Cancer Letters 543 (2022) 215767

Contents lists available at ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet

Clonality and timing of relapsing colorectal cancer metastasis revealed through whole-genome single-cell sequencing

Joao M. Alves [a,b], Sonia Prado-López [a,b,1], Laura Tomás [a,b], Monica Valecha [a,b], Nuria Estévez-Gómez [a,b], Pilar Alvariño [a,b], Dominik Geisel [c], Dominik Paul Modest [d], Igor M. Sauer [e], Johann Pratschke [e], Nathanael Raschzok [e,f], Christine Sers [g,h], Soulafa Mamlouk [g,h,**,2], David Posada [a,b,i,*,2]

[a] CINBIO, Universidade de Vigo, 36310, Vigo, Spain
[b] Galicia Sur Health Research Institute (IIS Galicia Sur), SERGAS-UVIGO, Spain
[c] Department of Radiology, Campus Virchow-Klinikum, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Germany
[d] Department of Hematology, Oncology, and Cancer Immunology, Campus Charité Mitte, Campus Virchow-Klinikum, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany
[e] Department of Surgery, Campus Charité Mitte, Campus Virchow-Klinikum, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu, Berlin, Germany
[f] Berlin Institute of Health at Charité – Universitätsmedizin Berlin, BIH Academy, Clinician Scientist Program, Berlin, Germany
[g] Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Pathology, Charitéplatz 1, 10117, Berlin, Germany
[h] German Cancer Consortium (DKTK), Partner Site Berlin, German Cancer Research Center (DKFZ), Heidelberg, Germany
[i] Department of Biochemistry, Genetics, and Immunology, Universidade de Vigo, 36310, Vigo, Spain

ARTICLE INFO

Keywords:
Single-cell genomics
Phylogenetics
Cancer evolution
Chemotherapy
Mutational signatures

ABSTRACT

Recurrence of tumor cells following local and systemic therapy is a significant hurdle in cancer. Most patients with metastatic colorectal cancer (mCRC) will relapse, despite resection of the metastatic lesions. A better understanding of the evolutionary history of recurrent lesions is required to identify the spatial and temporal patterns of metastatic progression and expose the genetic and evolutionary determinants of therapeutic resistance. With this goal in mind, here we leveraged a unique single-cell whole-genome sequencing dataset from recurrent hepatic lesions of an mCRC patient. Our phylogenetic analysis confirms that the treatment induced a severe demographic bottleneck in the liver metastasis but also that a previously diverged lineage survived this surgery, possibly after migration to a different site in the liver. This lineage evolved very slowly for two years under adjuvant drug therapy and diversified again in a very short period. We identified several non-silent mutations specific to this lineage and inferred a substantial contribution of chemotherapy to the overall, genome-wide mutational burden. All in all, our study suggests that mCRC subclones can migrate locally and evade resection, keep evolving despite rounds of chemotherapy, and re-expand explosively.

1. Introduction

Metastatic colorectal cancer (mCRC) remains one of the most common causes of cancer death [1]. While population-wide screening programs have led to significant improvements in the detection and treatment of early lesions [2], a large proportion of colorectal cancer patients present with distant metastatic disease at the time of diagnosis or after a disease-free period. In mCRC, the liver is the most frequent location of metastasis [3–5]. At this stage, resection of liver metastases

* Corresponding author. CINBIO, Universidade de Vigo, 36310, Vigo, Spain.
** Corresponding author. Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Pathology, Charitéplatz 1, 10117, Berlin, Germany.
E-mail addresses: soulafa.mamlouk@charite.de (S. Mamlouk), dposada@uvigo.es (D. Posada).
[1] Present address: Technische Universität Wien, Austria.
[2] S.M. and D.P. share senior authorship.

https://doi.org/10.1016/j.canlet.2022.215767
Received 21 March 2022; Received in revised form 26 May 2022; Accepted 29 May 2022
Available online 7 June 2022

can improve survival [6,7]. However, 50–75% of patients still suffer a relapse of the disease, primarily as a local recurrence in the liver [8,9].

List of abbreviations

CNV	Copy number variant
COSMIC	Catalog of somatic mutations in cancer
gDNA	Genomic DNA
HPD	Highest posterior density
MCC	Maximum clade credibility
MCMC	Markov chain Monte Carlo
mCRC	Metastatic colorectal cancer
scDNA-seq	Single-cell DNA sequencing
scWGA	Single-cell whole genome amplification
SNV	Single nucleotide variant
XML	Extensible markup language

Studies on the origin and mechanisms of metastatic recurrence have focused mainly on blood and bone marrow cancers, where acquiring longitudinal samples throughout a patient's disease is clinically feasible [10,11]. Studies of recurrence in solid tumors are less common due to the difficulty in obtaining longitudinal samples, although examples exist in triple-negative breast cancer, pancreatic ductal adenocarcinoma, and lung carcinoma [12–14]. The patterns of relapse vary between cancer types; however, most studies found recurrence after treatment to predominantly originate from clones already present in the pre-treated tumor. Despite numerous studies on the evolutionary relationship between primary CRC tumors and spatially or temporally distinct metastases [15–19], information on the subclonal pattern of recurring metastasis post-treatment in mCRC is lacking, and treatment continues to depend on routine clinical procedures.

Survival of tumor cells after surgical resection and therapy is a significant obstacle in treating most cancers. Understanding the evolution of recurrent lesions is critical to identifying the origin of resistant cells and the mechanisms behind cancer relapse, consequently helping to shape treatment strategies. This study investigates the evolution and potential determinants of metastatic liver relapse in mCRC by leveraging a unique single-cell dataset comprising whole-genome information from 60 cells collected from a liver metastasis and a recurrent liver lesion. Using a powerful phylogenetic approach, we were able to reconstruct the genomic history of the relapse, including its origins, age, and

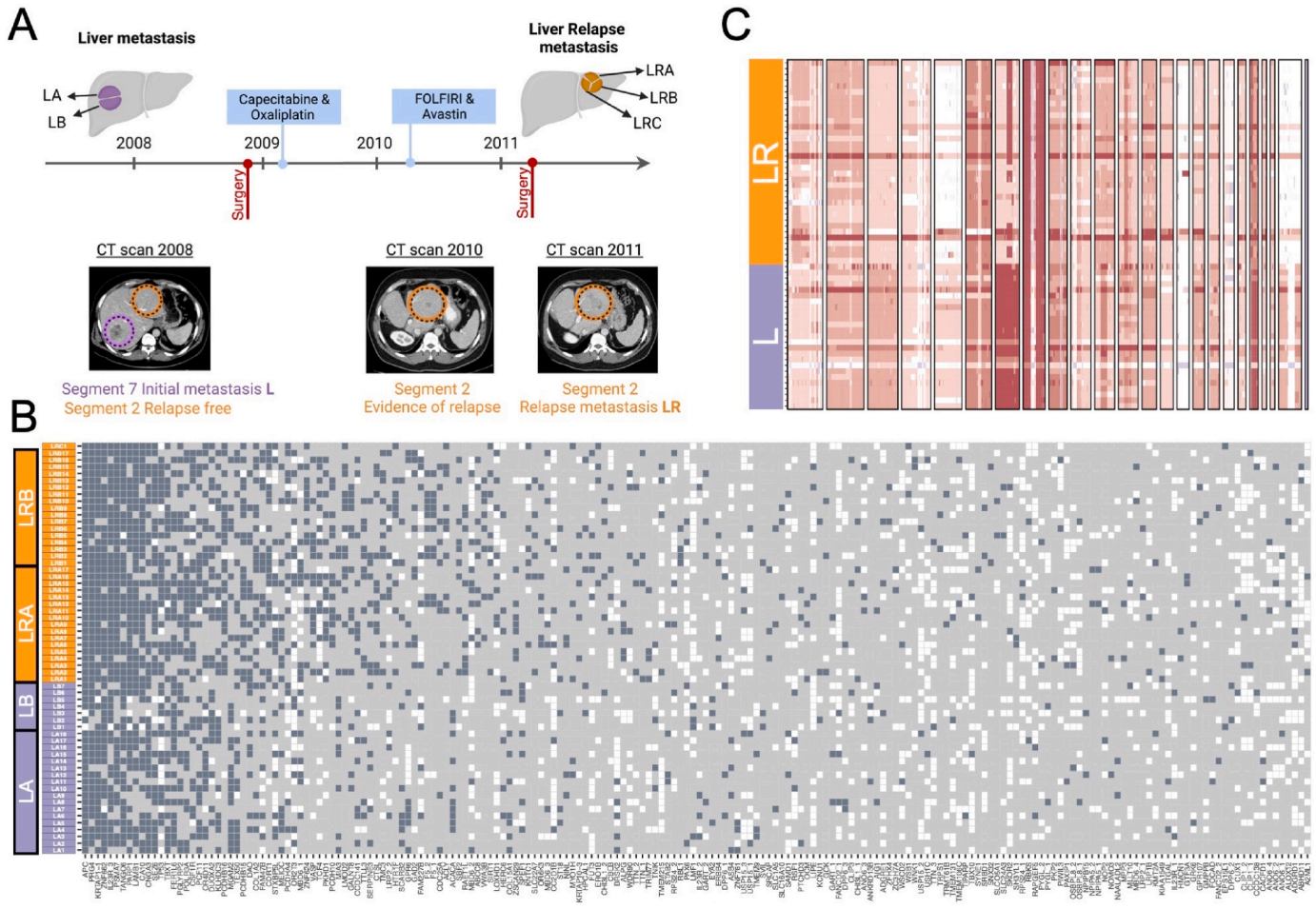

Fig. 1. Overview of clinical progression and genomic characterization of single-cell data obtained from longitudinal metastatic biopsies. A. Clinical course of disease with the timeline depicting the year at which the patient underwent surgical resection - 2008: Primary tumor (rectum) and metastasis (liver) resection; 2009–2010: chemotherapy administration; 2011: metastatic relapse (liver) resection. CT images of metastatic sites are shown below each time point. **B.** Genotype matrix of somatic mutations identified in the malignant single-cells (gray = reference, blue = mutated, white = missing). Only non-silent mutations are shown (n = 194), sorted according to their prevalence across the cell population. Gene names are displayed at the bottom of the map. Each row represents an individual cell. LRC1 is shown at the top of the matrix. Different colors highlight the tissue of origin - Purple: liver metastasis, Orange: liver relapse metastasis. **C.** Heatmap depicting genome-wide copy-number status (deletions: blue, amplifications: red) of tumor single-cells. Chromosomes ordered from 1 to 22, X and Y. Each row represents an individual cell, with the different colors highlighting the tissue of origin. Cell order is the same as in panel b.

clonality, and identified the putative evolutionary and molecular processes behind the reestablishment of the metastasis.

2. Material & methods

2.1. Patient history and sample collection

In November 2008, a male in his 40s diagnosed with a rectal adenocarcinoma with synchronous liver metastasis (pT4N0M1) underwent resection of the primary tumor, followed by a liver resection in December 2008 for removal of the liver metastasis. In February 2009, the patient was treated with eight cycles of adjuvant chemotherapy with capecitabine and oxaliplatin. In February 2010, a PET-CT scan revealed a new lesion in the liver, with histological assessment confirming a metastatic adenocarcinoma of colorectal origin. After that, seven cycles of FOLFIRI and Bevacizumab (VEGF inhibitor) were administered, from March to June 2010, leading to the stabilization of the disease. In February 2011, two years following the first liver resection, the recurrent liver lesion was surgically resected. An overview of the patient's clinical course during this period is depicted in Fig. 1a. The disease was further complicated by lung metastases, treated surgically in 2014, a local recurrence in 2015, and recurrent liver, lung, and bone metastases in 2018, which were treated with palliative radiation until the patient passed away in 2020.

We collected five tumor liver samples from the tumor and neighboring normal liver tissue at two different time points, from segment 7 in 2008 and segment 2 in 2011. All samples were snap-frozen in liquid nitrogen directly after surgical excision and stored at −80 °C. Specifically, we obtained two tumor samples before treatment (liver samples LA and LB), and three tumor samples after metastatic relapse resection (liver relapse samples LRA, LRB, and LRC). In addition, we collected one liver normal sample before treatment (liver normal sample NA).

2.2. Tissue disaggregation and single-cell sorting

We minced each sample into 1-mm^3 pieces with a scalpel and incubated them in Accutase (LINUS) for 1 h at 37 °C. We filtered the cell suspensions with a 70 μm cell strainer (Falcon, NY, USA) and assessed cell viability with Trypan Blue (Gibco, MA, USA). When the percentage of dead cells exceeded 30%, we performed a Ficoll-Paque density gradient centrifugation to discard dead cells and debris before sorting. We washed the cell pellets twice and re-suspended them in ice-cold phosphate-buffered saline (PBS). Then, we incubated them with antibodies against the epithelial cell adhesion molecule (EpCAM) (EBA1, FITC, 347,197) and stem cell markers CD44 (APC, 559,942), CD166 (PE, 559,263), and LGR5 (BV 421, 562925) (all from BD Biosciences, NJ, USA). DRAQ5 (Thermo Fisher Scientific, MA, USA) and 7AAD (Invitrogen, MA, USA) dyes were additionally used to select nucleated cells and exclude non-viable ones. Individual tumor cells were gated and sorted into stem cell (EpCAM+/CD44+/CD166+/Lgr5-; EpCAM-/CD44-/CD166-/Lgr5+) and non-stem cell (EpCAM+/CD44-/CD166-/Lgr5-) phenotypes (Fig. S1). We carried out cell sorting using a FACSAria III (BD Biosciences, NJ, USA) flow cytometer and analyzed the resulting data using FACSDiva (BD Biosciences, NJ, USA) and FlowLogic software (Miltenyi Biotec, Germany).

2.3. Single-cell whole-genome amplification

To obtain enough DNA for sequencing from each cell, we carried out single-cell whole genome amplification (scWGA) with the Ampli1 Kit (Menarini Silicon Biosystems, Italy). To avoid contamination, we worked under a biological safety cabinet, UV-irradiated all the plastic materials employed and used a dedicated set of pipettes. In addition to the patient cells, we included a positive (10 ng/μl REPLIg human control kit, QIAGEN, Netherlands) and negative (DNase/RNase free water) control in the amplification process. We assessed the quality of the amplified DNA with the Ampli1 QC Kit. For positive samples for the four Ampli1 QC PCR fragments, we used the Ampli1 ReAmp/ds kit to increase the amount of total double-stranded DNA. After this, we removed the kit adaptors adding 5 μl of NE Buffer 4 10X (New England Biolabs, MA, USA), 1 μl of *Mse*I 50U/μl (New England Biolabs, MA, USA), and 19 μl of nuclease-free water to every 25 μl of sample and incubated this mixture at 37 °C for 3 h, followed by 20 min at 65 °C for enzyme inactivation. After incubation, we purified the samples using 1.8X AMPure XP beads (Agencourt, Beckman Coulter, CA, USA). Finally, we measured the DNA yield with a Qubit 3.0 fluorometer (Thermo Fisher Scientific, MA, USA) and the amplicon size distribution with the D5000 ScreenTape assay in a 2200 TapeStation platform (Agilent Technologies, CA, USA).

2.4. Bulk genomic DNA (gDNA) isolation

We isolated the gDNA from the normal bulk sample using the QIAamp DNA Mini kit (QIAGEN, Netherlands). Next, we estimated the DNA yield using the Qubit 3.0 fluorometer (Thermo Fisher Scientific, MA, USA) and the DNA integrity with the Genomic DNA ScreenTape Assay (Agilent Technologies, CA, USA).

2.5. Next-generation sequencing

Single-cell and bulk libraries were constructed at the Spanish National Center for Genomic Analysis (CNAG; http://www.cnag.crg.eu) with the KAPA HyperPrep kit (Roche, Sweden). In total, 25 pre-treatment (18 LA + 7 LB) and 35 post-treatment (17 LRA + 17 LRB + 1 LRC) cells, plus one normal bulk sample (NA) were sequenced at ~6X and ~30X, respectively, on an Illumina NovaSeq 6000 (PE150) at CNAG.

2.6. Single-cell variant calling

We aligned the single-cell and bulk sequencing reads to the Genome Reference Consortium Human Build 37 (GRCh37) using the MEM algorithm in the BWA software [20]. Following a standardized best-practices pipeline [21], we filtered out reads displaying low mapping quality, performed local realignment around indels, and removed PCR duplicates. Under default settings, we identified somatic single nucleotide variants (SNVs) for each cell using SCcaller [22], a somatic variant caller specifically designed for scDNA-seq data. SNV calls were annotated using Annovar (v.20,200,608) [23]. In parallel, we called single-cell copy-number variants (CNVs) with GINKGO [24], using variable-length bins of around 500 kb; after binning, the data for each cell was normalized and segmented using default parameters. Afterward, we applied high-stringency thresholds to retain only high confidence calls, keeping mutations with a "True" somatic mutation tag and observed in a minimum of two single cells. Then, we removed non-biallelic sites and sites missing in more than 50% of the sampled cells. In addition, genotypes within genomic regions affected by deletion events were converted to missing data.

2.7. Phylogenetic analysis

We performed a Bayesian phylogenetic analysis using BEAST2 (2.6.4) [25] under a relaxed clock model with exponential growth and with tip dates as days before the present (with the present day being the last sampling date). Since our input consisted of a single-cell genotype matrix comprising variable genomic positions, we specified an SNV ascertainment bias correction by modifying the *constantPatterns* attribute in the input XML file. We ran a single MCMC chain during 500 million generations, recording samples every 1000 generations, and assessed convergence with Tracer v1.7.1. After discarding the first 10% of the samples as burn-in, we obtained point estimates for the different parameters using posterior means. Then, we constructed a maximum

clade credibility (MCC) tree using the median node heights. Finally, we mapped non-synonymous SNVs onto the MCC tree using the single-cell substitution model implemented in CellPhy [26].

2.8. Demographic inference

We inferred the demographic history of the metastatic cells using Bayesian nonparametric phylodynamic reconstruction with preferential sampling (BNPR_PS). This approach tries to avoid the bias introduced by situations in which the sampling times depend on the effective population size [27], as might be the case here. We used the MCC tree as input for the Phylodyn R package (https://github.com/mdkarcher/phylodyn).

2.9. Selection analyses

We obtained the coding sequences for every cell using the dndscv R package [28] and concatenated them into a multiple sequence

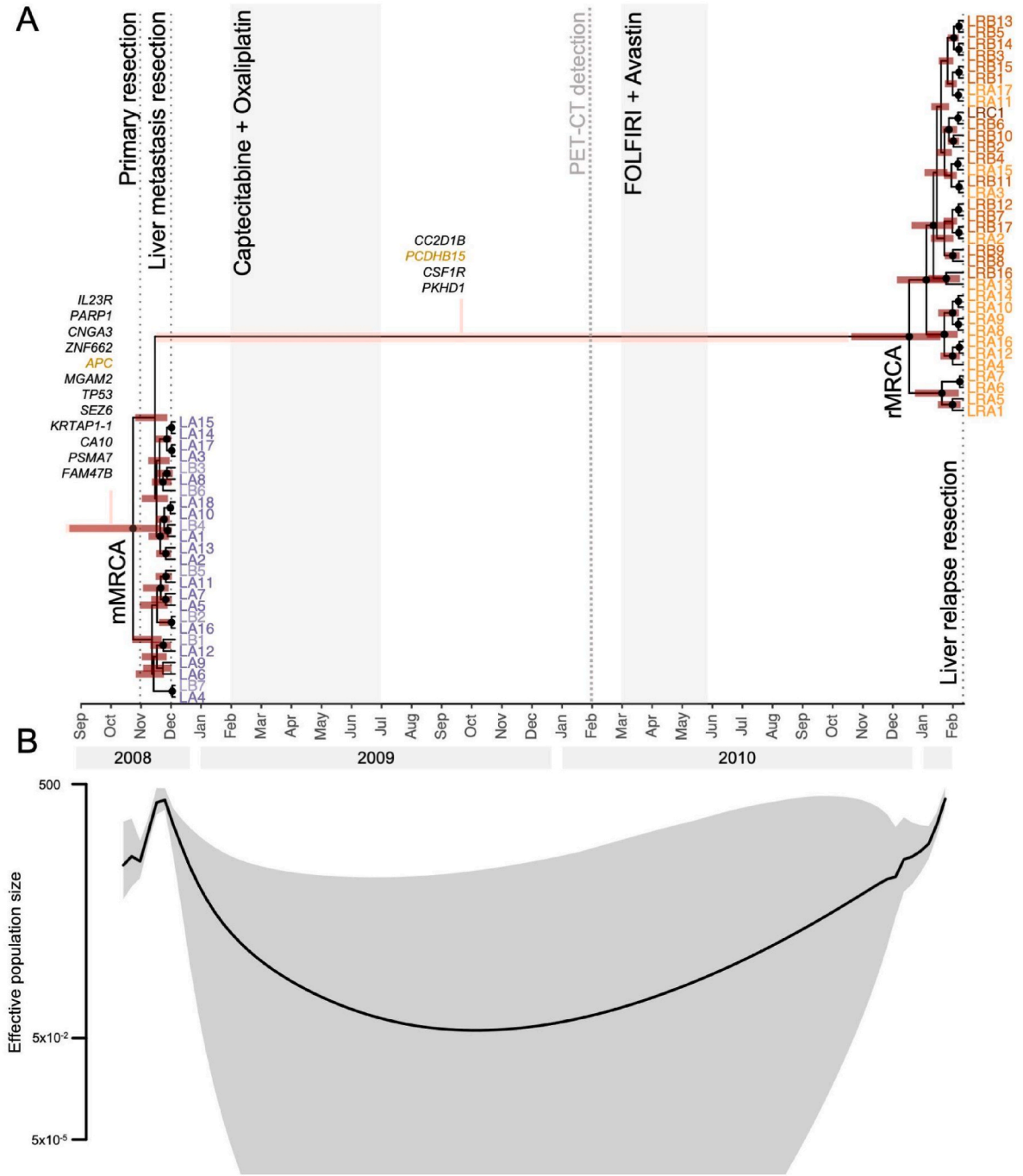

Fig. 2. Phylogenetic and phylodynamic inference. A. Bayesian time-calibrated phylogeny. Tree nodes with posterior probability values > 0.9 are indicated with solid black circles. Red horizontal bars correspond to the 95% highest posterior density intervals of the age estimates. Solid gray bars depict the duration of each chemotherapy regimen. Single-cell IDs are shown at the tips of the tree. Cell label colors highlight the sampled locations (LA, LB, LRA, LRB, LRC). The x-axis is scaled in months. Non-silent mutations (missense in black and nonsense in gold) are displayed on the branches leading to the most recent common ancestor of the metastasis (mMRCA) and relapse (rMRCA) clades **B.** Changes in the effective population size of the metastatic population over time. The shaded area represents the 95% credible interval. The y-axis is on a log scale.

alignment. We then used the RELAX framework [29] to identify shifts in the non-synonymous/synonymous substitution rate ratio (dN/dS) along the cell phylogeny. Specifically, we tested whether differences in dN/dS between the branch leading to the most recent common ancestor of the relapsed cells and the remaining branches of the cell phylogeny could be attributed to changes in the intensity of selection. For this purpose, RELAX estimates a selection intensity parameter (k) and performs a likelihood ratio test to measure the significance of the results. Additionally, we used PAML v4.9j [30] to obtain a maximum likelihood estimate of dN/dS for the whole phylogeny. The significance of this estimate was tested using a likelihood ratio test comparing an unconstrained codon model against a neutral model where the dN/dS ratio is fixed to 1 [31].

2.10. Mutational signature analysis

We ran sigProfilerExtractor [32] under default parameters to identify *de novo* SBS96 mutational signatures in the single cells and then assigned the decomposed signatures to the COSMICv3 ones (https://cancer.sanger.ac.uk/signatures). Moreover, in order to identify the mutational processes active during the diversification of the relapse, we re-ran sigProfilerExtractor using only the mutations mapping to the relapse cells (n = 6,502).

3. Results

3.1. The relapse population descends from a single lineage that survived treatment

Across all samples, we identified a total of 39,380 somatic SNVs, among which 194 were non-synonymous (Fig. 1b). We also found many CNVs across all malignant cells (Fig. 1c) with several marked differences in large-scale events between the pretreatment and relapse cells (e.g., gains in chr 7, 11p, 20, and X). The phylogenetic analysis of the complete set of SNVs clearly showed three distinct and well-supported clades, strongly supporting a monoclonal origin of the relapse (Fig. 2a). Inside each clade, cells did not group according to their sampling location, which suggests a lack of spatial genomic structure within the sampled regions. Notably, 12 non-synonymous SNVs appear to be clonal (i.e., they are shared by all sampled cells), including mutations in *APC*, *TP53* as well as several others affecting genes associated with increased tumor growth and poor prognosis in CRC and multiple cancer types, such as *IL23R* [33], *PARP1* [34], and *PSMA7* [35]. Moreover, our results suggest that the relapse population arose from one of the two major liver metastatic lineages, sharing with it 241 mutations, including five non-synonymous changes, *OR2T3* V117L, *STXBP5L* P783S, *TBXT* A381T, *COL1A2* G592S, and *MEIOC* E538G. Of these, only *TBXT* A381T is found in COSMIC (Genomic mutation ID COSV51616284), in a few samples. Interestingly, the liver relapsing lineage acquired a total of 516 exclusive mutations. Three of these mutations are non-synonymous (*CC2D1B* R611Q, *CSF1R* R41Q, *PKHD1* Q2620E), and one results in a stop codon (*PCDHB15* Q80*). *CC2D1B* R611Q appears in COSMIC (COSV52563677) in two glioblastoma samples. *CSF1R* R41Q can be seen in COSMIC (COSV53842868) in three skin and one kidney tumor samples. *PKHD1* Q2620E is not in COSMIC, but a mutation *PKHD1* Q2620L (COSV61877658) is reported in a single pituitary adenoma sample. *PCDHB15* Q80* is not in COSMIC but a *PCDHB15* Q80H (COSV50859624) mutation is described in a lung cancer sample. We estimated the age of the most recent common ancestor of all the metastatic cells sampled (mMRCA) to be 2.27 years before the relapse resection, with a 95% highest posterior density (HPD) of 2.44–2.19 years. Interestingly, our results suggest that the relapsing lineage started to diversify around 54 days (95% HPD: 22–112 days) before the surgical resection of the metastatic relapse.

3.2. The relapsing lineage evolved slowly after a strong demographic bottleneck induced by the treatment

We expect treatment to induce demographic changes in the tumor cell population. The Bayesian phylodynamic analysis (BNPR_PS) suggests that, following an initial expansion that led to the first metastasis, the tumor cell population underwent a sudden and sharp decline –coincident with the hepatic resection– reaching a flat stationary period after which the tumor cells began steadily expanding again, forming the relapse (Fig. 2b). On the other hand, the relaxed molecular clock analysis resulted in a global substitution rate of 4.31E-8 (median; per site per generation), falling close to previous estimates in CRC [19,36]. However, the substitution rate does not seem to have been constant; we inferred substitution rate changes along the inferred phylogeny spanning over three orders of magnitude (1.8E-10 – 2.3E-07) (Fig. 3). Remarkably, the substitution rate was markedly low on the branch leading to the relapse population. While the RELAX analysis pointed to this branch being under negative selection (dN/dS = 0.26), this result was not statistically significant (likelihood ratio test (LRT) p-value = 0.999). Besides, for the complete cell genealogy, we obtained a genome-wide dN/dS = 1.32. This estimate significantly deviates from neutrality (LRT p-value = 4.7E-05), pointing to positive selection driving the overall metastatic growth.

3.3. The mutational load of the relapse was affected by chemotherapy

We also explored the mutational signatures in the sampled cell populations to look for potential differences along the course of the disease. Of the four *de novo* mutational signatures that we identified (Fig. 4, Fig. S2), signature 96-B was only active in the relapse cells, suggesting differences in the mutational processes between the first metastasis and the relapse lesion (Fig. 4a). This signature, which is predominantly characterized by T > G mutations in a CTT trinucleotide context (Fig. 4b), was further decomposed into COSMIC SBS signatures and found to be enriched in "clock-like" (i.e., aging) signatures SBS1 and SBS5, as well as signatures SBS35 and SBS17b (Fig. 4c). While signature SBS35 has been linked to platinum-based therapies, signature SBS17b shows a clear resemblance to the "5-FU" signature, recently identified in CRC patients undergoing 5-Fluorouracil treatment [37]. Indeed, the appearance of chemotherapy-associated mutational signatures is coincident with Oxaliplatin and FOLFIRI exposure as there were no SBS35 and SBS17b signatures attributed to the pretreatment metastatic cell population. Importantly, we additionally found a strong decay in the activity of these signatures during the diversification of the relapsing lineage (i.e., after treatment was completed) (Fig. S3). During this period, somatic mutagenesis was primarily dominated (89.9% of total contribution) by the same signatures typically found in healthy colorectal cells - SBS1, SBS5, and SBS18 [38,39], thus suggesting that the mutational processes driving somatic variation changed rapidly after treatment ended. Three of the four non-silent mutations clonal for the relapse correspond to C > T mutations, those in *CC2D1B*, *PCDHB15,* and *CSF1R*. Even though SBS35 is mainly associated with C > T changes, these non-synonymous mutations occur in trinucleotide sequence contexts which are not the most common in this signature (TCG > TTG, GCA > GTA, and TCG > TTG, respectively).

4. Discussion

While primary and metastatic CRC tumors have been extensively characterized from a genomic point of view, there have been far fewer analyses of relapse CRC lesions, particularly with intervening systemic therapy in the course of the disease. Consequently, the time of onset and the molecular events underlying metastatic relapse and the associated drug resistance has so far been difficult to characterize. To address this question, here we leveraged a unique dataset comprising 60 single-cell whole-genomes derived from multiple samples, before and after

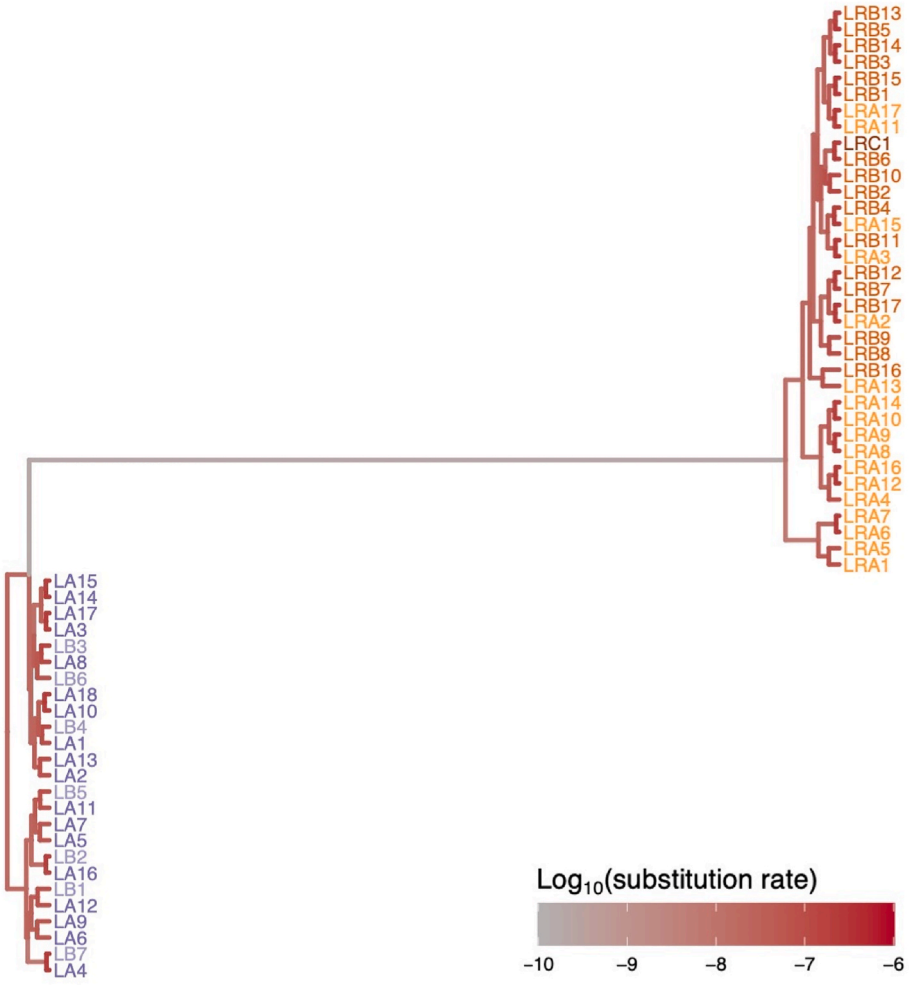

Fig. 3. Substitution rate variation during cancer evolution. Posterior substitution rate estimates per branch. Branch colors indicate the estimated substitution rate (slower = gray, faster = red).

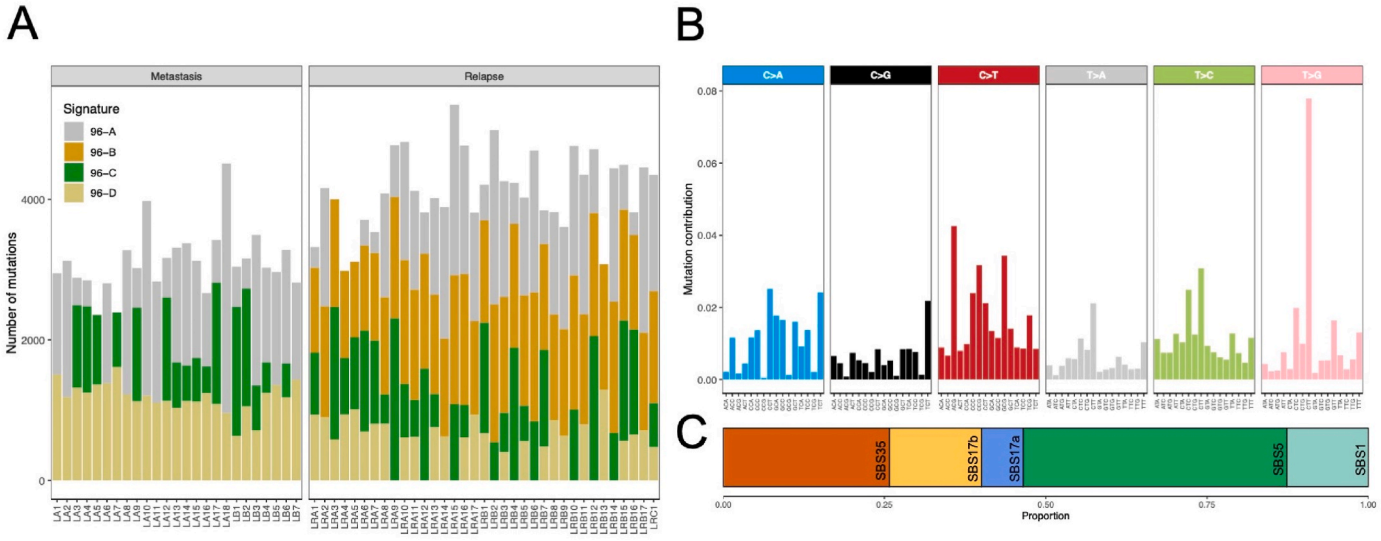

Fig. 4. Mutational signatures across sampled single-cells. A. Barplot depicting the number of mutations contributing to the different signatures/processes across the different single cells. Cells sorted according to the tissue of origin with Cell IDs are shown at the bottom. Different colors reflect the identified mutational signatures. **B.** 96 trinucleotide substitution profiles of the 96-B mutational signature identified in the sampled relapse cells. **C.** COSMIC SBS deconvolution results for signature 96-B.

treatment, from a liver metastasis of a CRC patient. In this study, the use of single-cell whole-genome data, coupled with a sophisticated phylogenetic methodology, enabled us to identify with high statistical support the clonality and the timing of the relapse and infer mutations specific to it. Note that the phylogenetic framework implemented is only available for SNVs, as current statistical models of evolution focus on independent, non-overlapping events, therefore precluding the use of CNVs for this type of quantitative analysis.

Single-cell sequencing offers several advantages over the typical bulk or gland sequencing approaches [19,40–42] for the study of CRC evolution. Bulk sequencing data entail mutational signals from millions of cells belonging to multiple subclones. In such cases, deconvolution algorithms are required to identify the distinct cell populations sampled. Unfortunately, these algorithms make stringent assumptions regarding ploidy, tumor purity, or mutational model, their performance is unclear, and, importantly, are not able to deal with datasets with thousands of SNVs [43–46]. The use of single-glands, mostly clonal [40,47,48], does circumvent some of these problems but restricts the analyses to CRC tumors still preserving a clear glandular structure, i.e., adenomas and early-stage adenocarcinomas. Single-cell sequencing, on the other hand, allows for a very detailed study of intratumoral heterogeneity [49], with an immediate assessment of clonality, in both primary tumors and metastases, and can be very useful to understand the evolution of resistant lineages.

Most of the models of CRC evolution proposed in the literature [reviewed in Ref. [50]] refer to primary tumor development or to the initial metastatic spread from the primary tumor. In the case of a relapse, the main questions of interest are whether only one or several resistant lineages are responsible for the relapse (i.e., it is monoclonal or polyclonal), and the geographical origins of this relapse (i.e., it migrated again from the primary tumor, or recurred *in situ*) [see also [18]]. In the presence of therapy, like in this case, it is also relevant to understand the molecular mechanisms that allowed the surviving lineage/s to relapse. Our phylogenetic and phylodynamic analyses suggest that the metastatic population underwent a strong bottleneck shortly after hepatic resection in November 2008, but that a single metastatic lineage survived the adjuvant therapy and expanded again two years later. If we assume that the relapsing lineage was still undetectable when the hepatic resection was carried out, one could speculate that this lineage likely migrated to another location in the liver (liver segment 2) before the resection. This would explain why only a single lineage survived the resection, and why the relapse lesion was already detected in February 2010 only in liver segment 2 in the PET CT scan.

Interestingly, the substitution rate was considerably reduced in the relapsing lineage. Such a slow pace of genetic change was likely the result of a strong drug-imposed selective bottleneck, which severely decreased the genetic diversity available. Nevertheless, the diversification of both the first liver metastasis and the relapse seems to have been driven by genome-wide positive selection, as seen in many primary tumors [28]. In our study, however, we did not identify CRC-specific driver genes responsible for this process.

We identified three non-synonymous and one stop codon mutations specific to the relapsing lineage in four different genes, CC2D1B, PCDHB15, CSF1R, and PKHD1. None of these genes seem to have an obvious role in drug resistance in mCRC, but some links have been established between them and CRC. Deshar et al. [51] showed that the double knockdown of CC2D1A and CC2D1B leads to accelerated degradation of both TLR4 and EGFR. EGFR is a known regulator of CRC progression and an established target for mCRC therapy [52]. However, in our study, while we find CC2D1B mutated, CC2D1A is not altered, at either the mutational or copy number level, in the relapsing tumor cells. In a recent study of 16 CRC primary and matched metastasis, PCDHB15 was one of the top mutated genes in the liver metastases [53]. These authors suggested that, together with other genes, PCDHB15 mutations might be associated with increased cell mobility in early CRC liver metastasis. Although CSF1R is mostly known as a transmembrane receptor on tumor-associated macrophages, an association between CSF1R mutations and shorter disease-free survival in stage III CRC patients has been reported [54]. Finally, PKDH1 mutations were found more frequently in CRC tumors which displayed loss of ARID1A expression and correlated with early-onset, lymphatic invasion, and lymph node metastasis [55]. Additional investigations would be required to assess the functional role of these genetic alterations in the relapse. Furthermore, recent studies increasingly point to the substantial contribution of non-genetic adaptations, such as DNA methylation, phenotypic plasticity, and immune evasion, to therapy resistance [56–59].

Moreover, we detected a unique mutational pattern affecting the cancer cells from the relapse lesion characterized by a substantial contribution of COSMIC signatures SBS35 and SBS17b. The noticeable contribution of these two mutational signatures to the overall mutational burden of the relapse cells appears to indicate a strong mutagenic effect associated with treatment [37,60]. However, we did not find evidence suggesting that the treatment-induced clinically relevant mutations, unlike other studies [61].

Translating tumor evolution concepts to tackle therapy resistance offers oncologists a conceptual framework for alternative strategies to treatment [62]. Routinely, cancer treatments, such as chemotherapy, are administered at the maximum tolerated dose. As late-stage CRC tumors typically contain extensive genomic heterogeneity [e.g. Refs. [63–66]], both sensitive and resistant subclones can be present. Systemic treatment eradicates sensitive CRC subclones, but subclones with no prior selective advantage can become refractive to treatment [67,68]. These resistance mechanisms can be driven by both genetic mutations and non-genetic determinants, such as cell plasticity and microenvironment influence [69–72]. In that sense, any type of therapy can eventually lead to resistant clones and subsequent disease relapse. In our study, we showed that the relapse originated from a single metastatic lineage that survived *in situ* the demographic bottleneck induced by therapy. This scenario creates an opportunity for local consolidative therapy to alter the trajectory of the subsequent relapsing disease and/or prolong the time to recurrence [61]. Unlike systematic treatment, local consolidative therapies, such as ablation, high-precision radiation, or repeated surgery, reduces the tumor-cell load without increasing the mutational burden. In this context, understanding cancer evolution becomes increasingly important for the development of therapy protocols as well as for patient stratification.

Declarations

4.1. Ethics approval and consent to participate

Tissue collection was performed at Charité - Universitätsmedizin Berlin and processed according to the Charité ethics vote EA1/187/16.

Availability of data and materials

We have deposited all single-cell somatic substitution and CNV calls on Mendeley Data under accession name “Clonality and timing of relapsing colorectal cancer metastasis revealed through whole-genome single-cell sequencing” (doi:10.17632/pbbx6gckck.1).

Funding

This work was supported by the European Research Council (ERC-617457-PHYLOCANCER awarded to D.P.) and by the Spanish Ministry of Science and Innovation - MICINN (PID2019-106247GB-I00 awarded to D.P.). D.P. receives further support from Xunta de Galicia. J.M.A. is supported by an AXA Research Fund Postdoctoral Fellowship and an AECC Investigador 2020 fellowship. M.V. is supported by an H2020/Marie Skłodowska-Curie Actions EU research framework programme grant (Project H2020 MSCA-ITN-2017-766030). L.T. received support

from a Ph.D. fellowship from Xunta de Galicia (ED481A-2018/303). This work was additionally funded by the German Ministry of Education and Research (DKTK to S.M., C.S.) and the DFG (DFG, German Research Foundation-Project MA 8222/2-1, to S.M.). Funding for open access charge: Universidade de Vigo/CISUG.

5.5. Author contributions

D.P. and S.M. conceived the study. D.P. and J.M.A. designed the analyses. C.S., N.R, D.G., J.P, and D.M obtained the tumor samples and patient information. S.P.L., N.E.G. and P.A. processed the samples, obtained the single-cells, and performed whole-genome amplification. J. M.A., L.T., and M.V. performed the analyses. D.G., N.K. and D.M. collected patient information and medical reports. All authors contributed to manuscript writing. All authors read and approved the final manuscript.

CRediT author statement

Joao M. Alves: Methodology, Formal analysis, Data Curation, Writing - Original Draft. Sonia Prado-López: Methodology, Investigation. Laura Tomás: Formal analysis, Data Curation. Monica Valecha: Formal analysis, Data Curation. Nuria Estévez-Gómez: Investigation, Data Curation. Pilar Alvariño: Investigation. Dominik Geisel: Resources, Data Curation. Dominik Paul Modest: Resources, Data Curation. Igor M. Sauer: Resources. Johann Pratschke: Resources. Nathanael Raschzok: Resources, Data Curation. Christine Sers: Resources. Soulafa Mamlouk: Conceptualization, Writing - Original Draft, Supervision, Project administration. David Posada: Conceptualization, Methodology, Writing - Original Draft, Supervision, Project administration, Funding acquisition.

Declaration of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank all members of the Phylogenomics lab for their comments on earlier versions of the manuscript. We also thank the Supercomputation Center of Galicia (CESGA) for providing all computational resources. We thank Dr. Med. Bert Hildebrandt (Klinik für Gastroenterologie, Hämatologie und Internistische Onkologie GLG Werner Forßmann Klinikum Eberswalde) for clinical feedback. We want to thank an anonymous reviewer for helpful suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.canlet.2022.215767.

References

[1] Y. Xi, P. Xu, Global colorectal cancer burden in 2020 and projections to 2040, Transl, Oncol. 14 (2021), 101174.

[2] T.R. Levin, D.A. Corley, C.D. Jensen, J.E. Schottinger, V.P. Quinn, A.G. Zauber, J. K. Lee, W.K. Zhao, N. Udaltsova, N.R. Ghai, A.T. Lee, C.P. Quesenberry, B. H. Fireman, C.A. Doubeni, Effects of organized colorectal cancer screening on cancer incidence and mortality in a large community-based population, Gastroenterology 155 (2018) 1383–1391, https://doi.org/10.1053/j.gastro.2018.07.017, e5.

[3] L.G.M. van der Geest, J. 't Lam-Boer, M. Koopman, C. Verhoef, M.A.G. Elferink, J. H.W. de Wilt, Nationwide trends in incidence, treatment and survival of colorectal cancer patients with synchronous metastases, Clin. Exp. Metastasis 32 (2015) 457–465.

[4] H. Sorbye, Recurrence patterns after resection of liver metastases from colorectal cancer, Recent Results Cancer Res. 203 (2014) 243–252.

[5] S. Filip, V. Vymetalkova, J. Petera, L. Vodickova, O. Kubecek, S. John, F. Cecka, M. Krupova, M. Manethova, K. Cervena, P. Vodicka, Distant metastasis in colorectal cancer patients-do we have new predicting clinicopathological and molecular biomarkers? A comprehensive review, Int. J. Mol. Sci. 21 (2020), https://doi.org/10.3390/ijms21155255.

[6] E. Van Cutsem, A. Cervantes, R. Adam, A. Sobrero, J.H. Van Krieken, D. Aderka, E. Aranda Aguilar, A. Bardelli, A. Benson, G. Bodoky, F. Ciardiello, A. D'Hoore, E. Diaz-Rubio, J.-Y. Douillard, M. Ducreux, A. Falcone, A. Grothey, T. Gruenberger, K. Haustermans, V. Heinemann, P. Hoff, C.-H. Köhne, R. Labianca, P. Laurent-Puig, B. Ma, T. Maughan, K. Muro, N. Normanno, P. Österlund, W.J.G. Oyen, D. Papamichael, G. Pentheroudakis, P. Pfeiffer, T.J. Price, C. Punt, J. Ricke, A. Roth, R. Salazar, W. Scheithauer, H.J. Schmoll, J. Tabernero, J. Taïeb, S. Tejpar, H. Wasan, T. Yoshino, A. Zaanan, D. Arnold, ESMO consensus guidelines for the management of patients with metastatic colorectal cancer, Ann. Oncol. 27 (2016) 1386–1422.

[7] S. Kopetz, G.J. Chang, M.J. Overman, C. Eng, D.J. Sargent, D.W. Larson, A. Grothey, J.-N. Vauthey, D.M. Nagorney, R.R. McWilliams, Improved survival in metastatic colorectal cancer is associated with adoption of hepatic resection and improved chemotherapy, J. Clin. Oncol. 27 (2009) 3677–3683.

[8] K. Imai, M.-A. Allard, C.C. Benitez, E. Vibert, A. Sa Cunha, D. Cherqui, D. Castaing, H. Bismuth, H. Baba, R. Adam, Early recurrence after hepatectomy for colorectal liver metastases: what optimal definition and what predictive factors? Oncol. 21 (2016) 887–894.

[9] M. D'Angelica, P. Kornprat, M. Gonen, R.P. DeMatteo, Y. Fong, L.H. Blumgart, W. R. Jarnagin, Effect on outcome of recurrence patterns after hepatectomy for colorectal metastases, Ann. Surg Oncol. 18 (2011) 1096–1103.

[10] E. Waanders, Z. Gu, S.M. Dobson, Ž. Antić, J.C. Crawford, X. Ma, M.N. Edmonson, D. Payne-Turner, M. van de Vorst, M.C.J. Jongmans, I. McGuire, X. Zhou, J. Wang, L. Shi, S. Pounds, D. Pei, C. Cheng, G. Song, Y. Fan, Y. Shao, M. Rusch, K. McCastlain, J. Yu, R. van Boxtel, F. Blokzijl, I. Iacobucci, K.G. Roberts, J. Wen, G. Wu, J. Ma, J. Easton, G. Neale, S.R. Olsen, K.E. Nichols, C.-H. Pui, J. Zhang, W. E. Evans, M.V. Relling, J.J. Yang, P.G. Thomas, J.E. Dick, R.P. Kuiper, C. G. Mullighan, Mutational landscape and patterns of clonal evolution in relapsed pediatric acute lymphoblastic leukemia, Blood Cancer Discov 1 (2020) 96–111.

[11] L.I. Shlush, A. Mitchell, L. Heisler, S. Abelson, S.W.K. Ng, A. Trotman-Grant, J.J. F. Medeiros, A. Rao-Bhatia, I. Jaciw-Zurakowsky, R. Marke, J.L. McLeod, M. Doedens, G. Bader, V. Voisin, C. Xu, J.D. McPherson, T.J. Hudson, J.C.Y. Wang, M.D. Minden, J.E. Dick, Tracing the origins of relapse in acute myeloid leukaemia to stem cells, Nature 547 (2017) 104–108.

[12] J. Chen, F. Facchinetti, F. Braye, A.A. Yurchenko, L. Bigot, S. Ponce, D. Planchard, A. Gazzah, S. Nikolaev, S. Michiels, D. Vasseur, L. Lacroix, L. Tselikas, C. Nobre, K. A. Olaussen, F. Andre, J.Y. Scoazec, F. Barlesi, J.C. Soria, Y. Loriot, B. Besse, L. Friboulet, Single cell DNA-seq depicts clonal evolution of multiple driver alterations in osimertinib resistant patients, Ann. Oncol. 33 (2022) 434–444, https://doi.org/10.1016/j.annonc.2022.01.004.

[13] H. Sakamoto, M.A. Attiyeh, J.M. Gerold, A.P. Makohon-Moore, A. Hayashi, J. Hong, R. Kappagantula, L. Zhang, J.P. Melchor, J.G. Reiter, A. Heyde, C. M. Bielski, A.V. Penson, M. Gönen, D. Chakravarty, E.M. O'Reilly, L.D. Wood, R. H. Hruban, M.A. Nowak, N.D. Socci, B.S. Taylor, C.A. Iacobuzio-Donahue, The evolutionary origins of recurrent pancreatic cancer, Cancer Discov. 10 (2020) 792–805.

[14] C. Kim, R. Gao, E. Sei, R. Brandt, J. Hartman, T. Hatschek, N. Crosetto, T. Foukakis, N.E. Navin, Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing, Cell 173 (2018) 879–893.e13.

[15] N. Ishaque, M.L. Abba, C. Hauser, N. Patil, N. Paramasivam, D. Huebschmann, J. H. Leupold, G.P. Balasubramanian, K. Kleinheinz, U.H. Toprak, B. Hutter, A. Benner, A. Shavinskaya, C. Zhou, Z. Gu, J. Kerssemakers, A. Marx, M. Moniuszko, M. Kozlowski, J. Reszec, J. Niklinski, J. Eils, M. Schlesner, R. Eils, B. Brors, H. Allgayer, Whole genome sequencing puts forward hypotheses on metastasis evolution and therapy in colorectal cancer, Nat. Commun. 9 (2018) 4782.

[16] H.-N. Chen, Y. Shu, F. Liao, X. Liao, H. Zhang, Y. Qin, Z. Wang, M. Luo, Q. Liu, Z. Xue, M. Cao, S. Zhang, W.-H. Zhang, Q. Hou, X. Xia, H. Luo, Y. Zhang, L. Yang, J.-K. Hu, X. Fu, B. Liu, H. Hu, C. Huang, Y. Peng, W. Cheng, L. Dai, L. Yang, W. Zhang, B. Dong, Y. Li, Y. Wei, H. Xu, Z.-G. Zhou, Genomic evolution and diverse models of systemic metastases in colorectal cancer, Gut 71 (2022) 322–332.

[17] J.M. Alves, S. Prado-López, J.M. Cameselle-Teijeiro, D. Posada, Rapid evolution and biogeographic spread in a colorectal cancer, Nat. Commun. 10 (2019) 5139.

[18] K. Naxerova, J.G. Reiter, E. Brachtel, J.K. Lennerz, M. van de Wetering, A. Rowan, T. Cai, H. Clevers, C. Swanton, M.A. Nowak, S.J. Elledge, R.K. Jain, Origins of lymphatic and distant metastases in human colorectal cancer, Science 357 (2017) 55–60.

[19] Z. Hu, J. Ding, Z. Ma, R. Sun, J.A. Seoane, J. Scott Shaffer, C.J. Suarez, A. S. Berghoff, C. Cremolini, A. Falcone, F. Loupakis, P. Birner, M. Preusser, H.-J. Lenz, C. Curtis, Quantitative evidence for early metastatic seeding in colorectal cancer, Nat. Genet. 51 (2019) 1113–1122.

[20] H. Li, Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM, 2013 arXiv [q-bio.GN], http://arxiv.org/abs/1303.3997.

[21] G.A. Van der Auwera, M.O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K.V. Garimella, D. Altshuler, S. Gabriel, M.A. DePristo, From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline, Curr. Protoc. Bioinformatics. 43 (2013) 11, 10.1–11.10.33.

[22] X. Dong, L. Zhang, B. Milholland, M. Lee, A.Y. Maslov, T. Wang, J. Vijg, Accurate identification of single-nucleotide variants in whole-genome-amplified single cells, Nat. Methods 14 (2017) 491–493.

[23] K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data, Nucleic Acids Res. 38 (2010) e164.

[24] T. Garvin, R. Aboukhalil, J. Kendall, T. Baslan, G.S. Atwal, J. Hicks, M. Wigler, M. C. Schatz, Interactive analysis and assessment of single-cell copy-number variations, Nat. Methods 12 (2015) 1058–1060.

[25] R. Bouckaert, T.G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, A. Gavryushkina, J. Heled, G. Jones, D. Kühnert, N. De Maio, M. Matschiner, F. K. Mendes, N.F. Müller, H.A. Ogilvie, L. du Plessis, A. Popinga, A. Rambaut, D. Rasmussen, I. Siveroni, M.A. Suchard, C.-H. Wu, D. Xie, C. Zhang, T. Stadler, A. J. Drummond, Beast 2.5: an advanced software platform for Bayesian evolutionary analysis, PLoS Comput. Biol. 15 (2019), e1006650.

[26] A.M. Kozlov, J.M. Alves, A. Stamatakis, D. Posada, CellPhy: Accurate and Fast Probabilistic Inference of Single-Cell Phylogenies from scDNA-Seq Data, bioRxiv, 2020, https://doi.org/10.1101/2020.07.31.230292.

[27] M.D. Karcher, J.A. Palacios, S. Lan, V.N. Minin, Phylodyn : an R package for phylodynamic simulation and inference, Molecular Ecology Resources 17 (2017) 96–100, https://doi.org/10.1111/1755-0998.12630.

[28] I. Martincorena, K.M. Raine, M. Gerstung, K.J. Dawson, K. Haase, P. Van Loo, H. Davies, M.R. Stratton, P.J. Campbell, Universal patterns of selection in cancer and somatic tissues, Cell 173 (2018) 1823.

[29] J.O. Wertheim, B. Murrell, M.D. Smith, S.L. Kosakovsky Pond, K. Scheffler, RELAX: detecting relaxed selection in a phylogenetic framework, Mol. Biol. Evol. 32 (2015) 820–832.

[30] Z. Yang, Paml 4: phylogenetic analysis by maximum likelihood, Mol. Biol. Evol. 24 (2007) 1586–1591.

[31] R. Nielsen, Z. Yang, Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene, Genetics 148 (1998) 929–936.

[32] S.M. Ashiqul Islam, Y. Wu, M. Díaz-Gay, E.N. Bergstrom, Y. He, M. Barnes, M. Vella, J. Wang, J.W. Teague, P. Clapham, S. Moody, S. Senkin, Y.R. Li, L. Riva, T. Zhang, A.J. Gruber, R. Vangara, C.D. Steele, B. Otlu, A. Khandekar, A. Abbasi, L. Humphreys, N. Syulyukina, S.W. Brady, B.S. Alexandrov, N. Pillay, J. Zhang, D. J. Adams, I. Martincorena, D.C. Wedge, M.T. Landi, P. Brennan, M.R. Stratton, S. G. Rozen, L.B. Alexandrov, Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor, bioRxiv (2021), https://doi.org/10.1101/2020.12.13.422570, 2020.12.13.422570.

[33] H. Suzuki, H. Ogawa, K. Miura, S. Haneda, K. Watanabe, S. Ohnuma, H. Sasaki, T. Sase, S. Kimura, T. Kajiwara, T. Komura, M. Toshima, Y. Matsuda, C. Shibata, I. Sasaki, IL-23 directly enhances the proliferative and invasive activities of colorectal carcinoma, Oncol. Lett. 4 (2012) 199.

[34] K. Nosho, H. Yamamoto, M. Mikami, H. Taniguchi, T. Takahashi, Y. Adachi, A. Imamura, K. Imai, Y. Shinomura, Overexpression of poly(ADP-ribose) polymerase-1 (PARP-1) in the early stage of colorectal carcinogenesis, Eur. J. Cancer 42 (2006) 2374–2381.

[35] L. Yang, Z. Tang, H. Zhang, W. Kou, Z. Lu, X. Li, Q. Li, Z. Miao, PSMA7 directly interacts with NOD1 and regulates its function, Cell. Physiol. Biochem. 31 (2013) 952–959.

[36] S. Jones, W.-D. Chen, G. Parmigiani, F. Diehl, N. Beerenwinkel, T. Antal, A. Traulsen, M.A. Nowak, C. Siegel, V.E. Velculescu, K.W. Kinzler, B. Vogelstein, J. Willis, S.D. Markowitz, Comparative lesion sequencing provides insights into tumor evolution, Proc. Natl. Acad. Sci. U. S. A 105 (2008) 4283–4288.

[37] S. Christensen, B. Van der Roest, N. Besselink, R. Janssen, S. Boymans, J.W. M. Martens, M.-L. Yaspo, P. Priestley, E. Kuijk, E. Cuppen, A. Van Hoeck, 5-Fluorouracil treatment induces characteristic T>G mutations in human cancer, Nat. Commun. 10 (2019) 1–11.

[38] H. Lee-Six, S. Olafsson, P. Ellis, R.J. Osborne, M.A. Sanders, L. Moore, N. Georgakopoulos, F. Torrente, A. Noorani, M. Goddard, P. Robinson, T.H. H. Coorens, L. O'Neill, C. Alder, J. Wang, R.C. Fitzgerald, M. Zilbauer, N. Coleman, K. Saeb-Parsy, I. Martincorena, P.J. Campbell, M.R. Stratton, The landscape of somatic mutation in normal colorectal epithelial cells, Nature 574 (2019) 532–537.

[39] A. Cagan, A. Baez-Ortega, N. Brzozowska, F. Abascal, T.H.H. Coorens, M. A. Sanders, A.R.J. Lawson, L.M.R. Harvey, S. Bhosle, D. Jones, R.E. Alcantara, T. M. Butler, Y. Hooks, K. Roberts, E. Anderson, S. Lunn, E. Flach, S. Spiro, I. Januszczak, E. Wrigglesworth, H. Jenkins, T. Dallas, N. Masters, M.W. Perkins, R. Deaville, M. Druce, R. Bogeska, M.D. Milsom, B. Neumann, F. Gorman, F. Constantino-Casas, L. Peachey, D. Bochynska, E.S.J. Smith, M. Gerstung, P. J. Campbell, E.P. Murchison, M.R. Stratton, I. Martincorena, Somatic mutation rates scale with lifespan across mammals, Nature 604 (2022) 517–524.

[40] A. Sottoriva, H. Kang, Z. Ma, T.A. Graham, M.P. Salomon, J. Zhao, P. Marjoram, K. Siegmund, M.F. Press, D. Shibata, C. Curtis, A Big Bang model of human colorectal tumor growth, Nat. Genet. 47 (2015) 209–216.

[41] S. Mamlouk, T. Simon, L. Tomás, D.C. Wedge, A. Arnold, A. Menne, D. Horst, D. Capper, M. Morkel, D. Posada, C. Sers, H. Bläker, Malignant transformation and genetic alterations are uncoupled in early colorectal cancer progression, BMC Biol. 18 (2020) 116.

[42] W. Cross, M. Kovac, V. Mustonen, D. Temko, H. Davis, A.-M. Baker, S. Biswas, R. Arnold, L. Chegwidden, C. Gatenbee, A.R. Anderson, V.H. Koelzer, P. Martinez, X. Jiang, E. Domingo, D.J. Woodcock, Y. Feng, M. Kovacova, T. Maughan, S.C.O.R. T. Consortium, M. Jansen, M. Rodriguez-Justo, S. Ashraf, R. Guy, C. Cunningham, J.E. East, D.C. Wedge, L.M. Wang, C. Palles, K. Heinimann, A. Sottoriva, S. J. Leedham, T.A. Graham, I.P.M. Tomlinson, The evolutionary landscape of colorectal tumorigenesis, Nat Ecol Evol 2 (2018) 1661–1672.

[43] S. Miura, T. Vu, J. Deng, T. Buturla, O. Oladeinde, J. Choi, S. Kumar, Power and pitfalls of computational methods for inferring clone phylogenies and mutation orders from bulk sequencing data, Sci. Rep. 10 (2020) 3498.

[44] G. Tanner, D.R. Westhead, A. Droop, L.F. Stead, Benchmarking pipelines for subclonal deconvolution of bulk tumour sequencing data, Nat. Commun. 12 (2021) 6396.

[45] A. Salcedo, M. Tarabichi, S.M.G. Espiritu, A.G. Deshwar, M. David, N.M. Wilson, S. Dentro, J.A. Wintersinger, L.Y. Liu, M. Ko, S. Sivanandan, H. Zhang, K. Zhu, T.-H. Ou Yang, J.M. Chilton, A. Buchanan, C.M. Lalansingh, C. P'ng, C.V. Anghel, I. Umar, B. Lo, W. Zou, D.R.E.A.M. Smc-Het Participants, J.T. Simpson, J.M. Stuart, D. Anastassiou, Y. Guan, A.D. Ewing, K. Ellrott, D.C. Wedge, Q. Morris, P. Van Loo, P.C. Boutros, A community effort to create standards for evaluating tumor subclonal reconstruction, Nat. Biotechnol. 38 (2020) 97–107.

[46] M. Tarabichi, A. Salcedo, A.G. Deshwar, M. Ni Leathlobhair, J. Wintersinger, D. C. Wedge, P. Van Loo, Q.D. Morris, P.C. Boutros, A practical guide to cancer subclonal reconstruction from DNA sequencing, Nat. Methods 18 (2021) 144–155.

[47] M.D. Ryser, B.-H. Min, K.D. Siegmund, D. Shibata, Spatial mutation patterns as markers of early colorectal tumor cell mobility, Proc. Natl. Acad. Sci. U. S. A 115 (2018) 5774–5779.

[48] M.D. Ryser, D. Mallo, A. Hall, T. Hardman, L.M. King, S. Tatishchev, I.C. Sorribes, C.C. Maley, J.R. Marks, E.S. Hwang, D. Shibata, Minimal barriers to invasion during human colorectal tumor growth, Nat. Commun. 11 (2020) 1280.

[49] A. Bowes, M. Tarabichi, N. Pillay, P. Van Loo, Leveraging single cell sequencing to unravel intra-tumour heterogeneity and tumour evolution in human cancers, J. Pathol. (2022), https://doi.org/10.1002/path.5914.

[50] A. Amaro, S. Chiara, U. Pfeffer, Molecular evolution of colorectal cancer: from multistep carcinogenesis to the big bang, Cancer Metastasis Rev. 35 (2016) 63–74.

[51] R. Deshar, E.-B. Cho, S.K. Yoon, J.-B. Yoon, CC2D1A and CC2D1B regulate degradation and signaling of EGFR and TLR4, Biochem. Biophys. Res. Commun. 480 (2016) 280–287.

[52] E. Martinelli, D. Ciardiello, G. Martini, T. Troiani, C. Cardone, P.P. Vitiello, N. Normanno, A.M. Rachiglio, E. Maiello, T. Latiano, F. De Vita, F. Ciardiello, Implementing anti-epidermal growth factor receptor (EGFR) therapy in metastatic colorectal cancer: challenges and future perspectives, Ann. Oncol. 31 (2020) 30–40.

[53] C. Li, J. Xu, X. Wang, C. Zhang, Z. Yu, J. Liu, Z. Tai, Z. Luo, X. Yi, Z. Zhong, Whole exome and transcriptome sequencing reveal clonal evolution and exhibit immune-related features in metastatic colorectal tumors, Cell Death Dis. 7 (2021) 222.

[54] P.-C. Lin, Y.-M. Yeh, B.-W. Lin, S.-C. Lin, R.-H. Chan, P.-C. Chen, M.-R. Shen, Intratumor heterogeneity of MYO18A and FBXW7 variants impact the clinical outcome of stage III colorectal cancer, Front. Oncol. 10 (2020), 588557.

[55] Y. Kishida, T. Oishi, T. Sugino, A. Shiomi, K. Urakami, M. Kusuhara, K. Yamaguchi, Y. Kitagawa, H. Ono, Associations between loss of ARID1A expression and clinicopathologic and genetic variables in T1 early colorectal cancer, Am. J. Clin. Pathol. 152 (2019) 463–470.

[56] S.M. Shaffer, M.C. Dunagin, S.R. Torborg, E.A. Torre, B. Emert, C. Krepler, M. Beqiri, K. Sproesser, P.A. Brafford, M. Xiao, E. Eggan, I.N. Anastopoulos, C. A. Vargas-Garcia, A. Singh, K.L. Nathanson, M. Herlyn, A. Raj, Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance, Nature 546 (2017) 431–435.

[57] C.C. Bell, K.A. Fennell, Y.-C. Chan, F. Rambow, M.M. Yeung, D. Vassiliadis, L. Lara, P. Yeh, L.G. Martelotto, A. Rogiers, B.E. Kremer, O. Barbash, H.P. Mohammad, T. M. Johanson, M.L. Burr, A. Dhar, N. Karpinich, L. Tian, D.S. Tyler, L. MacPherson, J. Shi, N. Pinnawala, C. Yew Fong, A.T. Papenfuss, S.M. Grimmond, S.-J. Dawson, R.S. Allan, R.G. Kruger, C.R. Vakoc, D.L. Goode, S.H. Naik, O. Gilan, E.Y.N. Lam, J.-C. Marine, R.K. Prinjha, M.A. Dawson, Targeting enhancer switching overcomes non-genetic drug resistance in acute myeloid leukaemia, Nat. Commun. 10 (2019) 2723.

[58] K. Kemper, P.L. de Goeje, D.S. Peeper, R. van Amerongen, Phenotype switching: tumor cell plasticity as a resistance mechanism and target for therapy, Cancer Res. 74 (2014) 5937–5941.

[59] Y. Oren, M. Tsabar, M.S. Cuoco, L. Amir-Zilberstein, H.F. Cabanos, J.-C. Hütter, B. Hu, P.I. Thakore, M. Tabaka, C.P. Fulco, W. Colgan, B.M. Cuevas, S.A. Hurvitz, D.J. Slamon, A. Deik, K.A. Pierce, C. Clish, A.N. Hata, E. Zaganjor, G. Lahav, K. Politi, J.S. Brugge, A. Regev, Cycling cancer persister cells arise from lineages with distinct programs, Nature 596 (2021) 576–582.

[60] O. Pich, F. Muiños, M.P. Lolkema, N. Steeghs, A. Gonzalez-Perez, N. Lopez-Bigas, The mutational footprints of cancer therapies, Nat. Genet. 51 (2019) 1732–1740.

[61] J.N. Fisk, A.R. Mahal, A. Dornburg, S.G. Gaffney, S. Aneja, J.N. Contessa, D. Rimm, J.B. Yu, J.P. Townsend, Premetastatic shifts of endogenous and exogenous mutational processes support consolidative therapy in EGFR-driven lung adenocarcinoma, Cancer Lett. 526 (2022) 346–351.

[62] S. Turajlic, C. Swanton, Implications of cancer evolution for drug development, Nat. Rev. Drug Discov. 16 (2017) 441–442.

[63] D.K.H. Chan, S.J.A. Buczacki, Tumour heterogeneity and evolutionary dynamics in colorectal cancer, Oncogenesis 10 (2021) 53.

[64] Y.-W. Lu, H.-F. Zhang, R. Liang, Z.-R. Xie, H.-Y. Luo, Y.-J. Zeng, Y. Xu, L.-M. Wang, X.-Y. Kong, K.-H. Wang, Colorectal cancer genetic heterogeneity delineated by multi-region sequencing, PLoS One 11 (2016), e0152673.

[65] T.-M. Kim, S.-H. Jung, C.H. An, S.H. Lee, I.-P. Baek, M.S. Kim, S.-W. Park, J.-K. Rhee, S.-H. Lee, Y.-J. Chung, Subclonal genomic architectures of primary and metastatic colorectal cancer based on intratumoral genetic heterogeneity, Clin. Cancer Res. 21 (2015) 4461–4472.

[66] Q. Wei, Z. Ye, X. Zhong, L. Li, C. Wang, R.E. Myers, J.P. Palazzo, D. Fortuna, A. Yan, S.A. Waldman, X. Chen, J.A. Posey, A. Basu-Mallick, B.H. Jiang, L. Hou, J. Shu, Y. Sun, J. Xing, B. Li, H. Yang, Multiregion whole-exome sequencing of matched primary and metastatic tumors revealed genomic heterogeneity and suggested polyclonal seeding in colorectal cancer metastasis, Ann. Oncol. 28 (2017) 2135–2141.

[67] M. Russo, G. Crisafulli, A. Sogari, N.M. Reilly, S. Arena, S. Lamba, A. Bartolini, V. Amodio, A. Magrì, L. Novara, I. Sarotto, Z.D. Nagel, C.G. Piett, A. Amatu, A. Sartore-Bianchi, S. Siena, A. Bertotti, L. Trusolino, M. Corigliano, M. Gherardi, M.C. Lagomarsino, F. Di Nicolantonio, A. Bardelli, Adaptive mutability of colorectal cancers in response to targeted therapies, Science 366 (2019) 1473–1480.

[68] A. Niida, K. Mimori, T. Shibata, S. Miyano, Modeling colorectal cancer evolution, J. Hum. Genet. 66 (2021) 869–878.

[69] J.-C. Marine, S.-J. Dawson, M.A. Dawson, Non-genetic mechanisms of therapeutic resistance in cancer, Nat. Rev. Cancer 20 (2020) 743–756.

[70] O. Marin-Bejar, A. Rogiers, M. Dewaele, J. Femel, P. Karras, J. Pozniak, G. Bervoets, N. Van Raemdonck, D. Pedri, T. Swings, J. Demeulemeester, S. V. Borght, S. Lehnert, F. Bosisio, J.J. van den Oord, I.V. Bempt, D. Lambrechts, T. Voet, O. Bechter, H. Rizos, M.P. Levesque, E. Leucci, A.W. Lund, F. Rambow, J.-C. Marine, Evolutionary predictability of genetic versus nongenetic resistance to anticancer drugs in melanoma, Cancer Cell 39 (2021) 1135–1149.e8.

[71] J.R.M. Black, N. McGranahan, Genetic and non-genetic clonal diversity in cancer evolution, Nat. Rev. Cancer 21 (2021) 379–392.

[72] M. Ramirez, S. Rajaram, R.J. Steininger, D. Osipchuk, M.A. Roth, L.S. Morinishi, L. Evans, W. Ji, C.-H. Hsu, K. Thurley, S. Wei, A. Zhou, P.R. Koduru, B.A. Posner, L. F. Wu, S.J. Altschuler, Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells, Nat. Commun. 7 (2016), 10690.