Contents lists available at ScienceDirect

Journal of Pathology Informatics

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

CNViz: An R/Shiny Application for Interactive Copy Number Variant Visualization in Cancer

Rebecca G. Ramesh [a,*], Ashkan Bigdeli [a], Chase Rushton [a], Jason N. Rosenbaum [a,b]

[a] Center for Personalized Diagnostics, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[b] Department of Pathology, Kaiser Permanente, Oakland, California, USA

ARTICLE INFO

ABSTRACT

Copy number variants (CNVs) comprise a class of mutation which includes deletion, duplication, or amplification events that range in size from smaller than a single-gene or exon, to the size of a full chromosome. These changes can affect gene expression levels and are thus implicated in disease, including cancer. Although a variety of tools and methodologies exist to detect CNVs using data from massively parallel sequencing (also referred to as next-generation sequencing), it can be difficult to appreciate the copy number profile in a list format or as a static image. CNViz is a freely accessible R/Bioconductor package that launches an interactive R/Shiny visualization tool to facilitate review of copy number data. As inputs, it requires genomic locations and corresponding copy number ratios for probe, gene, and/or segment-level data. If supplied, loss of heterozygosity (LOH), focal variant data [single nucleotide variants (SNVs) and small insertions and deletions (indels)], and metadata (e.g., specimen purity and ploidy) can also be incorporated into the visualization. The CNViz R/Bioconductor package is an easy-to-use tool built with the intent of encouraging visualization and exploration of copy number variation. CNViz can be used in a clinical setting as well as for research to study patterns in human cancers more broadly. The intuitive interface allows users to visualize the copy number profile of a specimen, dynamically change resolution to explore gene and probe-level copy number changes, and simultaneously integrate LOH, SNV, and indel findings. CNViz is available for download as an R package via Bioconductor. An example of the application is available at rebeccagreenblatt.shinyapps.io/cnviz_example.

Background

Copy number variants (CNVs) comprise a class of mutation which includes deletion, duplication, and amplification events that range in size from smaller than a single-gene or exon, to the size of a full chromosome. These changes can affect gene expression levels and are thus implicated in disease, including cancer.[1,2] For example, *ERBB2* (*Her2*) amplification is considered prognostic in breast cancer and *CDKN2A* deletion is considered diagnostic and prognostic in glioblastoma.[3]

Identifying amplification and deletion events that are correlated with differential gene expression can lead to the discovery of new genes and disease-related pathways.[4,5] As an example, by correlating CNV data with gene expression data, *FAM60A* expression was identified as a key prognostic factor in esophageal cancer, leading to a new biomarker and potential therapeutic target.[6]

Large genomic changes (such as trisomy 21) are still resolved by conventional karyotyping (microscopy).[7] Select CNVs can be identified directly by standalone, targeted assays [such as Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) or Fluorescence In Situ Hybridization (FISH)], whereas others are often not identified at all for clinical management.[8,9] Increasingly, massively parallel sequencing (MPS, also referred to as next generation sequencing or NGS) data are being used to estimate copy number across the whole exome or targeted panels of cancer-related genes.[10,11] Computational tools like CNVKit[10] and PureCN[11] assess targeted short read sequencing data and infer copy number. With appropriate assay design, this strategy offers the opportunity for more comprehensive assessment than standalone assays, higher resolution than conventional karyotyping, and the integration of parallel data detected or inferred from the sequencing. Although CNV information is extremely valuable, it is complex and can be more difficult to appreciate in list format than other types of genetic or genomic changes, such as single nucleotide variants (SNVs) or small insertions and deletions (indels). Visualization of CNV data is often limited to static scatterplots and diagrams.[10,11] While these approaches allow for visualization of large events at the resolution of chromosomes or chromosome bands, gene-level or sub-gene-level information is not easily accessible. Interrogation of gene-level information requires access to the raw data and a level of programming competency.[12]

Well-designed visualization tools empower end-users to appreciate and integrate complex information and discover new associations without

* Corresponding author at: 3020 Market St. Suite 220, Philadelphia, PA 19104, USA.
  E-mail addresses: Rebecca.Greenblatt@pennmedicine.upenn.edu (R.G. Ramesh), Ashkan.Bigdeli@pennmedicine.upenn.edu (A. Bigdeli), Chase.Rushton@pennmedicine.upenn.edu
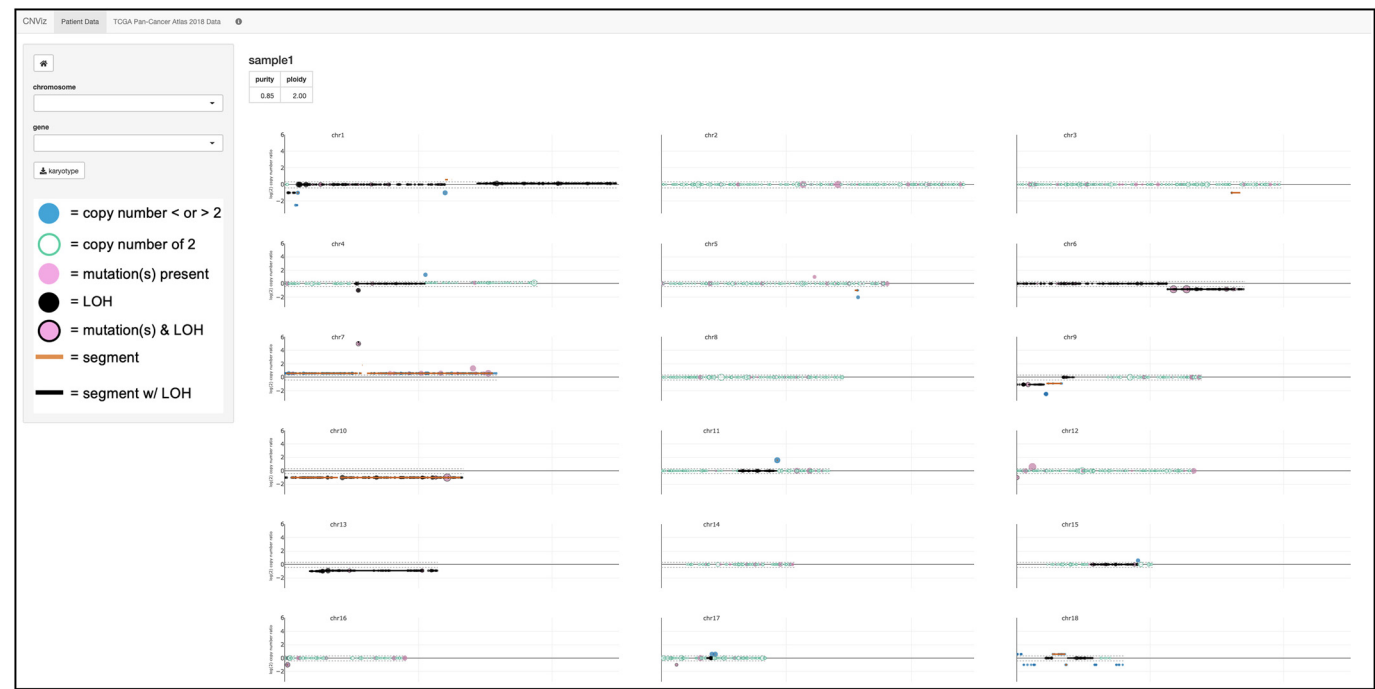(C. Rushton), Jason.N.Rosenbaum@KP.org (J.N. Rosenbaum).

Fig. 1. CNViz Homepage. The homepage displays the 'All Chromosomes' view. When hovering over a gene marker, the copy number is displayed. If a gene is selected, either by clicking on the marker or by selecting from the dropdown menu on the left, the user will be directed to the 'Gene View.'
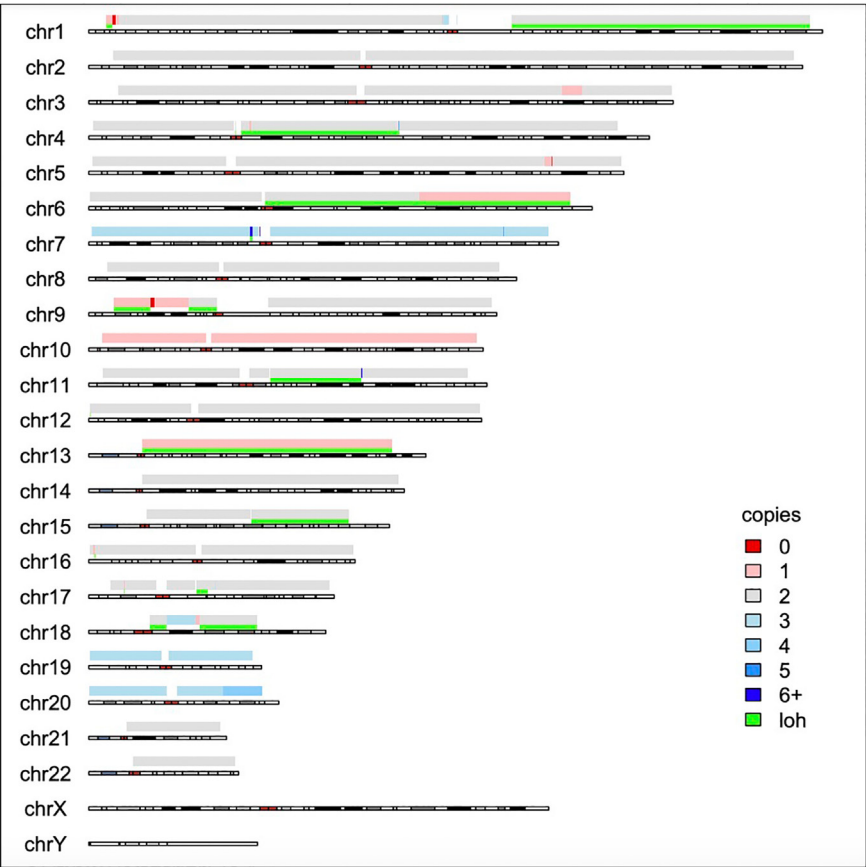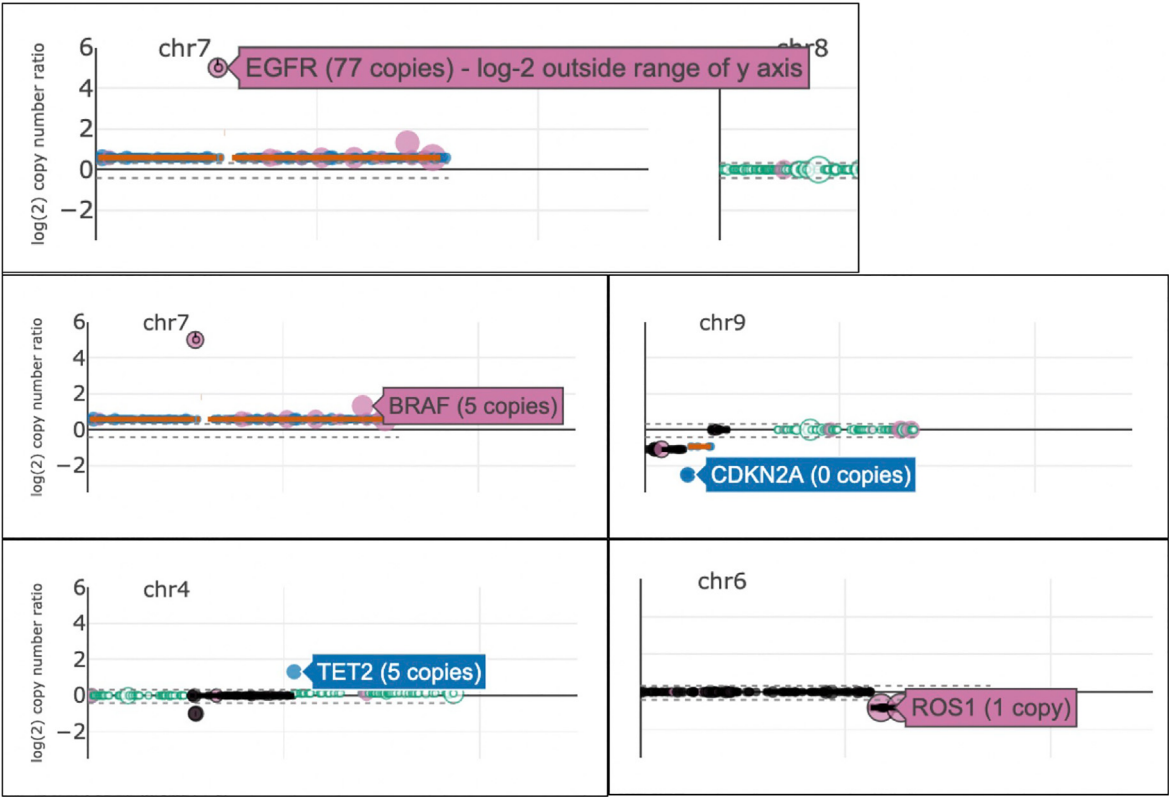

Fig. 2. Virtual Karyotype PDF. This image is a reflection of the information displayed in the 'All Chromosomes' view. Red indicates loss, blue indicates gain, and green indicates loss of heterozygosity. This karyotype will only be available if segment data is supplied by the user.

Fig. 3. Hover View. From the 'All Chromosomes' or 'Single Chromosome' view, when hovering over a gene marker, the copy number is displayed. The copy number is rounded to the nearest whole number. If the copy number is out of range (>64 copies), as is seen with *EGFR* in the figure, the correct copy number along with a warning will be displayed. If there are 0 copies, as is seen with *CDKN2A* in the figure, the marker will be displayed at a log-2 copy number ratio corresponding to −2.5.
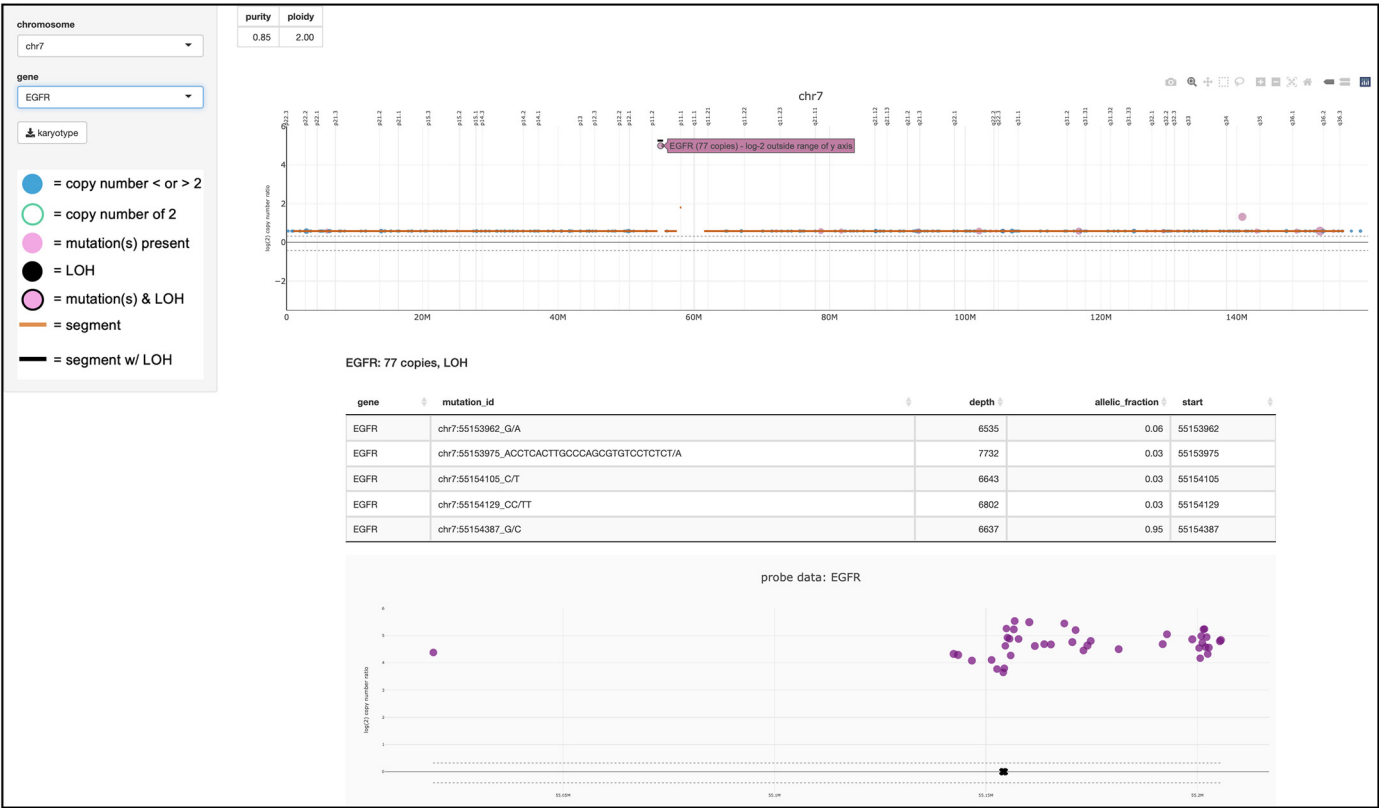
Fig. 4. Gene View. When a gene is selected, the user is taken to the relevant 'Single Chromosome' view, and gene-specific information is displayed underneath. Specifically, there is a table of focal variants and a plot with probe data. On the probe plot, an "X" on the x-axis indicates that a focal variant is present at that location.

requiring programming knowledge or extensive data-analytic skills.[13] Tools like the UCSC genome browser[14] and the Integrated-Genomics Viewer (IGV)[15] allow researchers and clinical practitioners to visually inspect variants in the context of their sequencing data. Visual inspection provides an opportunity to review variants identified by informatics algorithms, assess analytical validity of calls made by the algorithm, and characterize the potential biological (and sometimes clinical) consequence of the variants. The demonstrated utility and widespread adoption of these tools in the context of SNVs strongly implies potential utility for analogous tools designed for the visual inspection of CNVs.

The CNViz R/Bioconductor package,[16,17] outlined here, takes probe, gene, and segment-level copy number ratios and launches an interactive Shiny application,[18,19] enabling detailed visualization of the copy number profile of a given sample. Shiny is an R package that enables the creation of interactive web applications within the R environment, allowing the developer to take advantage of many open source data manipulation and bioinformatics packages. The resulting web applications can run on local servers and are therefore ideally suited to be used with patient-sensitive data. CNViz can integrate parallel data when available, such as loss of heterozygosity (LOH) data and focal variant data (SNVs and small indels)

identified by standard bioinformatic variant callers. CNViz also includes aggregated data from The Cancer Genome Atlas (TCGA) 2018 Pan-Cancer Atlas[20] studies, so the user can correlate findings with those seen in a variety of common cancer diagnoses. While CNViz was developed for use with targeted sequencing panels, the design makes it applicable to assay technologies from small amplicon sequencing to whole genome sequencing.

As an example, we walk through a case seen at the Center for Personalized Diagnostics at the University of Pennsylvania[21] using CNViz.

Methods

The launchCNViz function in the CNViz R package[16] takes properly formatted input data and launches a Shiny application[18] that visualizes a sample's copy number profile. Of note, CNViz is purely a visualization tool and thus no inference is performed; the data that are supplied by the user are displayed in an intuitive format that allows for interactive exploration. CNViz does not identify or display fusion genes or chromosomal rearrangements.

All plots were made using the Plotly R package,[19] which styles the plots and enables zoom functionality. Design of the application was informed by
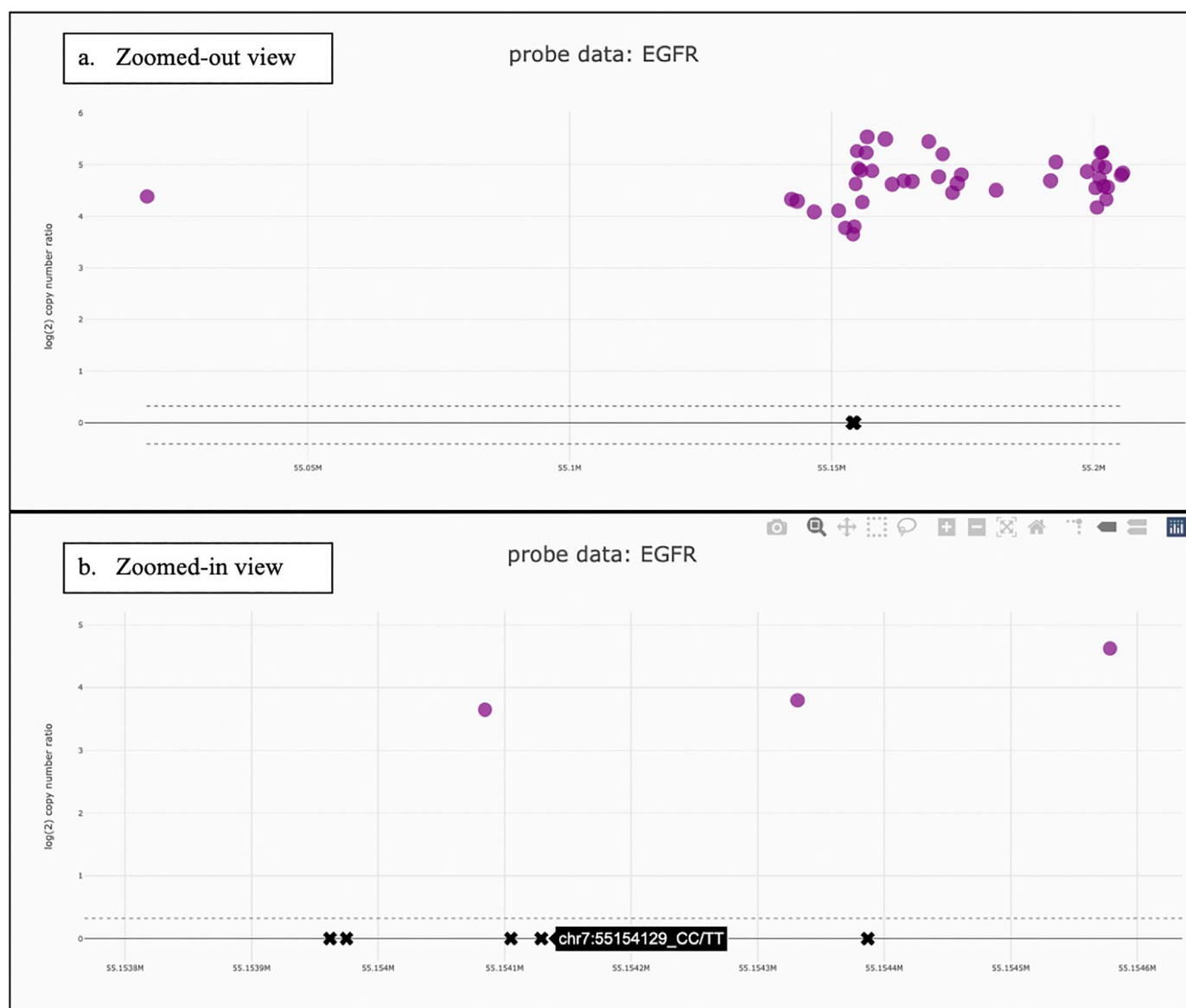
Fig. 5. Probe Data. (A) Zoomed-out View. The probe plot of EGFR reveals a possible relative deletion between positions 55.14 and 55.155M. (B) Zoomed-in View. Zooming-in reveals a cluster of variants around 55.154M. Hovering over each "X" displays the corresponding mutation i.d.

Shiny and Plotly's default themes, which are simple and place emphasis on the data represented. The palette was selected with the help of Color Universal Design in order to be colorblind accessible.[22] The user can view all chromosomes at once (Fig. 1), download a virtual karyotype (Fig. 2), hover to view gene copy number estimates (Fig. 3), view an individual chromosome, and select an individual gene to view probe-level and focal variant data (Fig. 4). These features are all available but will depend on the amount of data supplied to the function. For example, only one of probe-level, gene-level, and segment-level data is required to launch the application, whereas LOH, focal variant, and metadata are optional.

Probe, gene, and segment data inputs are structured similarly. At a minimum, they each must include columns for chromosome, start position, end position, and log-2 copy number ratio. Probe and gene data require the corresponding gene symbol. Genomic coordinates should be given with respect to GRCh38[23] to ensure the cytoband labels on the x-axis and the karyotype pdf are accurate.

LOH can be included as an optional column in gene and segment data and is represented by the gene marker or segment line colored in black. Focal variant data, specifically SNVs and small indels, can be included as well, such that when a gene is clicked, the corresponding variants are displayed in a table below the chromosome plot (Fig. 4) and are plotted on the probe plot (Fig. 5). If included, custom metadata, perhaps the sample's estimated purity and ploidy, is displayed underneath the sample name (Fig. 1). Ploidy is assumed to be two, unless it is specified in metadata, in which case it is rounded to the nearest whole number. Only segment lines differing from ploidy or with LOH are displayed. All segments from chromosomes X and Y are displayed.

The karyotype PDF is generated using the KaryoploteR,[24] CopyNumberPlots,[25] and GenomicRanges[26] R packages. The PDF will be generated only if segment data is supplied. If the LOH column is included in segment data, then this information will be incorporated into the diagram (Fig. 2).

This application was modeled based on data generated by CNVKit[10] (probe data) and PureCN[11] (gene, segment, LOH, focal variant, and metadata), but the aforementioned input data can be obtained through any means, such as a pre-packaged CNV software or a custom-built CNV algorithm. If data include integer copy number (as it does in PureCN), it can be converted to log-2 copy number ratio with $\log_2(C/2)$, where C is the integer copy number. Note that tools like PureCN[11] adjust integer copy number estimates for inferred purity and ploidy, whereas raw probe-level data might assume a ploidy of 2 and purity of 100%. For this reason, depending on the input data provided, gene and probe log-2 values may not align with one another.

An additional tab displays data from 2018 TCGA Pan-Cancer Atlas studies (Fig. 6).[20] These data were obtained through cBioPortal[27,28] using the cBioPortalData R package.[29] The purpose of this tab is to correlate sample findings with aggregated data from the same cancer type. The copy number data displayed were generated by the GISTIC algorithm[30]; they include gain, amplification, shallow deletion, and deep deletion prevalence for 468 cancer-related genes from the IMPACT-468 gene panel.[31]

Shown on rebeccagreenblatt.shinyapps.io/cnviz_example are simulated data; three samples sequenced at the University of Pennsylvania's Center for Personalized Diagnostics[21] were modified significantly, then combined to create this example. Specifically, the SNVs were shifted using a random number generator, so the reference nucleotide listed may not line-up with the GRCh38[23] reference genome. The application was deployed using the rsconnect R package[32] on the shinyapps.io server[33] and data are hosted on Amazon S3.[34]

The CNViz R package can be downloaded via Bioconductor[17] and the code is available at github.com/rebeccagreenblatt/cnviz. The launchCNViz function includes a more detailed description of the required structure of the input data and an example is provided in the corresponding vignette.

Fig. 6. TCGA Pan-Cancer Atlas 2018 Data. The user can select from 32 cancer types. Prevalence of deep deletion, shallow deletion, gain, and amplification for the IMPACT-468 gene panel is displayed. (A) Sort View. The sort feature allows the user to easily identify the most common copy number changes in a specific cancer type. In the figure, the 'Amplification' column is sorted, such that the most common amplifications in the Glioblastoma Multiforme cohort are displayed at the top. (B) Gene Search View. The search feature allows the user to see the data for a specific gene of interest.
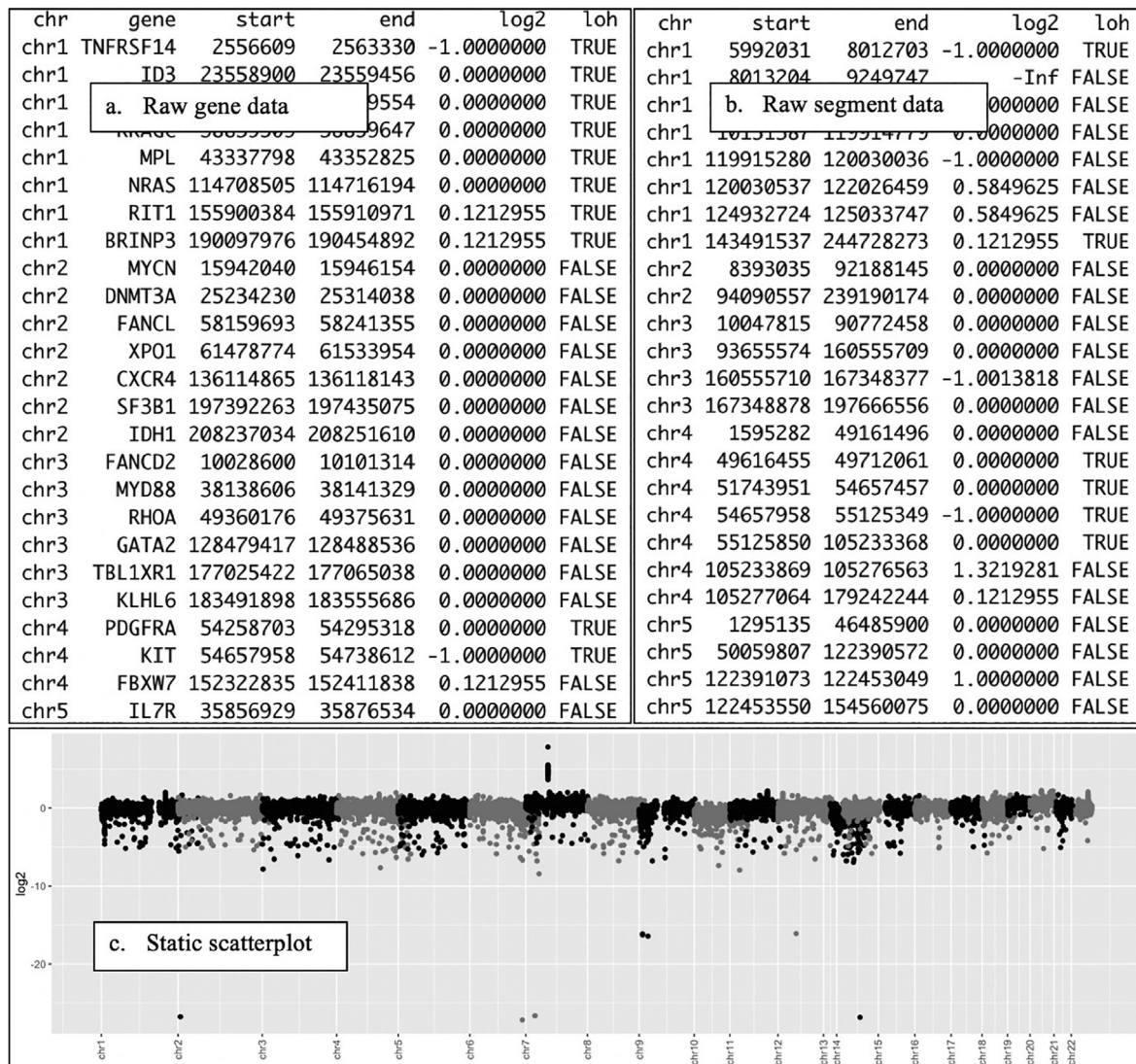
Results

CNViz displays copy number data in an intuitive and interactive format. It allows the user to resolve significant gene-level findings (amplifications, deletions, loss of heterozygosity, SNVs, and indels), as well as integrate these findings into the broader context of the tumor genome (i.e., larger duplication or deletion events). CNViz was modeled based on output from CNVKit[10] and PureCN,[11] but can take any input structured in the specified format; at a minimum, the input data must include probe, gene, or segment-level copy number ratios. Any number or combination of genes can be included, such as only the genes included in a predetermined subset, all the genes that were sequenced, or all genes (irrespective of whether they were sequenced or not).

For our example case, the tissue is brain, frontal lobe, and the neoplasm is identified by clinical histology review as glioblastoma multiforme. DNA was extracted from the tissue and sequenced on a clinically validated 152 MPS cancer panel. Probe-level data for 3226 genes was generated using CNVKit.[10] The PureCN[11] algorithm estimated gene-level integer copy number, segments, focal variants, and LOH, as well as sample purity and tumor ploidy. The data were transformed minimally in R[16] to meet the format

required for CNViz. For example, column names were standardized, integer copy number estimates were converted to log-2 copy number ratios for gene and segment data, and the LOH column values were simplified to true or false.

After installing the CNViz R/Bioconductor package, the application is launched using the launchCNViz function with the above-mentioned data as input. The homepage loads, which shows all chromosomes (Fig. 1). Clicking on the 'karyotype' button downloads a PDF virtual karyotype, which displays gain (blue), loss (red), and LOH (green) segments (Fig. 2). Immediately visualizable are copy-neutral LOH of chromosome 1 and chromosome arm 6p, gain of chromosome 7, and loss of arms 6q, 9p, and chromosome 10. Hovering over gene markers identifies amplification of *EGFR*, gains of *BRAF* and *TET2,* homozygous deletion of *CDKN2A*, and heterozygous deletion of *ROS1*, among other copy number changes (Fig. 3). Also identifiable are focal variants in *EGFR*, *BRAF*, and *ROS1* (indicated by pink markers), and LOH of *EGFR* and *ROS1* (indicated by markers outlined in black).

Clicking on *EGFR* opens the gene view (Fig. 4). This displays chromosome 7 at the top, with *EGFR*-specific information below. For this case, the gene view reveals 77 copies of *EGFR* and suspected LOH. Beneath the gene view

chr	gene	start	end	log2	loh
chr1	TNFRSF14	2556609	2563330	-1.0000000	TRUE
chr1	ID3	23558900	23559456	0.0000000	TRUE
chr1	*a. Raw gene data*		9554	0.0000000	TRUE
chr1			9647	0.0000000	TRUE
chr1	MPL	43337798	43352825	0.0000000	TRUE
chr1	NRAS	114708505	114716194	0.0000000	TRUE
chr1	RIT1	155900384	155910971	0.1212955	TRUE
chr1	BRINP3	190097976	190454892	0.1212955	TRUE
chr2	MYCN	15942040	15946154	0.0000000	FALSE
chr2	DNMT3A	25234230	25314038	0.0000000	FALSE
chr2	FANCL	58159693	58241355	0.0000000	FALSE
chr2	XPO1	61478774	61533954	0.0000000	FALSE
chr2	CXCR4	136114865	136118143	0.0000000	FALSE
chr2	SF3B1	197392263	197435075	0.0000000	FALSE
chr2	IDH1	208237034	208251610	0.0000000	FALSE
chr3	FANCD2	10028600	10101314	0.0000000	FALSE
chr3	MYD88	38138606	38141329	0.0000000	FALSE
chr3	RHOA	49360176	49375631	0.0000000	FALSE
chr3	GATA2	128479417	128488536	0.0000000	FALSE
chr3	TBL1XR1	177025422	177065038	0.0000000	FALSE
chr3	KLHL6	183491898	183555686	0.0000000	FALSE
chr4	PDGFRA	54258703	54295318	0.0000000	TRUE
chr4	KIT	54657958	54738612	-1.0000000	TRUE
chr4	FBXW7	152322835	152411838	0.1212955	FALSE
chr5	IL7R	35856929	35876534	0.0000000	FALSE

chr	start	end	log2	loh
chr1	5992031	8012703	-1.0000000	TRUE
chr1	8013204	9249747	-Inf	FALSE
chr1	*b. Raw segment data*		000000	FALSE
chr1			0.0000000	FALSE
chr1	119915280	120030036	-1.0000000	FALSE
chr1	120030537	122026459	0.5849625	FALSE
chr1	124932724	125033747	0.5849625	FALSE
chr1	143491537	244728273	0.1212955	TRUE
chr2	8393035	92188145	0.0000000	FALSE
chr2	94090557	239190174	0.0000000	FALSE
chr3	10047815	90772458	0.0000000	FALSE
chr3	93655574	160555709	0.0000000	FALSE
chr3	160555710	167348377	-1.0013818	FALSE
chr3	167348878	197666556	0.0000000	FALSE
chr4	1595282	49161496	0.0000000	FALSE
chr4	49616455	49712061	0.0000000	TRUE
chr4	51743951	54657457	0.0000000	TRUE
chr4	54657958	55125349	-1.0000000	TRUE
chr4	55125850	105233368	0.0000000	TRUE
chr4	105233869	105276563	1.3219281	FALSE
chr4	105277064	179242244	0.1212955	FALSE
chr5	1295135	46485900	0.0000000	FALSE
chr5	50059807	122390572	0.0000000	FALSE
chr5	122391073	122453049	1.0000000	FALSE
chr5	122453550	154560075	0.0000000	FALSE

Fig. 7. Static Comparison. These figures demonstrate the difficulty of appreciating copy number data in tabulated and static form. (A) Raw Gene Data. While individual gene data can be appreciated, only so many genes can be viewed at once, and it is difficult to integrate individual gene findings into the overall copy number profile of the specimen. (B) Raw Segment Data. Raw segment data does not provide the opportunity to identify which genes fall in each segment. (C) Static Scatterplot. While chromosome-level events can be visualized with a static scatterplot, probe and gene-level data cannot be appreciated.

is a table of focal variants and a plot of the probe data. The tool permits investigation of parallel or subclonal events, such as the presence of *EGFRvIII*, a variant relevant in tumors of the central nervous system which often co-occurs with *EGFR* amplification.[35] Visual inspection of the probe plot, reveals a possible relative deletion between positions 55.14M and 55.155M (Fig. 5a). Additionally, the tool displays a cluster of variants around 55.154M, which the tool can further magnify. Hovering over each "X" displays the corresponding mutation identification information (Fig. 5b). Collectively, the amplification, relative deletion, SNVs, and small indels can all be identified and analyzed individually, but also with respect to one another, permitting these variants to be appreciated in their genomic context.

Clicking on the TCGA Pan-Cancer Atlas 2018 Data tab enables comparison of these findings to other publicly available glioblastoma multiforme cases. Selection of glioblastoma multiforme reveals 575 cases in the TCGA 2018 cohort. Sorting by most common amplifications, shows *EGFR* amplification is present in 44.35% of this cohort (Fig. 6a). Searching for *ROS1* shows that 27.3% of cases in the TCGA Pan-Cancer Atlas 2018 cohort had a heterozygous deletion (Fig. 6b).

This workflow may be repeated for any other gene or finding of interest. Analysis and interpretation of CNV data are greatly facilitated by this visualization tool. By comparison, review of a tab-delimited file of gene or segment data (Figs 7a, 7b), makes it much more difficult to appreciate and integrate this information. Static scatterplots (Fig. 7c) are useful for identifying large chromosomal changes, but not for deeper analysis of gene and probe-level data. CNViz enables a more extensive and holistic case review, potentially reducing errors, overlooked data, and increasing the opportunity for integrating findings.

Although there are previously published tools that are improvements on the aforementioned static visualizations, CNViz has a number of attributes that we believe make it a novel contribution to the landscape. For example, while IGV[15] displays copy number data, it provides only segment-level information. Similar to CNViz, CNspector[36] and reconCNV[37] both provide interactive visualization of probe-level copy number data, however, CNViz displays the data in a nested format enabling the user to navigate between chromosome, gene and probe-level views with a single click. This aspect makes the application less visually overwhelming, enabling the user to more immediately hone-in on key gene copy number changes. With CNViz, users can easily click on or search for genes of interest to view more detailed information including probe data, suspected LOH status, and SNVs. Additionally, CNViz displays sample metadata (e.g., evaluated purity and estimated ploidy) as well as provides TCGA data for comparison; this allows reviewers to aggregate meaningful data to accurately assess validity of a putative copy number event.

Future versions of the application will integrate functionality to visualize chromosomal rearrangements and gene fusions, the ability to view variant allele fraction plots as a complement to LOH status, additional cohorts for data comparison beyond the TCGA 2018 data, and information on the level of evidence associated with a copy number event.

A working example of CNViz is hosted on rebeccagreenblatt.shinyapps. io/cnviz_example.[32] Here, the interactive nature of the application and zoom functionality can be appreciated. In this example, only genes from the University of Pennsylvania's Center for Personalized Diagnostics[21] hematologic malignancies panel are displayed, but any number and combination of genes can be included in the visualization.

Conclusion

The CNViz R/Bioconductor package is a freely accessible, easy-to-use tool built with the intent of encouraging visualization and exploration of copy number variation. CNViz can be used in a clinical setting as well as for research to study patterns in human cancers more broadly. The user-friendly interface allows users to visualize the copy number profile of a specimen, zoom-in to explore gene and probe-level copy number changes, and simultaneously integrate LOH and focal variant data. Furthermore, users can compare findings to those seen in a cohort of patients with a similar diagnosis. Ideally, CNViz will

facilitate the identification of clinically meaningful copy number changes in cancer.

References

1. Bhattacharya A, Bense RD, Urzúa-Traslaviña CG, de Vries EGE, van Vugt MATM, Fehrmann RSN. Transcriptional effects of copy number alterations in a large set of human cancers. Nat Commun 2020;11(1):715.
2. Shao X, Lv N, Liao J, et al. Copy number variation is highly correlated with differential gene expression: a pan-cancer study. BMC Med Genet 2019;20(1):175.
3. Mikhail FM, Biegel JA, Cooley LD, et al. Technical laboratory standards for interpretation and reporting of acquired copy-number abnormalities and copy-neutral loss of heterozygosity in neoplastic disorders: a joint consensus recommendation from the American College of Medical Genetics and Genomics (ACMG) and the Cancer Genomics Consortium (CGC). Genet Med 2019;21(9):1903–1916.
4. Ohshima K, Hatakeyama K, Nagashima T, et al. Integrated analysis of gene expression and copy number identified potential cancer driver genes with amplification-dependent overexpression in 1,454 solid tumors. Scient Rep 2017;7(1):641.
5. Ried T, Meijer GA, Harrison DJ, et al. The landscape of genomic copy number alterations in colorectal cancer and their consequences on gene expression levels and disease outcome. Mol Aspects Med 2019;69:48–61.
6. Dong G, Mao Q, Yu D, et al. Integrative analysis of copy number and transcriptional expression profiles in esophageal cancer to identify a novel driver gene for therapy. Scient Rep 2017;7(1):42060.
7. Wan TS, Au WY, Chan JC, Chan LC, Ma SK. Trisomy 21 as the sole acquired karyotypic abnormality in acute myeloid leukemia and myelodysplastic syndrome. Leuk Res 1999;23(11):1079–1083.
8. D'haene B, Vandesompele J, Hellemans J. Accurate and objective copy number profiling using real-time quantitative PCR. Methods 2010;50(4):262–270.
9. Hu L, Ru K, Zhang L, et al. Fluorescence in situ hybridization (FISH): an increasingly demanded tool for biomarker research and personalized medicine. Biomark Res 2014;2(1):1-13.
10. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. PLoS Comput Biol 2016;12(4), e1004873.
11. Riester M, Singh AP, Brannon AR, et al. PureCN: copy number calling and SNV classification using targeted short read sequencing. Source Code Biol Med 2016;11(1):13.
12. Chandramohan R, Kakkar N, Roy A, Parsons DW. reconCNV: interactive visualization of copy number data from high-throughput sequencing. Bioinformatics 2021;37(8):1164–1167.
13. Bertini E, Lalanne D. Surveying the complementary role of automatic data analysis and visualization in knowledge discovery. Proceedings of the ACM SIGKDD Workshop on Visual Analytics and Knowledge Discovery: Integrating Automated Analysis with Interactive Exploration; 2009.
14. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. Genome Res 2002;12(6):996-1006.
15. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol 2011;29(1):24–26.
16. R: A Language and Environment for Statistical Computing 2020 [Available from: https://www.R-project.org. Accessed on: May 18, 2021]
17. Bioconductor: Open Source Software For Bioinformatics 2021 [Available from: Bioconductor.org. Accessed on: May 18, 2021]
18. RStudio. *Easy web applications in R*. 2013.
19. Sievert C. *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC. 2020.
20. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. Cell. 2018;173(2):400-16. e11.
21. Penn Center for Personalized Diagnostics 2020 [Available from: https://www.pennmedicine.org/departments-and-centers/center-for-personalized-diagnostics. Accessed on: May 18, 2021]
22. Okabe MI, Kei. Color Universal Design (CUD): How to make figures and presentations that are friendly to Colorblind people Japan: Jfly: data depository for drosophila researchers; 2002 [Available from: https://jfly.uni-koeln.de/color/. Accessed on: January 3]
23. Schneider VA, Graves-Lindsay T, Howe K, et al. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. Genome research 2017;27(5):849–864.
24. Gel B, Serra E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 2017;33(19):3088–3090.
25. Gel B, Magallon M. *CopyNumberPlots: Create Copy-Number Plots using karyoploteR functionality*. 2020.
26. Lawrence M, Huber W, Pages H, et al. Software for computing and annotating genomic ranges. Computat Biol 2013;9(8).
27. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discovery 2012;2(5):401–404.
28. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013;6(269).pl1-pl.
29. Ramos M, LG, Oh S, et al. Multiomic integration of public oncology databases in bioconductor. JCO Clin Cancer Inform 2020;4:958–971.
30. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 2011;12(4):1-14.
31. MSK-IMPACT. *A Targeted Test for Mutations in Both Rare and Common Cancers*. 2021.
32. Allaire J. *rsconnect: Deployment Interface for R Mardown Documents an Shiny Applications*. 2019.

33. RStudio. shinyapps.io 2020 [Available from: https://www.shinyapps.io/. Accessed on: March 8, 2021]

34. Amazon S3 2021 [Available from: https://aws.amazon.com/s3/. Accessed on: March 8, 2021]

35. Gan HK, Cvrljevic AN, Johns TG. The epidermal growth factor receptor variant III (EGFRvIII): where wild things are altered. FEBS J 2013;280(21):5350–5370.

36. Markham JF, Yerneni S, Ryland GL, et al. CNspector: a web-based tool for visualisation and clinical diagnosis of copy number variation from next generation sequencing. Scient Rep 2019;9(1):1–9.

37. Chandramohan R, Kakkar N, Roy A, Parsons DW. reconCNV: interactive visualization of copy number data from high-throughput sequencing. Bioinformatics 2021;37(8):1164–1167.