Inferring chromosomal instability from copy number aberrations as a measure of chromosomal instability across human cancers

Sasha Thalluri[1*], Vishal H. Oza[1*], Tabea M. Soelter[1], Jennifer L. Fisher[1], and Brittany N. Lasseigne[1#]

[1] Department of Cell, Developmental and Integrative Biology, School of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA
*co-first author
#Corresponding author:
Brittany N. Lasseigne
THT 946 1900 University Blvd, Birmingham, Alabama, 35294, USA
Email address: bnp0001@uab.edu

Data Availability Statement
The data that support this study are openly available at the Genomic Data Commons (GDC) data portal (https://portal.gdc.cancer.gov/) as 'TCGA Level 3' data. In addition, all analyses associated with this paper are on GitHub (https://github.com/lasseignelab/CINmetrics_Cancer_Analysis) and available at https://zenodo.org/record/7942543#.ZGPu8OzMJ4A.

Funding Statement
ST, VHO, JLF, and BNL were supported by R00HG009678. ST, VHO, TMS, JLF, and BNL were supported by UAB funds to the Lasseigne Lab.

Conflicts of Interest
The authors have no conflicts to declare.

Ethics Approval Statement
This study was not classified as human subjects research or clinical investigation, and was determined to be exempt by the University of Alabama at Birmingham Institutional Review Board.

Patient Consent Statement
N/A

Permission to Reproduce Material From Other Sources
N/A

Clinical Trial Registration
N/A

Acknowledgments:
We would like to thank the members of the Lasseigne Lab for their critical and constructive feedback.

Author Contributions:
ST: Software, Formal Analysis, Data Curation, Writing-Original Draft, Writing-Review & Editing, Visualization; **VHO:** Methodology, Software, Validation, Data Curation, Writing-Review & Editing, Visualization, Project Administration; **TMS:** Software, Validation, Formal Analysis, Writing-Review & Editing, Visualization; **JLF:** Data Curation, Writing-Review & Editing; **BNL:** Conceptualization, Methodology, Resources, Writing-Original Draft, Writing-Review & Editing, Supervision, Project Administration, Funding Acquisition

Abstract:
Background
Cancer is a complex disease that is the second leading cause of death in the United States. Despite research efforts, the ability to manage cancer and select optimal therapeutic responses for each patient remains elusive. Chromosomal instability (CIN) is primarily a product of segregation errors wherein one or many chromosomes, in part or whole, vary in number. CIN is an enabling characteristic of cancer, contributes to tumor-cell heterogeneity, and plays a crucial role in the multistep tumorigenesis process, especially in tumor growth and initiation and in response to treatment.

Aims
Multiple studies have reported different metrics for analyzing copy number aberrations as surrogates of CIN from DNA copy number variation data. However, these metrics differ in how they are calculated with respect to the type of variation, the magnitude of change, and the inclusion of breakpoints. Here we compared metrics capturing CIN as either numerical aberrations, structural aberrations, or a combination of the two across 33 cancer data sets from The Cancer Genome Atlas (TCGA).

Methods and results
Using CIN inferred by methods in the CINmetrics R package, we evaluated how six copy number CIN surrogates compared across TCGA cohorts by assessing each across tumor types, as well as how they associate with tumor stage, metastasis, and nodal involvement, and with respect to patient sex.

Conclusions
We found that the tumor type impacts how well any two given CIN metrics correlate. While we also identified overlap between metrics regarding their association with clinical characteristics and patient sex, there was not complete agreement between metrics. We identified several cases where only one CIN metric was significantly associated with a clinical characteristic or patient sex for a given tumor type. Therefore, caution should be used when describing CIN based on a given metric or comparing to other studies.

1. Introduction

Genomic instabilities, molecular signatures of gross genomic alterations, are enabling characteristics of cancer etiology and pathogenesis (1,2), which result in chromosomal breakages and rearrangements that can develop into chromosomal instability (CIN). CIN is primarily caused by defective cell cycle quality control mechanisms including an elevated rate of segregation errors altering chromosomal content (3) and manifests as either numerical aberrations, structural aberrations, or a combination of the two. Numerical aberrations are whole chromosomal aberrations that lead to the loss of heterozygosity (i.e., where a chromosomal region is lost in one copy for a diploid genome) and variability in gene dosage effects, resulting in a phenotype that is a consequence of chromosome-wide altered expression patterns. Structural aberrations are sub-chromosomal and can lead to the fusing of gene products or amplified and/or deleted genes, specifically impacting the genes of the affected chromosomal regions (3). At the molecular level, CIN has been shown to represent distinct etiologies, promote disease progression, and metastases, as well as to be associated with patient prognosis, drug efficacy, and drug resistance across many cancers (4–7). As CIN has been associated with poor patient outcomes in some cancers but improved survival in others (8,9), CIN appears to have cell- and tissue-specific consequences associated with the originating tissue and tumor site (10,11). For example, while CIN has been shown to have a non-monotonic relationship with patient outcome in ER−/ERBB2− breast, gastric, ovarian, squamous non-small cell lung carcinomas (i.e., patients with the lowest or highest quartile of CIN have a significantly improved hazard ratio)(12), it has been associated with poorer prognosis in diffuse large B-cell lymphoma patients(13). Therefore, CIN is a promising biomarker of patient prognosis and drug response but requires tissue-specific evaluation.

Several CIN scores have been proposed for analyzing copy number aberrations (i.e., deletions or amplifications of segments of the genome) as surrogates of chromosomal instability. Still, these scores represent various aspects of CIN and have been associated with different clinical and biological phenotypes across several cancer types (14–17). It is critical to assess copy number aberration CIN surrogates across cancer types and tissue backgrounds to evaluate the role of CIN in cancer etiology and progression, in order to compare key results across studies and to identify robust scores for future biomarker development. We

recently published an R package, CINmetrics(18), for calculating six different copy number aberration CIN surrogate metrics, and here apply it to The Cancer Genome Atlas (TCGA). These metrics include total aberration index (TAI) (14), modified TAI, copy number abnormality (CNA) (15), number of break points (16), altered base segments (17), and fraction genome altered (FGA). In addition, they differ based on their ability to detect structural, numerical, or whole genome instability (discussed in depth in (18)). We determine these CIN scores across 33 cancers and 22,629 samples from 11,124 patients to examine how CIN metric scores vary across and within different cancers and with respect to clinical characteristics and patient sex.

2. Methods

We downloaded masked copy number variation (CNV) (Affymetrix SNP 6.0 array) data and associated patient clinical data from all 33 projects in the TCGA portal using the 'TCGAbiolinks' R package (2.18.0) (19) in August 2021 using R (Version 4.0.4) and RStudio (Version 1.4.1106) locally and stored within a CSV file. We also downloaded TCGA Level 3, normalized and aggregated RNA-seq count data in November 2021 from all 33 projects in the TCGA portal using the 'TCGAbiolinks' R package (Version 2.22.1) using R (Version 4.0.2) and RStudio (Version 1.1.463) with UAB's High-Performance Computing Cluster, Cheaha. All analyses associated with this paper are on GitHub (https://github.com/lasseignelab/CINmetrics_Cancer_Analysis) and available at https://zenodo.org/record/7942543#.ZGPu8OzMJ4A.

With the CINmetrics R package (Version 0.1.0), we calculated each CIN metric for the 22,629 non-tumor and tumor samples (18). CINmetrics analyzes six different copy number aberration CIN surrogate metrics from masked CNV data as previously described (14–17). Briefly, those metrics are TAI (Total Aberration Index), Modified TAI, CNA (Copy Number Abnormality), Base Segments (i.e., the number of altered bases), Break Points (i.e., the number of break points), and FGA (Fraction of Genome Altered). Each metric defines CIN by calculating numerical and/or structural aberrations as described (18) and can be grouped as numerical scores (base segments, FGA), structural scores (break points and CNA), and overall scores (TAI and Modified TAI). CNA and break points both consider the segmental abnormalities of the chromosome, but CNA requires that adjacent segments have a difference in segmentation mean values. TAI and modified TAI can both be interpreted as the absolute deviation from the normal copy number state averaged over all genomic locations, where modified TAI removes the directionality aspect of the TAI metric by taking the absolute value of the segment mean.

All cross-sample CIN metrics comparisons using Spearman's correlation (20) were conducted using the base 'stats' R package (21) (Version 4.0.4) 'cor' function between the non-tumor ("Blood Derived Normal," "Solid Tissue Normal," "Bone Marrow Normal," and "Buccal Cell Normal") and tumor ("Metastatic," "Primary Blood Derived Cancer," "Primary Tumor," "Recurrent Tumor," "Additional - New Primary," and "Primary Blood Derived Cancer - Peripheral Blood") samples. All heatmaps were generated using the 'ComplexHeatmap' R package (Version 2.9.3) (22) and clustered with the "complete' method by "Euclidean distance."

For analyses that compare CINmetrics by clinical characteristics, we used the TNM (tumor stage, node, and metastasis) staging provided by TCGA in the "ajcc_pathologic_m" (except for ACC where we used "ajcc_clinical_m"), "ajcc_pathologic_n," and "ajcc_pathologic_t" attributes using the 'GDCquery_clinic' function. These comparisons were made across the 22 cancer types with corresponding data (i.e., all but GBM, LAML, OV, THYM, UCS, DLBC, PCPG, UCEC, SARC, PRAD, and LGG). In addition, we conducted the Mann Whitney Wilcoxon Test (23,24) for comparisons between CIN metrics and the staging variables using the 'rstatix' R package (Version 0.7.0) (25) 'wilcox_test' function. We Bonferroni-corrected p-values to account for multiple hypothesis testing, and we considered corrected p-values of less than 0.05 to indicate significant CIN metric scores for the TNM staging analyses.

For analyses comparing CIN metrics by sex for each cancer, we used the 'gender' (i.e., biological male or female) information provided by TCGA for each sample. We compared 27 cancer types because TCGA projects OV, PRAD, UCEC, TGCT, CESC, and UCS occur predominantly in one biological sex. We conducted a Mann Whitney Wilcoxon test between CIN metrics and the sex variable using the 'rstatix' R package (Version 0.7.0) 'wilcox_test' function and multiple hypothesis tests corrected using the Bonferroni method. We plotted raincloud plots of these multiple hypothesis corrected p-values by CIN metric using the ggplot2 (Version 3.3.5), ggpubr (Version 0.4.0), PupillometryR (Version 0.0.4), and gghalves (Version 0.1.1) R packages for the top cancer identified by each CIN metric.

3. Results

We compiled masked CNV data for 11,124 patients across all TCGA projects (n=33 cancer types, 22,629 samples) and applied the six CIN inference calculations in the CINmetrics R package to each cancer data set (Figure 1).

Figure 1A shows the distribution of CIN scores across normal and tumor samples. Within each CIN "type" - structural (Base segments, FGA), numerical (CNA, Break Points), or overall (TAI, Modified TAI) - the distribution pattern is consistent. However, there are noticeable differences across these types. For example, structural CIN metrics show a more pronounced distinction between normal and tumor samples compared to numerical CIN metrics. Thus, one should be careful in evaluating genomic instability based on the choice of the metric. Subsequently, we conducted a Spearman's correlation to analyze the relationship between each CIN metric across all tumor patient samples irrespective of the cancer type and within each cancer type. The results are depicted in Figure 1B, which further emphasizes the point that different "types" of CIN metrics capture varying aspects or patterns of genomic instability. Overall, Base Segments and FGA showed the most separation between tumor and non-tumor samples by cancer type, and as expected, the directionality of TAI for tumor compared to non-tumor samples is opposite of the other metrics (Figure 1A). Additionally, the tumor type impacts how well each CIN metric correlates with the others. For example, TAI positively correlates with FGA in some cancer tissues (e.g., lung squamous cell carcinoma, LUSC) but not in others (e.g., acute myeloid leukemia, LAML).

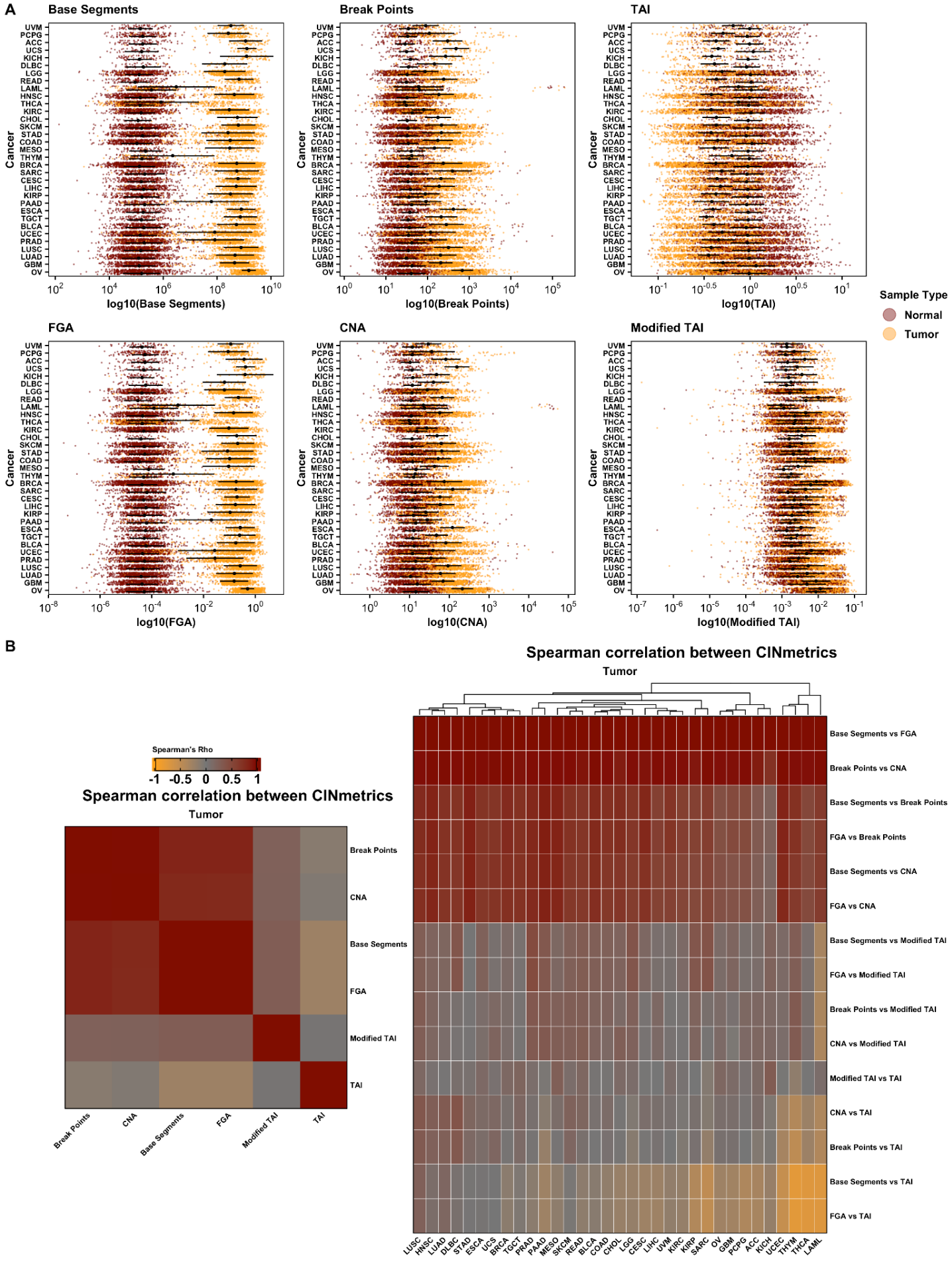
A
Base Segments
Break Points
TAI
FGA
CNA
Modified TAI
Cancer
UVM
PCPG
ACC
UCS
KICH
DLBC
LGG
READ
LAML
HNSC
THCA
KIRC
CHOL
SKCM
STAD
COAD
MESO
THYM
BRCA
SARC
CESC
LIHC
KIRP
PAAD
ESCA
TGCT
BLCA
UCEC
PRAD
LUSC
LUAD
GBM
OV
log10(Base Segments)
log10(Break Points)
log10(TAI)
log10(FGA)
log10(CNA)
log10(Modified TAI)
Sample Type
Normal
Tumor
B
Spearman's Rho
-1 -0.5 0 0.5 1
Spearman correlation between CINmetrics
Tumor
Break Points
CNA
Base Segments
FGA
Modified TAI
TAI
Spearman correlation between CINmetrics
Tumor
Base Segments vs FGA
Break Points vs CNA
Base Segments vs Break Points
FGA vs Break Points
Base Segments vs CNA
FGA vs CNA
Base Segments vs Modified TAI
FGA vs Modified TAI
Break Points vs Modified TAI
CNA vs Modified TAI
Modified TAI vs TAI
CNA vs TAI
Break Points vs TAI
Base Segments vs TAI
FGA vs TAI
LUSC HNSC LUAD DLBC STAD ESCA UCS BRCA TGCT PRAD PAAD MESO SKCM READ BLCA COAD CHOL LGG CESC LIHC UVM KIRC KIRP SARC OV GBM PCPG ACC KICH UCEC THYM THCA LAML

Figure 1.
CIN metrics by cancer type for tumor and matched non-tumor samples (A) and Spearman's correlation between tumor CIN metrics by cancer type (B). CIN, chromosomal instability.

Next, we determined if each CIN metric is significantly associated with clinical characteristics by tumor type. For these analyses, we used the TNM (tumor stage, node, and metastasis) staging provided by TCGA for each of the 22 cancer types with corresponding TNM data (Figure 2). We found that breast invasive carcinoma (BRCA; significant for base segments, FGA, break points, and CNA; Figure 2A-D), rectum adenocarcinoma (READ; significant for break points, and CNA; Figure 2C-D), head and neck squamous cell carcinoma (HNSC; significant for CNA; Figure 2D), and LUSC (significant for TAI; Figure 1E) each had at least one CIN metric significantly associated with tumor stage (Mann Whitney Wilcoxon Test <0.05 after Bonferroni correction). However, only colon adenocarcinoma (COAD; significant for base segments, FGA, break points, and CNA; Figure 2A-D) had CIN metrics significantly associated with metastases, while COAD (significant for base segments and FGA; Figure 2A-B), kidney renal papillary cell carcinoma (KIRP; significant for Break Points and CNA; Figure 2C-D), and lung adenocarcinoma (LUAD; significant for modified TAI; Figure 2F) all had at least one CIN metric associated with nodal involvement (Mann Whitney Wilcoxon Test <0.05 after Bonferroni correction).

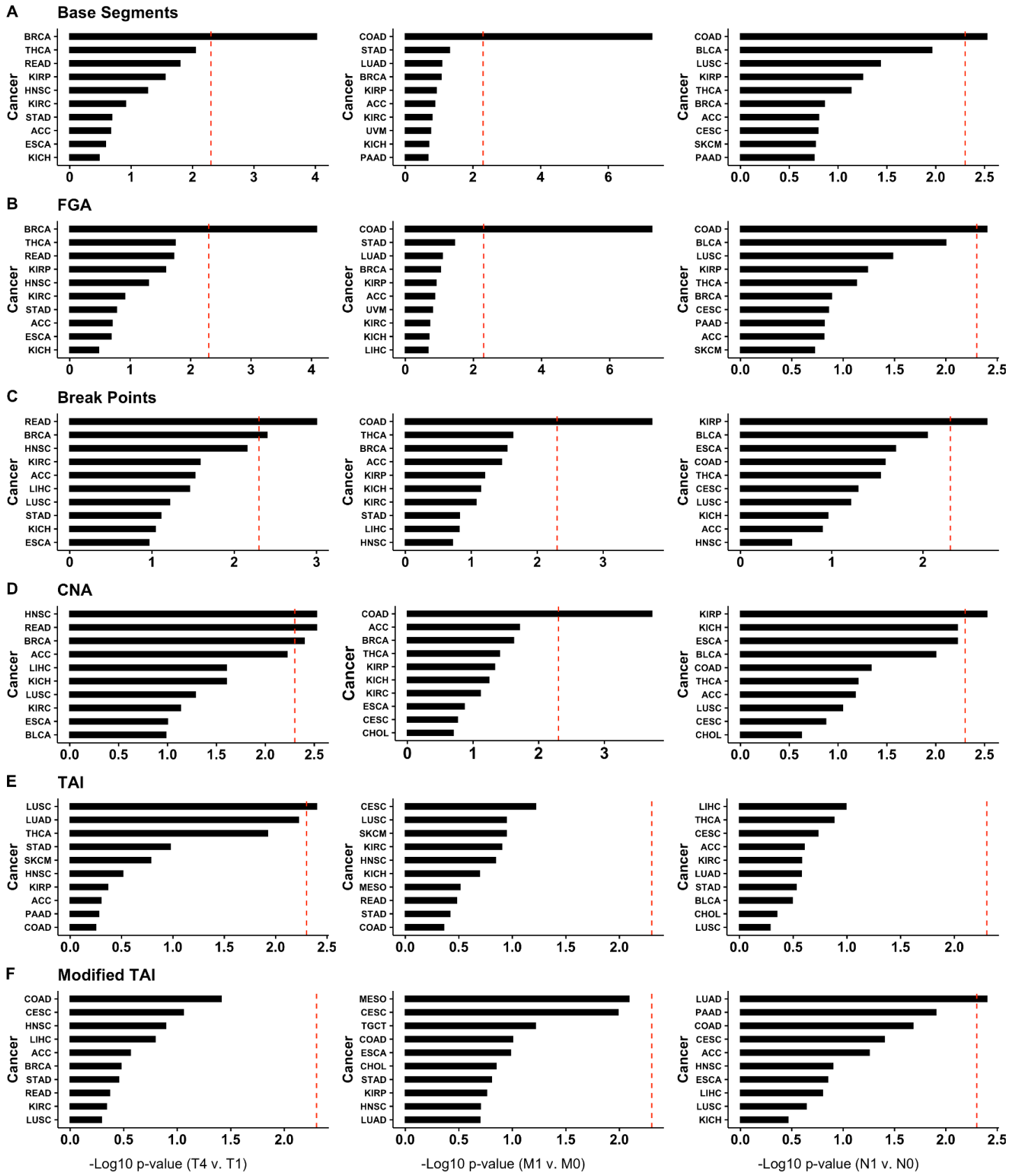

Figure 2.

Top ten cancers with the lowest Bonferroni-corrected p-values for (A) base segments, (B) FGA, (C) break points, (D) CNA, (E) TAI, and (F) modified TAI association with tumor stage (T4 compared to T1), metastasis (M1 compared to M0), and nodal involvement (N1 compared to N0). TAI, total aberration index; CNA, copy number abnormality; FGA, fraction genome altered.

Finally, we compared CIN metrics by sex for each cancer by using the ‘gender’ (i.e., biological male or female) information provided by TCGA for each patient for the 27 cancer types with cases in both sexes (Figure 3). We found that HNSC CIN was significantly different between the sexes based on the base segments, FGA, break points, CNA, and modified TAI metrics, esophageal carcinoma (ESCA) was significantly different between the sexes based on the break points and CNA metrics (Figure 3A-B), thyroid carcinoma (THCA) based on the CNA metric (Figure 3A), and glioblastoma (GBM; Figure 3A-B), COAD (Figure 3A-B), kidney renal clear cell carcinoma (KIRC), READ, LUAD, and LUSC based on the modified TAI metric (Mann Whitney Wilcoxon Test <0.05 after Bonferroni correction, Figure 3A).

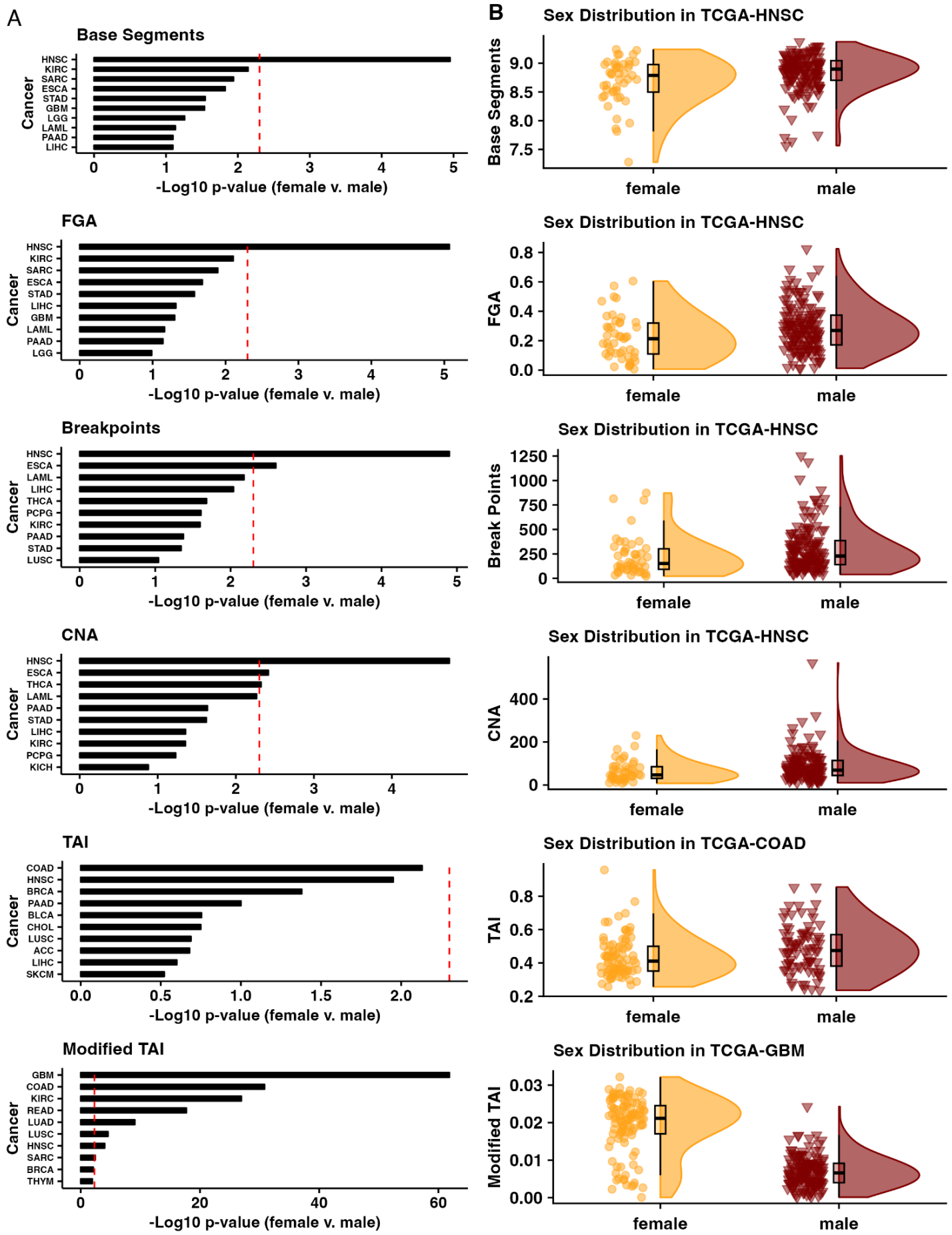
A
Base Segments
Cancer
HNSC
KIRC
SARC
ESCA
STAD
GBM
LGG
LAML
PAAD
LIHC
-Log10 p-value (female v. male)
FGA
Cancer
HNSC
KIRC
SARC
ESCA
STAD
LIHC
GBM
LAML
PAAD
LGG
-Log10 p-value (female v. male)
Breakpoints
Cancer
HNSC
ESCA
LAML
LIHC
THCA
PCPG
KIRC
PAAD
STAD
LUSC
-Log10 p-value (female v. male)
CNA
Cancer
HNSC
ESCA
THCA
LAML
PAAD
STAD
LIHC
KIRC
PCPG
KICH
-Log10 p-value (female v. male)
TAI
Cancer
COAD
HNSC
BRCA
PAAD
BLCA
CHOL
LUSC
ACC
LIHC
SKCM
-Log10 p-value (female v. male)
Modified TAI
Cancer
GBM
COAD
KIRC
READ
LUAD
LUSC
HNSC
SARC
BRCA
THYM
-Log10 p-value (female v. male)
B
Sex Distribution in TCGA-HNSC
Base Segments
female
male
Sex Distribution in TCGA-HNSC
FGA
female
male
Sex Distribution in TCGA-HNSC
Break Points
female
male
Sex Distribution in TCGA-HNSC
CNA
female
male
Sex Distribution in TCGA-COAD
TAI
female
male
Sex Distribution in TCGA-GBM
Modified TAI
female
male

Figure 3.

(A) Top ten cancers with the lowest Bonferroni-corrected p-values between the sexes for each CIN metric and the (B) distribution of CIN metric scores for the most different tumor cohort by sex for each CIN metric. CIN, chromosomal instability.

4. Discussion

Here we evaluated how the TAI, modified TAI, CNA, base segments, break points, and FGA copy number aberration CIN surrogates (ie., names..) compared across TCGA cohorts by assessing each across tumor types, how they associate with tumor stage, metastasis, and nodal involvement, and with respect to patient sex. We found that the tumor type impacts how well any two given CIN metrics correlate. While we also identified overlap between CIN metrics regarding their association with clinical characteristics (e.g., CIN was significantly associated with tumor stage in BRCA for 4 of the 6 metrics) and patient sex (e.g., CIN was significantly different between the sexes in HNSC for 5 of the 6 metrics), there was not complete agreement between metrics. We identified several cases where only one CIN metric was significantly associated with a clinical characteristic or patient sex for a given tumor type (e.g., modified TAI was the only CIN metric significantly associated with nodal involvement for LUAD). Therefore, caution should be used when describing CIN based on any one metric or comparing across studies.

A recent study of 1,421 samples from 394 tumors across 22 tumor types demonstrates that somatic copy number aberrations in cancer are both pervasive (i.e., occurring at least once in 99% of tumors) and dynamic (i.e., more than 20% of the genome was subject to subclonal somatic copy number aberrations in 45% of tumors) (26). Generally, CIN has been associated with distinct cancer etiologies and progression, patient prognosis, drug efficacy, and drug resistance in tissue- and tumor-specific manners (4–11,27,28). For example, van Dijk et al. (28) found that variation in the chromosomal copy number within a tumor (CIN heterogeneity) was strongly associated with poor survival in patients with solid tumors such as breast cancer, lung cancer, and colorectal cancer. However, as we show, the choice of metric used to calculate CIN can affect the measurement of such variation. In (27), Lukowet al. discuss the role of aneuploidy in cancer drug resistance by either overexpression of genes that suppress DNA repair, promote cell growth, or through the accumulation of mutations in genes that encode DNA repair enzymes. Thus, in such cases measuring not only the CIN but the region where it happens might be more useful for predicting drug resistance. While this indicates that CIN may be a promising biomarker of patient prognosis and drug response, a thorough understanding of how to measure and interpret CIN is critical. Our study further underscores the need to be specific about how and when during the disease course CIN is calculated and that patient characteristics like sex may impact or be associated with such metrics.

There are several limitations to this study. The first is that these CIN metrics are calculated based on one genomic profile generated from a tumor, or a tumor sample, at a given time. Tumors are often very heterogeneous, including across time, so this provides only a snapshot of a dynamic system. Recent studies have underscored that some tumor types have a strong correlation between CIN and metastasis that may be associated with when during tumor development copy number aberrations happen as well as the tissue of origin (29). Additionally, the data used to calculate the scores in this study are array-based intensity scores from bulk profiles (18). High-throughput sequencing and single-cell technologies will continue to allow for more comprehensive profiling and provide an opportunity for future studies to characterize cancer CIN more precisely (30). Finally, this brief report does not investigate how different causes of CIN may influence these metrics.

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar 4;144(5):646–74.

2. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022 Jan;12(1):31–46.

3. McGranahan N, Burrell RA, Endesfelder D, Novelli MR, Swanton C. Cancer chromosomal instability: therapeutic and diagnostic challenges. EMBO Rep. 2012 Jun 1;13(6):528–38.

4. Tijhuis AE, Johnson SC, McClelland SE. The emerging links between chromosomal instability (CIN), metastasis, inflammation and tumour immunity. Mol Cytogenet. 2019 May 14;12:17.

5. Lukow DA, Sausville EL, Suri P, Chunduri NK, Wieland A, Leu J, et al. Chromosomal instability accelerates the evolution of resistance to anti-cancer therapies. Dev Cell [Internet]. 2021 Aug 2; Available from: http://dx.doi.org/10.1016/j.devcel.2021.07.009

6. Bakhoum SF, Compton DA. Chromosomal instability and cancer: a complex relationship with therapeutic potential. J Clin Invest. 2012 Apr;122(4):1138–43.

7. Ferguson LR, Chen H, Collins AR, Connell M, Damia G, Dasgupta S, et al. Genomic instability in human cancer: Molecular insights and opportunities for therapeutic attack and prevention through diet and nutrition. Semin Cancer Biol. 2015 Dec;35 Suppl:S5–24.

8. van Jaarsveld RH, Kops GJPL. Difference Makers: Chromosomal Instability versus Aneuploidy in Cancer. Trends Cancer Res. 2016 Oct;2(10):561–71.

9. Ben-David U, Amon A. Context is everything: aneuploidy in cancer. Nat Rev Genet. 2020 Jan;21(1):44–62.

10. Thompson SL, Bakhoum SF, Compton DA. Mechanisms of chromosomal instability. Curr Biol. 2010 Mar 23;20(6):R285–95.

11. Weaver BAA, Silk AD, Montagna C, Verdier-Pinard P, Cleveland DW. Aneuploidy acts both oncogenically and as a tumor suppressor. Cancer Cell. 2007 Jan;11(1):25–36.

12. Birkbak NJ, Eklund AC, Li Q, McClelland SE, Endesfelder D, Tan P, et al. Paradoxical relationship between chromosomal instability and survival outcome in cancer. Cancer Res. 2011 May 15;71(10):3447–52.

13. Bakhoum SF, Danilova OV, Kaur P, Levy NB, Compton DA. Chromosomal instability substantiates poor prognosis in patients with diffuse large B-cell lymphoma. Clin Cancer Res. 2011 Dec 15;17(24):7704–11.

14. Baumbusch LO, Helland Å, Wang Y, Liestøl K, Schaner ME, Holm R, et al. High levels of genomic aberrations in serous ovarian cancers are associated with better survival. PLoS One. 2013 Jan 23;8(1):e54356.

15. Davison JM, Yee M, Krill-Burger JM, Lyons-Weiler MA, Kelly LA, Sciulli CM, et al. The degree of segmental aneuploidy measured by total copy number abnormalities predicts survival and recurrence in superficial gastroesophageal adenocarcinoma. PLoS One. 2014 Jan 16;9(1):e79079.

16. Lee AJX, Endesfelder D, Rowan AJ, Walther A, Birkbak NJ, Futreal PA, et al. Chromosomal instability confers intrinsic multidrug resistance. Cancer Res. 2011 Mar 1;71(5):1858–70.

17. Chin SF, Teschendorff AE, Marioni JC, Wang Y, Barbosa-Morais NL, Thorne NP, et al. High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. Genome Biol. 2007;8(10):R215.

18. Oza VH, Fisher JL, Darji R, Lasseigne BN. CINmetrics: an R package for analyzing copy number aberrations as a measure of chromosomal instability. PeerJ. 2023 Apr 25;11:e15244.

19. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016 May 5;44(8):e71.

20. Spearman C. nThe Proof and Measurement of Association Between Two Things, oAmerican J. Psychol; 1904.

21. Computing R, Others. R: A language and environment for statistical computing. Vienna: R Core Team [Internet]. 2013; Available from: https://www.yumpu.com/en/document/view/6853895/r-a-language-and-environment-for-statistical-computing

22. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016 Sep 15;32(18):2847–9.

23. Mann HB, Whitney DR. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Ann Math Stat. 1947;18(1):50–60.

24. Wilcoxon F. Some Uses of Statistics in Plant Pathology. Biometrics Bulletin. 1945;1(4):41–5.

25. Kassambara A. Pipe-Friendly Framework for Basic Statistical Tests [R package rstatix version 0.7.0]. 2021 Feb 13 [cited 2022 Feb 1]; Available from: https://CRAN.R-project.org/package=rstatix

26. Watkins TBK, Lim EL, Petkovic M, Elizalde S, Birkbak NJ, Wilson GA, et al. Pervasive chromosomal instability and karyotype order in tumour evolution. Nature. 2020 Nov;587(7832):126–32.

27. Lukow DA, Sheltzer JM. Chromosomal instability and aneuploidy as causes of cancer drug resistance. Trends Cancer Res. 2022 Jan;8(1):43–53.

28. van Dijk E, van den Bosch T, Lenos KJ, El Makrini K, Nijman LE, van Essen HFB, et al. Chromosomal copy number heterogeneity predicts survival rates across cancers. Nat Commun. 2021 May 27;12(1):3188.

29. Nguyen B, Fong C, Luthra A, Smith SA, DiNatale RG, Nandakumar S, et al. Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. Cell. 2022 Feb 3;185(3):563–75.e11.

30. Funnell T, O'Flanagan CH, Williams MJ, McPherson A, McKinney S, Kabeer F, et al. Single-cell genomic variation induced by mutational processes in cancer. Nature. 2022 Dec;612(7938):106–15.