bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Copy number footprints of platinum-based anticancer therapies

Santiago Gonzalez[1], Nuria Lopez-Bigas[1,2,3,^], Abel Gonzalez-Perez[1,2,^]

Affiliations:

1. Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac, 10, 08028 Barcelona, Spain.
2. Centro de Investigación Biomédica en Red en Cáncer (CIBERONC), Instituto de Salud Carlos III, Madrid, Spain.
3. Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

^Corresponding authors
Abel Gonzalez-Perez: abel.gonzalez@irbbarcelona.org
Nuria Lopez-Bigas: nuria.lopez@irbbarcelona.org

Abstract

Recently, distinct mutational footprints observed in metastatic tumors, secondary malignancies and normal human tissues have been demonstrated to be caused by the exposure to several chemotherapeutic drugs. These characteristic mutations originate from specific lesions caused by these chemicals to the DNA of exposed cells. However, it is unknown whether the exposure to these chemotherapies leads to a specific footprint of larger chromosomal aberrations. Here, we address this question exploiting whole genome sequencing data of metastatic tumors obtained from patients exposed to different chemotherapeutic drugs. As a result, we discovered a specific copy number footprint across tumors from patients previously exposed to platinum-based therapies. This footprint is characterized by a significant increase in the number of chromosomal fragments of copy number 1-4 and size smaller than 10 Mb in exposed tumors with respect to their unexposed counterparts (median 14-387% greater across tumor types). The number of chromosomal fragments characteristic of the platinum-associated CN footprint increases significantly with the activity of the well known platinum-related footprint of single nucleotide variants across exposed tumors.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Chemotherapies and radiotherapy (in conjunction with surgery) constitute the first line of cancer treatment for many tumor types. In the past half century they have contributed to an important reduction of the mortality related to these malignancies (Siegel *et al*, 2022). These treatments are designed to directly damage the DNA, or to interfere with basic processes (such as DNA replication or chromosome separation during mitosis) of cancer cells (Rottenberg *et al*, 2021; Hennenfent & Govindan, 2006; Pommier *et al*, 2010; Longley *et al*, 2003). Exposure to some of them, thus contributes to shaping the genome of cancer cells throughout their evolution. For example, some drugs that directly damage the DNA may, in consequence, leave unique mutational footprints, composed of specific patterns of single base substitutions (SBS) or double base substitutions (DBS) along the genome of exposed tumors (Pich *et al*, 2019; Christensen *et al*, 2019; Landau *et al*, 2020; Maura *et al*, 2021; Pleasance *et al*, 2020; Pich *et al*, 2021).

The type of DNA damage introduced by some of these anti-cancer therapies may also lead to structural variants in the tumor genome. For example, the exposure to radiation therapy is known to cause the introduction of indels and certain types of translocations (Behjati *et al*, 2016; Kocakavuk *et al*, 2021). However, the footprint of large (starting at the kilobase range) structural variants (SVs) of anti-cancer therapies has not been systematically studied to date. Unlike single base substitutions, hundreds or thousands of which may be contributed to a tumor genome by a mutational process, and the overwhelming majority of which bear no significant consequence to the cell, fewer large SVs occur in a cancer cell. Their lower prevalence is constrained both by their sheer size and by their functional impact and it poses an important hurdle to study the footprint of large SVs that may be associated with anticancer treatments.

Throughout the evolution of tumors, the genome of their cells gain and lose chromosomal fragments of different length. As a consequence, the tumor genome appears like a collage of fragments of the chromosomes of its ancestral normal cell with different copy number and zygosity. One of the most dramatic events affecting the structure of a cell's genetic material is the duplication of its entire genome. This event of whole-genome doubling (WGD) constitutes a hallmark of cancer genome evolution (Bielski *et al*, 2018; Gerstung *et al*, 2020; Priestley *et al*, 2019). By doubling the size of the genome, the WGD is thought to relax the selective pressure on large SVs, in particular deletions, effectively increasing genome instability and plasticity in the evolution of cancer cells (López *et al*, 2020; Prasad *et al*, 2021; Dewhurst *et al*, 2014).

Therefore, we reasoned that metastatic WGD tumors could be exploited to overcome the difficulties outlined above to study the influence of anticancer therapies on the landscape of SVs of tumor cells. We carried out a systematic analysis of 2709 WGD metastatic tumors from patients who, as part of the treatment of their primary malignancies, had been exposed to 85 anticancer therapies, sequenced by the Hartwig Medical Foundation (HMF cohort) (Priestley *et al*, 2019). We identified a specific footprint of large loss of heterozygosity (LoH) events associated with the exposure of tumors to platinum-based therapies. A lower prevalence of the same footprints was observed in non-WGD metastatic tumors exposed to the same drugs, prompting us to speculate that WGD may provide tumors with an advantage to withstand their effect.

Results

High prevalence of WGD across metastatic tumors

Our starting hypothesis is that WGD tumors exposed to certain chemotherapies accumulate large SVs with specific features. Therefore, we first set out to identify tumors with WGD across

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

two cohorts of cancer samples, one comprising 2778 primary tumors (obtained from the PCAWG consortium; henceforth, PCAWG cohort) (Campbell *et al*, 2020) and the other 4902 metastatic tumors (obtained from the HMF) (Priestley *et al*, 2019).

The tumors in these cohorts clearly separate into two clusters when represented in a plane defined by their ploidy and their level of homozygosity (Fig. 1a and Fig. S1a). These two clusters, which may be separated by a straight line, (as represented in the Figure) group tumors with (orange) or without (gray) WGD. It is known that WGD is more prevalent across metastatic tumors than their primary counterparts (Priestley *et al*, 2019). We corroborated this across different cancer types (Fig. 1b; Table S1).

Moreover, while the interquartile range (that is, between quartiles 1 and 3) of the ploidy of tumors with WGD corresponds to a range between 2.9 and 3.6 (Fig. 1a), there are no significant differences between primary and metastatic WGD tumors in the distribution of their ploidy (Fig. 1c).

However, as mentioned in the introduction, by relaxing the selective constraints on large SVs, the occurrence of WGD may provide the tumor with an indirect advantage: a more plastic genome. Large chromosomal segments are lost by tumors that experience a WGD, resulting in an average ploidy of 3.2 for both primary and metastatic malignancies (Fig. 1a,c). However, metastatic tumors bearing WGD exhibit significantly larger fractions of their genome with LoH than their WGD primary counterparts (p-value $3x10^{-16,}$ Fig. 1c; Fig. S1b). This could be explained as more (or larger) chromosomal segments being lost by WGD metastatic tumors than by WGD primary tumors. We thus hypothesized that the exposure to certain anticancer treatments may leave a footprint of chromosomal fragments with different copy numbers in WGD exposed tumors.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 1

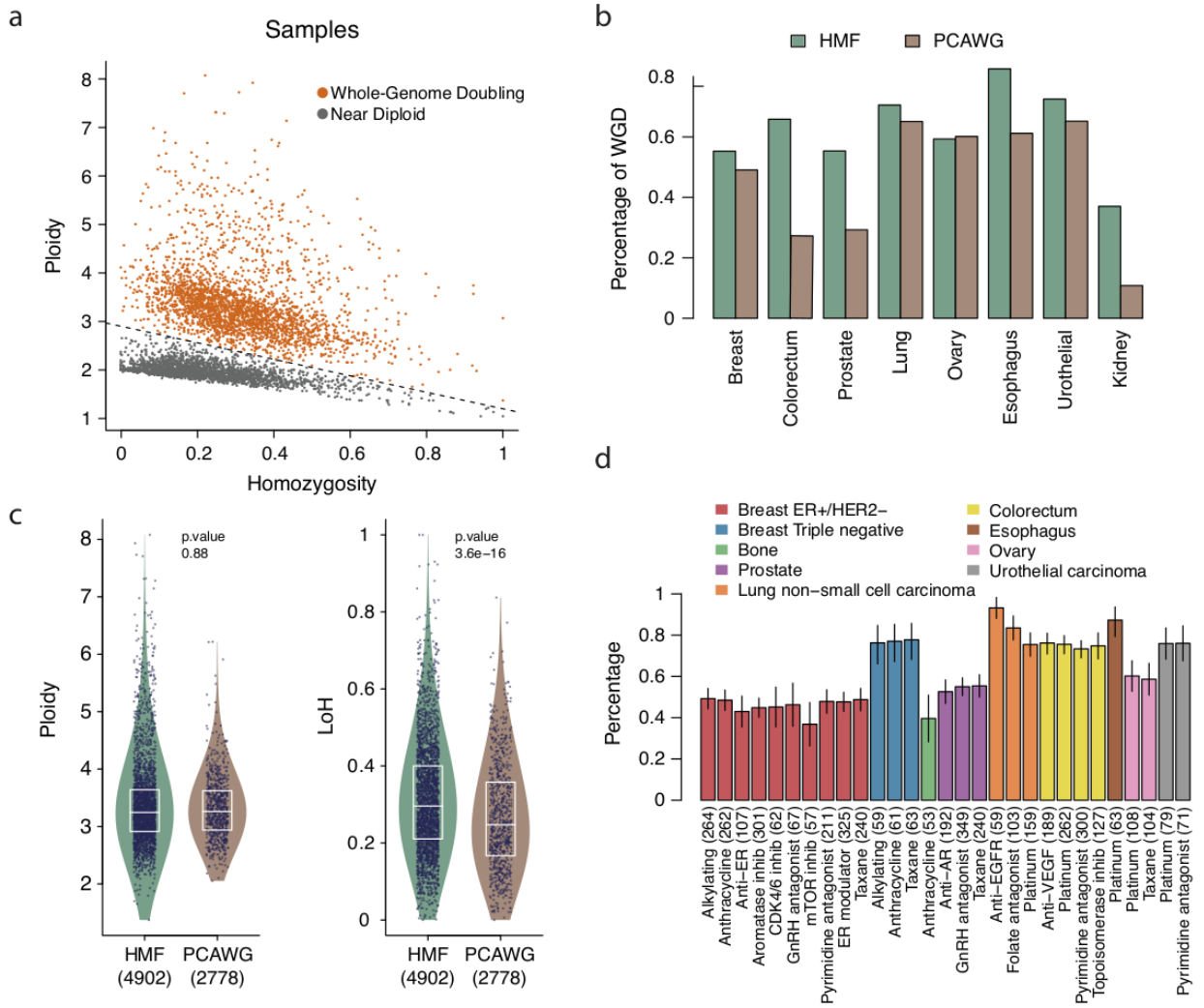

Figure 1. Features of WGD tumors exposed and unexposed to anticancer therapies

a) Tumors across one primary (PCAWG) and one metastatic (HMF) cohort may be classified as WGD and non WGD (or nearly diploid) based on their ploidy and level of heterozygosity.

b) Different fractions of WGD tumors are observed across different types of cancer in primary and metastatic cohorts.

c) Distribution of ploidy (left) and fraction of the genome with LoOH (right) of WGD tumors across primary and metastatic cohorts. The boxes inside the violin plots delimit the first, second and third quartiles of the distribution. All tumors in each group are represented as dots. P-values were derived from a two-tailed Wilcoxon-Mann-Whitney test.

d) Fraction of WGD metastatic tumors among those exposed to different therapies in the HMF cohort. The number of tumors exposed to each treatment appear in parentheses. The bars represent the mean fraction for tumors in each group (group size below each bar) and the short vertical line, the standard deviation.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Profiles of chromosomal fragments associated with anticancer therapies

To test this hypothesis, we leveraged whole-genome sequencing data obtained by the Hartwig Medical Foundation (Priestley *et al*, 2019) on 1093 WGD metastatic tumors originating from 10 primary sites which had been exposed to 15 families of anticancer therapies (Fig. 2a; Table S2; methods). The most widely used chemotherapy families were platinum-based therapies (498 exposed tumors), pyrimidine antagonists (445 exposed tumors) and taxanes (432 exposed tumors). No apparent differences are observed in the prevalence of WGD across tumors of the same origin exposed to each type of therapy (Fig. 1d)

We classified the chromosomal segments in each cancer genome according to their copy number, zygosity and length into 48 groups (Fig. 2b) that have been used previously to describe the large-scale organization of tumor genomes with the aim of extracting signatures of copy number variants (Steele *et al*, 2022). Specifically, these features reflect the number of chromosomal segments (ranging in size from less than 1 kilobase to several megabases) with different copy number (CN), between 0 (that is, homozygously deleted) to 9 or more, and their zygosity. For example, a comparison of the two mock examples represented in Figure 2b shows an increase of heterozygous chromosomal segments (across the entire length range) with 3-4 copies in the bottom case, with a complementary decrease of segments with 2 copies and LoOH. For simplicity, henceforth we refer to these features as CN categories. If a given pattern of increased and decreased CN categories is significantly associated with the exposure to a treatment family, we can call it the CN footprint of the treatment in the genome, by analogy with previously identified mutational (SBS and DBS) footprints of anticancer treatments (Pich *et al*, 2019).We computed the prevalence of the 48 CN categories across WGD tumors of every cancer type that had been exposed to different treatments. (Tumors of each cancer type were analyzed separately.) Then, we used a Wilcoxon signed-rank test to assess which features significantly differed between tumors (of a given site of origin) exposed and unexposed to a given treatment. We computed one test per CN category, site of origin of the tumor and treatment family, which amounted to 1782 combinations with at least 20 samples in each group (exposed and unexposed). We also computed a fold change between the median prevalence of each category across WGD tumors exposed to the treatments of the family and across those unexposed to them.

We identified 10 CN categories that change significantly (q-value < 0.05) between WGD tumors exposed or unexposed to four types of anticancer therapies (Fig. 2c; Table S3). These are platinum-based therapies (across tumors originating in the colorectum or the lung), taxanes (across breast and prostate tumors), anti-EGFR drugs (across lung tumors) and pyrimidine analogs (across colorectal tumors). These associations may be driven by co-variables other than the treatment in question, such as. a second treatment (i.e., two treatments given in combination). Thus, to restrict these results to CN categories change more likely associated to the treatment, we required that: i) the change of CN category continues to be observed when tumors exposed to the second treatment are removed from the comparison; ii) the change of CN cannot be explained by another clinical variable (e.g., sex); and iii) the change is consistently observed across tumor types.

CN categories associated with pyrimidine analogs failed to fulfill the first condition. Specifically, a high fraction of all metastatic colorectal tumors exposed to pyrimidine analogs (N=194, or 88%) were also exposed to platinum-based therapies. When the comparison is restricted to those exposed only to pyrimidine analogs (N=26) and those unexposed to pyrimidine analogs or platinum (N=120), the significance is lost (p-value=0.23 vs p-value=2.6×10^{-4} in the comparison for the most significant CN category). Across lung tumors, all 22 WGD samples exposed to pyrimidine analogs are also exposed to platinum-based therapies (N=120). Overlaps between

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

all pairs of treatments across tumor types are presented in Table S4. With respect to anti-EGFR therapies across lung tumors, the signal is lost after the overlap with platinum-based drugs (16, 29%) and a gender effect (see Methods) are removed. In detail, the p-value (before correction) of the most significant CN category changes from $6x10^{-5}$ to 0.034). We thus discarded the associations with anti-EGFR drugs or pyrimidine analogs. In the case of taxanes, CN categories that appear significantly different across exposed or unexposed tumors, were not replicated across tumor types, i.e., different CN categories appear across breast tumors and prostate tumors (Fig. 2e). Expanding the analysis to other tumor types treated with taxanes did not yield any common pattern (Fig. S2). It is possible that diverse mechanisms underlie these observed differences across tumor types. It is also plausible that more numerous cohorts of exposed tumors are required to understand whether the observed differences are indeed robust. We also decided to discard the case of taxanes for further analysis within this work.

Unlike the previous cases, the associations between CN categories and the exposure to platinum-based drugs fulfill the three conditions set to consider them significant. Specifically, 11 CN categories –representing diploid LOH fragments below 100 Kb, heterozygous fragments smaller than 100 Kb and larger than 10 Mb, heterozygous fragments with copy number 3-4 smaller than 40 Mb, heterozygous fragments with copy number 5-8 larger than 10 Mb and heterozygous fragments with higher copy number smaller than 1 Mb– appeared significantly increased across 98 lung tumors solely exposed to platinum-based drugs (uncorrected p-value; Table S4). Moreover, 65 WGD ovarian tumors, which were exposed only to carboplatin (Table S4) exhibit differences (with respect to their unexposed counterparts) in the same CN categories, although these do not reach significance (Fig. S6; Table S2).

In summary, CN categories significantly associated with the exposure to platinum-based therapies are detected even in the absence of another treatment, and we are able to observe them across cancer types (see below and Fig. S6). In the next section, we describe and discuss in detail this CN footprint of platinum-based therapies.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 2

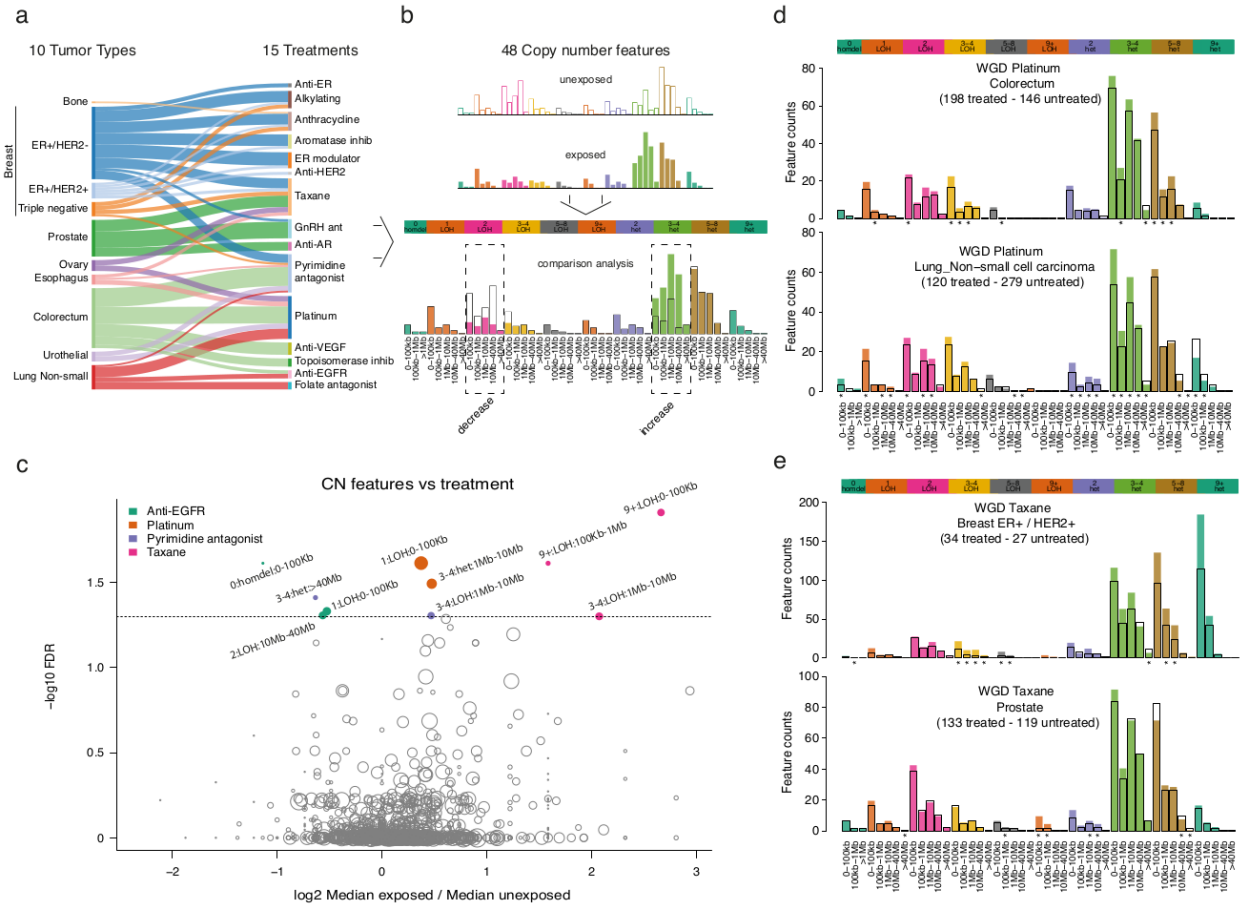

Figure 2. The profile of chromosomal fragments associated with anticancer therapies

a) Cancer types (sites of origin) represented in the HMF metastatic cohort and anticancer treatments to which they were exposed. Numbers of tumors of different cancer types in the HMF cohort exposed to anticancer therapies appear in Table S2.

b) Toy representation of CN categories computed for every tumor and their values for exposed and unexposed tumors (top). Bars represent the count of chromosomal fragments in each CN category. Comparison across CN categories reveals several that appear significantly increased or decreased across exposed tumors with respect to their unexposed counterparts. The CN and ploidy of each category are always shown above the graph, while the range of sizes of chromosomal fragments represented by each bar appear below. For each comparison (one per CN category), a p-value and fold-change can be computed.

c) Results (minus logarithm corrected p-value of the Wilcoxon signed-rank test vs fold-change) of 1782 comparisons carried out between CN categories of tumors within a cancer type that were exposed or unexposed to a given anticancer therapy. Significant comparisons (False Discovery Rate corrected p-value below 0.05)

d) Results of the comparison of CN categories across WGD colorectal (top) and lung (bottom) tumors exposed or unexposed to platinum-based drugs in the HMF cohort. Significant comparisons (uncorrected p-value below 0.05) are highlighted with asterisks below each graph.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

e) Results of the comparison of CN categories across WGD breast ER+/HER2+ (top) and prostate (bottom) tumors exposed or unexposed to taxanes in the HMF cohort. Significant comparisons (uncorrected p-value below 0.05) are highlighted with asterisks below each graph.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The CN footprint of platinum-based therapies

To study the CN footprint of platinum-based therapies we first analyzed all CN categories with a p-value < 0.05 (before FDR correction) across colorectal and lung tumors (Fig. 2d). These CN categories comprise fragments shorter than 40 Mbases across the spectrum of copy number, but mostly heterozygous fragments with CN between 3 and 8 and LoH fragments with CN 3-4 that appear significantly increased across exposed metastatic WGD tumors. Conversely, heterozygous chromosomal fragments longer than 40 Mbases with 3-4 copies appear significantly decreased across exposed tumors originating in either organ (in lung tumors, also those with 5-8 copies).

To better capture this CN footprint of platinum-based therapies, we summed the number of fragments in each CN category in platinum-exposed and unexposed WGD tumors across cancer types, and computed the CN category-wise difference. The CN footprint of platinum-based therapies thus appears as an increase in the number of heterozygous and LoH chromosomal fragments of copy number 1-4 and size smaller than 10 Mb across platinum-exposed WGD tumors (Fig. 3a). This footprint is better captured by this array of CN features than by CN signatures identified de novo using a state-of-the-art extraction method or by previously extracted CN signatures across primary tumors (Fig. S3a-c). This is probably due to intrinsic hurdles in defining signatures of CN, in contrast with commonly defined SBS (or DBS) signatures (Steele *et al*, 2022).

Platinum-based drugs (cisplatin, carboplatin and oxaliplatin) are bifunctional agents that create adducts in the DNA by establishing covalent links with one or two closeby bases. Adducts connecting two bases may be intra or interstrand (Siddik, 2003). Although the latter account for less than 15% of all adducts, they may impose a physical obstacle to the separation of the two strands that is required for DNA replication and transcription (Vare *et al*, 2012; Deans & West, 2011). While some of these interstrand bridges may be repaired by a combination of translesion synthesis and nucleotide excision repair, others result in double strand breaks created by the Fanconi Anemia (FANC) complex (Deans & West, 2011). The resulting double strand breaks may then be solved via the homologous recombination (HR), the non-homologous end-joining (NHEJ) repair or other DNA repair pathways (Vare *et al*, 2012; Deans & West, 2011; Chang *et al*, 2017). The activity of both, HR and NHEJ in the repair of double strand breaks can result in the deletion of chromosomal fragments of varying sizes. In particular, the use of non-sister chromatids as template for HR repair may lead to loss of heterozygosity (Westphalen *et al*, 2022; Vare *et al*, 2012; Chang *et al*, 2017; Deans & West, 2011). Moreover, other mechanisms of repair of double strand breaks which are more error-prone, such as single strand annealing or alternative end joining may also result in the introduction of short insertions and deletions around microhomology elements (Deans & West, 2011).

Since intrastrand links generated by platinum-based drugs are expected to distribute randomly along the genome of WGD tumors during exposure, we expected that chromosomal fragments resulting from their repair would also distribute evenly across the genome. To test this hypothesis, we computed the difference in the number of chromosomal fragments of copy number 1-4 and size smaller than 10Mb between WGD tumors exposed and unexposed to platinum-based drugs along each chromosome. As anticipated, we observed that chromosomal fragments smaller than 10 Mb (representative of the platinum CN signature) are evenly distributed along the genomes of WGD colorectal and lung tumors (Fig. 3b; Fig. S4a,b). Had this increase in the number of platinum-related chromosomal fragments across exposed tumors been constrained to one or few genomic regions, it would point to positive selection of one or more resistance-associated genes. A recent systematic analysis of the HMF cohort revealed that only mutations in TP53 across stomach adenocarcinomas appear as a potential bona fide

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

driver event associated with the exposure to platinum in the HMF cohort (Martínez-Jiménez *et al*, 2022).

Next, we asked if an imbalance between insertions and deletions occurred during the error-prone repair of these interstrand links. To answer this question, we looked at the overall ploidy of platinum exposed and unexposed WGD tumors, and observed no significant differences between these two groups of cases, with the notable exception of lung tumors (Fig. S5). This supports the notion that across colorectal, esophageal, urothelial and ovarian tumors no large imbalances between insertions and deletions occur in the process of repair of platinum interstrand links. While purifying selection on likely damaging chromosomal deletions could in principle limit their observed number in diploid genomes, this effect is likely removed from this analysis through the study of WGD tumors. Further analyses with larger cohorts are required to clarify this point, which appears especially to clarify whether a significant imbalance in favor of deleted chromosomal fragments does occur across platinum-exposed lung tumors.

We then hypothesized that the same CN footprint will be apparent across WGD tumors from other primary sites, even if they don't reach significance due to fewer numbers of exposed samples. Indeed, the same trends of increase of LoH and heterozygous chromosomal fragments smaller than 10 Mbase observed across WGD tumors of ovarian (Fig. 3c), esophageal or urothelial (Fig. S6) origin.

While the occurrence of the WGD results in a decrease of the selective pressure on chromosome fragmentation induced by platinum-based drugs –via repair of interstrand bridges–, we expect that such fragmentation also occurs across tumors without WGD. We thus focused on non-WGD tumors exposed to platinum-based therapies. As with WGD tumors, a trend towards the increase of chromosomal fragments below 10 Mb is observed across non-WGD metastatic lung tumors (Fig. 3d). Interestingly, this CN footprint is shifted towards smaller CN values (i.e., showing the highest increase for LoH fragments with CN 1; (Fig. 3d). Less starting DNA in non-WGD tumors, and the stronger purifying selection resulting from the diploid rather than tetraploid genome are likely the reasons behind this shift towards fragments of smaller CN. However, the overall higher rate of chromosomal fragments across platinum-exposed tumors appears significant across both WGD and non-WGD tumors (Fig. 4a). Specifically, the median number of chromosomal fragments of size smaller than 10 Mb across exposed tumors is between 14% (colorectum WGD tumors) and 387% (ovarian non-WGD tumors) greater than the median of their counterpart unexposed samples.

To validate the platinum CN footprint we tapped onto an independent smaller cohort of advanced tumors (POG507) (Pleasance *et al*, 2020). The POG507 cohort has important differences with respect to the HMF cohort, the most important of which are the smaller sample size and the shorter time elapsed between the end of the treatment and the biopsy of the metastasis (Fig. S7a), which is known to decrease the detection of the platinum related mutations through bulk sequencing (Pich *et al*, 2021). Nevertheless, a significant increase in the number of chromosomal fragments below 10 Mb was indeed apparent across 144 (31 platinum exposed) breast and 42 (30 exposed) lung tumors (Fig. S7b). The increase of these short chromosomal fragments, however, was not significant across 65 colorectal tumors (40 exposed).

We also reasoned that the more intense the exposure to platinum-based therapies, the greater the observed CN footprint will be. We previously demonstrated that higher activity of the platinum-related mutational (SBS and DBS) footprint results from longer times of exposure to these drugs (Pich *et al*, 2019). We thus reasoned that the activity of this mutational footprint can

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

be used as a proxy of the intensity of the exposure to platinum-based therapies. We found that the rate of chromosomal fragments smaller than 10Mb increases significantly with the increase of the activity of the platinum mutational footprint across metastatic colorectal and lung tumors (Fig. 4b).

In summary, we postulate that a platinum CN footprint characterized by increased number of short (below tens of megabases) LoH and heterozygous chromosomal fragments may result from the increase in the number of double strand breaks stemming from platinum-induced interstrand bridges. While it is possible that the presence of this CN footprint increases the aggressiveness of platinum-exposed tumors, the mortality curves are not significantly distinct, Fig. S8). Therefore, further analyses with larger cohorts will be needed to clarify this.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 3

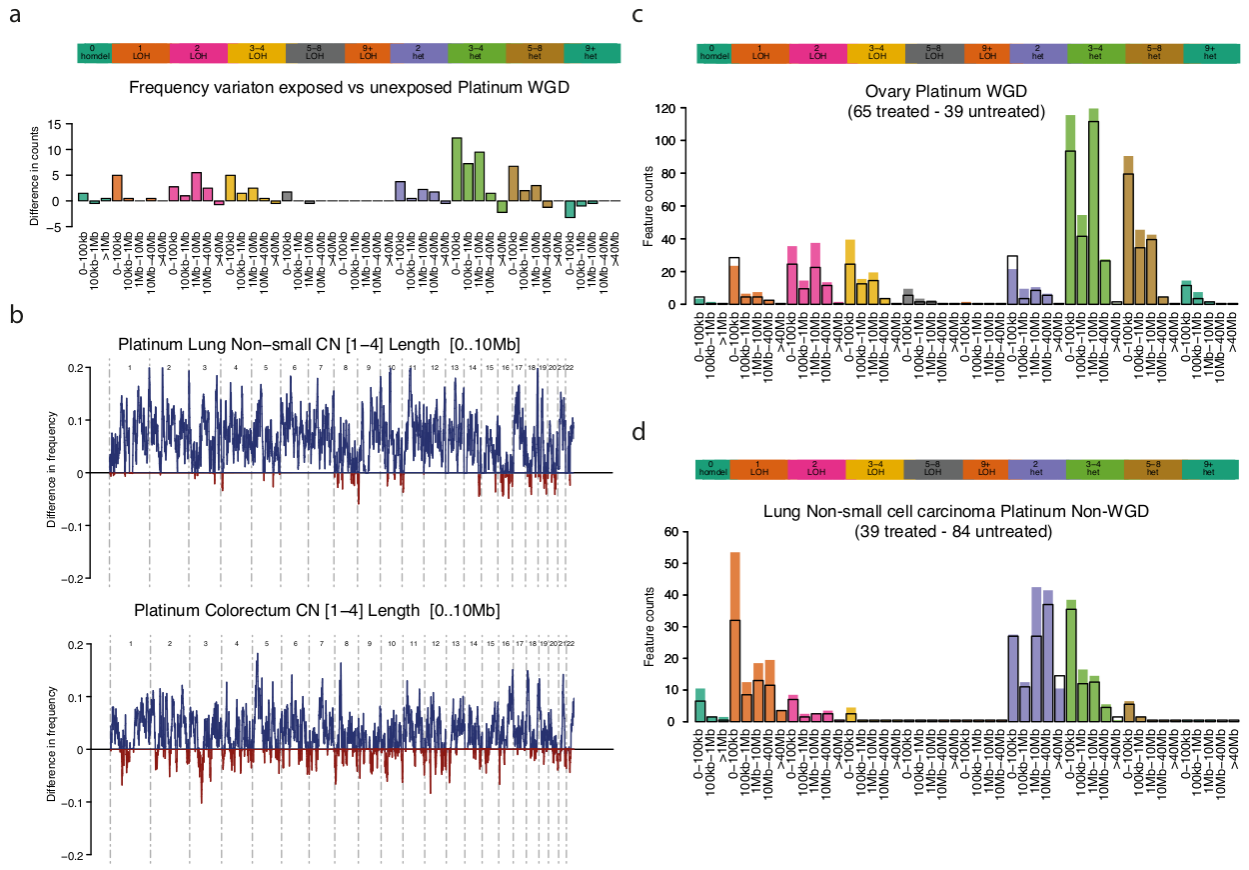

Figure 3. CN footprint of platinum-based therapies

a) Representation of the CN footprint of platinum based therapies. The bars represent the average (across WGD lung and colorectal tumors) number of chromosomal fragments in each CN category.

b) Frequency of occurrence of chromosomal fragments with copy number 1-4 and length smaller than between 10 Kb and 10 Mb across WGD lung (left) and colorectal (right) tumors in the HMF cohort of tumors exposed (blue, positive) or unexposed (red, negative) to platinum-based drugs.

c) Results of the comparison of CN categories across WGD ovarian tumors exposed or unexposed to platinum-based drugs in the HMF cohort.

d) Results of the comparison of CN categories across non-WGD lung tumors exposed or unexposed to platinum-based drugs in the HMF cohort.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4

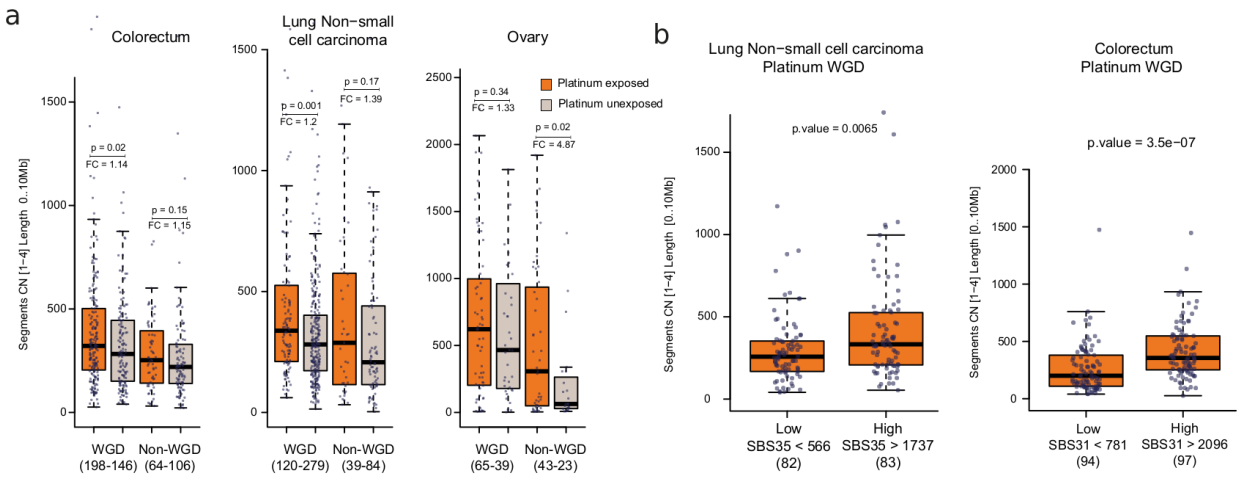

Figure 4. Relationship between the intensity of the treatment its CN footprint

a) Distribution of the number of chromosomal fragments with copy number 1-4 and length smaller than 10 Mbwith length between 10 Kb and 10 Mb across WGD and non-WGD colorectal (left), lung (center) or ovarian (right) metastatic tumors exposed or unexposed to platinum-based drugs. The p-values were derived from a two-tailed Wilcoxon-Mann-Whitney test. Fold-change values (median number of CN fragments of size smaller than 10 Mb across exposed tumors vs its median across unexposed tumors) are shown. Boxes in the boxplot represent the limits of the first and third quartile of the distribution and the whiskers extend between its minimum and maximum, excluding the outliers. Tumors in each group are represented as dots.

b) Distribution of the number of chromosomal fragments with copy number 1-4 and length smaller than 10 Mbwith length between 10 Kb and 10 Mb across WGD lung (left) and colorectal (right) metastatic tumors in the lower or upper tertile of activity of the platinum SBS mutational signature. Boxes in the boxplot represent the limits of the first and third quartile of the distribution and the whiskers extend between its minimum and maximum, excluding the outliers. Tumors in each group are represented as dots. The p-values were derived from a two-tailed Wilcoxon-Mann-Whitney test.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Discussion

Identifying and measuring the genetic footprints of chemotherapies in exposed cells makes an important contribution to the study of their effects on both tumor and normal cells. We recently tapped on our ability to detect the mutational footprint of several chemotherapies to estimate their mutational toxicity on the basis of regularly used treatment schemes(Pich *et al*, 2019, 2021). We also studied this toxicity in vivo in cells that were normal at the time of exposure to the chemotherapies. Specifically, we demonstrated that in the emergence of tumors that are secondary to the exposure to chemotherapy this exposure precedes the start of the clonal expansion that gives rise to the secondary malignancy (Pich *et al*, 2021). These studies were made possible by our previous identification of the mutational footprints of these anticancer treatments.

Here, we identify the CN footprint of platinum-based drugs in metastatic tumors, as a result of the exposure to these drugs. The CN profile of platinum-based drugs appears as an increase in the fragmentation of the tumor genome, likely due to the gain of double strand breaks resulting from interstrand links. A clear connection between the ability to repair double strand breaks via homologous recombination repair and the effectiveness of platinum-based therapies has been shown by recent studies that segment patients on the basis of the observed activity of inferred chromosomal instability signatures (Drews *et al*, 2022; Macintyre *et al*, 2018). Furthermore, we found that there is a direct relationship between the intensity of the platinum-based treatment and the degree of the footprint signal.

Though these footprints are apparent across tumors originating in different organs, and across both WGD and non-WGD tumors, their signal is still weaker than that of SBS mutational footprints (Pich *et al*, 2019; Christensen *et al*, 2019; Alexandrov *et al*, 2013, 2020). The smaller number of CN fragments –orders of magnitude below that of SBS– available to detect the signal is probably the main reason behind this. It is interesting to point out that platinum-based and taxane therapies (which come out significant from the CN categories analysis) are two of the most abundant across the cohort of patients studied. Larger cohorts will thus be needed to corroborate these CN footprints and to potentially detect others associated to different anticancer treatments which may still be below the statistical power afforded by the HMF cohort. The relatively small number of tumors of different origins exposed to platinum-based drugs has also prevented us from exploring in more depth the influence of these treatment families on the aggressiveness of the resulting metastatic tumors, although at least platinum-based drugs appear to increase them. If this trend is confirmed in larger cohorts, this could be important to decide further rounds of treatment on relapse patients.

Since chemotherapies are systemic treatments, their effects –not only the immediate ones, but also those occurring at longer terms– may also affect normal cells of the exposed person. We recently showed that the platinum-related mutational signature is detectable in hematopoietic cells that were normal at the time of exposure to the treatment, and which subsequently suffered a full clonal expansion to produce a full blown acute myeloid leukemia (Pich *et al*, 2021). It seems reasonable to assume that as a result of the increased number of double strand breaks resulting from the exposure to these drugs, normal cells may also suffer an increased number of SVs. Moreover, these SVs –as described by the platinum CN footprint– are bound to affect much larger genomic portions than platinum-contributed point mutations (Pich *et al*, 2019, 2021). Therefore, their impact on exposed healthy cells and on the development of late effects of the chemotherapy could potentially be greater than those caused by previously recognized platinum-related SBS footprints. Identifying these SVs –a task that may be supported by the CN

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

footprint described here– and understanding if they have a role in the emergence of these secondary malignancies is thus an outstanding goal.

Methods

Collection of genomic and clinical data

Somatic copy number (CN) per chromosomal fragment (CN fragments) calculated from whole genome sequencing (and the clinical data for each patient) were obtained from 2 different studies. We obtained primary tumors from the Pancancer Analysis of Whole-Genomes (PCAWG(Campbell *et al*, 2020)) upon request from dbGAP (ph). A cohort of metastatic tumors was obtained upon request from the Harwtig Medical Foundation (HMF(Priestley *et al*, 2019)); https://www.hartwigmedicalfoundation.nl. The PCAWG cohort comprises 2778 whole-genome samples from 47 different primary tumor types. Similarly, HMF comprises 4902 whole-genomes of metastatic tumors across 40 primary anatomical sites. We also obtained single base substitutions (SBS) across HMF tumors.

For each of the samples, genome ploidy represents the mean of the total ploidy of each of the called CN fragments weighted by their genomic length. The fraction of the genome under a loss of heterozygosity (LoH) in PCAWG is computed as the sum of the length of all CN fragments with a minor copy number allele equal to zero divided by the fraction of the genome with discernible CN (i.e., with a value provided by the calling algorithm). In the case of HMF, the LoH estimation is equivalent, but since the CN is not provided as an absolute number, but rather as a continuous value, LoH segments are those with a minor CN smaller than 0.5.

The clinical data obtained for tumors in the HMF cohort included the primary cancer site and the information regarding the pre-treatment (treatment in this manuscript) received by each patient as part of the management of their primary malignancy. This treatment data (see Table S2) was used as the basis for the identification of CN categories (see below) associated with the exposure to anticancer therapies.

Whole-genome doubling status and treatments comparison

The whole-genome doubling status (WGD) of each sample is determined based on the average ploidy of the genome and its fraction of LoH (Gerstung *et al*, 2020). In both cohorts, primary and metastatic, the samples cluster into 2 distinct groups, which allows a clear differentiation between mostly diploid samples and those carrying a WGD (Fig. 1a and Fig. S1). To perform the classification we have employed the formula:

$$\text{WGD: } 2.9 - 1.7 * \text{LoH} <= \text{Ploidy}.$$

As a representation of a variety of tumors and treatments, we selected all the combinations of tumor types and treatment comprising at least 50 samples (Table S2) in order to investigate the impact of the treatment on the fraction of WGD tumors in a cancer type or exposed to a given treatment. For each pair tumor-treatment, the percentage of samples carrying a WGD was calculated and the confidence intervals were estimated using the 5 and 95 percentiles of the binomial distribution estimated based on their prevalence and cohort size.

Selection and evaluation of CN features

This study uses a total of 48 CN categories that classify each chromosomal CN segment according to its absolute copy number (0, 1, 2, 3-4, 5-8, or 9+), zygosity (homologous deletions, heterozygous or loss of heterozygosity), and length (0-100Kb, 100Kb-1Mb, 1Mb-10Mb, 10-Mb-40Mb, or 40Mb+) following the classification implemented by Christopher D. Steele and

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

colleagues (Steele *et al*, 2022) (Fig. 2b). The CN categories present no overlap, and all chromosomal segments in a genome are classified into one of them.

We filtered out samples with no treatment record and those tumor type-treatment combinations with less than 20 WGD cases. As a result, we obtained a total of 15 treatments and 10 tumor types based on their primary location. The analysis results in 39 combinations of tumor type-treatment to interrogate the impact of anticancer therapies on the selected CN category. For each CN category across each tumor type-treatment group, an independent Wilcoxon's signed-rank test was used to estimate a p-value of the difference of the prevalence of the CN category between exposed and unexposed WGD tumors (Table S3).

In order to avoid the overlap present across many tumor-treatment pairs, only one treatment is selected for the multi-test correction from those with an overlap of more than 90% of the samples in either treated or untreated groups. All obtained values were corrected by the effect of the multiple tests using the Benjamini & Hochberg false discovery rate.

We performed a manual curation of significant results to determine the appropriate treatment for those considered as overlapping (see previous paragraph). For colorectal cancer, pyrimidine analogs and platinum cohorts present an overlap of 88 and 97% of cases respectively, and show significant associations between some CN categories and the treatments (Table S4). We selected platinum as the main contributor due to the consistency in treatment footprint across other tumor types, including some with less or none exposure to pyrimidine analogs (Fig 2, Fig3 and Fig S5). Similarly, anti-EGFR drugs lose their significance once the overal with platinum and a sex bias effect are accounted for.

Copy Number signature extraction

With the number of the 48 CN categories in each tumor (i.e., their CN feature profile), we first generated a matrix of counts using the SigProfilerMatrixGenerator version 1.2 (Bergstrom et al., 2019). Then, CN signatures were extracted using the sigProfilerExtractor version 1.1 (Islam et al., 2020) and the default set of parameters. The solution recommended by the tool based on signature stability and the average error of reconstruction of the CN profiles comprised 9 CN signatures. To determine if any of these CN signatures was associated with the exposure to any treatment, we carried out a Wilcoxon signed-rank test of the activity of each signature (also provided by the sigProfilerExtractor as part of the CN signatures extraction de novo) across exposed and unexposed WGD tumors, for each tumor type separately (Fig. S3a).

Then, we analyzed (part of the sigProfiler package) the equivalence between these 9 CN signatures and the CN signatures extracted across primary tumors in a recent study (Steele *et al*, 2022). Briefly, the tool explains the CN profile of each de novo signature as a linear combination of the profiles of one or several signatures in the reference set. In this case, the sigProfiler decomposes the 7 CN signatures extracted into linear combinations of 10 CN reference signatures, although the cosine similarity of the reconstructed and original de novo signatures is low in general (Fig. S3b). In any case, no significant association (using the approach described above) exists between the activity of any of these CN reference signatures and with the exposure to any anticancer therapy (Fig. S3c).

Platinum CN footprint and distribution along the genome

To integrate all CN categories with different prevalence between platinum exposed and unexposed tumors into a platinum CN footprint, we first combined the prevalence (count) of chromosomal fragments in each CN category across all platinum exposed and unexposed WGD tumors, irrespective of their tumor type. Then, we summed over each CN feature the prevalence

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

observed for exposed tumors, and separately that observed across unexposed tumors. Finally, we subtracted the vector representing the cumulative platinum CN profile of unexposed tumors from the vector representing the cumulative platinum CN profile of exposed tumors. Thus, in the resulting vector, positive components correspond to CN categories with higher prevalence across platinum exposed tumors. We selected these to form the platinum CN footprint (Fig. 3a). For practical purposes, to directly measure the intensity of this footprint in tumors, we directly count the number of heterozygous or LoH chromosomal fragments with CN between 1 and 4, and of length smaller than 10 Mb. This definition was used to study the distribution of the intensity of the platinum CN footprint along the genome (see below), to compare that intensity in treated/untreated and WGD/non-WGD tumors (Fig. 4a) and to study its correlation with the platinum SBS footprint (Fig. 4b).

To determine the intensity of the platinum CN footprint along the genome, we divided the genome into regions of 1Mb and calculated the intensity of the CN footprint per WGD tumor in each region Then, we subtracted intensity in each region for platinum non-exposed tumors from that of platinum exposed tumors. The resulting exposed-to-unexposed differences in intensity per genomic region are presented in Figure 3b.

Correlation between Platinum SBS signature and CN footprint

We extracted a new set of SBS signatures on its 3 base nucleotide context (SBS96) across HMF and PCAWG tumors. In order to increase the number of samples from the same tumor type and to perform a more accurate attribution of signatures, we aggregated metastatic and primary samples from the same cancer types. The matrix of counts of SBS in their trinucleotide contexts (SBS96 context) was constructed using SigprofilerMatrixGenerator version 1.1. (Bergstrom *et al*, 2019), and the extraction of signature profiles and their activities across tumors was carried out using the SigProfilerExtractor version 1.1 (Islam *et al*, 2022) for a minimum of 2 and a maximum of 20 signatures per tumor type.

The final signatures' extraction result corresponds to the one recommended by the SigProfilerExtractor. Across the suggested sets of signatures, the platinum signature was selected as the one with a cosine similarity higher than 0.9 with any of the previously described platinum signatures (SBS31 and SBS35) in COSMIC version 3.2.

The presence of the platinum SBS signature on exposed samples is limited to those tumors that underwent a clonal expansion fixing the mutations of the treatment (Pich *et al*, 2019; Landau *et al*, 2020; Pich *et al*, 2021). We selected exposed samples with platinum SBS signature activity.

Finally, we compared the number of mutations represented by that activity with the intensity of the platinum CN footprint (see above). The samples were divided into tertiles according to their amount of detected platinum SBS (quantiles 0.33, 0.66, 1) and the significant differences in the intensity of the platinum CN footprint across groups was estimated with a two-tailed Wilcoxon-Mann-Whitney test.

Platinum survival model

For each of the patients, the time to death represents the time between the date of the annotated biopsy and the date of death. Those patients without a date of death are considered alive and the last follow-up is fixed at the date when the current version of the data was published (October 21th, 2021). The survival analysis follows a Cox proportional hazards regression model for each tumor type of 4 different groups of samples: (I) Platinum-exposed with WGD, (II) Platinum-exposed without WGD, (III) Platinum-unexposed with WGD, and (IV) Platinum-unexposed without WGD.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Analysis of POG507 cohort

We obtained files containing the clinical and treatment data and CN calls of 507 metastatic tumors in the POG507 cohort (Pleasance *et al*, 2020) from a repository set up by the team that sequenced them (https://www.bcgsc.ca/downloads/POG570/). Patients who received platinum-based drugs (cisplatin, oxaliplatin and/or carboplatin) as part of the treatment of the primary tumor were identified. Only the colon, breast or lung cohorts satisfied the requirement of at least 10 platinum exposed and unexposed tumors, and were thus selected for analysis. The length of chromosomal fragments identified in each tumor were computed as the difference between their start and end genomic coordinates. Then, the chromosomal fragments of CN 1-4 and length below 10 Mb in each tumor (platinum CN footprint) were counted and the counts of exposed and unexposed tumors compared. To compute the number of days elapsed between the end of the platinum treatment and the biopsy across exposed tumors, the relative date of treatment end was subtracted from the relative biopsy date.

Data and code availability

This study uses data previously generated by three main projects: (I) Metastatic tumor cohort data (DR-110_update7) from the Hartwig Medical Foundation (HMF cohort) for academic research upon request (https://www.hartwigmedicalfoundation.nl/en). (II) Primary tumor cohort data from PCAWG is publicly available upon request (https://dcc.icgc.org/pcawg). (III) Chromosomal segments and clinical data of tumors in the POG507 cohort were obtained from the public domain at https://www.bcgsc.ca/downloads/POG570/. All the code necessary for running the analyses described in the manuscript will be available at https://bitbucket.org/bbglab/cn_platinum at the time of publication.

Acknowledgements

The authors wish to acknowledge the contribution of patients, families and biomedical researchers who shared, processed and sequenced the data used within the study. The results published here are in part based on the data generated by the Pan-Cancer Analysis of Whole Genomes. This publication and the underlying study have been made possible partly on the basis of the data that Hartwig Medical Foundation has made available to the study. We acknowledge the contribution of Claudia Arnedo-Pac of signatures extraction in the HMF cohort. N.L.-B. acknowledges funding from the European Research Council (consolidator grant 682398) and ERDF/Spanish Ministry of Science, Innovation and Universities – Spanish State Research Agency/DamReMap Project (RTI2018-094095-B-I00) and Asociación Española Contra el Cáncer (AECC) (GC16173697BIGA). S. G work is supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 754510. IRB Barcelona is a recipient of a Severo Ochoa Centre of Excellence Award from the Spanish Ministry of Economy and Competitiveness (MINECO; Government of Spain) and is supported by CERCA (Generalitat de Catalunya).

Author contributions

SG, NL-B and AG-P conceived and designed the study. SG carried out analyses of PCAWG and HMF cohorts and created main and supplementary figures. AG-P analyzed the POG570 cohort. SG, NL-B and AG-P discussed and interpreted the results, drafted and edited the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, *et al* (2020) The repertoire of mutational signatures in human cancer. *Nature* 578: 94–101

Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale A-L, *et al* (2013) Signatures of mutational processes in human cancer. *Nature* 500: 415–21

Behjati S, Gundem G, Wedge DC, Roberts ND, Tarpey PS, Cooke SL, Van Loo P, Alexandrov LB, Ramakrishna M, Davies H, *et al* (2016) Mutational signatures of ionizing radiation in second malignancies. *Nat Commun* 7: 12605

Bergstrom EN, Huang MN, Mahto U, Barnes M, Stratton MR, Rozen SG & Alexandrov LB (2019) SigProfilerMatrixGenerator: a tool for visualizing and exploring patterns of small mutational events. *BMC Genomics* 20: 685

Bielski CM, Zehir A, Penson AV, Donoghue MTA, Chatila W, Armenia J, Chang MT, Schram AM, Jonsson P, Bandlamudi C, *et al* (2018) Genome doubling shapes the evolution and prognosis of advanced cancers. *Nat Genet* 50: 1189–1195

Campbell PJ, Getz G, Korbel JO, Stuart JM, Jennings JL, Stein LD, Perry MD, Nahal-Bose HK, Ouellette BFF, Li CH, *et al* (2020) Pan-cancer analysis of whole genomes. *Nature* 578: 82–93

Chang HHY, Pannunzio NR, Adachi N & Lieber MR (2017) Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* 18: 495–506

Christensen S, Van der Roest B, Besselink N, Janssen R, Boymans S, Martens JWM, Yaspo M-L, Priestley P, Kuijk E, Cuppen E, *et al* (2019) 5-Fluorouracil treatment induces characteristic T>G mutations in human cancer. *Nat Commun* 10: 4571

Deans AJ & West SC (2011) DNA interstrand crosslink repair and cancer. *Nat Rev Cancer* 11: 467–480

Dewhurst SM, McGranahan N, Burrell RA, Rowan AJ, Grönroos E, Endesfelder D, Joshi T, Mouradov D, Gibbs P, Ward RL, *et al* (2014) Tolerance of Whole-Genome Doubling Propagates Chromosomal Instability and Accelerates Cancer Genome Evolution. *Cancer Discov* 4: 175–185

Drews RM, Hernando B, Tarabichi M, Haase K, Lesluyes T, Smith PS, Morrill Gavarró L, Couturier D-L, Liu L, Schneider M, *et al* (2022) A pan-cancer compendium of chromosomal instability. *Nature*: 1–8

Gerstung M, Jolly C, Leshchiner I, Dentro SC, Gonzalez S, Rosebrock D, Mitchell TJ, Rubanova Y, Anur P, Yu K, *et al* (2020) The evolutionary history of 2,658 cancers. *Nature* 578: 122–128

Hennenfent KL & Govindan R (2006) Novel formulations of taxanes: a review. Old wine in a new bottle? *Ann Oncol* 17: 735–749

Islam SMA, Díaz-Gay M, Wu Y, Barnes M, Vangara R, Bergstrom EN, He Y, Vella M, Wang J, Teague JW, *et al* (2022) Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor. 2020.12.13.422570 doi:10.1101/2020.12.13.422570 [PREPRINT]

Kocakavuk E, Anderson KJ, Varn FS, Johnson KC, Amin SB, Sulman EP, Lolkema MP, Barthel FP & Verhaak RGW (2021) Radiotherapy is associated with a deletion signature that contributes to poor outcomes in patients with cancer. *Nat Genet* 53: 1088–1096

Landau HJ, Yellapantula V, Diamond BT, Rustad EH, Maclachlan KH, Gundem G, Medina-Martinez J, Ossa JA, Levine MF, Zhou Y, *et al* (2020) Accelerated single cell seeding in relapsed multiple myeloma. *Nat Commun* 11: 3617

Longley DB, Harkin DP & Johnston PG (2003) 5-Fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer* 3: 330–338

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.24.517816; this version posted November 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

López S, Lim EL, Horswell S, Haase K, Huebner A, Dietzen M, Mourikis TP, Watkins TBK, Rowan A, Dewhurst SM, *et al* (2020) Interplay between whole-genome doubling and the accumulation of deleterious alterations in cancer evolution. *Nat Genet* 52: 283–293

Macintyre G, Goranova TE, De Silva D, Ennis D, Piskorz AM, Eldridge M, Sie D, Lewsley L-A, Hanif A, Wilson C, *et al* (2018) Copy number signatures and mutational processes in ovarian carcinoma. *Nat Genet* 50: 1262–1270

Martínez-Jiménez F, Movasati A, Brunner S, Nguyen L, Priestley P, Cuppen E & Hoeck AV (2022) Pan-cancer whole genome comparison of primary and metastatic solid tumors. 2022.06.17.496528 doi:10.1101/2022.06.17.496528 [PREPRINT]

Maura F, Weinhold N, Diamond B, Kazandjian D, Rasche L, Morgan G & Landgren O (2021) The mutagenic impact of melphalan in multiple myeloma. *Leukemia* 35: 2145–2150

Pich O, Cortes-Bullich A, Muiños F, Pratcorona M, Gonzalez-Perez A & Lopez-Bigas N (2021) The evolution of hematopoietic cells under cancer therapy. *Nat Commun* 12: 4803

Pich O, Muiños F, Lolkema MP, Steeghs N, Gonzalez-Perez A & Lopez-Bigas N (2019) The mutational footprints of cancer therapies. *Nat Genet* 51: 1732–1740

Pleasance E, Titmuss E, Williamson L, Kwan H, Culibrk L, Zhao EY, Dixon K, Fan K, Bowlby R, Jones MR, *et al* (2020) Pan-cancer analysis of advanced patient tumors reveals interactions between therapy and genomic landscapes. *Nat Cancer* 1: 452–468

Pommier Y, Leo E, Zhang H & Marchand C (2010) DNA Topoisomerases and Their Poisoning by Anticancer and Antibacterial Drugs. *Chem Biol* 17: 421–433

Prasad K, Bloomfield M, Levi H, Keuper K, Bernhard SV, Baudoin NC, Leor G, Giam M, Wong CK, Rancati G, *et al* (2021) Whole-genome duplication shapes the aneuploidy landscape of human cancers. 2021.05.05.442712 doi:10.1101/2021.05.05.442712 [PREPRINT]

Priestley P, Baber J, Lolkema MP, Steeghs N, de Bruijn E, Shale C, Duyvesteyn K, Haidari S, van Hoeck A, Onstenk W, *et al* (2019) Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* 575: 210–216

Rottenberg S, Disler C & Perego P (2021) The rediscovery of platinum-based cancer therapy. *Nat Rev Cancer* 21: 37–50

Siddik ZH (2003) Cisplatin: mode of cytotoxic action and molecular basis of resistance. *Oncogene* 22: 7265–7279

Siegel RL, Miller KD, Fuchs HE & Jemal A (2022) Cancer statistics, 2022. *CA Cancer J Clin* 72: 7–33

Steele CD, Abbasi A, Islam SMA, Bowes AL, Khandekar A, Haase K, Hames-Fathi S, Ajayi D, Verfaillie A, Dhami P, *et al* (2022) Signatures of copy number alterations in human cancer. *Nature* 606: 984–991

Vare D, Groth P, Carlsson R, Johansson F, Erixon K & Jenssen D (2012) DNA interstrand crosslinks induce a potent replication block followed by formation and repair of double strand breaks in intact mammalian cells. *DNA Repair* 11: 976–985

Westphalen CB, Fine AD, André F, Ganesan S, Heinemann V, Rouleau E, Turnbull C, Garcia Palacios L, Lopez J-A, Sokol ES, *et al* (2022) Pan-cancer Analysis of Homologous Recombination Repair–associated Gene Alterations and Genome-wide Loss-of-Heterozygosity Score. *Clin Cancer Res* 28: 1412–1421