Ploidetect enables pan-cancer analysis of the causes and impacts of chromosomal instability

Luka Culibrk[1,2], Jasleen K. Grewal[1,2], Erin D. Pleasance[1], Laura Williamson[1], Karen Mungall[1], Janessa Laskin[3], Marco A. Marra[1,4], and Steven J.M. Jones[1,4,✉]

[1]Canada's Michael Smith Genome Sciences Center at BC Cancer, Vancouver, British Columbia, Canada
[2]Bioinformatics training program, University of British Columbia, Vancouver, British Columbia, Canada
[3]Department of Medical Oncology, BC Cancer, Vancouver, British Columbia, Canada
[4]Department of Medical Genetics, Faculty of Medicine, Vancouver, British Columbia, Canada

Cancers routinely exhibit chromosomal instability, resulting in the accumulation of changes in the abundance of genomic material, known as copy number variants (CNVs). Unfortunately, the detection of these variants in cancer genomes is difficult. We developed Ploidetect, a software package that effectively identifies CNVs within whole-genome sequenced tumors. Ploidetect was more sensitive to CNVs in cancer related genes within advanced, pre-treated metastatic cancers than other tools, while also segmenting the most contiguously. Chromosomal instability, as measured by segment contiguity, was associated with several biological and clinical variables, including tumor mutation burden, tumor type, duration of therapy and immune microenvironment, highlighting the relevance of measuring CNV across the cancer genome. Investigation of gene mutations in samples revealed and the mutation status of several genes including ROCK2 and AC074391.1. Leveraging our heightened ability to detect CNVs, we identified 282 genes which were recurrently homozygously deleted in metastatic tumors. Further analysis of one recurrently deleted gene, MACROD2, identified a putative fragile tumor suppressor locus associated with response to chromosomal instability and chemotherapeutic agents. Our results outline the multifaceted impacts of CNVs in cancer by providing evidence of their involvement in tumorigenic behaviors and their utility as biomarkers for biological processes. We propose that increasingly accurate determination of CNVs is critical for their productive study in cancer, and our work demonstrates advances made possible by progress in this regard.

Correspondence: ***sjones@bcgsc.ca***

DRAFT

Introduction

Copy number variation (CNV) is a major class of somatic mutation often observed in the context of cancer. CNV has been shown to impact the majority of cancer genomes (1) and is associated with poor clinical outcomes (2). Homozygous deletions are CNVs which result in the complete loss of genomic loci, and recurrent homozygous deletions are known to be enriched for tumor suppressor genes (3). Homologous recombination deficiency (HRD), a cause of CNV, has been positively associated with platinum therapy response (4), and other phenotypes such as an increased prevalence of tandem duplications have been described with relevance to biological processes (5, 6). Overall, chromosomal instability (CIN) driving the accumulation of CNVs is associated with poor patient outcome in multiple cancer types (7–9). Although CNV is a highly pervasive mechanism by which tumors mutate, these variants are considerably more difficult to detect accurately compared to other types of mutations and consequently they may represent an under-explored facet of tumor biology.

While small mutations can be determined through base changes embedded within aligned sequence reads, CNVs are variations in DNA quantity and are typically determined through measurement of relative DNA abundance by way of sequence read depth (10, 11). Due to random noise and sequence bias, read depth data is unreliable at the granular level and therefore must be noise corrected. Segmentation, the process of clustering genomic loci into contiguous regions of constant copy number is used for CNV analysis as the aggregation of numerous sequential read depth observations aids in overcoming the inherent noise in the data. There have been a number of segmentation methods developed to tackle this issue in the context of both next-generation sequencing data and array CGH (10, 12, 13). Although CNVs are arguably comparable in importance with small variants in cancer, tools for somatic CNV detection tend to make errors considerably more often than small variant callers (14). While highly sensitive CNV detection for identifying tumor suppressor deletions and oncogene amplifications is desirable, it is vital to also ensure that CNV segmentation is precise. Highly fragmented segmentation results preclude the use of metrics such as CNV burden as phenotypic features related to chromosomal instability (CIN) in cancer biology because false positive breakpoints cause an inflated estimate of the magnitude of CIN. Inherently noisy data impedes the ability to separate signal from noise, necessarily limiting the degree of positive, publishable findings that can be generated from CNV data (15).

Clinical genomics is a rising area of interest in cancer research. Studies that evaluate the utility of a personalized approach to identifying potential cancer therapeutics have been reported in recent years (16, 17). Studies that broadly attempt to genomically characterize large cohorts of tumor biopsies are also becoming increasingly common-place (18, 19). As the cost of sequencing continues to decrease, large-scale sequencing studies of tumor biopsies will become more commonplace, highly precise and accurate tools to interrogate these genomes will be required for their effective interpretation.

Here we present Ploidetect, a tool that performs copy number variation analysis of cancer genomes. Ploidetect contributes to the field of somatic CNV detection by using a novel segmentation approach to better segment genomes compared to contemporary methods. We demonstrate its performance on a cohort of 725 metastatic tumor biopsies from patients with treatment-resistant cancers sequenced as part of the Personalized OncoGenomic (POG) program at BC Cancer (20). Further, we find that by using Ploidetect's CNV estimates we can reconstruct existing knowledge of the CIN phenotype in cancer, discover novelties regarding CIN, and elucidate putative gene associations with CIN. As a unique advantange of Ploidetect, parameter tuning for tool usage is entirely data-driven. Ploidetect utilizes a two-step process of joint purity-ploidy estimation followed by CNV calling using a novel segmentation procedure. In a cohort of 725 real-world metastatic tumor biopsies, Ploidetect outperforms existing methods for detection of CNVs, while also limiting oversegmentation. Using Ploidetect's CNV determination, we found that CIN correlates with a number of genomic and clinical features in cancer, including site of origin, immune infiltration, and response to chemotherapy. Further analysis identified known and putative gene associations with CIN. Lastly, a permutation analysis of recurrently deleted regions (RDRs) within the genome enabled prioritization of candidate fragile tumor suppressor genes. This work represents our contribution to improving CNV detection in paired tumor-normal genomic analyses and the ability of our method to enable CNV biomarker research and discovery.

Methods

Sample collection, sequencing and POG analysis. Samples were obtained from the Personalized OncoGenomics (POG) project, with inclusion criteria, enrolment details, clinical background, and specific details on sequencing methodology, analysis, and software previously described (17). Briefly, POG tumor biopsies are obtained from advanced, treatment-resistant metastatic cancer patients. The biopsies are sequenced using short-read paired-end sequencing technology to approximately 80-fold genomic coverage, with a matched germline genome from blood sequenced to approximately 40-fold coverage. Expression data are also obtained using RNA-seq with a target depth of 200 million paired-end reads. Data from 725 total patients were used. Binary alignment map files were procesed using samtools 1.11 (21), bedtools 1.11 (21), and BEDtools 2.30 (22). SNVs, isoform-specific expression, and clinical data were processed as described previously and accessed from our internal database, vardb (Lewis *et al.,* in preparation.).

Ploidetect package. The Ploidetect R package is described in detail in the Supplementary Methods.

Statistical analysis. Statistical tests were the two-sided Wilcoxon rank-sum test unless otherwise specified. Analyses were performed in R version 3.6.3.

Benchmarking. We compared Ploidetect with 3 other CNV tools, namely Sequenza (11), BIC-seq2 (10) and PURPLE (23). We used the latest versions of each software at the time of analysis, corresponding to Sequenza version 3.0.0, BIC-seq2 version 0.7.2, and AMBER/COBALT/PURPLE versions 3.5/1.11/2.51, respectively. Our test set was 725 cases of the Personalized Oncogenomics (POG) cohort (17), including recent cases sequenced after the publication of Pleasance *et al*. Purity and ploidy results were obtained from Ploidetect, Sequenza and PURPLE. While BIC-seq2 included a program, purityEM, which appeared to perform tumor purity and ploidy estimation, which we were unable to successfully implement due in part to a lack of documentation. We instead used Ploidetect's tumor purity and ploidy estimates to assign copy number to the BIC-seq2 segments. Each tool was run using its respective default settings using a Snakemake (24) workflow. Homozygous deletions were defined as a copy number below 0.25, and amplifications were defined as a copy number $\geq$ ploidy + 3.

Detection of significantly deleted regions. We performed a permutation analysis by shuffling the midpoint of all homozygous deletions on each chromosome 10^7 times. For each interval, we recorded the number of permutations where more shuffled deletions were more numerous than actual deletions. P-values were obtained by dividing this count by the number of permutations overall. The p-values were adjusted for multiple testing using the Benjamini-Hochberg method (25). For analysis we considered regions with an FDR <= 0.01 and regions which contained deletions in at least five separate paired tumor-normal cases.

Identification of variants associated with CIN. We obtained the set of all single nucleotide variants (SNVs) in POG and removed all intergenic, synonymous, upstream, downstream, intronic, and untranslated region SNVs. We selected genes which had SNVs in at least 10 cases and assigned a binary variable for each case and gene combination, where 1 indicated an SNV was present in the gene, and a 0 indicated the absence. Segment N50 values were scaled by dividing each value by the maximum. We performed recursive feature elimination (RFE) by training a support vector regression (SVR) model using scikit-learn (26). We used a linear kernel, a C value of 1 and a gamma of 0.001. We eliminated the 10% least informative positively signed coefficients and the 10% least informative negatively signed coefficients, measured by the absolute value of the coefficients at each iteration. At each iteration, 5-fold cross validation was performed to assess the training and test error, measured as the root-mean-squared concatenated difference between predicted scaled N50 values and the real scaled N50 values for all folds. Using the set of genes from RFE with the lowest test error, we performed a one-sided Wilcoxon rank-sum test for each gene using their SVM coefficients as the prior to inform the directionality of each test. We tested the difference in segment N50 between variant and non-variant cases

DRAFT

for each gene selected by RFE, and corrected the resulting P-values for multiple testing using the FDR method.

Ploidetect software availability. Ploidetect is implemented in R and made available freely at github.com/lculibrk/Ploidetect. Analyses in this paper were conducted using the version with commit master@98197bc.

Results

Accurate estimation of tumor purity in a diverse cohort of advanced cancers. Firstly, we assessed the ability of Ploidetect to estimate tumor purity in a cohort of samples from advanced cancer patients enrolled in the POG program (17, 20). In the context of the POG project, tumor purity was already assessed by manual review of CNV data from CNAseq and APOLLOH (27) and ranged between 10 to 100%. We found that Ploidetect tumor purity estimates correlated well with those obtained by manual review (Pearson's $\rho = 0.807$, Figure 1A). Because Ploidetect provides visualizations of alternate model fits, it enabled the consideration and review of alternate solutions for tumor purity. We found a modest improvement in tumor purity estimates after manual review of Ploidetect output (Figure 1B). To assess its ability to detect a wide range of tumor purity, we applied Ploidetect to an *in-silico* dilution of the COLO829 cell line (28), simulating tumor purities of 10% to 100%. Ploidetect was able to accurately estimate tumor purity in all cases (Figure 1C).

Ploidetect is a highly sensitive and precise copy number variation detector. We next tested Ploidetect's ability to conservatively segment cancer genomes while retaining the ability to detect focal somatic CNVs of interest. As a proxy for a specificity metric (since there is no appropriate gold standard data for this task), we assessed the segment N50 of the CNV profiles, defined as the size of the smallest segment for which all segments that size and below comprise half of the genome. If two methods are equally capable of detecting CNVs, a higher segment N50 implies fewer false-positive breakpoints. Sensitivity was assessed by the ability of each tool to detect a number of manually-curated CNVs which are known to commonly occur in cancer (cancer-related CNVs, crCNVs) within the cohort, measured as the proportion of crCNVs called by a tool compared to the crCNVs called by all tools for a given case. Genes labeled as crCNVs in addition to their associated CNV state are listed in Table 1. Of all the assessed tools, Ploidetect identified more cancer-related CNVs than the other tools, and demonstrated the highest observed segment contiguity (Figure 2A). While Sequenza scored second-highest in CNV sensitivity, the segment contiguity was consistently low. BIC-seq2 tended to segment more contiguously than Sequenza but had the overall worst sensitivity for crCNVs. PURPLE was the only CNV caller with contiguity comparable to Ploidetect, although it was less sensitive in detection of crCNVs, being unable to detect any cancer CNVs found by other tools in at least 25% of cases. Of the called crCNVs, we randomly selected 50 amplification CNVs and 50 homozygous deletion CNVs from each tool for a total of 400 CNV events, and manually reviewed each of them to determine whether the CNVs actually existed or not. Ploidetect demonstrated the second-highest precision for amplifications, behind only BIC-seq2, although it was considerably more sensitive. When assessing homozygous deletions, Ploidetect demonstrated 100% precision (Figure 2B). Ploidetect was the only tool that did not make even a single false homozygous deletion call in the manual review cohort.

Chromosomal instability in metastatic cancer. Leveraging our improved ability to segment tumor genomes, we sought to characterize the interaction of CIN with tumor mutation burden (TMB), cancer type, treatment response, immune infiltration and mutation status using CNV data. Since segment N50 is a measure of the CNV burden of the genome, we used it as an estimate of CIN. We found that segment N50 is inversely correlated with both the HRDscore and the genomic instability index (GII) (Supplemental Figure S1). Reflecting previous research on CIN (29), we found that TMB correlated negatively with CNV segment N50 in POG (spearman's rho = -0.39, log-linear model p = 1.5 x 10^{-15}), suggesting that CIN is correlated with TMB (Figure 3A). Furthermore, we noted tissue of origin specific variation in segment N50, with breast, stomach, and ovarian tumors as being the most chromosomally unstable (Figure 3B). Time to treatment discontinuation (TTD) has been previously used in analyses of POG data as a proxy for treatment outcome, as therapies that elicit a positive response would naturally be used for longer durations (17, 30). Since HRD causes CIN and predicts platinum treatment response (31), we tested whether segment N50 was a sufficiently sensitive metric to detect a difference in TTD of platinum therapy. We discretized the segment N50 into CIN-extreme (1st quartile N50 cohort-wide) and CIN-low (2nd-4th quartile N50 cohort-wide) groups and found that patients with CIN-extreme tumors had significantly elevated TTD on platinum chemotherapy (p = 0.019, Wilcoxon rank-sum test) (Figure 3C). This association was not significant when using either HRDscore or GII instead of segment n50 (Supplemental Figure S2). CIN has been reported as having a suppressive impact on the tumor immune microenvironment, possibly as a consequence of negative selection by the cGAS-STING pathway (32). We investigated the potential impact of CIN on the presence of immune cells, and in particular CD8+ T cells, in the bulk tumour samples as measured using CIBERSORT (33). Both T-cell infiltration and $CD8^{+}$ T-cell scores were modestly associated with segment N50 after correcting for tumor purity (p = 9.82 x $10^{-}3$ and 8.92 x $10^{-}3$ respectively, multiple regression using segment N50 and tumor purity) (Figure 3D). Finally, we attempted to uncover genes with novel relevance to CIN in metastatic cancer. Having established the biological relevance of our CIN measure, we identified 212 genes whose somatic small mutation status could predict the segment N50 through training support vector machines (SVMs) using recursive feature elimination (RFE) (Supplemental Table S1). Six genes were significantly associated with a change in segment N50 (FDR < 0.05, Wilcoxon rank-sum test with Benjamini-Hochberg FDR correction) (Table

DRAFT

2). Among the significant genes was TP53, mutations of which have been recurrently implicated in CIN (6, 34). Also present were AC074391.1, TTN, ROCK2, ZNF804B, and DMBT1. We selected TP53, ROCK2 and AC074391.1 for further analysis. Given the highly recurrent nature of TP53 mutations in cancer, we further investigated the independent association of ROCK2 and AC074391.1 with CIN. We found that ROCK2 mutations led to a decrease in segment N50 independent of TP53 mutation status with a lack of additive impact in the double-mutant case (Figure 3E). AC074391.1 mutated cases demonstrated a subtractive impact on the segment N50 when TP53 was also mutated (Figure 3F). These results imply a potential functional association of ROCK2 and AC074391.1 with CIN, and further investigation into mechanisms that may underlie this association are required.

Recurrent homozygous deletions in metastatic tumors. We evaluated whether we could use Ploidetect to identify recurrent homozygously deleted regions in metastatic tumor genomes. Using a permutation strategy similar to that reported by Cheng *et al* (3), we identified 131 recurrently deleted regions (RDRs) in the POG cohort covering 282 genes (False discovery rate (FDR) (25) ≤ 0.01, # cases ≥ 5) (Figure 4A). The mean size of the RDRs was 560.8Kb long, with the smallest RDR being 5080bp and the longest being 9.65Mb. Overall we identified 24.21Mb of the genome as being recurrently deleted within the POG cohort. For each affected gene, we profiled the proportion of exons which were overlapped by each RDR (Figure 4B). The presence of known tumor suppressor genes, such as TP53, PTEN, and CDKN2A as well as known fragile sites, including FHIT and WWOX in our analysis demonstrates the effectiveness of this approach for identifying recurrent deletions in tumours (Table 3) (Supplemental Table S2).

MACROD2 RDR deletions promote CIN and confer resistance to chemotherapy. Analysis of frequently deleted regions in our metastatic cohort revealed recurrent deletion of the gene MACROD2 as having relevance relevance to both CIN and treatment outcome. We identified 28 homozygous deletions within the RDR overlapping MACROD2 (Figure 5A). These deletions were found to be proximal to an alternative promoter for MACROD2 (Supplemental Figure S3) (35). While the MACROD2 RDR did overlap at least one exon for each MACROD2 transcript, only 14% of the deletions overlapped an exon and perturbed the protein coding sequence (Supplemental Figure S4). Because of the proximal alternate promoter and the observation that this RDR only impacts a restricted region of the gene rather than the entire gene, we investigated the impact of MACROD2 RDR deletions on the expression of MACROD2's transcripts. MACROD2 RNA expression was significantly decreased in tumors with this deletion ($p<2x10^{-9}$, two-sided Wilcoxon rank-sum test), although the balance of transcript isoforms was unaffected (Figure 5B & Figure 5C). Since MACROD2 is involved in PARP-mediated single stranded break (SSB) repair, we assessed whether deletions of this locus would be associated with CIN as well as TTD on therapeutic agents that induce SSBs (36). Tumors with deletions in MACROD2 had a higher degree of CIN; their genomes had significantly reduced segment N50s as well as a significant increase in the number of segments called overall (Figure 5D). . Deletions within this MACROD2 segment were significantly associated with a reduced TTD in patients treated with TOP2A inhibitors or cyclophosphamide (Figure 5E), but not platinum-based therapeutics, suggesting MACROD2 may play a role in response to certain DNA damaging chemotherapies (Figure 5E). These trends were observed irrespective of tumor type (Supplemental Figure S5).

Discussion

Copy number variation detection in cancer is a challenging problem (37). Poor segmentation of CNVs leads to incorrect assignment of CNV breakpoints. This results in a failure to detect important CNVs, an inability to assess CIN due to an increase in false positive breakpoints, or both. In the present study, we have described a novel CNV detection method which performed demonstrably better in these regards compared to other CNV software. Due in large part to our ability to obtain accurately CNV segmented tumor genomes, we were able to effectively estimate CIN from bulk CNV burden. We showed that our measure of CIN, the segment N50 associated with TMB, tissue of origin, platinum treatment response, and immune infiltration in our cohort. We used our measures of CIN to identify genes with potential impact on CIN, and functionally contrasted their impact alongside TP53 deficiency-induced CIN. Ploidetect's unparalleled sensitivity for CNVs in our cohort allowed for an in-depth analysis of recurrently deleted loci in advanced cancers, including one within the fragile tumor suppressor gene MACROD2 with implications on both CIN and chemotherapy response.

By leveraging accurate segmentation of CNVs, we were able to directly use genome contiguity to measure CIN. Our results agree with previous research on TMB and CIN (29), and we find that different tumor types are differentially susceptible to CIN. Ovarian cancers in particular have been reported to frequently exhibit CIN (6, 38), although stomach and breast tumors have also been observed to have consistently high CNV burden (39). Cancer type and site of origin are well known to affect patterns of both small mutations (40, 41) and CNVs (1). We observed a broad increase in CIN within different tumor types, consistent with hypotheses that different tumor types enjoy different degrees of fitness improvement from CIN (42). More broadly, CIN appears to be broadly connected with genomic instability in cancers (43). Broader phenotypic features of cancer genomes, such as HRD, manifest alongside and in combination with CIN. HRD, measured through the HRDscore has been previously used as a measure of CIN (44), and, when combined with five other metrics into HRDetect, was found to predict platinum treatment response (31). Compared to the HRDscore, the

DRAFT

metric of segment N50 is comparatively simple and was associated with platinum TTD in our cohort. While this observation in our study is encouraging, the observed impact was substantively small. The incorporation of accurate segment contiguity metrics into ensemble classifiers such as HRDetect could further improve the predictive potential of these classifiers and better refine the characterization of this phenotype.

In addition to HRD, we investigated the relationship between CIN and the tumor immune microenvironment. Although activation of innate immunity by CIN through the cGAS-STING pathway is known (45), paradoxical correlation of suppression of immune activation with CIN has been noted in the literature (32, 46). Particularly, it has been shown that the absence of $CD8^+$ T-cells has a causal relationship with CIN in mouse xenograft models of gliomas (47). Accordingly, we found that tumors with higher CIN had reduced levels of both T-cell infiltration and $CD8^+$ T-cells in our study. Research investigating a possible connection between CIN and immunotherapy efficacy are as yet conflicted on whether CIN alone can predict immunotherapy response (29, 48).Taken together with the literature, our data suggest that the immune microenvironment exerts a selective pressure on tumors to limit CIN during tumor development, and consequently immunoediting becomes increasingly necessary to overcome this pressure in tumors which depend on CIN. Loss of specific chromosomal loci, such as the region harboring the interferon gene clusters on 9p, may partly be the cause of this immuno-modulatory effect (32), although further research will be needed to investigate this.

Numerous genes are involved in regulating the structural stability of genomes, and we have identified a number of genes that may also be involved. ROCK2 is a protein kinase that has been reported with oncological relevance recurrently in the literature (49, 50), including one study which identified some functional impact on DNA repair, specifically demonstrating an impact on radiotherapy resistance. (51). Unfortunately, we lack sufficient radiotherapy-treated patients in the POG study to test this hypothesis. The association of ROCK2 mutations with CIN in our study appears to be independent of TP53 mutations, since there was no evidence of any combinatorial impact of these mutations on CIN in our data. We also identified mutations in AC074391.1, a long non-coding RNA gene as being significantly associated with an increase in CIN. While long non-coding RNAs have been observed to impact chromosome stability (52), we could find no specific mention of AC074391.1 in literature. Unlike ROCK2, AC074391.1 demonstrated an additive or epistatic interaction with TP53 mutation status on CIN, suggesting that any mechanism by which AC074391.1 mutations confer CIN is distinct from mechanisms impacted by TP53 variants. While we did also identify TTN in our analysis, we consider this likely to be noise due to TTN having a large coding region and our observed correlation between TMB and CIN. We also identified ZNF804B and DBMT1 mutations as being significantly associated with CIN. Specific genetic and biochemical study will be required to better understand ROCK2, AC074391.1, and the other identified genes' potential mechanistic interactions with TP53 and CIN.

Ploidetect enabled a pan-cancer analysis of frequently homozygously deleted tumor genome loci. We identified nearly 300 significantly deleted genes within the POG cohort. Although many of these observations are likely to be passenger mutations on segments with tumor suppressors, such as the 57 genes found within the segment containing CDKN2A, we identified numerous tumor suppressor genes and fragile loci recurrently found in cancer (3). Many of these well-known tumor suppressor genes such as CDKN2A, PTEN and TP53 were found to be recurrently deleted without preference for any specific exonic or intronic regions, indicating that all positions of each gene are equally fragile. We identified a number of genes which overlapped RDRs at only a few or even none of their exons. Many of these genes have been previously annotated as fragile tumor suppressors, such as FHIT (53), WWOX (54), and RBFOX1 (55). We have further delineated the fragile loci within each gene at exonic resolution. Due to their high prevalence within our cohort, we decided to investigate MACROD2 RDR deletions further. We observed MACROD2 RDR deletions to be primarily intronic and upstream of an alternate promoter site. Consequently, RDR deletion-induced loss of MACROD2 function probably stems from a splicing impairment or a hypothetical cis-acting regulatory site which has not been described. While MACROD2 deletions have previously been reported to cause CIN in colorectal tumors (36), we identified them in numerous other cancer types and report their impact, both on CIN and on treatment resistance. MACROD2 is necessary for PARP1 function through the removal of mono-ADP ribose from PARP1 (36). Cyclophosphamide and TOP2A inhibitors induce cytotoxicity through PARP1 hyperactivation and consequential depletion of NAD^+ stores (56, 57). MACROD2 RDR deletions may dampen PARP1 hyperactivation, preserving cellular NAD^+, thereby preventing cytotoxicity and allowing tumor cells to survive. While we identified numerous other promising deletions with lower incidence, heterogeneity of cancer type, demographics, and treatment strategies preclude deeper investigation into their impacts on treatment status and genomic phenotype. Larger datasets will be required to properly interrogate the lower frequency deletions we have identified to identify potential implications for treatment response and patient outcome.

While we have identified a number of enticing observations in our data, there are a number of factors to consider in their interpretation. Our use of the segment N50 metric as a measure of CIN was based in part on the intuitive concept that genomes broken at multiple positions would be considered chromosomally unstable. The use of CNVs for CIN research is likely to be enabled by continued improvement in CNV segmentation methods, as we have described here. Ploidetect's contribution to the accurate detection of CNVs has demonstrably provided a practical benefit to CNV research by providing improved segmentation performance relative to other methods. The cohort that we have chosen to study is a unique cohort of pre-treated, advanced tumors with a diverse

DRAFT

range of patient backgrounds, tumor types, and lines of prior therapy. While many previous large-scale studies focus on primary, treatment-naive tumors, here we examine metastatic lesions where CIN serves as a source of genetic variation for further tumor evolution, including the development of treatment resistance (45, 58, 59). Because of the advanced nature of our cohort, survivorship bias (referring here to tumor survival) may also impact our results by selecting for patients whose tumors lack features that would confer sensitivity to common therapies.

Emerging technologies for assessing patterns of CNV in whole genomes enable comprehensive examination of the interplay between this class of mutation and cancer. Iterative improvement in our ability to determine accurately the suite of CNVs within sequenced tumors will provide the opportunity to ask nuanced questions of CNV data with greater power and without the confounding effects of persistent false positives. To enable the most effective research, the goal of the field must be to eventually approach the level of performance routinely achieved by small variant callers, a process for which we have herein provided our contribution. Interplay between CIN and and the many genomic and phenotypic features of cancer, including HRD and immune escape are beginning to be understood. The further elucidation of these patterns are likely to give rise to a greater understanding of the intricacies of this disease as well as potential means to combat it.

ACKNOWLEDGEMENTS

We would like to gratefully acknowledge the patients of the POG program and their families, without whom we would have never been able to conduct this study. We also gratefully thank the oncologists, nurses, analysts, and volunteers involved with POG. We thank the BC Cancer Foundation and Genome British Columbia for their support (project B20POG). We also acknowledge contributions towards equipment and infrastructure from Genome Canada and Genome British Columbia (projects 202SEQ, 212SEQ, 12002), Canada Foundation for Innovation (projects 20070, 30981, 30198, 33408 and 35444) and the BC Knowledge Development Fund. LC is supported by a Canadian Institutes for Health Research (CIHR) Frederick Banting and Charles Best Canada Graduate Scholarship GSD-164207 and a University of British Columbia 4-year fellowship. JG is supported by a University of British Columbia 4-year fellowship. SJMJ and MAM acknowledge funding from the Canadian Research Chairs Program.

Bibliography

1. Travis I. Zack, Steven E. Schumacher, Scott L. Carter, Andrew D. Cherniack, Gordon Saksena, Barbara Tabak, Michael S. Lawrence, Cheng Zhong Zhang, Jeremiah Wala, Craig H. Mermel, Carrie Sougnez, Stacey B. Gabriel, Bryan Hernandez, Hui Shen, Peter W. Laird, Gad Getz, Matthew Meyerson, and Rameen Beroukhim. Pan-cancer patterns of somatic copy number alteration. *Nature Genetics*, 2013. ISSN 10614036. doi: 10.1038/ng.2760.
2. Haley Hieronymus, Rajmohan Murali, Amy Tin, Kamlesh Yadav, Wassim Abida, Henrik Moller, Daniel Berney, Howard Scher, Brett Carver, Peter Scardino, Nikolaus Schultz, Barry Taylor, Andrew Vickers, Jack Cuzick, and Charles L. Sawyers. Tumor copy number alteration burden is a pan-cancer prognostic factor associated with recurrence and death. *eLife*, 2018. ISSN 2050084X. doi: 10.7554/eLife.37294.
3. Jiqiu Cheng, Jonas Demeulemeester, David C. Wedge, Hans Kristian M. Vollan, Jason J. Pitt, Hege G. Russnes, Bina P. Pandey, Gro Nilsen, Silje Nord, Graham R. Bignell, Kevin P. White, Anne Lise Børresen-Dale, Peter J. Campbell, Vessela N. Kristensen, Michael R. Stratton, Ole Christian Lingjærde, Yves Moreau, and Peter Van Loo. Pan-cancer analysis of homozygous deletions in primary tumours uncovers rare tumour suppressors. *Nature Communications*, 2017. ISSN 20411723. doi: 10.1038/s41467-017-01355-0.
4. L. Telli Melinda, M. Timms Kirsten, Reid Julia, Hennessy Bryan, B. Mills Gordon, C. Jensen Kristin, Szallasi Zoltan, T. Barry William, P. Winer Eric, M. Tung Nadine, J. Isakoff Steven, D. Ryan Paula, Greene Colozzi April, Gutin Alexander, Sangale Zaina, Iliev Diana, Neff Chris, Abkevich Victor, T. Jones Joshua, S. Lanchbury Jerry, Hartman Anne-Renee, E. Garber Judy, M. Ford James, P. Silver Daniel, and L. Richardson Andrea. Homologous recombination deficiency (hrd) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer. *Clinical Cancer Research*, 2016. ISSN 15573265. doi: 10.1158/1078-0432.CCR-15-2477.
5. Francesca Menghi, Floris P. Barthel, Vinod Yadav, Ming Tang, Bo Ji, Zhonghui Tang, Gregory W. Carter, Yijun Ruan, Ralph Scully, Roel G.W. Verhaak, Jos Jonkers, and Edison T. Liu. The Tandem Duplicator Phenotype Is a Prevalent Genome-Wide Cancer Configuration Driven by Distinct Gene Mutations. *Cancer Cell*, 2018. ISSN 18783686. doi: 10.1016/j.ccell.2018.06.008.
6. Geoff Macintyre, Teodora E. Goranova, Dilrini De Silva, Darren Ennis, Anna M. Piskorz, Matthew Eldridge, Daoud Sie, Liz Anne Lewsley, Aishah Hanif, Cheryl Wilson, Suzanne Dowson, Rosalind M. Glasspool, Michelle Lockley, Elly Brockbank, Ana Montes, Axel Walther, Sudha Sundar, Richard Edmondson, Geoff D. Hall, Andrew Clamp, Charlie Gourley, Marcia Hall, Christina Fotopoulou, Hani Gabra, James Paul, Anna Supernat, David Millan, Aoisha Hoyle, Gareth Bryson, Craig Nourse, Laura Mincarelli, Luis Navarro Sanchez, Bauke Ylstra, Mercedes Jimenez-Linan, Luiza Moore, Oliver Hofmann, Florian Markowetz, Iain A. McNeish, and James D. Brenton. Copy number signatures and mutational processes in ovarian carcinoma. *Nature Genetics*, 2018. ISSN 15461718. doi: 10.1038/s41588-018-0179-8.
7. Christoph E. Heilig, Harald Löffler, Ulrich Mahlknecht, Johannes W.G. Janssen, Anthony D. Ho, Anna Jauch, and Alwin Krämer. Chromosomal instability correlates with poor outcome in patients with myelodysplastic syndromes irrespectively of the cytogenetic risk group. *Journal of Cellular and Molecular Medicine*, 2010. ISSN 15821838. doi: 10.1111/j.1582-4934.2009.00905.x.
8. Béatrice Orsetti, Janick Selves, Caroline Bascoul-Mollevi, Laurence Lasorsa, Karine Gordien, Frédéric Bibeau, Blandine Massemin, François Paraf, Isabelle Soubeyran, Isabelle Hostein, Valérie Dapremont, Rosine Guimbaud, Christophe Cazaux, Michel Longy, and Charles Theillet. Impact of chromosomal instability on colorectal cancer progression and outcome. *BMC Cancer*, 2014. ISSN 14712407. doi: 10.1186/1471-2407-14-121.
9. Chang Min Choi, Kwang Won Seo, Se Jin Jang, Yeon Mok Oh, Tae Sun Shim, Woo Sung Kim, Dong Soon Lee, and Sang Do Lee. Chromosomal instability is a risk factor for poor prognosis of adenocarcinoma of the lung: Fluorescence in situ hybridization analysis of paraffin-embedded tissue from Korean patients. *Lung Cancer*, 2009. ISSN 01695002. doi: 10.1016/j.lungcan.2008.07.016.
10. Ruibin Xi, Semin Lee, Yuchao Xia, Tae Min Kim, and Peter J. Park. Copy number analysis of whole-genome data using BIC-seq2 and its application to detection of cancer susceptibility variants. *Nucleic Acids Research*, 2016. ISSN 13624962. doi: 10.1093/nar/gkw491.
11. F. Favero, T. Joshi, A. M. Marquard, N. J. Birkbak, M. Krzystanek, Q. Li, Z. Szallasi, and Aron C. Eklund. Sequenza: Allele-specific copy number and mutation profiles from tumor sequencing data. *Annals of Oncology*, 26(1):64–70, 2015. ISSN 15698041. doi: 10.1093/annonc/mdu479.
12. Adam B. Olshen, E. S. Venkatraman, Robert Lucito, and Michael Wigler. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*, 2004. ISSN 14654644. doi: 10.1093/biostatistics/kxh008.
13. Gro Nilsen, Knut Liestøl, Peter Van Loo, Hans Kristian Moen Vollan, Marianne B. Eide, Oscar M. Rueda, Suet Feung Chin, Roslin Russell, Lars O. Baumbusch, Carlos Caldas, Anne Lise Børresen-Dale, and Ole Christian Lingjærde. Copynumber: Efficient algorithms for single- and multi-track copy number segmentation. *BMC Genomics*, 2012. ISSN 14712164. doi: 10.1186/1471-2164-13-591.
14. Le Zhang, Wanyu Bai, Na Yuan, and Zhenglin Du. Comprehensively benchmarking applications for detecting copy number variation. *PLoS Computational Biology*, 2019. ISSN 15537358. doi: 10.1371/journal.pcbi.1007069.
15. The importance of no evidence. *Nature Human Behaviour*, 3(3):197, 2019. ISSN 23973374. doi: 10.1038/s41562-019-0569-7.
16. Dan R. Robinson, Yi Mi Wu, Robert J. Lonigro, Pankaj Vats, Erin Cobain, Jessica Everett, Xuhong Cao, Erica Rabban, Chandan Kumar-Sinha, Victoria Raymond, Scott Schuetze, Ajjai Alva, Javed Siddiqui, Rashmi Chugh, Francis Worden, Mark M. Zalupski, Jeffrey Innis, Rajen J. Mody, Scott A. Tomlins, David Lucas, Laurence H. Baker, Nithya Ramnath, Ann F. Schott, Daniel F. Hayes, Joseph Vijai, Kenneth Offit, Elena M. Stoffel, J. Scott Roberts, David C. Smith, Lakshmi P. Kunju, Moshe Talpaz, Marcin Cieślik, and Arul M. Chinnaiyan. Integrative clinical genomics of metastatic cancer. *Nature*, 548(7667):297–303, 2017. ISSN 14764687. doi: 10.1038/nature23306.
17. Erin Pleasance, Emma Titmuss, Laura Williamson, Harwood Kwan, Luka Culibrk, Eric Y. Zhao, Katherine Dixon, Kevin Fan, Reanne Bowlby, Martin R. Jones, Yaoqing Shen, Jasleen K. Grewal, Jahanshah Ashkani, Kathleen Wee, Cameron J. Grisdale, My Linh Thibodeau, Zoltan Bozoky, Hillary Pearson, Elisa Majounie, Tariq Vira, Reva Shenwai, Karen L. Mungall, Eric Chuah, Anna Davies, Mya Warren, Caralyn Reisle, Melika Bonakdar, Gregory A. Taylor, Veronika Csizmok, Simon K. Chan, Zusheng Zong, Steve Bilobram, Amir Muhammadzadeh, Darryl D'Souza, Richard D. Corbett, Daniel MacMillan, Marcus Carreira, Caleb Choo, Dustin Bleile, Sara Sadeghi, Wei Zhang, Tina Wong, Dean Cheng, Scott D. Brown, Robert A. Holt, Richard A. Moore, Andrew J. Mungall, Yongjun Zhao, Jessica Nelson, Alexandra Fok, Yussanne Ma, Michael K. C. Lee, Jean-Michel Lavoie, Shehara Mendis, Joanna M. Karasinska, Balvir Deol, Ana Fisic, David F. Schaeffer, Stephen Yip, Kasmintan Schrader, Dean A. Regier, Deirdre Weymann, Stephen Chia, Karen Gelmon, Anna Tinker, Sophie Sun, Howard Lim, Daniel J. Renouf, Janessa Laskin, Steven J. M. Jones, and Marco A. Marra. Pan-cancer analysis of advanced patient tumors reveals interactions between therapy and genomic landscapes. *Nature Cancer*, 2020. doi: 10.1038/s43018-020-0050-6.
18. Serena Nik-Zainal, Helen Davies, Johan Staaf, Manasa Ramakrishna, Dominik Glodzik, Xueqing Zou, Inigo Martincorena, Ludmil B. Alexandrov, Sancha Martin, David C. Wedge, Peter Van Loo, Young Seok Ju, Marcel Smid, Arie B. Brinkman, Sandro Morganella, Miriam R. Aure, Ole Christian Lingjærde, Anita Langerød, Markus Ringnér, Sung Min Ahn, Sandrine Boyault, Jane E. Brock, Annegien Broeks, Adam Butler, Christine Desmedt, Luc Dirix, Serge Dronov, Aquila Fatima, John A. Foekens, Moritz Gerstung, Gerrit K.J. Hooijer, Se Jin Jang, David R. Jones, Hyung Yong Kim, Tari A. King, Savitri Krishnamurthy, Hee Jin Lee, Jeong Yeon Lee, Yilong Li, Stuart McLaren, Andrew Menzies, Ville Mustonen, Sarah O'Meara, Iris Pauporté, Xavier Pivot, Colin A. Purdie, Keiran Raine, Kamna Ramakrishnan, F. Germán Rodríguez-González, Gilles Romieu, Anieta M. Sieuwerts, Peter T. Simpson, Rebecca Shepherd, Lucy Stebbings, Olafur A. Stefansson, Jon Teague, Stefania Tommasi, Isabelle Treilleux, Gert G. Van Den Eynden, Peter Vermeulen, Anne Vincent-Salomon, Lucy Yates, Carlos Caldas, Laura Van t. Veer, Andrew Tutt, Stian Knappskog, Benita Kiat Tee Tan, Jos Jonkers, Åke Borg, Naoto T. Ueno, Christos Sotiriou, Alain Viari, P. Andrew Futreal, Pe-

DRAFT

ter J. Campbell, Paul N. Span, Steven Van Laere, Sunil R. Lakhani, Jorunn E. Eyfjord, Alastair M. Thompson, Ewan Birney, Hendrik G. Stunnenberg, Marc J. Van De Vijver, John W.M. Martens, Anne Lise Børresen-Dale, Andrea L. Richardson, Gu Kong, Gilles Thomas, and Michael R. Stratton. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*, 2016. ISSN 14764687. doi: 10.1038/nature17676.

19. Peter J. Campbell, Gad Getz, Jan O. Korbel, Joshua M. Stuart, Jennifer L. Jennings, Lincoln D. Stein, Marc D. Perry, Hardeep K. Nahal-Bose, B. F.Francis Ouellette, Constance H. Li, Esther Rheinbay, G. Petur Nielsen, Dennis C. Sgroi, Chin Lee Wu, William C. Faquin, Vikram Deshpande, Paul C. Boutros, Alexander J. Lazar, Katherine A. Hoadley, David N. Louis, L. Jonathan Dursi, Christina K. Yung, Matthew H. Bailey, Gordon Saksena, Keiran M. Raine, Ivo Buchhalter, Kortine Kleinheinz, Matthias Schlesner, Junjun Zhang, Wenyi Wang, David A. Wheeler, Li Ding, Jared T. Simpson, Brian D. O'Connor, Sergei Yakneen, Kyle Ellrott, Naoki Miyoshi, Adam P. Butler, Romina Royo, Solomon I. Shorser, Miguel Vazquez, Tobias Rausch, Grace Tiao, Sebastian M. Waszak, Bernardo Rodriguez-Martin, Suyash Shringarpure, Dai Ying Wu, German M. Demidov, Olivier Delaneau, Shuto Hayashi, Seiya Imoto, Nina Habermann, Ayellet V. Segre, Erik Garrison, Andy Cafferkey, Eva G. Alvarez, José María Heredia-Genestar, Francesc Muyas, Oliver Drechsel, Alicia L. Bruzos, Javier Temes, Jorge Zamora, Adrian Baez-Ortega, Hyung Lae Kim, R. Jay Mashl, Kai Ye, Anthony DiBiase, Kuan lin Huang, Ivica Letunic, Michael D. McLellan, Steven J. Newhouse, Tal Shmaya, Sushant Kumar, David C. Wedge, Mark H. Wright, Venkata D. Yellapantula, Mark Gerstein, Ekta Khurana, Tomas Marques-Bonet, Arcadi Navarro, Carlos D. Bustamante, Reiner Siebert, Hidewaki Nakagawa, Douglas F. Easton, Stephan Ossowski, Jose M.C. Tubio, Francisco M. De La Vega, Xavier Estivill, Denis Yuen, George L. Mihaiescu, Larsson Omberg, Vincent Ferretti, Radhakrishnan Sabarinathan, Oriol Pich, Abel Gonzalez-Perez, Amaro Taylor-Weiner, Matthew W. Fittall, Jonas Demeulemeester, Maxime Tarabichi, Nicola D. Roberts, Peter Van Loo, Isidro Cortés-Ciriano, Lara Urban, Peter Park, Bin Zhu, Esa Pitkänen, Yilong Li, Natalie Saini, Leszek J. Klimczak, Joachim Weischenfeldt, Nikos Sidiropoulos, Ludmil B. Alexandrov, Raquel Rabionet, Georgia Escaramis, Mattia Bosio, Aliaksei Z. Holik, Hana Susak, Aparna Prasad, Serap Erkek, Claudia Calabrese, Benjamin Raeder, Eoghan Harrington, Simon Mayes, Daniel Turner, Sissel Juul, Steven A. Roberts, Lei Song, Roelof Koster, Lisa Mirabello, Xing Hua, Tomas J. Tanskanen, Marta Tojo, Jieming Chen, Lauri A. Aaltonen, Gunnar Rätsch, Roland F. Schwarz, Atul J. Butte, Alvis Brazma, Stephen J. Chanock, Nilanjan Chatterjee, Oliver Stegle, Olivier Harismendy, G. Steven Bova, Dmitry A. Gordenin, David Haan, Lina Sieverling, Lars Feuerbach, Don Chalmers, Yann Joly, Bartha Knoppers, Fruzsina Molnár-Gábor, Mark Phillips, Adrian Thorogood, David Townend, Mary Goldman, Nuno A. Fonseca, Qian Xiang, Brian Craft, Elena Piñeiro-Yáñez, Alfonso Muñoz, Robert Petryszak, Anja Füllgrabe, Fatima Al-Shahrour, Maria Keays, David Haussler, John Weinstein, Wolfgang Huber, Alfonso Valencia, Irene Papatheodorou, Jingchun Zhu, Yu Fan, David Torrents, Matthias Bieg, Ken Chen, Zechen Chong, Kristian Cibulskis, Roland Eils, Robert S. Fulton, Josep L. Gelpi, Santiago Gonzalez, Ivo G. Gut, Faraz Hach, Michael Heinold, Taobo Hu, Vincent Huang, Barbara Hutter, Natalie Jäger, Jongsun Jung, Yogesh Kumar, Christopher Lalansingh, Ignaty Leshchiner, Dimitri Livitz, Eric Z. Ma, Yosef E. Maruvka, Ana Milovanovic, Morten Muhlig Nielsen, Nagarajan Paramasivam, Jakob Skou Pedersen, Montserrat Puiggròs, S. Cenk Sahinalp, Iman Sarrafi, Chip Stewart, Miranda D. Stobbe, Jeremiah A. Wala, Jiayin Wang, Michael Wendl, Johannes Werner, Zhenggang Wu, Hong Xue, Takafumi N. Yamaguchi, Venkata Yellapantula, Brandi N. Davis-Dusenbery, Robert L. Grossman, Youngwook Kim, Michael C. Heinold, Jonathan Hinton, David R. Jones, Andrew Menzies, Lucy Stebbings, Julian M. Hess, Mara Rosenberg, Andrew J. Dunford, Manaswi Gupta, Marcin Imielinski, Matthew Meyerson, Rameen Beroukhim, Jüri Reimand, Priyanka Dhingra, Francesco Favero, Stefan Dentro, Jeff Wintersinger, Vasilisa Rudneva, Ji Wan Park, Eun Pyo Hong, Seong Gu Heo, André Kahles, Kjong Van Lehmann, Cameron M. Soulette, Yuichi Shiraishi, Fenglin Liu, Yao He, Deniz Demircioğlu, Natalie R. Davidson, Liliana Greger, Siliang Li, Dongbing Liu, Stefan G. Stark, Fan Zhang, Samirkumar B. Amin, Peter Bailey, Aurélien Chateigner, Milana Frenkel-Morgenstern, Yong Hou, Matthew R. Huska, Helena Kilpinen, Fabien C. Lamaze, Chang Li, Xiaobo Li, Xinyue Li, Xingmin Liu, Maximillian G. Marin, Julia Markowski, Tannistha Nandi, Akinyemi I. Ojesina, Qiang Pan-Hammarström, Peter J. Park, Chandra Sekhar Pedamallu, Hong Su, Patrick Tan, Bin Tean Teh, Jian Wang, Heng Xiong, Chen Ye, Christina Yung, Xiuqing Zhang, Liangtao Zheng, Shida Zhu, Philip Awadalla, Chad J. Creighton, Kui Wu, Huanming Yang, Jonathan Göke, Zemin Zhang, Angela N. Brooks, Matthew W. Fittall, Iñigo Martincorena, Carlota Rubio-Perez, Malene Juul, Steven Schumacher, Ofer Shapira, David Tamborero, Loris Mularoni, Henrik Hornshøj, Jordi Deu-Pons, Ferran Muiños, Johanna Bertl, Qianyun Guo, Abel Gonzalez-Perez, Qian Xiang, Wojciech Bazant, Elisabet Barrera, Sultan T. Al-Sedairy, Axel Aretz, Cindy Bell, Miguel Betancourt, Christiane Buchholz, Fabien Calvo, Christine Chomienne, Michael Dunn, Stuart Edmonds, Eric Green, Shailja Gupta, Carolyn M. Hutter, Karine Jegalian, Nic Jones, Youyong Lu, Hitoshi Nakagama, Gerd Nettekoven, Laura Planko, David Scott, Tatsuhiro Shibata, Kiyo Shimizu, Michael R. Stratton, Takashi Yugawa, Giampaolo Tortora, K. VijayRaghavan, Jean C. Zenklusen, David Townend, Bartha M. Knoppers, Brice Aminou, Javier Bartolome, Keith A. Boroevich, Rich Boyce, Alex Buchanan, Niall J. Byrne, Zhaohong Chen, Sunghoon Cho, Wan Choi, Peter Clapham, Michelle T. Dow, Lewis Jonathan Dursi, Juergen Eils, Claudiu Farcas, Nodirjon Fayzullaev, Paul Flicek, Allison P. Heath, Oliver Hofmann, Jongwhi H. Hong, Thomas J. Hudson, Daniel Hübschmann, Sinisa Ivkovic, Seung Hyup Jeon, Wei Jiao, Rolf Kabbe, Andre Kahles, Jules N.A. Kerssemakers, Hyunghwan Kim, Jihoon Kim, Michael Koscher, Antonios Koures, Milena Kovacevic, Chris Lawerenz, Jia Liu, Sanja Mijalkovic, Ana Mijalkovic Mijalkovic-Lazic, Satoru Miyano, Mia Nastic, Jonathan Nicholson, David Ocana, Kazuhiro Ohi, Lucila Ohno-Machado, Todd D. Pihl, Manuel Prinz, Petar Radovic, Charles Short, Heidi J. Sofia, Jonathan Spring, Adam J. Struck, Nebojsa Tijanic, David Vicente, Zhining Wang, Ashley Williams, Youngchoon Woo, Adam J. Wright, Liming Yang, Mark P. Hamilton, Todd A. Johnson, Abdullah Kahraman, Manolis Kellis, Paz Polak, Richard Sallari, Nasa Sinnott-Armstrong, Christian von Mering, Sergi Beltran, Daniela S. Gerhard, Marta Gut, Jean Rémi Trotta, Justin P. Whalley, Beifang Niu, Shadrielle M.G. Espiritu, Shengjie Gao, Yi Huang, Christopher M. Lalansingh, Jon W. Teague, Michael C. Wendl, Federico Abascal, Gary D. Bader, Pratiti Bandopadhayay, Jonathan Barenboim, Søren Brunak, Joana Carlevaro-Fita, Dimple Chakravarty, Calvin Wing Yiu Chan, Jung Kyoon Choi, Klev Diamanti, J. Lynn Fink, Joan Frigola, Carlo Gambacorti-Passerini, Dale W. Garsed, Nicholas J. Haradhvala, Arif O. Harmanci, Mohamed Helmy, Carl Herrmann, Asger Hobolth, Ermin Hodzic, Chen Hong, Keren Isaev, Jose M.G. Izarzugaza, Rory Johnson, Randi Istrup Juul, Jaegil Kim, Jong K. Kim, Jan Komorowski, Andrés Lanzós, Erik Larsson, Donghoon Lee, Shantao Li, Xiaotong Li, Ziao Lin, Eric Minwei Liu, Lucas Lochovsky, Shaoke Lou, Tobias Madsen, Kathleen Marchal, Alexander Martinez-Fundichely, Patrick D. McGillivray, William Meyerson, Marta Paczkowska, Keunchil Park, Kiejung Park, Tirso Pons, Sergio Pulido-Tamayo, Iker Reyes-Salazar, Matthew A. Reyna, Mark A. Rubin, Leonidas Salichos, Chris Sander, Steven E. Schumacher, Mark Shackleton, Ciyue Shen, Raunak Shrestha, Shimin Shuai, Tatsuhiko Tsunoda, Husen M. Umer, Liis Uusküla-Reimand, Lieven P.C. Verbeke, Claes Wadelius, Lina Wadi, Jonathan Warrell, Guanming Wu, Jun Yu, Jing Zhang, Xuanping Zhang, Yan Zhang, Zhongming Zhao, Lihua Zou, Michael S. Lawrence, Benjamin J. Raphael, Peter J. Bailey, David Craft, Mary J. Goldman, Hiroyuki Aburatani, Hans Binder, Huy Q. Dinh, Simon C. Heath, Steve Hoffmann, Charles David Imbusch, Helene Kretzmer, Peter W. Laird, Jose I. Martin-Subero, Genta Nagae, Hui Shen, Qi Wang, Dieter Weichenhan, Wanding Zhou, Benjamin P. Berman, Benedikt Brors, Christoph Plass, Kadir C. Akdemir, David D.L. Bowtell, Kathleen H. Burns, John Busanovich, Kin Chan, Ana Dueso-Barroso, Paul A. Edwards, Dariush Etemadmoghadam, James E. Haber, David T.W. Jones, Young Seok Ju, Marat D. Kazanov, Youngil Koh, Kiran Kumar, Eunjung Alice Lee, Jake June Koo Lee, Andy G. Lynch, Geoff Macintyre, Florian Markowetz, Fabio C.P. Navarro, John V. Pearson, Karsten Rippe, Ralph Scully, Izar Villasante, Nicola Waddell, Lixing Yang, Xiaotong Yao, Sung Soo Yoon, Cheng Zhong Zhang, Erik N. Bergstrom, Arnoud Boot, Kyle Covington, Akihiro Fujimoto, Mi Ni Huang, S. M.Ashiqul Islam, John R. McPherson, Sandro Morganella, Ville Mustonen, Alvin Wei Tian Ng, Stephenie D. Prokopec, Ignacio Vázquez-García, Yang Wu, Fouad Yousif, Willie Yu, Steven G. Rozen, Vasilisa A. Rudneva, Suyash S. Shringarpure, Daniel J. Turner, Tian Xia, Gurnit Atwal, David K. Chang, Susanna L. Cooke, Bishoy M. Faltas, Syed Haider, Vera B. Kaiser, Rosa Karlić, Mamoru Kato, Kirsten Kübler, Adam Margolin, Sancha Martin, Serena Nik-Zainal, Christine P'ng, Colin A. Semple, Jaclyn Smith, Ren X. Sun, Kevin Thai, Derek W. Wright, Ke Yuan, Andrew V. Biankin, Levi Garraway, Sean M. Grimmond, David J. Adams, Pavana Anur, Shaolong Cao, Elizabeth L. Christie, Marek Cmero, Yupeng Cun, Kevin J. Dawson, Stefan C. Dentro, Amit G. Deshwar, Nilgun Donmez, Ruben M. Drews, Moritz Gerstung, Gavin Ha, Kerstin Haase, Lara Jerman, Yuan Ji, Clemency Jolly, Juhee Lee, Henry Lee-Six, Salem Malikic, Thomas J. Mitchell, Quaid D. Morris, Layla Oesper, Martin Peifer, Myron Peto, Daniel Rosebrock, Yulia Rubanova, Adriana Salcedo, Subhajit Sengupta, Ruian Shi, Seung Jun Shin, Oliver Spiro, Shankar Vembu, Jeffrey A. Wintersinger, Tsun Po Yang, Kaixian Yu, Hongtu Zhu, Paul T. Spellman, John N. Weinstein, Yiwen Chen, Masashi Fujita, Leng Han, Takanori Hasegawa, Mitsuhiro Komura, Jun Li, Shinichi Mizuno, Eigo Shimizu, Yumeng Wang, Yanxun Xu, Rui Yamaguchi, Fan Yang, Yang Yang, Christopher J. Yoon, Yuan Yuan, Han Liang, Malik Alawi, Ivan Borozan, Daniel S. Brewer, Colin S. Cooper, Nikita Desai, Adam Grundhoff, Murat Iskar, Xiaoping Su, Marc Zapatka, Peter Lichter, Kathryn Alsop, Timothy J.C. Bruxner, Angelika N. Christ, Stephen M. Cordner, Prue A. Cowin, Ronny Drapkin, Sian Fereday, Joshy George, Anne Hamilton, Oliver Holmes, Jillian A. Hung, Karin S. Kassahn, Stephen H. Kazakoff, Catherine J. Kennedy, Conrad R. Leonard, Linda Mileshkin, David K. Miller, Gisela Mir Arnau, Chris Mitchell, Felicity Newell, Katia Nones, Ann Marie Patch, Michael C. Quinn, Darrin F. Taylor, Heather Thorne, Nadia Traficante, Ravikiran Vedururu, Nick M. Waddell, Paul M. Waring, Scott Wood, Qinying Xu, Anna deFazio, Matthew J. Anderson, Davide Antonello, Andrew P. Barbour, Claudio Bassi, Samantha Bersani, Ivana Cataldo, Lorraine A. Chantrill, Yoke Eng Chiew, Angela Chou, Sara Cingarlini, Nicole Cloonan, Vincenzo Corbo, Maria Vittoria Davi, Fraser R. Duthie, Anthony J. Gill, Janet S. Graham, Ivon Harliwong, Nigel B. Jamieson, Amber L. Johns, James G. Kench, Luca Landoni, Rita T. Lawlor, Andrea Mafficini, Neil D. Merrett, Marco Miotto, Elizabeth A. Musgrove, Adnan M. Nagrial, Karin A. Oien, Marina Pajic, Mark Pinese, Alan J. Robertson, Ilse Rooman, Borislav C. Rusev, Jaswinder S. Samra, Maria Scardoni, Christopher J. Scarlett, Aldo Scarpa, Elisabetta Sereni, Katarzyna O. Sikora, Michele Simbolo, Morgan L. Taschuk, Christopher W. Toon, Caterina Vicentini, Jianmin Wu, Nikolajs Zeps, Andreas Behren, Hazel Burke, Jonathan Cebon, Rebecca A. Dagg, Ricardo De Paoli-Iseppi, Ken Dutton-Regester, Matthew A. Field, Anna Fitzgerald, Peter Hersey, Valerie Jakrot, Peter A. Johansson, Hojabr Kakavand, Richard F. Kefford, Loretta M.S. Lau, Georgina V. Long, Hilda A. Pickett, Antonia L. Pritchard, Gulietta M. Pupo, Robyn P.M. Saw, Sarah Jane Schramm, Catherine A. Shang, Ping Shang, Andrew J. Spillane, Jonathan R. Stretch, Varsha Tembe, John F. Thompson, Ricardo E. Vilain, James S. Wilmott, Jean Y. Yang, Nicholas K. Hayward, Graham J. Mann, Richard A. Scolyer, John Bartlett, Prashant Bavi, Dianne E. Chadwick, Michelle Chan-Seng-Yue, Sean Cleary, Ashton A. Connor, Karolina Czajka, Robert E. Denroche, Neesha C. Dhani, Jenna Eagles, Steven Gallinger, Robert C. Grant, David Hedley, Michael A. Hollingsworth, Gun Ho Jang, Jeremy Johns, Sangeetha Kalimuthu, Sheng Ben Liang, Ilinca Lungu, Xuemei Luo, Faridah Mbabaali, Treasa A. McPherson, Jessica K. Miller, Malcolm J. Moore, Faiyaz Notta, Danielle Pasternack, Gloria M. Petersen, Michael H.A. Roehrl, Michelle Sam, Iris Selander, Stefano Serra, Sagedeh Shahabi, Sarah P. Thayer, Lee E. Timms, Gavin W. Wilson, Julie M. Wilson, Bradly G. Wouters, John D. McPherson, Timothy A. Beck, Vinayak Bhandari, Colin C. Collins, Neil E. Fleshner, Natalie S. Fox, Michael Fraser, Lawrence E. Heisler, Emilie Lalonde, Julie Livingstone, Alice Meng, Veronica Y. Sabelnykova, Yu Jia Shiah, Theodorus Van der Kwast, Robert G. Bristow, Shuai Ding, Daiming Fan, Lin Li, Yongzhan Nie, Xiao Xiao, Rui Xing, Shanlin Yang, Yingyan Yu, Yong Zhou, Rosamonde E. Banks, Guillaume Bourque, Paul Brennan, Louis Letourneau, Yasser Riazalhosseini, Ghislaine Scelo, Naveen Vasudev, Juris Viksna, Mark Lathrop, Jörg Tost, Sung Min Ahn, Samuel Aparicio, Laurent Arnould, M. R. Aure, Shriram G. Bhosle, Ewan Birney, Ake Borg, Sandrine Boyault, Arie B. Brinkman, Jane E. Brock, Annegien Broeks, Anne Lise Børresen-Dale, Carlos Caldas, Suet Feung Chin, Helen Davies, Christine Desmedt, Luc Dirix, Serge Dronov, Anna Ehinger, Jorunn E. Eyfjord, Aquila Fatima, John A. Foekens, P. Andrew Futreal, Øystein Garred, Dilip D. Giri, Dominik Glodzik, Dorthe Grabau, Holmfridur Hilmarsdottir, Gerrit K. Hooijer, Jocelyne Jacquemier, Se Jin Jang, Jon G. Jonasson, Jos Jonkers, Hyung Yong Kim, Tari A. King, Stian Knappskog, Gu Kong, Savitri Krishnamurthy, Sunil R. Lakhani, Anita Langerød, Denis Larsimont, Hee Jin Lee, Jeong Yeon Lee, Ming Ta Michael Lee, Ole Christian Lingjærde, Gaetan MacGrogan, John W.M. Martens, Sarah O'Meara, Iris Pauporté, Sarah Pinder, Xavier Pivot, Elena Provenzano, Colin A. Purdie, Manasa Ramakrishna, Kamna Ramakrishnan, Jorge Reis-Filho, Andrea L.

Richardson, Markus Ringnér, Javier Bartolomé Rodriguez, F. Germán Rodríguez-González, Gilles Romieu, Roberto Salgado, Torill Sauer, Rebecca Shepherd, Anieta M. Sieuwerts, Peter T. Simpson, Marcel Smid, Christos Sotiriou, Paul N. Span, Ólafur Andri Stefánsson, Alasdair Stenhouse, Henk G. Stunnenberg, Fred Sweep, Benita Kiat Tee Tan, Gilles Thomas, Alastair M. Thompson, Stefania Tommasi, Isabelle Treilleux, Andrew Tutt, Naoto T. Ueno, Steven Van Laere, Gert G. Van den Eynden, Peter Vermeulen, Alain Viari, Anne Vincent-Salomon, Bernice H. Wong, Lucy Yates, Xueqing Zou, Carolien H.M. van Deurzen, Marc J. van de Vijver, Laura van't Veer, Ole Ammerpohl, Sietse Aukema, Anke K. Bergmann, Stephan H. Bernhart, Arndt Borkhardt, Christoph Borst, Birgit Burkhardt, Alexander Claviez, Maria Elisabeth Goebler, Andrea Haake, Siegfried Haas, Martin Hansmann, Jessica I. Hoell, Michael Hummel, Dennis Karsch, Wolfram Klapper, Michael Kneba, Markus Kreuz, Dieter Kube, Ralf Küppers, Dido Lenze, Markus Loeffler, Cristina López, Luisa Mantovani-Löffler, Peter Möller, German Ott, Bernhard Radlwimmer, Julia Richter, Marius Rohde, Philip C. Rosenstiel, Andreas Rosenwald, Markus B. Schilhabel, Stefan Schreiber, Peter F. Stadler, Peter Staib, Stephan Stilgenbauer, Stephanie Sungalee, Monika Szczepanowski, Umut H. Toprak, Lorenz H.P. Trümper, Rabea Wagener, Thorsten Zenz, Volker Hovestadt, Christof von Kalle, Marcel Kool, Andrey Korshunov, Pablo Landgraf, Hans Lehrach, Paul A. Northcott, Stefan M. Pfister, Guido Reifenberger, Hans Jörg Warnatz, Stephan Wolf, Marie Laure Yaspo, Yassen Assenov, Clarissa Gerhauser, Sarah Minner, Thorsten Schlomm, Ronald Simon, Guido Sauter, Holger Sültmann, Nidhan K. Biswas, Arindam Maitra, Partha P. Majumder, Rajiv Sarin, Stefano Barbi, Giada Bonizzato, Cinzia Cantù, Angelo P. Dei Tos, Matteo Fassan, Sonia Grimaldi, Claudio Luchini, Giuseppe Malleo, Giovanni Marchegiani, Michele Milella, Salvatore Paiella, Antonio Pea, Paolo Pederzoli, Andrea Ruzzenente, Roberto Salvia, Nicola Sperandio, Yasuhito Arai, Natsuko Hama, Nobuyoshi Hiraoka, Fumie Hosoda, Hiromi Nakamura, Hidenori Ojima, Takuji Okusaka, Yasushi Totoki, Tomoko Urushidate, Masashi Fukayama, Shumpei Ishikawa, Hitoshi Katai, Hiroto Katoh, Daisuke Komura, Hirofumi Rokutan, Mihoko Saito-Adachi, Akihiro Suzuki, Hirokazu Taniguchi, Kenji Tatsuno, Tetsuo Ushiku, Shinichi Yachida, Shogo Yamamoto, Hiroshi Aikata, Koji Arihiro, Shun ichi Ariizumi, Kazuaki Chayama, Mayuko Furuta, Kunihito Gotoh, Shinya Hayami, Satoshi Hirano, Yoshiiku Kawakami, Kazuhiro Maejima, Toru Nakamura, Kaoru Nakano, Hideki Ohdan, Aya Sasaki-Oku, Hiroko Tanaka, Masaki Ueno, Masakazu Yamamoto, Hiroki Yamaue, Su Pin Choo, Ioana Cutcutache, Narong Khuntikeo, Choon Kiat Ong, Chawalit Pairojkul, Irinel Popescu, Keun Soo Ahn, Marta Aymerich, Armando Lopez-Guillermo, Carlos López-Otín, Xose S. Puente, Elias Campo, Fernanda Amary, Daniel Baumhoer, Sam Behjati, Bodil Bjerkehagen, P. A. Futreal, Ola Myklebost, Nischalan Pillay, Patrick Tarpey, Roberto Tirabosco, Olga Zaikova, Adrienne M. Flanagan, Jacqueline Boultwood, David T. Bowen, Mario Cazzola, Anthony R. Green, Eva Hellstrom-Lindberg, Luca Malcovati, Jyoti Nangalia, Elli Papaemmanuil, Paresh Vyas, Yeng Ang, Hugh Barr, Duncan Beardsmore, Matthew Eldridge, James Gossage, Nicola Grehan, George B. Hanna, Stephen J. Hayes, Ted R. Hupp, David Khoo, Jesper Lagergren, Laurence B. Lovat, Shona MacRae, Maria O'Donovan, J. Robert O'Neill, Simon L. Parsons, Shaun R. Preston, Sonia Puig, Tom Roques, Grant Sanders, Sharmila Sothi, Simon Tavaré, Olga Tucker, Richard Turkington, Timothy J. Underwood, Ian Welch, Rebecca C. Fitzgerald, Daniel M. Berney, Johann S. De Bono, Declan Cahill, Niedzica Camacho, Nening M. Dennis, Tim Dudderidge, Sandra E. Edwards, Cyril Fisher, Christopher S. Foster, Mohammed Ghori, Pelvender Gill, Vincent J. Gnanapragasam, Gunes Gundem, Freddie C. Hamdy, Steve Hawkins, Steven Hazell, William Howat, William B. Isaacs, Katalin Karaszi, Jonathan D. Kay, Vincent Khoo, Zsofia Kote-Jarai, Barbara Kremeyer, Pardeep Kumar, Adam Lambert, Daniel A. Leongamornlert, Naomi Livni, Yong Jie Lu, Hayley J. Luxton, Luke Marsden, Charlie E. Massie, Lucy Matthews, Erik Mayer, Ultan McDermott, Sue Merson, David E. Neal, Anthony Ng, David Nicol, Christopher Ogden, Edward W. Rowe, Nimish C. Shah, Sarah Thomas, Alan Thompson, Clare Verrill, Tapio Visakorpi, Anne Y. Warren, Hayley C. Whitaker, Hongwei Zhang, Nicholas van As, Rosalind A. Eeles, Adam Abeshouse, Nishant Agrawal, Rehan Akbani, Hikmat Al-Ahmadie, Monique Albert, Kenneth Aldape, Adrian Ally, Elizabeth L. Appelbaum, Joshua Armenia, Sylvia Asa, J. Todd Auman, Miruna Balasundaram, Saianand Balu, Jill Barnholtz-Sloan, Oliver F. Bathe, Stephen B. Baylin, Christopher Benz, Andrew Berchuck, Mario Berrios, Darell Bigner, Michael Birrer, Tom Bodenheimer, Lori Boice, Moiz S. Bootwalla, Marcus Bosenberg, Reanne Bowlby, Jeffrey Boyd, Russell R. Broaddus, Malcolm Brock, Denise Brooks, Susan Bullman, Samantha J. Caesar-Johnson, Thomas E. Carey, Rebecca Carlsen, Robert Cerfolio, Vishal S. Chandan, Hsiao Wei Chen, Andrew D. Cherniack, Jeremy Chien, Juok Cho, Eric Chuah, Carrie Cibulskis, Leslie Cope, Matthew G. Cordes, Erin Curley, Bogdan Czerniak, Ludmila Danilova, Ian J. Davis, Timothy Defreitas, John A. Demchok, Noreen Dhalla, Rajiv Dhir, Harsha Vardh. Pan-cancer analysis of whole genomes. *Nature*, 2020. ISSN 14764687. doi: 10.1038/s41586-020-1969-6.

20. Janessa Laskin, Steven Jones, Samuel Aparicio, Stephen Chia, Carolyn Ch'ng, Rebecca Deyell, Peter Eirew, Alexandra Fok, Karen Gelmon, Cheryl Ho, David Huntsman, Martin Jones, Katayoon Kasaian, Aly Karsan, Sreeja Leelakumari, Yvonne Li, Howard Lim, Yussanne Ma, Colin Mar, Monty Martin, Richard Moore, Andrew Mungall, Karen Mungall, Erin Pleasance, S. Rod Rassekh, Daniel Renouf, Yaoqing Shen, Jacqueline Schein, Kasmintan Schrader, Sophie Sun, Anna Tinker, Eric Zhao, Stephen Yip, and Marco A. Marra. Lessons learned from the application of whole-genome analysis to the treatment of patients with advanced cancers. *Molecular Case Studies*, 2015. ISSN 2373-2865. doi: 10.1101/mcs.a000570.
21. Petr Danecek, James K. Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O. Pollard, Andrew Whitwham, Thomas Keane, Shane A. McCarthy, Robert M. Davies, and Heng Li. Twelve years of SAMtools and BCFtools. *GigaScience*, 2021. ISSN 2047217X. doi: 10.1093/gigascience/giab008.
22. Aaron R. Quinlan and Ira M. Hall. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, 2010. ISSN 13674803. doi: 10.1093/bioinformatics/btq033.
23. Daniel L Cameron, Jonathan Baber, Charles Shale, Anthony T Papenfuss, Jose Espejo Valle-Inclan, Nicolle Besselink, Edwin Cuppen, and Peter Priestley. GRIDSS, PURPLE, LINX: Unscrambling the tumor genome via integrated analysis of structural variation and copy number. *bioRxiv*, page 781013, 1 2019. doi: 10.1101/781013.
24. Johannes Köster and Sven Rahmann. Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics*, 2012. ISSN 14602059. doi: 10.1093/bioinformatics/bts480.
25. Yoav Benjamini and Yosef Hochberg. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995. ISSN 00359246. doi: 10.2307/2346101.
26. Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011. ISSN 15324435.
27. Gavin Ha, Andrew Roth, Daniel Lai, Ali Bashashati, Jiarui Ding, Rodrigo Goya, Ryan Giuliany, Jamie Rosner, Arusha Oloumi, Karey Shumansky, Suet Feung Chin, Gulisa Turashvili, Martin Hirst, Carlos Caldas, Marco A. Marra, Samuel Aparicio, and Sohrab P. Shah. Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. *Genome Research*, 2012. ISSN 10889051. doi: 10.1101/gr.137570.112.
28. Erin D Pleasance, R Keira Cheetham, Philip J Stephens, David J McBride, Sean J Humphray, Chris D Greenman, Ignacio Varela, Meng-Lay Lin, Gonzalo R Ordóñez, Graham R Bignell, Kai Ye, Julie Alipaz, Markus J Bauer, David Beare, Adam Butler, Richard J Carter, Lina Chen, Anthony J Cox, Sarah Edkins, Paula I Kokko-Gonzales, Niall A Gormley, Russell J Grocock, Christian D Haudenschild, Matthew M Hims, Terena James, Mingming Jia, Zoya Kingsbury, Catherine Leroy, John Marshall, Andrew Menzies, Laura J Mudie, Zemin Ning, Tom Royce, Ole B Schulz-Trieglaff, Anastassia Spiridou, Lucy A Stebbings, Lukasz Szajkowski, Jon Teague, David Williamson, Lynda Chin, Mark T Ross, Peter J Campbell, David R Bentley, P Andrew Futreal, and Michael R Stratton. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature*, 463:191–196, 2009.
29. Chiao En Wu, Da Wei Yeh, Yi Ru Pan, Wen Kuan Huang, Ming Huang Chen, John Wen Cheng Chang, Jen Shi Chen, Yu Chao Wang, and Chun Nan Yeh. Chromosomal instability may not be a predictor for immune checkpoint inhibitors from a comprehensive bioinformatics analysis. *Life*, 2020. ISSN 20751729. doi: 10.3390/life10110276.
30. G. M. Blumenthal, Y. Gong, K. Kehl, P. Mishra-Kalyani, K. B. Goldberg, S. Khozin, P. G. Kluetz, G. R. Oxnard, and R. Pazdur. Analysis of time-To-Treatment discontinuation of targeted therapy, immunotherapy, and chemotherapy in clinical trials of patients with non-small-cell lung cancer. *Annals of Oncology*, 2019. ISSN 15698041. doi: 10.1093/annonc/mdz060.
31. Eric Y. Zhao, Yaoqing Shen, Erin Pleasance, Katayoon Kasaian, Sreeja Leelakumari, Martin Jones, Pinaki Bose, Carolyn Ch'ng, Caralyn Reisle, Peter Eirew, Richard Corbett, Karen L. Mungall, Nina Thiessen, Yussanne Ma, Jacqueline E. Schein, Andrew J. Mungall, Yongjun Zhao, Richard A. Moore, Wendie Den Brok, Sheridan Wilson, Diego Villa, Tamara Shenkier, Caroline Lohrisch, Stephen Chia, Stephen Yip, Karen Gelmon, Howard Lim, Daniel Renouf, Sophie Sun, Kasmintan A. Schrader, Sean Young, Ian Bosdet, Aly Karsan, Janessa Laskin, Marco A. Marra, and Steven J.M. Jones. Homologous recombination deficiency and platinum-based therapy outcomes in advanced breast cancer. *Clinical Cancer Research*, 2017. ISSN 15573265. doi: 10.1158/1078-0432.CCR-17-1941.
32. Samuel F. Bakhoum and Lewis C. Cantley. The Multifaceted Role of Chromosomal Instability in Cancer and Its Microenvironment, 2018. ISSN 10974172.
33. Aaron M. Newman, Chih Long Liu, Michael R. Green, Andrew J. Gentles, Weiguo Feng, Yue Xu, Chuong D. Hoang, Maximilian Diehn, and Ash A. Alizadeh. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods*, 2015. ISSN 15487105. doi: 10.1038/nmeth.3337.
34. Ann Kathrin Schmidt, Karoline Pudelko, Jan Eric Boekenkamp, Katharina Berger, Maik Kschischo, and Holger Bastians. The p53/p73 - p21CIP1 tumor suppressor axis guards against chromosomal instability by restraining CDK1 in human cancer cells. *Oncogene*, 2021. ISSN 14765594. doi: 10.1038/s41388-020-01524-4.
35. Rene Dreos, Giovanna Ambrosini, Romain Groux, Rouaida Cavin Perier, and Philipp Bucher. The eukaryotic promoter database in its 30th year: Focus on non-vertebrate organisms. *Nucleic Acids Research*, 2017. ISSN 13624962. doi: 10.1093/nar/gkw1069.
36. Anuratha Sakthianandeswaren, Marie J. Parsons, Dmitri Mouradov, Ruth N. Mackinnon, Bruno Catimel, Sheng Liu, Michelle Palmieri, Christopher Love, Robert N. Jorissen, Shan Li, Lachlan Whitehead, Tracy L. Putoczki, Adele Preaudet, Cary Tsui, Cameron J. Nowell, Robyn L. Ward, Nicholas J. Hawkins, Jayesh Desai, Peter Gibbs, Matthias Ernst, Ian Street, Michael Buchert, and Oliver M. Sieber. MACROD2 haploinsufficiency impairs catalytic activity of PARP1 and promotes chromosome instability and growth of intestinal tumors. *Cancer Discovery*, 2018. ISSN 21598290. doi: 10.1158/2159-8290.CD-17-0909.
37. Shunichi Kosugi, Yukihide Momozawa, Xiaoxi Liu, Chikashi Terao, Michiaki Kubo, and Yoichiro Kamatani. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. *Genome Biology*, 2019. ISSN 1474760X. doi: 10.1186/s13059-019-1720-5.
38. Andrea M Marquard, Aron C Eklund, Tejal Joshi, Marcin Krzystanek, Francesco Favero, Zhigang C Wang, Andrea L Richardson, Daniel P Silver, Zoltan Szallasi, and Nicolai J Birkbak. Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. *Biomarker Research*, 2015. ISSN 2050-7771. doi: 10.1186/s40364-015-0033-4.
39. Duc Hiep Bach, Wei Zhang, and Anil K. Sood. Chromosomal instability in tumor initiation and development, 2019. ISSN 15387445.
40. Ludmil B. Alexandrov, Serena Nik-Zainal, David C. Wedge, Samuel A.J.R. Aparicio, Sam Behjati, Andrew V. Biankin, Graham R. Bignell, Niccolò Bolli, Ake Borg, Anne Lise Børresen-Dale, Sandrine Boyault, Birgit Burkhardt, Adam P. Butler, Carlos Caldas, Helen R. Davies, Christine Desmedt, Roland Eils, Jórunn Erla Eyfjörd, John A. Foekens, Mel Greaves, Fumie Hosoda, Barbara Hutter, Tomislav Ilicic, Sandrine Imbeaud, Marcin Imielinsk, Natalie Jäger, David T.W. Jones, David Jonas, Stian Knappskog, Marcel Koo, Sunil R. Lakhani, Carlos López-Otín, Sancha Martin, Nikhil C. Munshi, Hiromi Nakamura, Paul A. Northcott, Marina Pajic, Elli Papaemmanuil, Angelo Paradiso, John V. Pearson, Xose S. Puente, Keiran Raine, Manasa Ramakrishna, Andrea L. Richardson, Julia Richter, Philip Rosenstiel, Matthias Schlesner, Ton N. Schumacher, Paul N. Span, Jon W. Teague, Yasushi Totoki, Andrew N.J. Tutt, Rafael Valdés-Mas, Marit M. Van Buuren, Laura Van 'T Veer, Anne Vincent-Salomon, Nicola Waddell, Lucy R. Yates, Jessica Zucman-Rossi, P. Andrew Futreal, Ultan McDermott, Peter Lichter, Matthew Meyerson, Sean M. Grimmond, Reiner Siebert, Elías

Campo, Tatsuhiro Shibata, Stefan M. Pfister, Peter J. Campbell, and Michael R. Stratton. Signatures of mutational processes in human cancer. *Nature*, 2013. ISSN 14764687. doi: 10.1038/nature12477.

41. Matthew H. Bailey, Collin Tokheim, Eduard Porta-Pardo, Sohini Sengupta, Denis Bertrand, Amila Weerasinghe, Antonio Colaprico, Michael C. Wendl, Jaegil Kim, Brendan Reardon, Patrick Kwok Shing Ng, Kang Jin Jeong, Song Cao, Zhining Wang, Jianjiong Gao, Qingsong Gao, Fang Wang, Eric Minwei Liu, Loris Mularoni, Carlota Rubio-Perez, Niranjan Nagarajan, Isidro Cortés-Ciriano, Daniel Cui Zhou, Wen Wei Liang, Julian M. Hess, Venkata D. Yellapantula, David Tamborero, Abel Gonzalez-Perez, Chayaporn Suphavilai, Jia Yu Ko, Ekta Khurana, Peter J. Park, Eliezer M. Van Allen, Han Liang, Samantha J. Caesar-Johnson, John A. Demchok, Ina Felau, Melpomeni Kasapi, Martin L. Ferguson, Carolyn M. Hutter, Heidi J. Sofia, Roy Tarnuzzer, Liming Yang, Jean C. Zenklusen, Jiashan (Julia) Zhang, Sudha Chudamani, Jia Liu, Laxmi Lolla, Rashi Naresh, Todd Pihl, Qiang Sun, Yunhu Wan, Ye Wu, Juok Cho, Timothy DeFreitas, Scott Frazer, Nils Gehlenborg, Gad Getz, David I. Heiman, Michael S. Lawrence, Pei Lin, Sam Meier, Michael S. Noble, Gordon Saksena, Doug Voet, Hailei Zhang, Brady Bernard, Nyasha Chambwe, Varsha Dhankani, Theo Knijnenburg, Roger Kramer, Kalle Leinonen, Yuexin Liu, Michael Miller, Sheila Reynolds, Ilya Shmulevich, Vesteinn Thorsson, Wei Zhang, Rehan Akbani, Bradley M. Broom, Apurva M. Hegde, Zhenlin Ju, Rupa S. Kanchi, Anil Korkut, Jun Li, Shiyun Ling, Wenbin Liu, Yiling Lu, Gordon B. Mills, Kwok Shing Ng, Arvind Rao, Michael Ryan, Jing Wang, John N. Weinstein, Jiexin Zhang, Adam Abeshouse, Joshua Armenia, Debyani Chakravarty, Walid K. Chatila, Ino de Bruijn, Benjamin E. Gross, Zachary J. Heins, Ritika Kundra, Konnor La, Marc Ladanyi, Augustin Luna, Moriah G. Nissan, Angelica Ochoa, Sarah M. Phillips, Ed Reznik, Francisco Sanchez-Vega, Chris Sander, Nikolaus Schultz, Robert Sheridan, S. Onur Sumer, Yichao Sun, Barry S. Taylor, Pavana Anur, Myron Peto, Paul Spellman, Christopher Benz, Joshua M. Stuart, Christopher K. Wong, Christina Yau, D. Neil Hayes, Joel S. Parker, Matthew D. Wilkerson, Adrian Ally, Miruna Balasundaram, Reanne Bowlby, Denise Brooks, Rebecca Carlsen, Eric Chuah, Noreen Dhalla, Robert Holt, Steven J.M. Jones, Katayoon Kasaian, Darlene Lee, Yussanne Ma, Marco A. Marra, Michael Mayo, Richard A. Moore, Andrew J. Mungall, Karen Mungall, A. Gordon Robertson, Sara Sadeghi, Jacqueline E. Schein, Payal Sipahimalani, Angela Tam, Nina Thiessen, Kane Tse, Tina Wong, Ashton C. Berger, Rameen Beroukhim, Andrew D. Cherniack, Carrie Cibulskis, Stacey B. Gabriel, Galen F. Gao, Gavin Ha, Matthew Meyerson, Steven E. Schumacher, Juliann Shih, Melanie H. Kucherlapati, Raju S. Kucherlapati, Stephen Baylin, Leslie Cope, Ludmila Danilova, Moiz S. Bootwalla, Phillip H. Lai, Dennis T. Maglinte, David J. Van Den Berg, Daniel J. Weisenberger, J. Todd Auman, Saianand Balu, Tom Bodenheimer, Cheng Fan, Katherine A. Hoadley, Alan P. Hoyle, Stuart R. Jefferys, Corbin D. Jones, Shaowu Meng, Piotr A. Mieczkowski, Lisle E. Mose, Amy H. Perou, Charles M. Perou, Jeffrey Roach, Yan Shi, Janae V. Simons, Tara Skelly, Matthew G. Soloway, Donghui Tan, Umadevi Veluvolu, Huihui Fan, Toshinori Hinoue, Peter W. Laird, Hui Shen, Wanding Zhou, Michelle Bellair, Kyle Chang, Kyle Covington, Chad J. Creighton, Huyen Dinh, Harsha Vardhan Doddapaneni, Lawrence A. Donehower, Jennifer Drummond, Richard A. Gibbs, Robert Glenn, Walker Hale, Yi Han, Jianhong Hu, Viktoriya Korchina, Sandra Lee, Lora Lewis, Wei Li, Xiuping Liu, Margaret Morgan, Donna Morton, Donna Muzny, Jireh Santibanez, Margi Sheth, Eve Shinbrot, Linghua Wang, Min Wang, David A. Wheeler, Liu Xi, Fengmei Zhao, Julian Hess, Elizabeth L. Appelbaum, Matthew Bailey, Matthew G. Cordes, Li Ding, Catrina C. Fronick, Lucinda A. Fulton, Robert S. Fulton, Cyriac Kandoth, Elaine R. Mardis, Michael D. McLellan, Christopher A. Miller, Heather K. Schmidt, Richard K. Wilson, Daniel Crain, Erin Curley, Johanna Gardner, Kevin Lau, David Mallery, Scott Morris, Joseph Paulauskis, Robert Penny, Candace Shelton, Troy Shelton, Mark Sherman, Eric Thompson, Peggy Yena, Jay Bowen, Julie M. Gastier-Foster, Mark Gerken, Kristen M. Leraas, Tara M. Lichtenberg, Nilsa C. Ramirez, Lisa Wise, Erik Zmuda, Niall Corcoran, Tony Costello, Christopher Hovens, Andre L. Carvalho, Ana C. de Carvalho, José H. Fregnani, Adhemar Longatto-Filho, Rui M. Reis, Cristovam Scapulatempo-Neto, Henrique C.S. Silveira, Daniel O. Vidal, Andrew Burnette, Jennifer Eschbacher, Beth Hermes, Ardene Noss, Rosy Singh, Matthew L. Anderson, Patricia D. Castro, Michael Ittmann, David Huntsman, Bernard Kohl, Xuan Le, Richard Thorp, Chris Andry, Elizabeth R. Duffy, Vladimir Lyadov, Oxana Paklina, Galiya Setdikova, Alexey Shabunin, Mikhail Tavobilov, Christopher McPherson, Ronald Warnick, Ross Berkowitz, Daniel Cramer, Colleen Feltmate, Neil Horowitz, Adam Kibel, Michael Muto, Chandrajit P. Raut, Andrei Malykh, Jill S. Barnholtz-Sloan, Wendi Barrett, Karen Devine, Jordonna Fulop, Quinn T. Ostrom, Kristen Shimmel, Yingli Wolinsky, Andrew E. Sloan, Agostino De Rose, Felice Giuliante, Marc Goodman, Beth Y. Karlan, Curt H. Hagedorn, John Eckman, Jodi Harr, Jerome Myers, Kelinda Tucker, Leigh Anne Zach, Brenda Deyarmin, Hai Hu, Leonid Kvecher, Caroline Larson, Richard J. Mural, Stella Somiari, Ales Vicha, Tomas Zelinka, Joseph Bennett, Mary Iacocca, Brenda Rabeno, Patricia Swanson, Mathieu Latour, Louis Lacombe, Bernard Têtu, Alain Bergeron, Mary McGraw, Susan M. Staugaitis, John Chabot, Hanina Hibshoosh, Antonia Sepulveda, Tao Su, Timothy Wang, Olga Potapova, Olga Voronina, Laurence Desjardins, Odette Mariani, Sergio Roman-Roman, Xavier Sastre, Marc Henri Stern, Feixiong Cheng, Sabina Signoretti, Andrew Berchuck, Darell Bigner, Eric Lipp, Jeffrey Marks, Shannon McCall, Roger McLendon, Angeles Secord, Alexis Sharp, Madhusmita Behera, Daniel J. Brat, Amy Chen, Keith Delman, Seth Force, Fadlo Khuri, Kelly Magliocca, Shishir Maithel, Jeffrey J. Olson, Taofeek Owonikoko, Alan Pickens, Suresh Ramalingam, Dong M. Shin, Gabriel Sica, Erwin G. Van Meir, Wil Eijckenboom, Ad Gillis, Esther Korpershoek, Leendert Looijenga, Wolter Oosterhuis, Hans Stoop, Kim E. van Kessel, Ellen C. Zwarthoff, Chiara Calatozzolo, Lucia Cuppini, Stefania Cuzzubbo, Francesco DiMeco, Gaetano Finocchiaro, Luca Mattei, Alessandro Perin, Bianca Pollo, Chu Chen, John Houck, Pawadee Lohavanichbutr, Arndt Hartmann, Christine Stoehr, Robert Stoehr, Helge Taubert, Sven Wach, Bernd Wullich, Witold Kycler, Dawid Murawa, Maciej Wiznerowicz, Ki Chung, W. Jeffrey Edenfield, Julie Martin, Eric Baudin, Glenn Bubley, Raphael Bueno, Assunta De Rienzo, William G. Richards, Steven Kalkanis, Tom Mikkelsen, Houtan Noushmehr, Lisa Scarpace, Nicolas Girard, Marta Aymerich, Elias Campo, Eva Giné, Armando López Guillermo, Nguyen Van Bang, Phan Thi Hanh, Bui Duc Phu, Yufang Tang, Howard Colman, Kimberley Evason, Peter R. Dottino, John A. Martignetti, Hani Gabra, Hartmut Juhl, Teniola Akeredolu, Serghei Stepa, Dave Hoon, Keunsoo Ahn, Koo Jeong Kang, Felix Beuschlein, Anne Breggia, Michael Birrer, Debra Bell, Mitesh Borad, Alan H. Bryce, Erik Castle, Vishal Chandan, John Cheville, John A. Copland, Michael Farnell, Thomas Flotte, Nasra Giama, Thai Ho, Michael Kendrick, Jean Pierre Kocher, Karla Kopp, Catherine Moser, David Nagorney, Daniel O'Brien, Brian Patrick O'Neill, Tushar Patel, Gloria Petersen, Florencia Que, Michael Rivera, Lewis Roberts, Robert Smallridge, Thomas Smyrk, Melissa Stanton, R. Houston Thompson, Michael Torbenson, Ju Dong Yang, Lizhi Zhang, Fadi Brimo, Jaffer A. Ajani, Ana Maria Angulo Gonzalez, Carmen Behrens, Jolanta Bondaruk, Russell Broaddus, Bogdan Czerniak, Bita Esmaeli, Junya Fujimoto, Jeffrey Gershenwald, Charles Guo, Alexander J. Lazar, Christopher Logothetis, Funda Meric-Bernstam, Cesar Moran, Lois Ramondetta, David Rice, Anil Sood, Pheroze Tamboli, Timothy Thompson, Patricia Troncoso, Anne Tsao, Ignacio Wistuba, Candace Carter, Lauren Haydu, Peter Hersey, Valerie Jakrot, Hojabr Kakavand, Richard Kefford, Kenneth Lee, Georgina Long, Graham Mann, Michael Quinn, Robyn Saw, Richard Scolyer, Kerwin Shannon, Andrew Spillane, Jonathan Stretch, Maria Synott, John Thompson, James Wilmott, Hikmat Al-Ahmadie, Timothy A. Chan, Ronald Ghossein, Anuradha Gopalan, Douglas A. Levine, Victor Reuter, Samuel Singer, Bhuvanesh Singh, Nguyen Viet Tien, Thomas Broudy, Cyrus Mirsaidi, Praveen Nair, Paul Drwiega, Judy Miller, Jennifer Smith, Howard Zaren, Joong Won Park, Nguyen Phi Hung, Electron Kebebew, W. Marston Linehan, Adam R. Metwalli, Karel Pacak, Peter A. Pinto, Mark Schiffman, Laura S. Schmidt, Cathy D. Vocke, Nicolas Wentzensen, Robert Worrell, Hannah Yang, Marc Moncrieff, Chandra Goparaju, Jonathan Melamed, Harvey Pass, Natalia Botnariuc, Irina Caraman, Mircea Cernat, Inga Chemencedji, Adrian Clipca, Serghei Doruc, Ghenadie Gorincioi, Sergiu Mura, Maria Pirtac, Irina Stancul, Diana Tcaciuc, Monique Albert, Iakovina Alexopoulou, Angel Arnaout, John Bartlett, Jay Engel, Sebastien Gilbert, Jeremy Parfitt, Harman Sekhon, George Thomas, Doris M. Rassl, Robert C. Rintoul, Carlo Bifulco, Raina Tamakawa, Walter Urba, Nicholas Hayward, Henri Timmers, Anna Antenucci, Francesco Facciolo, Gianluca Grazi, Mirella Marino, Roberta Merola, Ronald de Krijger, Anne Paule Gimenez-Roqueplo, Alain Piché, Simone Chevalier, Ginette McKercher, Kivanc Birsoy, Gene Barnett, Cathy Brewer, Carol Farver, Theresa Naska, Nathan A. Pennell, Daniel Raymond, Cathy Schilero, Kathy Smolenski, Felicia Williams, Carl Morrison, Jeffrey A. Borgia, Michael J. Liptay, Mark Pool, Christopher W. Seder, Kerstin Junker, Larsson Omberg, Mikhail Dinkin, George Manikhas, Domenico Alvaro, Maria Consiglia Bragazzi, Vincenzo Cardinale, Guido Carpino, Eugenio Gaudio, David Chesla, Sandra Cottingham, Michael Dubina, Fedor Moiseenko, Renumathy Dhanasekaran, Karl Friedrich Becker, Klaus Peter Janssen, Julia Slotta-Huspenina, Mohamed H. Abdel-Rahman, Dina Aziz, Sue Bell, Colleen M. Cebulla, Amy Davis, Rebecca Duell, J. Bradley Elder, Joe Hilty, Bahavna Kumar, James Lang, Norman L. Lehman, Randy Mandt, Phuong Nguyen, Robert Pilarski, Karan Rai, Lynn Schoenfield, Kelly Senecal, Paul Wakely, Paul Hansen, Ronald Lechan, James Powers, Arthur Tischler, William E. Grizzle, Katherine C. Sexton, Alison Kastl, Joel Henderson, Sima Porten, Jens Waldmann, Martin Fassnacht, Sylvia L. Asa, Dirk Schadendorf, Marta Couce, Markus Graefen, Hartwig Huland, Guido Sauter, Thorsten Schlomm, Ronald Simon, Pierre Tennstedt, Oluwole Olabode, Mark Nelson, Oliver Bathe, Peter R. Carroll, June M. Chan, Philip Disaia, Pat Glenn, Robin K. Kelley, Charles N. Landen, Joanna Phillips, Michael Prados, Jeffry Simko, Karen Smith-McCune, Scott VandenBerg, Kevin Roggin, Ashley Fehrenbach, Ady Kendler, Suzanne Sifri, Ruth Steele, Antonio Jimeno, Francis Carey, Ian Forgie, Massimo Mannelli, Michael Carney, Brenda Hernandez, Benito Campos, Christel Herold-Mende, Christin Jungk, Andreas Unterberg, Andreas von Deimling, Aaron Bossler, Joseph Galbraith, Laura Jacobus, Michael Knudson, Tina Knutson, Deqin Ma, Mohammed Milhem, Rita Sigmund, Andrew K. Godwin, Rashna Madan, Howard G. Rosenthal, Clement Adebamowo, Sally N. Adebamowo, Alex Boussioutas, David Beer, Thomas Giordano, Anne Marie Mes-Masson, Fred Saad, Therese Bocklage, Lisa Landrum, Robert Mannel, Kathleen Moore, Katherine Moxley, Russel Postier, Joan Walker, Rosemary Zuna, Michael Feldman, Federico Valdivieso, Rajiv Dhir, James Luketich, Edna M.Mora Pinero, Mario Quintero-Aguilo, Carlos Gilberto Carlotti, Jose Sebastião Dos Santos, Rafael Kemp, Ajith Sankarankuty, Daniela Tirapelli, James Catto, Kathy Agnew, Elizabeth Swisher, Jenette Creaney, Bruce Robinson, Carl Simon Shelley, Eryn M. Godwin, Sara Kendall, Cassaundra Shipman, Carol Bradford, Thomas Carey, Andrea Haddad, Jeffey Moyer, Lisa Peterson, Mark Prince, Laura Rozek, Gregory Wolf, Rayleen Bowman, Kwun M. Fong, Ian Yang, Robert Korst, W. Kimryn Rathmell, J. Leigh Fantacone-Campbell, Jeffrey A. Hooke, Albert J. Kovatich, Craig D. Shriver, John DiPersio, Bettina Drake, Ramaswamy Govindan, Sharon Heath, Timothy Ley, Brian Van Tine, Peter Westervelt, Mark A. Rubin, Jung Il Lee, Natália D. Aredes, Armaz Mariamidze, Adam Godzik, Nuria Lopez-Bigas, Josh Stuart, David Wheeler, Ken Chen, and Rachel Karchin. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell*, 2018. ISSN 10974172. doi: 10.1016/j.cell.2018.02.060.
42. Wilma H.M. Hoevenaar, Aniek Janssen, Ajit I. Quirindongo, Huiying Ma, Sjoerd J. Klaasen, Antoinette Teixeira, Bastiaan van Gerwen, Nico Lansu, Folkert H.M. Morsink, G. Johan A. Offerhaus, René H. Medema, Geert J.P.L. Kops, and Nannette Jelluma. Degree and site of chromosomal instability define its oncogenic potential. *Nature Communications*, 2020. ISSN 20411723. doi: 10.1038/s41467-020-15279-9.
43. Christine J. Ye, Zachary Sharpe, and Henry H. Heng. Origins and consequences of chromosomal instability: From cellular adaptation to genome chaos-mediated system survival. *Genes*, 2020. ISSN 20734425. doi: 10.3390/genes11101162.
44. Luisa Matos Do Canto, Simon J. Larsen, Bruna E.Catin Kupper, Maria Dirlei Ferreira De Souza Begnami, Cristóvam Scapulatempo-Neto, Annabeth Høgh Petersen, Mads M. Aagaard, Jan Baumbach, Samuel Aguiar, and Silvia R. Rogatto. Increased levels of genomic instability and mutations in homologous recombination genes in locally advanced rectal carcinomas. *Frontiers in Oncology*, 2019. ISSN 2234943X. doi: 10.3389/fonc.2019.00395.
45. Samuel F. Bakhoum, Bryan Ngo, Ashley M. Laughney, Julie Ann Cavallo, Charles J. Murphy, Peter Ly, Pragya Shah, Roshan K. Sriram, Thomas B.K. Watkins, Neil K. Taunk, Mercedes Duran, Chantal Pauli, Christine Shaw, Kalyani Chadalavada, Vinagolu K. Rajasekhar, Giulio Genovese, Subramanian Venkatesan, Nicolai J. Birkbak, Nicholas McGranahan, Mark Lundquist, Quincey LaPlant, John H. Healey, Olivier Elemento, Christine H. Chung, Nancy Y. Lee, Marcin Imielenski, Gouri Nanjangud, Dana Pe'er, Don W. Cleveland, Simon N. Powell, Jan Lammerding, Charles Swanton, and Lewis C. Cantley. Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature*, 2018. ISSN 14764687. doi: 10.1038/nature25432.
46. Teresa Davoli, Hajime Uno, Eric C. Wooten, and Stephen J. Elledge. Tumor aneuploidy

correlates with markers of immune evasion and with reduced response to immunotherapy. *Science*, 2017. ISSN 10959203. doi: 10.1126/science.aaf8399.

47. Joshua R. Kane, Junfei Zhao, Takashi Tsujiuchi, Brice Laffleur, Víctor A. Arrieta, Aayushi Mahajan, Ganesh Rao, Angeliki Mela, Crismita Dmello, Li Chen, Daniel Y. Zhang, Edgar Gonzalez-Buendia, Catalina Lee-Chang, Ting Xiao, Gerson Rothschild, Uttiya Basu, Craig Horbinski, Maciej S. Lesniak, Amy B. Heimberger, Raul Rabadan, Peter Canoll, and Adam M. Sonabend. CD8+ T-cell–Mediated Immunoediting Influences Genomic Evolution and Immune Evasion in Murine Gliomas. *Clinical Cancer Research*, 2020. ISSN 15573265. doi: 10.1158/1078-0432.CCR-19-3104.
48. Glen J. Weiss, Julia Beck, Donald P. Braun, Kristen Bornemann-Kolatzki, Heather Barilla, Rhiannon Cubello, Walter Quan, Ashish Sangal, Vivek Khemka, Jordan Waypa, William M. Mitchell, Howard Urnovitz, and Ekkehard Schutz. Tumor cell-free DNA copy number instability predicts therapeutic response to immunotherapy. *Clinical Cancer Research*, 2017. ISSN 15573265. doi: 10.1158/1078-0432.CCR-17-0231.
49. Junlin Chi, Shuang Liu, Zhizhong Wu, Yanqiang Shi, Chengmin Shi, Tong Zhang, Binghong Xiong, Yujian Zeng, and Xiangqian Dong. circNSUN2 promotes the malignant biological behavior of colorectal cancer cells via the miR-181a-5p/ROCK2 axis. *Oncology Reports*, 46 (1):1–10, 2021. ISSN 1021-335X. doi: 10.3892/or.2021.8093.
50. Ping Pui Wong, José M. Muñoz-Félix, Maruan Hijazi, Hyojin Kim, Stephen D. Robinson, Beatriz De Luxán-Delgado, Irene Rodríguez-Hernández, Oscar Maiques, Ya Ming Meng, Qiong Meng, Natalia Bodrug, Matthew Scott Dukinfield, Louise E. Reynolds, George Elia, Andrew Clear, Catherine Harwood, Yu Wang, James J. Campbell, Rajinder Singh, Penglie Zhang, Thomas J. Schall, Kylie P. Matchett, Neil C. Henderson, Peter W. Szlosarek, Sally A. Dreger, Sally Smith, J. Louise Jones, John G. Gribben, Pedro R. Cutillas, Pascal Meier, Victoria Sanz-Moreno, and Kairbaan M. Hodivala-Dilke. Cancer Burden Is Controlled by Mural Cell-β3-Integrin Regulated Crosstalk with Tumor Cells. *Cell*, 2020. ISSN 10974172. doi: 10.1016/j.cell.2020.02.003.
51. Annapurna Pranatharthi, Pavana Thomas, Avinash H. Udayashankar, Chandra Bhavani, Srinag Bangalore Suresh, Sudhir Krishna, Jayashree Thatte, Nirmala Srikantia, Cecil R. Ross, and Sweta Srivastava. RhoC regulates radioresistance via crosstalk of ROCK2 with the DNA repair machinery in cervical cancer. *Journal of Experimental and Clinical Cancer Research*, 2019. ISSN 17569966. doi: 10.1186/s13046-019-1385-7.
52. Jasbeer S. Khanduja, Isabel A. Calvo, Richard I. Joh, Ian T. Hill, and Mo Motamedi. Nuclear Noncoding RNAs and Genome Stability, 2016. ISSN 10974164.
53. Catherine E. Waters, Joshua C. Saldivar, Seyed Ali Hosseini, and Kay Huebner. The FHIT gene product: Tumor suppressor and genome "caretaker", 2014. ISSN 14209071.
54. Ge Gao and David I. Smith. WWOX, large common fragile site genes, and cancer. *Experimental Biology and Medicine*, 2015. ISSN 15353699. doi: 10.1177/1535370214565992.
55. Thomas W. Glover, Thomas E. Wilson, and Martin F. Arlt. Fragile sites in cancer: More than meets the eye, 2017. ISSN 14741768.
56. Marcie K. Weil and Alice P. Chen. PARP Inhibitor Treatment in Ovarian and Breast Cancer. *Current Problems in Cancer*, 2011. ISSN 01470272. doi: 10.1016/j.currproblcancer.2010.12.002.
57. Hyeon Jun Shin, Hyuk Kwon Kwon, Jae Hyeok Lee, Xiangai Gui, Asma Achek, Jae Ho Kim, and Sangdun Choi. Doxorubicin-induced necrosis is mediated by poly-(ADP-ribose) polymerase 1 (PARP1) but is independent of p53. *Scientific Reports*, 2015. ISSN 20452322. doi: 10.1038/srep15798.
58. Andréa E. Tijhuis, Sarah C. Johnson, and Sarah E. McClelland. The emerging links between chromosomal instability (CIN), metastasis, inflammation and tumour immunity. *Molecular Cytogenetics*, 2019. ISSN 17558166. doi: 10.1186/s13039-019-0429-1.
59. Alvin J.X. Lee, David Endesfelder, Andrew J. Rowan, Axel Walther, Nicolai J. Birkbak, P. Andrew Futreal, Julian Downward, Zoltan Szallasi, Ian P.M. Tomlinson, Michael Howell, Maik Kschischo, and Charles Swanton. Chromosomal instability confers intrinsic multidrug resistance. *Cancer Research*, 2011. ISSN 00085472. doi: 10.1158/0008-5472.CAN-10-3604.

DRAFT

Gene ID	Expected CNV
TP53	Homozygous Deletion
CDKN2A	Homozygous Deletion
CDKN2B	Homozygous Deletion
PTEN	Homozygous Deletion
NF1	Homozygous Deletion
RB1	Homozygous Deletion
ARID1A	Homozygous Deletion
RAD51B	Homozygous Deletion
SMAD4	Homozygous Deletion
STK11	Homozygous Deletion
BRCA1	Homozygous Deletion
BRCA2	Homozygous Deletion
FANCD2	Homozygous Deletion
FAT1	Homozygous Deletion
SMARCB1	Homozygous Deletion
MYC	Amplification
AURKA	Amplification
ERBB2	Amplification
CCND1	Amplification
EGFR	Amplification
FGFR1	Amplification
CDK6	Amplification

Table 1. Known commonly deleted and commonly amplified genes. CNV results from Ploidetect, BIC-seq2, Sequenza, and PURPLE in the POG cohort were examined for the frequency of the indicated mutations in the listed genes.

DRAFT

Gene ID	log2(fold change)	Alternative Hypothesis	P value	FDR
TP53	-0.516191547388182	greater	4.24589052523666e-16	9.00128791350172e-14
AC074391.1	-0.38844586709978	greater	2.66268315220876e-10	2.82244414134128e-08
TTN	-0.36287482231269	greater	1.93173893254297e-07	1.36509551233036e-05
ROCK2	-1.18705337317961	greater	4.23855363982281e-05	0.00224643342910609
ZNF804B	-0.991509216655021	greater	9.36865772155271e-05	0.00397231087393835
DMBT1	0.495820509150211	less	0.00123655861800198	0.0436917378360699

Table 2. Genes with mutation status associated with CIN. Gene IDs, log2(fold change) of segment N50, alternative hypothesis (for significance testing), nominal p-value and false discovery rate are shown. All significant (FDR < 0.05) genes are shown.

DRAFT

Chr	Start	End	Gene ID	P value	FDR	Segment	Deleted Exons	Total Exons	Exon Deleted Proportion	Cases affected
2	141691930	142220754	LRP1B	0	0	4204	17	91	0.1868	13
3	30673341	30749712	TGFBR2	0	0	4540	6	9	0.6667	5
3	59760906	61026602	FHIT	0	0	2934	4	22	0.1818	44
8	12929151	12946264	DLC1	0.0035	0.0087	6477	4	25	0.16	5
8	39225787	39346345	ADAM5	0	0	3034	7	15	0.4667	16
9	9364129	9717470	PTPRD	0	0	6830	2	50	0.04	13
9	21965764	21999582	CDKN2A	0	0	3045	7	7	1	114
10	89618825	89781058	PTEN	0	0	3048	12	12	1	22
13	48875658	49057901	RB1	0	0	3059	27	27	1	17
16	6288694	7092253	RBFOX1	0	0	1809	3	24	0.125	17
16	78367828	78908173	WWOX	0	0	1950	1	12	0.0833	23
17	7570798	7593227	TP53	0	0	3077	18	18	1	9
17	11896639	12077866	MAP2K4	0	0	3078	12	12	1	13
17	29388421	29705018	NF1	0	0	3081	61	61	1	11
18	48553523	48611791	SMAD4	0	0	3085	12	12	1	20
20	14619797	15226720	MACROD2	0	0	3093	3	21	0.1429	29
22	23409931	23409931	RSPH14	7e-04	0.0021	3755			0	5
22	49989769	50043752	C22orf34	0	1e-04	4006	5	11	0.4545	6

Table 3. RDRs labeled in Figure 45. Genomic coordinates of RDRs overlapping each gene, nominal P-value, FDR, RDR number (segment), count of exons covered by RDRs, total exon count per gene, percentage of exons affected by the deletions, and number of cases affected by the RDRs are listed. Decimal values are rounded to 4 decimal places.

DRAFT

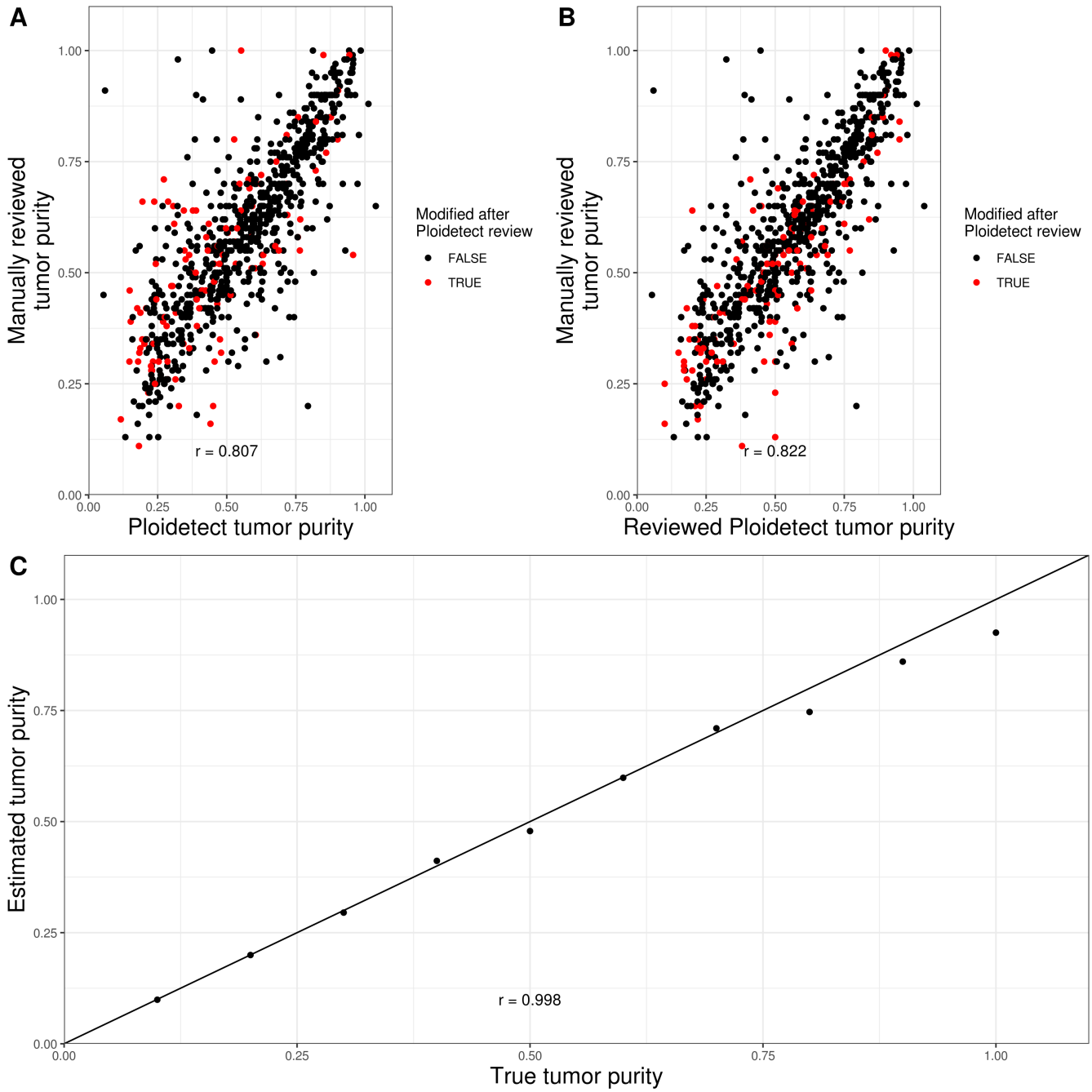

Fig. 1. Ploidetect accurately determines tumor purity from whole-genome sequence data. For all plots, Pearson's r is shown. **A.** Tumor purity detected automatically by Ploidetect is compared with estimates obtained by manual review of CNV data by an expert bioinformatician. Estimates found to be incorrect by manual review of Ploidetect outputs are labeled in red. **B.** Alternate models for explaining tumor purity were considered for Ploidetect results, and if one was available, the better-fitting model was chosen. Estimates which changed after review are labeled in red. **C.** Reads from a COLO829 cell line library were mixed with matched normal at known proportions to obtain *in-silico* dilutions with known tumor purities. Ploidetect estimates are plotted against the true values.

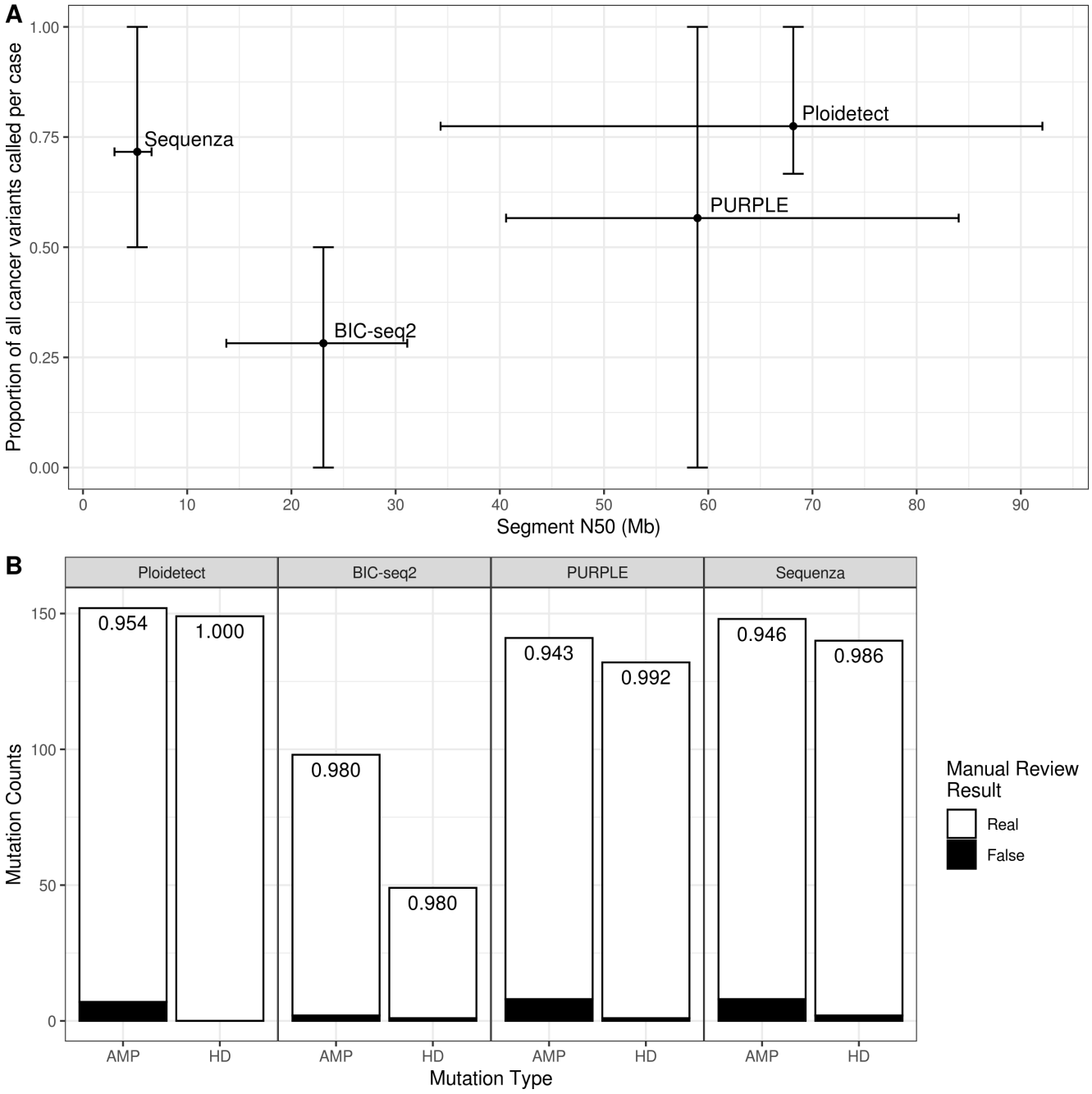

Fig. 2. Ploidetect enables detection of CNVs with greater sensitivity and specificity than other methods. **A.** Segment N50 and crCNVs within POG were quantified for each tool. For each case, the proportion of crCNVs called by each tool was compared to the union of crCNVs called by all tools. Bars indicate the IQR. **B.** Results of manual review of crCNVs. Each panel denotes results from a single CNV caller. White area indicates crCNV calls identified as correct whereas black area indicates calls found by manual review to be incorrect. Numbers in bars indicate observed precision of calls. AMP = amplification, HD = homozygous deletion.

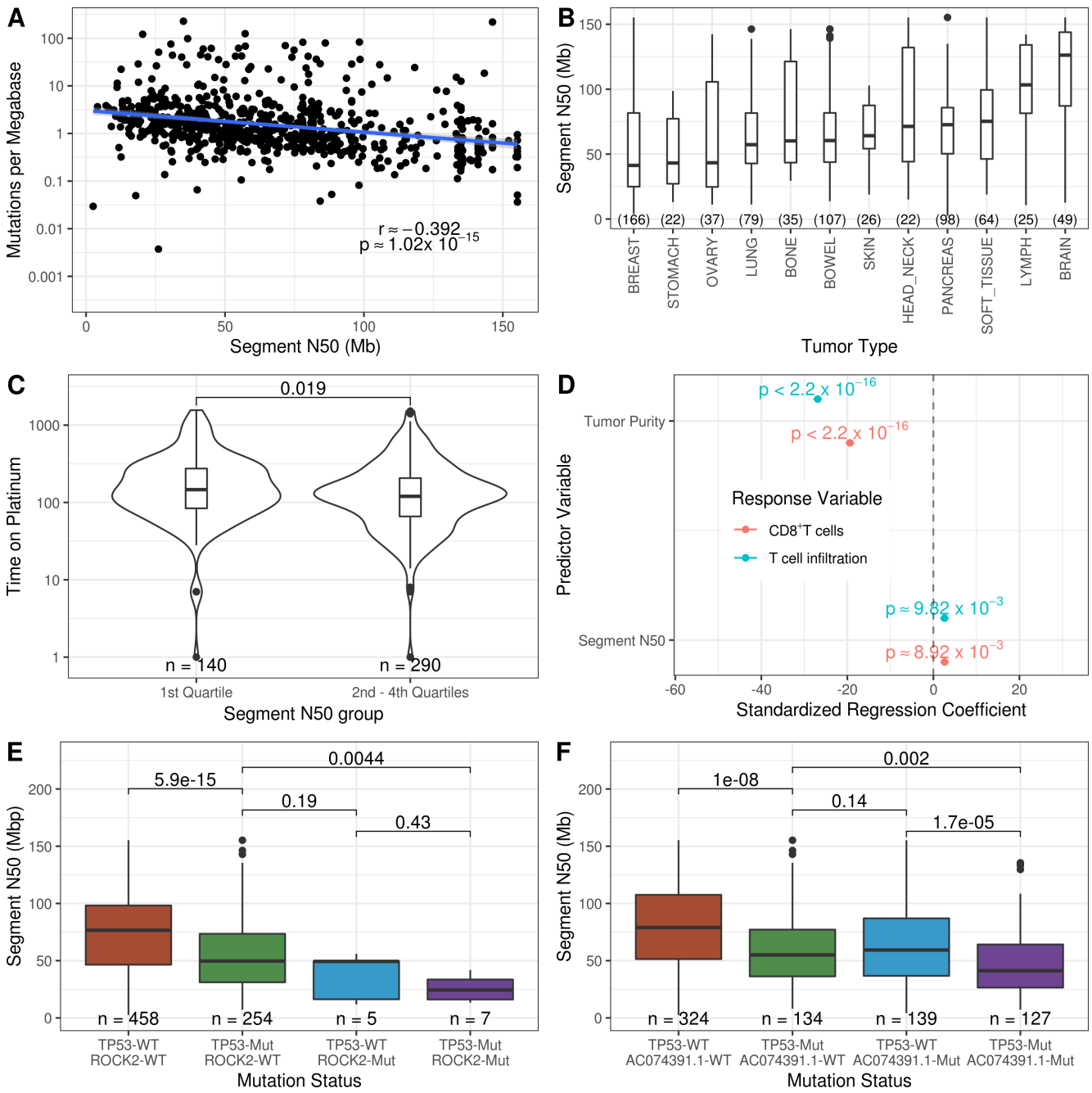

Fig. 3. CIN is implicated in multiple facets of cancer biology. **A.** Correlation of log-scaled SNV and indel mutations per megabase with CNV segment N50. Spearman correlation r and linear model p-value are shown. **B.** Boxplots of segment N50 in OncoTree tissue types. Tissue types represented in at least 20 samples are shown. Sample sizes for each tumor type are indicated within brackets. **C.** Total TTD on platinum chemotherapy for samples in the 1st quartile of segment N50 compared to all other samples. **D.** Multiple regression was used to assess the impact of immune infiltration on segment n50, accounting for tumor purity. Standardized regression coefficients were calculated by dividing the coefficients and standard deviations for each feature by their standard deviations. Error bars corresponding to the standard deviation are shown. **E.** Segment N50s of each combination of wild type TP53 (TP53-WT), wild type ROCK2 (ROCK2-WT), mutant TP53 (TP53-Mut), mutant ROCK2 (ROCK2-Mut) are compared. **F.** Segment N50s of each combination of wild type TP53 (TP53-WT), wild type AC074391.1 (AC074391.1-WT), mutant TP53 (TP53-Mut), mutant AC074391.1 (AC074391.1-Mut) are compared. P-values for all panels are from two-sided Wilcoxon rank-sum tests.

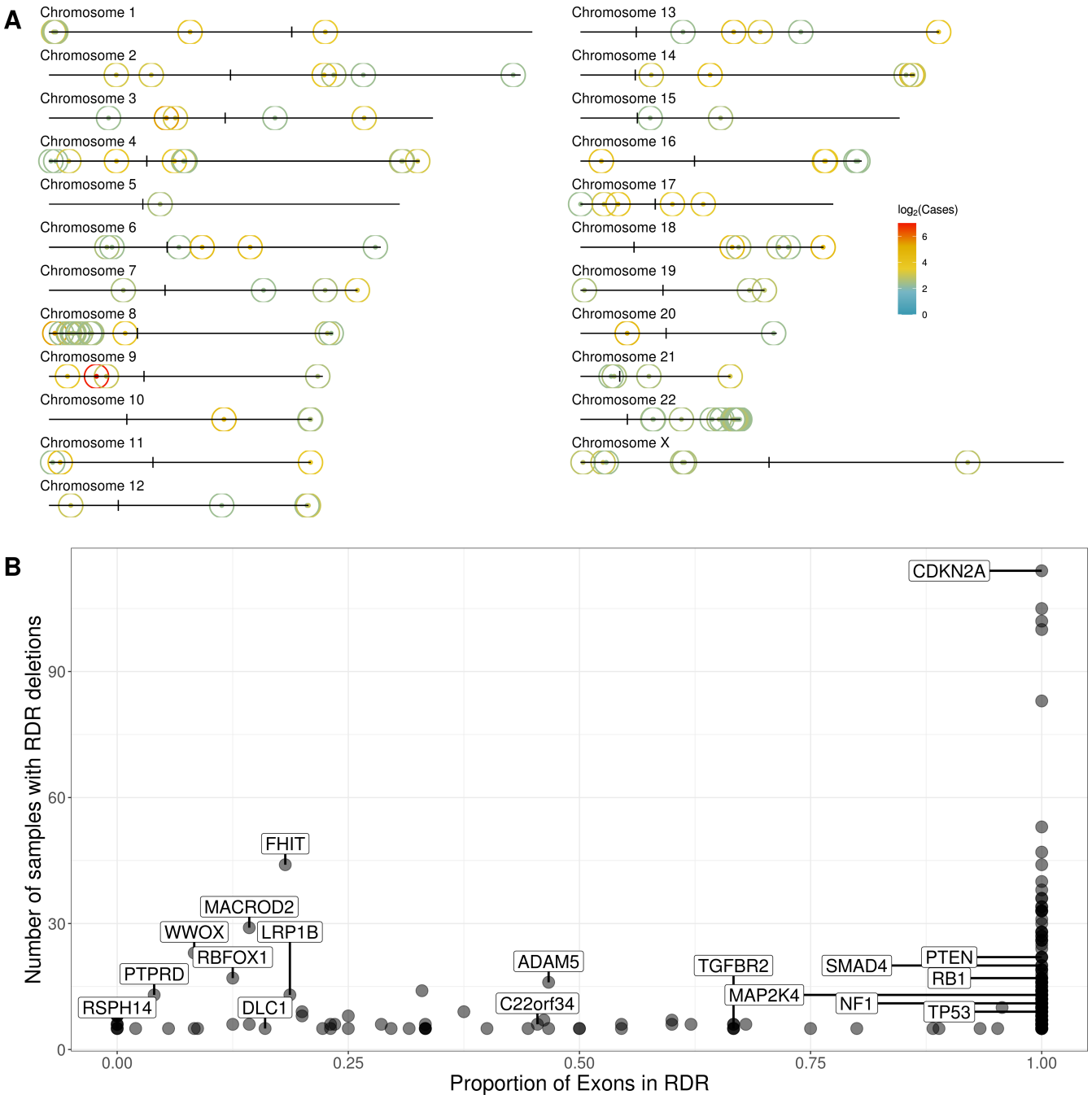

Fig. 4. Recurrent homozygous deletions in the metastatic cancer genome. **A.** Positions of RDRs are indicated on each chromosome with a circle. Colors indicate the number of cases with deletions contributing to that RDR. **B.** For each gene within an RDR, the number of cases with a deletion in the gene were calculated. These values are plotted against the proportion of exons for that gene which were affected by any RDR. Selected genes with previously described oncological properties are labeled.

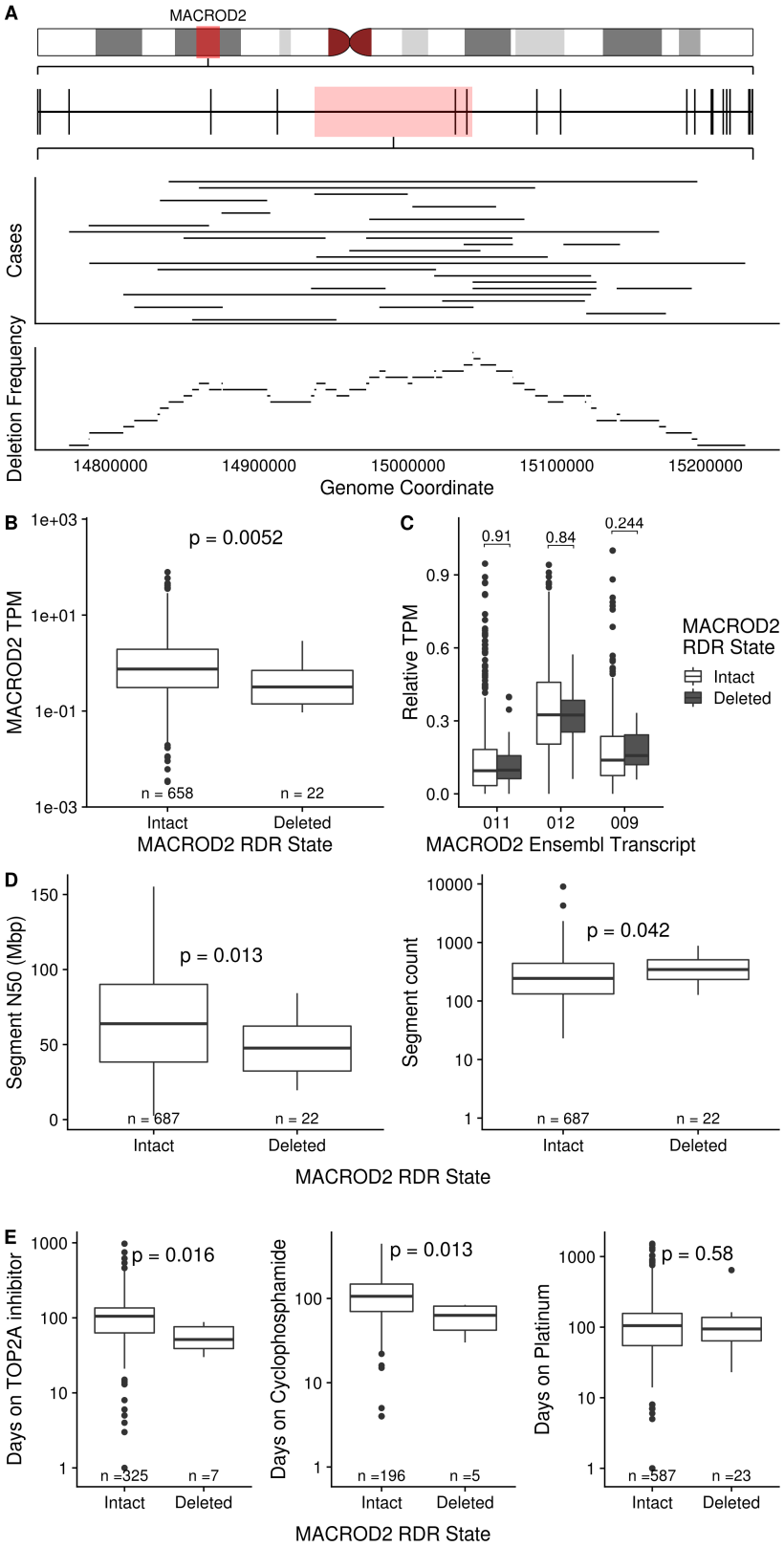

Fig. 5. MACROD2 RDR deletions lead to reduced expression, increased CIN and worsened single-strand break drug response **A.** MACROD2 location on chromosome 20 indicated. Ideogram of MACROD2 with the MACROD2 RDR highlighted. Vertical lines on ideogram indicate annotated exons. Individual deletions contributing to the RDR are shown. Cumulative coverage plot of the individual deletions on bottom. **B.** Comparison of MACROD2 expression in MACROD2-RDR-intact and MACROD2-RDR-intact cases. **C.** Differential splicing comparison in MACROD2-RDR-intact and MACROD2-RDR-deleted cases. Relative TPMs were obtained by dividing each case's transcript TPM values by the total transcript TPMs per case. The three most expressed MACROD2 transcripts are shown. **D.** Comparison of CIN metrics (segment n50, segment count) in MACROD2-RDR-intact cases and MACROD2-RDR-deleted cases. **E.** Comparison of TTDs in MACROD2-RDR-intact cases and MACROD2-RDR-Deleted cases for TOP2A inhibitors (Doxorubicin, Epirubicin), Cyclophosphamide, or Platinum (Cisplatin, Carboplatin, Oxaliplatin). Two-sided Wilcoxon rank-sum tests were used for all comparisons.