The evolutionary landscape of colorectal tumorigenesis

William Cross [1,2], Michal Kovac [2,3], Ville Mustonen [4], Daniel Temko [1,5], Hayley Davis [6], Ann-Marie Baker [1], Sujata Biswas [6], Roland Arnold [7],Laura Chegwidden [8], Chandler Gatenbee [9], Alexander R Anderson [9], Viktor H Koelzer [2,10], Pierre Martinez [1], Xiaowei Jiang [11], Enric Domingo [2], Dan Woodcock [12], Yun Feng [2], Monica Kovacova [13], Tim Maughan [14], The S:CORT Consortium [15], Marnix Jansen [16], Manuel Rodriguez-Justo [16], Shazad Ashraf [17], Richard Guy [18], Christopher Cunningham [18], James E East [19], David C Wedge [12], Lai Mun Wang [20], Claire Palles [8], Karl Heinimann [21], Andrea Sottoriva [2], Simon J Leedham [6,19], Trevor A Graham[1#] and Ian PM Tomlinson[11,23#]

[1] Evolution and Cancer Laboratory, Bart's Cancer Institute, Queen Mary University of London, EC1M 6BQ, UK; [2] Molecular and Population Genetics Laboratory and [6] Cancer Stem Cell Biology Laboratory, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford, OX3 7BN, UK; [3] Bone Tumour Reference Center at the Institute of Pathology, University Hospital Basel, Schönbeinstrasse 40, 4031 Basel, Switzerland; [4] Department of Biosciences, Department of Computer Science, Institute of Biotechnology, University of Helsinki, 00014, Helsinki, Finland; [5] University College London, Gower St, London, WC1E 6BT, UK; [7] Cancer Bioinfomatics Group, [8] Gastrointestinal Cancer Genetics Laboratory and [11] Cancer Genetics abd Evolution Laboratory, Institute of Cancer and Genomic Sciences, College of Medical and Dental Sciences, University of Birmingham, Vincent Drive, Edgbaston, Birmingham B15 2TT, UK; [9] Integrated Mathematical Oncology Department, Moffitt Comprehensive Cancer Centre, Tampa, Florida, USA; [10] Institute of Pathology, University of Bern, Murtenstrasse 31, CH 3008 Bern, Switzerland; [12] Big Data Institute, Old Road Campus, University of Oxford, Oxford OX3 7LF, UK; [13] Institute of Mathematics and Physics, Slovak University of Technology, 84248 Bratislava, Slovak Republic; [14] Department of Oncology, Old Road Campus Research Building, University of Oxford, Oxford OX3 7DQ, UK; [15] A list of members is provided in the Supplementary Information; [16] Department of Research Pathology, Cancer Institute, University College London, London WC1E 6JJ, UK; [17] Department of Surgery and [23] Department of Histopathology, University Hospitals Birmingham, Birmingham B15 2TH; [18] Department of Colorectal Surgery, Cancer Centre, Churchill Hospital, Oxford University Hospital NHS Foundation Trust, Old Road, Headington, Oxford OX3 7LE, UK; [19] Translational Gastroenterology Unit and [20] Department of Cellular Pathology, John Radcliffe Hospital, Headington, Oxford OX3

9DU, UK; [21] Medical Genetics, University Hospital Basel, Burgfelderstrasse 101, 4055 Basel, Switzerland; [22] Centre for Evolution and Cancer, The Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK.

co-corresponding authors

Abstract

The evolutionary events that cause colorectal adenomas to progress to carcinomas remain largely undetermined. Using multi-region genome/exome sequencing of 24 benign and malignant colorectal tumours, we probe the evolutionary fitness landscape occupied by these neoplasms. Unlike carcinomas, advanced adenomas frequently harbour sub-clonal driver mutations that have not swept to fixation, and have relatively high genetic heterogeneity. The burden of single nucleotide alterations, including drivers, are similar between the two types of tumour. Carcinomas are instead distinguished from adenomas by widespread aneusomies that are usually clonal and often accrue in a "punctuated" fashion. Adenomas evolve across an undulating fitness landscape, whereas carcinomas occupy a sharper fitness peak, probably owing to stabilising selection.

Introduction

The classical adenoma-carcinoma sequence of colorectal tumorigenesis[1] postulates that a conventional colorectal adenoma (CRA) is initiated by "two hits" at APC[2,3], and typically progresses to colorectal cancer (CRC) through a stepwise accumulation of driver mutations such as KRAS and TP53 and deletion of chromosome 18q[4]. The evolutionary dynamics presumed to underlie this process comprise a series of selective sweeps to (near) fixation, each triggered by an elevation in sub-clone fitness through the occurrence of a new, positively-selected driver mutation[5]. In this model, progression to an invasive lesion (carcinoma) is postulated to be prompted by the acquisition of a critical driver mutation burden, implying that adenomas and carcinomas should be distinguishable by specific driver mutations. CRCs can, however, develop without the full complement of driver mutations[6,7], and some studies have suggested that sub-clonal evolution within established tumours is 'effectively neutral'[8,9], questioning whether selective sweeps occur at all, especially in established CRCs.

As part of a comprehensive assessment of colorectal tumour evolution, here we have attempted to re-assess the classical model and outline the evolutionary 'fitness landscape' of CRAs and CRCs. The fitness landscape, a concept, first introduced by Sewall Wright in 1932[10], is an abstraction to help visualise the relationship between

genotypes and reproductive success (sub-clone fitness in this context). The X- and Y-axes can be thought of as the genotype 'space' (simplified to 2 dimensions) that can be occupied by adenomas and carcinomas. The Z-axis or height is proportional to genotype fitness: peaks represent particularly fit genotypes, valleys less fit genotypes, and ridges/plateaux equally fit genotypes. Individuals sampled from a population are likely to occupy (local) fitness peaks, because less fit individuals have been removed by negative (purifying or stabilising) selection. Herein we search for the genotypes associated with the fitness peaks occupied by CRAs and CRCs and probe peak shapes by quantifying intra-tumour heterogeneity (ITH). Transitions around the landscape are measured using phylogenetic and molecular clock analyses. These data provide a comprehensive understanding of the evolutionary trajectories underpinning the development of CRAs and CRCs.

Results

To map the evolutionary landscape of CRAs and CRCs, we performed multi-region whole-genome sequencing (WGS) or whole-exome sequencing (WES) on 2-16 regions (total 118) from 9 CRAs and 15 CRCs, each with constitutional DNA (Table S1 for sample details and S2 for sequencing statistics). Five CRCs, including four from Lynch syndrome patients, had microsatellite instability (MSI) owing to defective DNA mismatch repair, and these tumours were analysed as a distinct group unless otherwise stated. The remaining ten CRCs were microsatellite-stable (MSS) and of these, two were synchronous lesions from a single patient. Mutations in a subset of genes were validated using targeted molecular inversion probe sequencing (Online Methods).

Somatic single nucleotide alterations do not define CRC fitness peaks

We first assessed how somatic single nucleotide alterations (SNAs) defined the co-ordinates of CRAs and CRCs in the fitness landscape. CRAs tended to have only slightly fewer SNAs than MSS CRCs CRAs: median exonic burden=94, 95% range [51-146]; MSS CRCs: median=130, 95% range [98-171]; p=0.29 Wilcoxon test; Figure 1A, Table S2). After sequencing coverage normalisation, the mutational frequency in CRAs

remained very similar to that of MSS CRCs (CRA; 4.1/Mb [3.3-4.9], MSS CRC; 4.2/Mb [2.9-6.4], p=0.9).

Next we compared the burden of driver mutations across CRAs and CRCs, and included SNAs and indels, and also cnLOH and monosomy (chromosome loss) events that are known to act as 'second hits' to the tumour suppressor genes APC and TP53, and also 18q allelic loss/imbalance (Figure 1C-D and Table S3). The burden of tier 1 mutations, which we defined as likely pathogenic changes in known CRC driver genes (see Online Methods, Table S4), was not significantly different (CRAs: median=5 [2-9]; MSS CRCs: median=6 [2-8], p=0.9). We noted that the difference remained non-significant when comparing drivers across individual biopsies (p=0.19; Wilcoxon test). Individual tier 1 driver mutations were detected at similar frequencies across CRAs and CRCs, with the exception of TP53, which was more commonly mutated (possessing least one SNA, indel or copy change) in CRCs (Fisher test; p=0.005, see Figure 1D and Table S5). The frequency of tier 2 driver mutations (uncertain pathogenicity changes in CRC or pan-cancer driver genes) was also similar in CRAs and MSS CRCs (CRAs: median=3 [2-4]; MSS CRCs: median=3 [1-7]; p=0.8). Several tier 2 driver mutations were specific to CRAs or CRCs, but most occurred infrequently; only KMT2C was notable, being mutated in 4 CRAs and no CRCs. The total driver mutation burdens (tier 1 and 2 combined; medians CRAs=7 [2-12]; CRCs=8 [3-15]; p=0.6; Figure 1C, Tables S3, S4) were also similar. Furthermore, when using an alternative definition of driver genes (the top 15 genes mutated in MSS CRCs [excluding TTN], according to the TCGA publication[7], Table S4b), the burdens remained not statistically different in CRAs and CRCs (CRAs: median=5 [2-7]; MSS CRCs: median=5.5 [2-9], p=0.7).

In order to confirm the somewhat surprising finding that our CRAs and CRCs had similar mutational burdens, we accessed data from 481 primary CRCs (stages II-IV) and 55 adenomas of similar type and size to those in our MSeq study. A single sample of each tumour had been sequenced by the S:CORT project (https://www.s-cort.org/) to a median depth of 150X using a panel that comprised the major CRC drivers and a number of other putative driver genes. Of the top 7 CRC driver genes by mutation frequency (APC, KRAS, BRAF, TP53, SMAD4, FBXW7, PIK3CA), the mean SNV burden was 2.1 in adenomas and 1.9 in CRCs (p=0.58, Wilcoxon test). In the complete set of 73 CRC driver genes (https://www.intogen.org/search?cancer=COREAD) present in the panel, there was similarly no significant different between mutation burden in adenomas

(mean=11.1) and CRCs (mean=10.9; p=0.28). Thus, even in this large cohort analysed by different methods, the driver and total SNV burdens were indistinguishable between CRAs and CRCs.

On the reasonable assumption that CRCs occupy a higher fitness peak than CRAs, our mutation burden data suggest that the relative co-ordinates of CRAs and CRCs in the evolutionary landscape are not principally determined by SNAs, and that the evolutionary 'progress' from adenoma towards cancer is not simply measured by accumulated SNAs, including tier 1 driver mutations.

Intra-tumour heterogeneity and phylogenetic analyses suggest that CRCs occupy sharper fitness peaks than CRAs

To broadly assess the shape of the fitness peaks occupied by CRAs and MSS CRCs, we measured the degree of ITH in each tumour. Excluding tumours with only two regions sampled (see Online Methods), a median 56% [53-70%] of all CRA SNAs were "sub-clonal" (variant not detected in all sampled regions). MSS CRCs had a significantly lower proportion of sub-clonal SNAs (45% [23-77%]; p=0.04; Figure 2 inset) than CRAs. The average pairwise genetic divergence between the regions of each tumour was then assessed following normalisation of sequencing coverage (Online Methods). CRAs showed significantly more divergence between biopsies than CRCs (CRA mean=2.0 versus CRC=1.7, divergent SNAs/Mb; $p<2x10^{-16}$; Figure S1A), despite having the same average mutation burden. The measured values of ITH were unaffected by the number of biopsies available from each neoplasm (Figure S1B-C).

To further quantify ITH, we used SNAs to construct maximum parsimony phylogenetic trees (Figure 2). CRC topologies were often characterised by long trunks (variants ubiquitous across biopsies) with comparatively short branches and leaves (relatively few sub-clonal variants), thus appearing 'palm tree-shaped'[11]. CRAs had proportionally shorter trunks, and thus longer branches/leaves than CRCs, albeit at borderline significance (average branch and leaf length as a proportion of the trunk: CRAs 82% versus MSS CRCs 50%; p=0.06; Figure 2, Table S6). The difference remained when the MSI+ CRCs were included in the analysis (CRAs 82% versus all CRCs 45%; p=0.05). CRAs are thus more genetically diverse than CRCs.

To investigate whether individual CRAs and CRCs occupied single or multiple fitness peaks, we compared the lengths of the phylogenetic tree branches/leaves. Large variations in branch length indicate that mutations accrue faster in some tumour regions than others, which can potentially be caused by selection on a new fitness peak. Average intra-tumour variation in relative branch/leaf length was generally low and similar across CRAs and CRCs (mean standard deviation: CRA 0.14 [0.06-0.24] versus CRC 0.2 [0.06-0.47]; p=0.68; Figure 3, Table S6). Formal assessment of unbalanced tree topologies could only be performed on one tumour (carcinoma 6) as high numbers of samples are needed for sufficient power[12]. Unbalanced trees occur when some ancestor clones produce more surviving lineages than another, another potential indicator of sub-clonal selection. We did not find any significant asymmetry in this single tumour analysis (Colless' test, Yule model, p=0.3). Thus the available data were consistent with the idea that tumours occupied a single, potentially broad, fitness peak.

The SNA-based ITH and phylogenetic analyses suggested that CRAs were more heterogeneous than CRCs, consistent with the former occupying a broader fitness peak, under which several distinct genotype-phenotype combinations could co-exist. The lower ITH in CRCs could also, however, reflect a more recent selective sweep with a genetic bottleneck during the transition from an adenoma, and/or that CRCs were more spatially mixed than CRAs, causing variants at sub-clonal frequency in multiple samples to appear truncal. We therefore directly sought evidence of stronger selection in CRCs by examining the ratio of non-synonymous to synonymous mutations on tumour trunks and branches/leaves. This showed a reduction in non-synonymous mutations on the branches/leaves of CRCs relative to their trunks (Wilcoxon sign rank test, p=0.01; Figure S2), but no such reduction for CRAs (p=0.9), possibly representing on-going positive subclonal selection in CRAs. On the reasonable assumption that positive selection acted on the phylogenetic trunk - the location of almost all tier 1 driver mutations – together these results indicate that subclonal selection is absent (neutral dynamics) or weak within the established carcinoma, with possible negative (stabilising) selection also at play.

Mutational processes are not detectably associated with fitness advantages

Mutation signatures were identified de novo using the EMu program[13]. We recovered ageing, MSI-associated and molecular clock signatures[14] (our Signatures A, B and C respectively), as expected (Figure S3A). Our Signature D, which resembles COSMIC Signature 17 (unknown aetiology, high CTT>CGT frequency[15]) was present at appreciable levels within carcinomas 2, 7, 9P and 10, with its activity often differing between the trunks and branches/leaves of the same lesion (Figure S3B,C). We explored whether signature D had any effect on sub-clonal evolution in CRCs with WGS. It appeared to increase the mutation burden in two CRCs 2 and 9P, but had no discernible effect on their evolution (details in Figure S3D). Carcinoma 9D, the synchronous partner of carcinoma 9P, showed low signature 17 activity, despite being located only 10cm apart in the bowel. These cancers also had different driver mutations, confirming that they essentially behaved as independent neoplasms, with no detectable effect of any shared microenvironment on mutagenic processes (Figure S4).

Major driver mutations can be sub-clonal in CRAs, but are very rarely so in MSS CRCs

Tier 1 driver mutations (defined above) were typically, but not always, clonal in CRAs, whereas in MSS CRCs drivers were more commonly clonal. However, these distributions were not significantly different between tumour types (CRAs=39/49, 80% versus CRCs=49/55, 89%; p=0.3). The clonal distributions of tier 2 clonal driver mutations were however, different; CRAs had significantly less clonal drivers than MSS CRCs (CRAs=7/15, 47% versus CRCs=21/26, 80%; Fisher's exact test, p=0.03, Table S3). We noted that the clonality of tier 1 driver mutations was the same when using the second definition of driver mutations based on the TCGA publication[7] (37/43, 86% versus 45/53, 85%; Fisher's exact test, p>0.9, see Online Methods). The findings are consistent with a scarcity of sub-clonal expansions after the most recent common ancestor (MRCA) in CRCs. This trend seems to be similar in CRAs, though CRAs do show some evidence of sub-clonal driver mutations.

We additionally noted, however, that the most frequently mutated CRC driver genes, apart from the probable tumour-initiating mutations in APC, were sub-clonal in at least one CRA. Notably in adenoma 2, KRAS Q61H and an ARID2 frameshift mutation were present in one region, which was separate from the three regions of this tumour that contained a TP53 E219X mutation. Adenoma 3 had a PIK3CA E545K mutation in two

tumour regions, GNAS R201H in another, and an AKAP9 frameshift in another. SMAD4 R496H in adenoma 4 was also present in a single region. There was no evidence from the phylogenetic analysis that these proven driver mutations were associated with differential sub-clonal expansion, suggesting that their selective benefits were relatively modest (Figure 2). NRAS G60V and PIK3CA H1047R were present in both regions of adenomas 7 and 8 respectively, but were putatively sub-clonal since their corrected allele frequencies were significantly lower ($p<0.05$) than those of other driver mutations, suggesting that biopsies crossed sub-clonal boundaries. By contrast, only one sub-clonal mutation with high-confidence pathogenicity (CHD1 R619X in carcinoma 1) was found in the MSS carcinomas. There was no evidence for parallel evolution of sub-clones based on recurrent known or novel drivers (details not shown).

We next we wished to relate the heterogeneity of mutational burdens to fundamental molecular processes. Immunohistochemistry for Ki-67 (proliferation) and β-catenin (activated Wnt-signalling; Figure S5; Online Methods) showed positive cell fractions of 53% [2-80%] and 82% [3-97%] respectively, with considerable variability between and within CRCs (Table S7). Neither Ki67 nor β-catenin expression was associated with regional SNA burden or ploidy (SNA burden, $R^2=0.2$, $p=0.2$, ploidy, $R^2=0.9$, $p=0.08$; Figure S5).

Genetic and spatial relationships between CRC sub-clones

In all CRCs, physical and phylogenetic distances between biopsies were strongly correlated ($R^2=0.81$-0.93, $p<10^{-4}$ for all carcinomas measured; Figure 4). The invasive edge of CRCs and central regions had similar mutational burdens (exonic SNAs, edge versus central; $p=0.76$). We looked further for sub-clonal mixing within the sampled regions of the MSS CRCs with WGS data by clustering of SNA cancer cell fractions across related samples, using a Dirichlet process-based model (Figure S6). Only 10% of biopsy samples showed evidence of ≥1 sub-clonal population. Whilst we do not exclude a degree of sub-clonal intermingling, these results suggest that, given the depth of the sequencing data, sub-clonal expansions broadly occurred in a spatially contiguous, uniform and discrete fashion.

Copy number changes differ between CRAs and CRCs

We next assessed whether somatic copy number alterations (CNAs) might define the fitness peaks occupied by CRAs and MSS CRCs. Every region of every tumour carried at least one CNA, including cnLOH (see Online Methods). In a combined analysis of all regions from each tumour, as expected, adenomas had fewer CNAs (number of discrete CNA segments >1Mb) than carcinomas[16,17] (CRAs median=13 [7-11] versus CRCs median=40 [15-42], p=0.003; Figure 1B). Correspondingly, the overall average proportion of the genome disrupted by CNAs (copy number ≠ 2, allelic ratio ≠ 1) was higher in CRCs (CRCs, 72% versus CRAs, 40%; p=0.05; Figure 3). These data show that despite carrying similar SNA burdens, CRCs display higher CNA levels than CRAs.

Driver CNAs are currently hard to identify with certainty in cancer[18]. In colorectal tumours, losses (deletions or cnLOH) on chromosomes 5q, 17p and 18q are often thought to be second hits involving tumour suppressors APC, TP53 and SMAD4 respectively (although 18q loss is more common than SMAD4 mutation). The status of other recurrent changes – such as 1q gain, 7 gain, 8p deletion, 13q gain and 20 gain – as drivers or passengers is less clear. Many recurrent, and hence potential driver[7], CNAs were present at significantly higher frequencies in CRCs compared to CRAs (Figure 3). Notably, 17p loss occurred in 9/10 MSS CRCs, but only 2/9 CRAs (Fisher's exact test, p=0.005), paralleling the TP53 SNA data. By comparison, loss at the APC locus (8/10 CRCs versus 5/9 CRAs; p=0.35) and the SMAD4 locus (7/10 CRCs versus 4/9 CRAs; p=0.37; Figure 2) occurred at similar frequencies in both lesion types.

Every tumour had at least 2 clearly sub-clonal CNAs (non-ubiquitous, present versus absent changes; Figure 3A) and no chromosome aberration was exclusively ubiquitous or sub-clonal across the tumours. Overall, 75% and 48% of gains were sub-clonal in CRAs and CRCs respectively (p=0.002), compared with 57% and 27% of losses/cnLOH (p=0.007; Figure S7). Thus, a greater proportion of CNAs were sub-clonal in CRAs than in CRCs.

We compared the size distribution of large (>1Mb) CNAs in early (truncal) versus late (sub-clonal) tumour evolution. In CRCs, sub-clonal CNAs were smaller than ubiquitous CNAs (p<0.001 ; Figure S3C), but this difference was not present in CRAs (p=0.45). The lower frequency of large CNAs later in evolutionary time in CRCs suggests that the

cancers have obtained a near-optimal level of aneuploidy, with further large-scale CNAs subjected to negative/stabilising selection. In adenomas, since the overall CNA burden is lower, new large CNAs may still be tolerated.

MSS CRC evolution can involve either *"punctuated"* or more gradual CNA acquisition

Since CNAs were the principal genetic feature distinguishing CRAs and CRCs, we investigated their role in the transition between the benign and malignant fitness peaks. Utilising a similar strategy to Durink[19] and Newman[20] (details in Online Methods), we used the SNAs within informative chromosomal segments (copy number gains and cnLOH) as a molecular clock to time the occurrence of that CNA. SNAs present on a chromosome prior to gain, cnLOH or amplification increase in frequency (VAF) following the copy number change, whereas SNAs that accrue after the gain remain at their original, lower VAF. The ratio of higher to lower VAF SNAs therefore estimates the time of CNA occurrence.

Sufficient SNAs for molecular clock analysis were only present in WGS data. Of the five MSS CRCs analysed by WGS, carcinomas 3, 9P and 10 showed a clustering of CNA timings shortly before the MRCA (Kolmogorov-Smirnov test against a uniform distribution of CNA timings, $p<0.02$ for all, Figure 5). A similar, borderline significant CNA cluster occurred in carcinoma 9D. Carcinoma 5 showed a more gradual accumulation of CNAs.

Since the timing method demonstrated a form of "punctuated" CNA evolution (rejection of null hypothesis of uniform accumulation), but did not distinguish between multiple gains of individual chromosomes and genome doubling followed by chromosomal gain or loss, we searched heuristically for evidence of genome doubling using a score based on the number of chromosome centromeres present at copy number 4 or above, with extra weight for allelic balance (Figure S8). Based on this measure, all of the CRCs with significantly or borderline significantly clustered CNA timings (n=4) were genome-doubled on this measure, as was the untimed carcinoma 8. The CNAs in these tumours are typically trisomies, judged to have arisen by chromosome (arm) loss subsequent to allele-balanced genome doubling. The other tumours (including carcinomas 1, 2, 4, 5, 6, 7) were scored as non-genome-doubled. Sub-clonal genome doubling was present in one CRA (adenoma 2). This tumour carried a TP53 mutation in its genome-doubled

regions, and overall TP53 mutations (SNAs and/or CNAs) were associated with genome-doubling in MSS CRCs (Fisher's Exact Test p=0.018). In addition, genome-doubled cancer regions had higher Ki67 expression (see above; p=0.04; Figure S5), hinting at the existence of a selective benefit of doubling.

The evolutionary landscape of microsatellite-unstable CRCs

The overall SNA burden of the 5 MSI+ CRCs was, as expected, far higher than in MSS CRCs (Figure S9A). More pointedly, the number of tier 1 CRC driver mutations was also higher (median=12 [4-14]) than in MSS CRCs, median=3, p=0.042, Figure S9B), whilst CNA burden was lower (Figure S9C). Of note, in MSI+ CRCs, the great majority of driver SNAs were truncal, the number of sub-clonal tier 1 drivers was only a little greater (median=1 [0-7]) than in MSS CRCs, and the proportion of all sub-clonal SNAs was not significantly increased (median MSI+ CRCs 34% versus MSS CRCs 42%; p=0.13; Figure S9D). In phylogenetic analysis, neither the average branch/leaf length as a proportion of the trunk nor its variability differed significantly between MSI+ and MSS CRCs (Figure S9E). Our signature B (COSMIC signature 6) predominated in MSI+ CRCs, especially on the branches/leaves, but the other COSMIC MSI-associated signatures[14] were not detected. Overall, the data suggest that MSI+ CRCs evolve in a similar way to MSS CRCs, albeit with some limited evidence of sub-clonal selection.

Discussion

Here we have contrasted the patterns of evolution in colorectal carcinomas and their classical adenomatous precursor lesion, and our data begin to reveal the shape of the fitness landscape over which CRCs grow. CRAs tend to evolve through acquisition of major driver mutations in genes such as APC, KRAS, TP53 and via 18q loss as per the Kinzler and Vogelstein model[4]. More recently discovered cancer driver mutations are also present in many adenomas (Table S3 & S4). In fact, CRAs can harbour mutations in any of the major CRC driver genes, but those mutations do not necessarily occur in a stereotypic order. Driver mutation acquisition also does not necessarily cause selective sweeps (leading to 'stepwise' evolution of the tumour cell population), since sub-clones with additional major driver mutations may not displace sub-clones lacking those

mutations, but instead may co-exist in spatially discrete areas. It follows that many driver mutations probably confer a relatively small selective advantage. This is reflected in several observations in CRAs, including a relatively high level of genetic diversity (both SNAs and CNAs), variation in the major driver mutation complement in different regions of individual tumours, and phylogenetic trees with relatively long branches/leaves. It is even possible that SNA accumulation is not an essential feature of tumorigenesis prior to malignancy, and we speculate that carcinomas need not arise from the sub-clone with the greatest number of driver mutations, thus explaining why some CRCs have a very small driver mutation complement[7].

MSS CRCs have similar overall and major driver SNA burdens to precursor CRAs, but are less diverse, with longer phylogenetic tree trunks than branches/leaves. These findings may reflect the influence of several factors, including not only selective constraints, but also time from the MRCA after an additional selective sweep, ploidy, sample purity and genomic instability. They are, nevertheless, consistent with the notion that carcinomas are not simply ‘old adenomas’ that have accumulated addition SNAs during a relatively protracted evolution. Overall, the lack of sub-clonal driver SNAs and reduction in non-synonymous SNAs on the branches and leaves of CRCs suggest that there is not strong positive sub-clonal selection for SNAs after the MRCA. CRAs on the other hand do show subclonal drivers and relatively high ITH together providing evidence of (perhaps relatively weak) subclonal selection.

Although present in CRAs, large CNAs and genome doubling are much more common in CRCs. CNAs on CRC tree branches/leaves are smaller than those on trunks. Whilst negative or stabilising selection remains difficult to measure, this is consistent with the relatively low genetic diversity in CRCs, based on SNAs and large CNAs. For most MSS CRCs, a near-triploid karyotype seems optimal, either through genome doubling followed by loss of some chromosomes, or through a gain of chromosomes that mostly occurs within a putatively short time window between malignant progression and the MRCA. In each case, one or more selective sweeps seem to occur, rendering the driver SNAs and most CNAs clonal. We do not exclude additional positive selection for specific sub-clonal CNAs in CRCs, but this remains unproven and indeed our data showed no evidence of sub-clonal selection. Although every CRC had at least one sub-clonal CNA, we found no evidence of parallel CNA evolution.

In all our MSI+ cases, defective MMR and most major driver mutations arose on the phylogenetic trunk, and the relative branch/leaf length was similar to that of MSS CRCs. Although the sporadic MSI+ cancer had a low driver mutation burden, as expected if driven in part by a methylator phenotype[21], its evolution was otherwise similar to the Lynch syndrome CRCs. We speculate that MSI+ CRCs experience either multiple selective sweeps driven by individual SNAs, or, more intriguingly, by co-occurring or epistatically acting non-canonical driver SNAs (such as CTNNB1, SOX9, NF1 and CASP8).

A small number of ITH studies have been undertaken previously in CRC. Kim et al[22] and Uchi et al[23] performed multi-region WES of 5 primary and metastatic CRCs and 9 CRCs respectively, and Suzuki et al[24] performed deep targeted sequencing of 799 genes in four CRCs. Similar to our study, these studies reported that major driver mutations, affecting APC, KRAS and TP53, were truncal, with the exception of PIK3CA. Uchi et al[23] also reported that large copy gains were common on the trunk of the evolutionary tree, with focal deletions on branches.

Fewer studies of CRA evolution exist. Kim et al[25] used WES to compare malignant and benign regions of 4 mixed cancer-in-adenoma polyps. They reported similar SNA burdens in cancer and adenoma regions, and thus suggested that the regions evolved in parallel, rather than the carcinoma progressing from a late adenoma. We note, however, that it is extremely hard to distinguish benign and malignant components of these lesions, since malignancy is defined not by cytology, but by invasion and hence the location of tumour cells. The different neoplastic components of such polyps may therefore, in reality, both be ‘cancerous’. For these reasons, in this study, we based our comparison between advanced CRAs and CRCs that were distinct lesions.

Previous work from our group[10] examined single glands from 11 CRCs and 4 CRAs for CNAs, and for Ampliseq panels of SNAs that had been derived from bulk tumour WES. Although that manuscript and our present study had very different focuses, some of the findings are consistent. For example, one feature of the “Big Bang” model of sub-clonal intermixing expounded in the previous study is that after the MRCA, CRC sub-clones radiate outwards without notable differential sub-clonal expansion or selection of further advantageous variants; our present study is broadly consistent with those data. There are also, however, some differences between the studies that allow refinement of the

"Big Bang" model. For example, our current study, which benefits from the significantly increased genomic resolution of WGS/WES, emphasises that large sub-clones after the MRCA remain spatially restricted in CRCs and consequently that the previously observed, widespread clonal 'intermixing' in some CRCs[10] may reflect the shape, size, and boundary location of discrete sub-clones as well as outwards radiation of low frequency clones. Furthermore, whilst the "Big Bang" was broadly consistent with ITH measured in CRAs, our present study finds that sub-clonal driver mutations in the absence of selective sweeps occur commonly in these tumours.

In a study analogous to ours, Stachler et al[26] exome-sequenced 5-11 samples of oesophageal carcinoma and its precursor, Barrett's oesophagus (BE) from 5 patients. Comparing the two studies reveals both similarities and differences. BE is not a discrete tumour and is generally a highly polyclonal lesion, reflected in multiple "initiating" deletion mutations in CDKN2A and a series of clonal expansions without selective sweeps. By comparison, CRAs are discrete and probably have monoclonal origins usually caused by bi-allelic APC mutation, followed either by selective sweeps, or by polyclonal expansions reminiscent of BE. We note that in both BE[27,28] and CRA, data are consistent with malignant progression sometimes occurring from a sub-clone that does not have the largest driver mutation burden.

In summary, we have used measurements of intra-tumour heterogeneity to reveal the evolutionary trajectories of colorectal tumour cell populations across what appears to be a rather flat fitness landscape for adenomas, with a higher, sharper peak occupied by cancers. Our data refine the Fearon and Vogelstein model[4] of CRC progression by showing that driver mutations do not necessarily lead to hard selective sweeps and that progression to CRC can involve punctuated evolution.

Acknowledgements

WCHC, SL, TAG (A19771) and IT (A27327) are funded by Cancer Research UK. We acknowledge core funding to the Wellcome Trust Centre for Human Genetics from the Wellcome Trust (090532/Z/09/Z). TG and SL were also supported by Bowel and Cancer Research Charity small grant scheme. VM was supported in part by funding from The

Wellcome Trust (098051). DCW is supported by the Li Ka Shing foundation. XJ and IT are supported by an ERC advanced grant (EVOCAN-340560). The S:CORT study is funded by MRC and Cancer Research UK.

References

1. Morson, B. C. Evolution of cancer of the colon and rectum. Cancer **34,** 845–849 (1974).
2. Ashton-Rickardt, P. G. et al. High frequency of APC loss in sporadic colorectal carcinoma due to breaks clustered in 5q21-22. Oncogene **4,** 1169–1174 (1989).
3. Powell, S. M. et al. APC mutations occur early during colorectal tumorigenesis. Nature **359,** 235–237 (1992).
4. Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. Cell **61,** 759–767 (1990).
5. Jones, S. et al. Comparative lesion sequencing provides insights into tumor evolution. Proc. Natl. Acad. Sci. **105,** 4283–4288 (2008).
6. Smith, G. et al. Mutations in APC, Kirsten-ras, and p53—alternative genetic pathways to colorectal cancer. Proc. Natl. Acad. Sci. U. S. A. **99,** 9433–9438 (2002).
7. Muzny, D. M. et al. Comprehensive molecular characterization of human colon and rectal cancer. Nature **487,** 330–337 (2012).
8. Sottoriva, A. et al. A Big Bang model of human colorectal tumor growth. Nat. Genet. **47,** 209–216 (2015).
9. 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).
10. Wright, S. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. Sixth Int. Congr. Genet. **1,** 356–366 (1932).
11. Yap, T. A., Gerlinger, M., Futreal, P. A., Pusztai, L. & Swanton, C. Intratumor heterogeneity: seeing the wood for the trees. Sci. Transl. Med. **4,** 127ps10–127ps10 (2012).
12. Blum, M. G. B. & François, O. On statistical tests of phylogenetic tree imbalance: The Sackin and other indices revisited. Math. Biosci. **195,** 141–153 (2005).

13. Fischer, A., Illingworth, C. J., Campbell, P. J. & Mustonen, V. EMu: probabilistic inference of mutational processes and their localization in the cancer genome. Genome Biol **14,** R39 (2013).

14. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. Nature **500,** 415–421 (2013).

15. Katainen, R. et al. CTCF/cohesin-binding sites are frequently mutated in cancer. Nat. Genet. **47,** 818–821 (2015).

16. Quirke, P. et al. DNA aneuploidy in colorectal adenomas. Br. J. Cancer **53,** 477 (1986).

17. Jones, A. M. et al. Analysis of copy number changes suggests chromosomal instability in a minority of large colorectal adenomas. J. Pathol. **213,** 249–256 (2007).

18. Wang, H., Liang, L., Fang, J.-Y. & Xu, J. Somatic gene copy number alterations in colorectal cancer: new quest for cancer drivers and biomarkers. Oncogene **35,** 2011–2019 (2016).

19. Durinck, S. et al. Temporal Dissection of Tumorigenesis in Primary Cancers. Cancer Discov. **1,** 137–143 (2011).

20. Newman, S. et al. The relative timing of mutations in a breast cancer genome. PloS One **8,** e64991 (2013).

21. Toyota, M. et al. CpG island methylator phenotype in colorectal cancer. Proc. Natl. Acad. Sci. U. S. A. **96,** 8681–8686 (1999).

22. Kim, T.-M. et al. Subclonal Genomic Architectures of Primary and Metastatic Colorectal Cancer Based on Intratumoral Genetic Heterogeneity. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. **21,** 4461–4472 (2015).

23. Uchi, R. et al. Integrated Multiregional Analysis Proposing a New Model of Colorectal Cancer Evolution. PLOS Genet. **12,** e1005778 (2016).

24. Suzuki, Y. et al. Multiregion ultra-deep sequencing reveals early intermixing and variable levels of intratumoral heterogeneity in colorectal cancer. Mol. Oncol. **11,** 124–139 (2017).

25. Kim, T.-M. et al. Clonal origins and parallel evolution of regionally synchronous colorectal adenoma and carcinoma. Oncotarget **6,** 27725–27735 (2015).

26. Stachler, M. D. et al. Paired exome analysis of Barrett's esophagus and adenocarcinoma. Nat. Genet. **47,** 1047–1055 (2015).

27. Maley, C. C. et al. Genetic clonal diversity predicts progression to esophageal adenocarcinoma. Nat. Genet. **38,** 468–473 (2006).

28. Ross-Innes, C. S. et al. Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. Nat. Genet. **47,** 1038–1046 (2015).

Figure legends

Figure 1. Mutation burdens in CRAs and CRCs

A. CRAs tended to have slightly fewer exonic SNAs than CRCs but the difference was not significant. The average burden and 95% range across these different tumours is shown by the rightmost bars. **B.** The number of individual CNAs (as measured by the number of segmentations) is significantly greater in CRCs than CRAs (p=0.003, 95% range shown by bars). **C.** SNA driver mutation burdens and allelic loss of 5q, 17p and 18q, are shown for each tumour. A comparison of all events is show by the red bars, while tier 1 driver changes exclusively are shown in dark grey, with tier 2 in light grey. **D.** Distribution of canonical driver mutations across tumours. APC is the only ubiquitous driver event. There is no significant enrichment of cnLOH mutations as second hits to APC or TP53 mutations in adenomas compared to carcinomas (though TP53 is borderline).

Figure 2. Phylogenetic analysis of CRAs and MSS CRCs

Maximum parsimony construction of evolutionary trees. For tumours with only two regional biopsies, truncal mutations were simply those shared between the regions. Tier 1 driver mutations (Table S3) are shown, illustrating their enrichment on the trunks, especially in CRCs, indicating they are acquired early in evolutionary time. Phylogenetic trees showed were produced using all available SNAs. Tree shape robustness (branch support) was confirmed by bootstrapping. Branches had greater than 95% support unless otherwise stated (44/55 (80%) of branches had >95% support) . The most parsimonious trees are shown except in carcinoma 6, where one clade could not be resolved (**A:** green box). **Left Bar chart:** Ubiquitous SNAs (found in all regional biopsies and on the trunk of the phylogenetic tree) are compared with sub-clonal SNAs on the phylogenetic tree branches (non-ubiquitous, but present in >1 region) and leaf (present in only one region). CRAs have a smaller proportion of ubiquitous variants than CRCs.

Figure 3. Copy number alterations in CRAs and MSS CRCs

A. A genome-wide view of CNAs is shown for each region of CRAs (top) and CRCs (bottom). Cancers show a greater CNA burden than adenomas, and most CNAs are clonal in cancers, whereas CRAs show more frequent sub-clonal CNAs. Copy number ≥5 is shown as “polysomy”. **B.** The figure shows estimated ploidy and summarises the proportion of each tumour at different copy-states. Black bars show the range of biopsy

copy-numbers. **C**. Size distributions of ubiquitous and sub-clonal (branch and leaf) CNAs demonstrate the preference of CRCs to have larger events. The colour-coding of copy number states (top right) applies to all panels.

Figure 4. Geography of CRCs

Photographs of the tumour specimens from histopathology departments are shown, with biopsy locations marked. The sporadic MSI+ cancer 4 is included here. The corresponding phylogenetic relationship between tumour regions is shown below the photograph of each tumour. The regression plots show pairwise physical and genetic separation for each biopsy from that cancer. There was a significant positive correlation between the phylogenetic (mutational) distance and physical distance in every case.

Figure 5. CNA timing

The plots show the CNA timing results for the six neoplasms with WGS data. For each tumour, the X-axis represents inferred evolutionary time to the MRCA, since tumour initiation (unit of measurement is SNAs accrued per unit time). Green dashed line is inferred from the “second hit” at APC (and thus likely represents the time of initiation of the adenoma). The upper panels show the accumulation of CNAs (red, arrowed line) relative to a steady accumulation (black, dashed line); p-values are derived from Kolmogorov-Smirnov tests of inferred CNA time versus a uniform accumulation. The lower panel shows the estimated times of driver mutations, where these could be derived, for individual CNAs by chromosome arm and type of change. Bars indicate 95% confidence intervals for CNA timing estimates.