have applied mathematical modelling to calculate the time-lapse between the onset of primary and metastatic spread [13, 14]. CRC modelling suggests that metastases can develop in less
than two years from the primary cancer [14]. However, these types of analysis remain difficult to interpret and hard to generalise due to the lack of temporal references. Here we exploited the occurrence of a needle tract metastasis to define the exact chronology of primary onset and metastatic spread. Our observations suggest that this process might be faster than expected, occurring within the initial twelve months from the primary onset. Similar patterns of early metastatic divergence have been reported in other tumour types using an analogous phylogenomic approach, where in the case of metachronous lesions, the time between primary and metastatic resection was used as a reference [15].

Despite rapid metastatic transformation, our patient showed an indolent course of disease. How can we explain the discordance between the early metastatic dissemination and a relatively favourable prognosis in a patient with synchronous metastasis? The pattern of disease progression we observed is quite uncommon for a CRC patient but, in our opinion, does not account for the favourable outcome. Solitary lung metastases are observed in 11% of colon cancer patients; surgical resection can be curative in some cases but prognosis remains modest [16]. Thyroid and urinary tract metastasis are uncommon and are usually associated with poor prognosis and multi-metastatic disease. CRC metastases to a thyroid adenoma (tumour-to-tumour metastasis) are even less frequent and have always been associated with poor outcome [17]. Although the resection of both lung and the thyroid metastases might have contributed to the improved survival of our patient, the analysis of his molecular makeup suggests that a rather unusual pattern of low genomic instability may have contributed to the indolent course of disease.

In keeping with previous evidence from large-scale genomic studies [18], targeted sequencing showed clonal mutations in genes accounting for the “breakthrough” and “expansion” phase of the primary CRC [19] such as APC, CTNNB1, KRAS and P53. As expected, all these mutations were also detected in metastatic sites. In contrast, discrepancies between the primary and metastatic sites were observed for ADAMTS20 and AKAP9. In line with our observation, ADAMTS20 mutations were detected in a large series of primary CRC but not in their matching synchronous metastasis [6] suggesting this mutation might promote invasion of the primary cancer but might not be critical for cell fitness in metastatic sites. The anchor protein-coding gene AKAP9 has also been investigated in CRC and found to be up-regulated in CRC with nodal metastasis [20] suggesting a pro-metastatic role for this gene. The molecular pathology of our patient appears common to many other sporadic CRCs. Although the clonal or private mutations described above may have allowed our patient’s cancer cells
to metastasise and survive in different environments, this does not explain the slow pattern of progression, suggesting the reason for our patient’s indolent disease may lie elsewhere.

Chromosomal (CIN) and Microsatellite (MSI) instability can be used to classify CRCs with different molecular and prognostic features [21]. Interestingly our patient showed no MSI and relatively few and clonal chromosomal aberrations, with a focal amplification in 13p12.2-12.3 encompassing the VEGFR2 and CDX2 loci, genes previously found over-expressed in synchronous versus metachronous metastatic CRC [22]. Different lines of evidence suggest that high CIN occurring at early stages of neoplastic transformation might represent a hallmark of aggressive behaviour [23, 24] thus the lack of it might explain the indolent course of disease observed in our patient.

A final interesting point relates to the contribution of the tumour micro-environment to CRC evolution: low neo-antigen burden has been associated with worse prognosis in early-stage lung cancer [25]. In our case, the primary cancer displayed low number of neo-antigens and the lack of anti-tumour immunity against the primary cancer might explain the early metastatic spread.

Genomic instability [24] and tumour-stroma interactions [26] represent synergic drivers of cancer progression. The genomic data suggest that in this patient these two hallmarks of cancer evolution might have been uncoupled: while the lack of immune response against the primary tumour might have promoted early cancer metastasis, the relative stability in genomic aberrations might account for the slow rate of progression.
Figure Legends

Figure 1. Overview of clinical, radiological and pathological progression of a CRC patient. The patient underwent anterior resection for a \( pT_4pN_2 \) sigmoid cancer in November 2008. A lung nodule of unknown significance at the time of diagnosis was kept under surveillance and was subsequently proven to be metastatic in nature (biopsy in December 2011) and was resected in January 2012. An FDG-avid thyroid nodule was resected in August 2012 and showed a Hurte adenoma containing foci of metastatic CRC. Two years after the lung biopsy, a chest wall mass on the site of previous lung biopsy was histologically confirmed and treated with chemo-radiotherapy. Six months after treatment, the chest wall metastasis showed disease progression and a new site of metastasis along the ureter was histologically confirmed during the insertion of a urinary stent. The timeline shows dates that metastatic sites became apparent on imaging. Haematoxylin and Eosin (H&E) staining of primary and metastatic sites are shown.

Figure 2. Histopathology and molecular profiling of the primary CRC and metastatic sites. (A) microscopy of the primary and metastatic sites including H&E staining, proliferation (nuclear Ki-67) and apoptosis (Caspase-3). (B) microscopy of the CRC lung metastasis that led to the needle tract seeding to the chest wall. Cells stained negative for TTF1 and positive for CDK20, CDX2 and CEA. (C) gross pathology and microscopy of the Hurtle Adenoma containing a CRC metastasis. Macroscopic examination showed the Hurtle adenoma containing a haemorrhagic or cystic area. Different magnifications show the Hurtle adenomatous cells (background large eosinophilic cells) with an invasive front of CRC metastasis. The dotted line indicates the border between the normal thyroid and the Hurtle adenoma. The black rectangle indicates the area of invasive CRC within the adenoma. (D) shows microscopy of the urothelial metastasis. H&E and CDK20 staining confirmed the colonic origin of the metastasis. (E) Drivers of CRC detected by WGS and validated by targeted exome sequencing. (F) Copy number changes in primary and metastatic sites.

Figure 3. Chronology of the patient’s CRC evolution. (A) Whole-genome sequencing of multiple lesions from the patient’s malignancy allowed phylogenetic reconstruction of the tumour tree. In the phylogenetic tree, dates within brackets indicate the time of clinical diagnosis whereas dates in bold highlight the estimated times of the different lesions. (B) Illustrative cartoon of the patient disease
progression and samples taken. At time $t_1$, the first colorectal cancer cell arose, giving rise to the primary tumour. At time $t_1$ the first metastatic clone emerged, giving rise to the lung metastasis, quickly followed by a second metastasis to the thyroid emerging at time $t_2$. At time $t_2^p$ the sample from the primary tumour was collected (resection) and analysed. During the lung biopsy, the needle tract seeding event spread cancer cells in the chest wall at time $t_{cw}$. A few days later, at time $t_{cw}^F$ the lung metastasis was resected and profiled. Finally at time $t_{cw}^F$ the chest wall metastasis was also sampled.
REFERENCES


ACKNOWLEDGMENTS AND FUNDING

This work was supported by Cancer Research UK (CEA A18052), European Union FP7 (CIG 334261) and the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) at The Royal Marsden NHS Foundation Trust and The Institute of Cancer Research (grants A62, A100, A101 and A159) to NV; Italian Cancer Genome Project – Ministry of University [FIRB RBAP10AHJB]; Associazione Italiana Ricerca Cancro [AIRC 12182] to AScarpa; Wellcome Trust [105104/Z/14/Z] to the Centre for Evolution and Cancer. We acknowledge support from the NIHR BRC at The Royal Marsden NHS Foundation Trust and The Institute of Cancer Research. ASottoriva is supported by The Chris Rokos Fellowship in Evolution and Cancer. BW is supported by the Geoffrey W Lewis Post-Doctoral Training fellowship.

We thank the patient and his family.

Disclosure of Potential Conflicts of Interest: No potential conflicts of interest were disclosed.