

# Shatterseek: an R package for the detection of chromothripsis from Next-Generation Sequencing (NGS) data

Isidro Cortes-Ciriano<sup>1,2,3</sup>, Ruibin Xi<sup>4</sup>, and Peter J. Park<sup>\*1,2</sup>

<sup>1</sup>*Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA*

<sup>2</sup>*Ludwig Center at Harvard, Boston, MA 02115, USA*

<sup>3</sup>*Centre for Molecular Science Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*

<sup>4</sup>*School of Mathematical Sciences and Center for Statistical Science, Peking University, Beijing 100871, China*

February 13, 2018

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Workflow implemented in Shatterseek</b>	<b>2</b>
2.1	Detection of clusters of interleaved SVs . . . . .	3
2.2	Statistical criteria implemented in Shatterseek . . . . .	4
2.2.1	Equal distribution of fragment joints . . . . .	4
2.2.2	Chromosomal enrichment of breakpoints . . . . .	4
2.2.3	Random distribution of breakpoints . . . . .	4
2.2.4	Number of oscillating copy number segments . . . . .	5

---

\*peter\_park@hms.harvard.edu

2.2.5	Interspersed loss of heterozygosity (LOH) . . . . .	5
2.3	Detection of multichromosomal chromothripsis events . . . . .	5
<b>3</b>	<b>Recommended cut-off values to intepret the output of Shatterseek</b>	<b>6</b>
<b>4</b>	<b>Shatterseek tutorial</b>	<b>7</b>
4.1	Installation . . . . .	7
4.2	Load SVs and CN data into R . . . . .	7
<b>5</b>	<b>Visualization of chromothripsis regions</b>	<b>11</b>
<b>6</b>	<b>Visualization of multichromosomal chromothripsis events</b>	<b>12</b>
<b>7</b>	<b>Bibliography</b>	<b>12</b>

## 1 Introduction

Chromothripsis refers to the genomic alterations characterized by massive de novo rearrangements, often generated in a single catastrophic event, where the DNA is shattered into a number of fragments that are subsequently stitched together in random order and orientation. Chromothripsis can be confined to small chromosomal regions or can affect multiple chromosomes, and can involve from tens to hundreds of rearrangements [1]. Therefore, chromothripsis represents a mechanism for the accrual of tens to hundreds of rearrangements in a few cell divisions.

Chromothripsis regions are characterized by copy number profiles oscillating between two or three states, (in cases where partial duplications follow chromothripsis), interspersed loss of heterozygosity (LOH), and clusters of interleaved SVs, where the proportions of fragment joints (i.e., duplication-like, deletion-like, head-to-head and tail-to-tail inversions) are roughly equal, consistent with the random stitching of genomic fragments through mostly non-homologous end-joining (NHEJ) DNA repair.

Initial studies used array-derived copy number profiles to detect chromothripsis, as the amount

of whole-genome sequencing data sets was limited [2], [3]. SNP arrays do not permit to fully characterize the genome-wide landscape of structural rearrangements at single-base resolution. Nevertheless, the detection of chromothripsis is more accurate if structural rearrangements and copy number data are integrated [4]–[7]. The increasing amount of whole-genome sequencing data sets thus requires the development of easy-to-use pipelines to detect chromothripsis from NGS data. Korbelt and Campbell [6] proposed a set of statistical criteria for the detection of chromothripsis that have been widely used in the literature. There exist two publicly available packages for the detection of chromothripsis, namely: CTLP scanner [3], [8] and Shatterproof [7]. The former uses SNP array data, whereas the latter CN and SV data.

We have developed Shatterseek, an R package that integrates copy number and SV data for the detection and visualization of chromothripsis events from NGS data. Shatterseek implements a custom graph-based approach to identify candidate chromothripsis regions, and then applies a set of statistical criteria based on Korbelt and Campbell [6] to detect both intra- and interchromosomal chromothripsis events. In addition, Shatterseek provides functionalities for the easy visualization of SVs, as well as CN and LOH profiles. Visual inspection of candidate chromothripsis regions is still required in a number of cases due to the complexity of the observed events, and the overlapping features of chromothripsis and other complex events. We have recently validated Shatterseek in a large-scale study of ca. 2,600 cancer genomes and its higher performance with respect to Shatterproof. We refer the reader to this work for further details about the rates and characteristics of chromothripsis across diverse human cancers. The chromothripsis calls for all these tumors generated using Shatterseek can be accessed at <http://compbio.med.harvard.edu/chromothripsis/>

In the next sections, we explain in detail the statistical criteria implemented in Shatterseek to detect chromothripsis events, and illustrate its functionalities using data from X cancer patients.

## 2 Workflow implemented in Shatterseek

To identify chromothripsis-like patterns in cancer genomes, we implemented and extended the set of criteria proposed by Campbell and Korbelt [6]. The pipeline to detect chromothripsis consists of three major steps, namely: (i) discovery of clusters of interleaved SVs using a graph-based approach, (ii) evaluation of a set of statistical criteria in the genomic regions spanned by these clusters, and (iii) evaluation of whether chromothripsis is confined to a single or to multiple chromosomes.

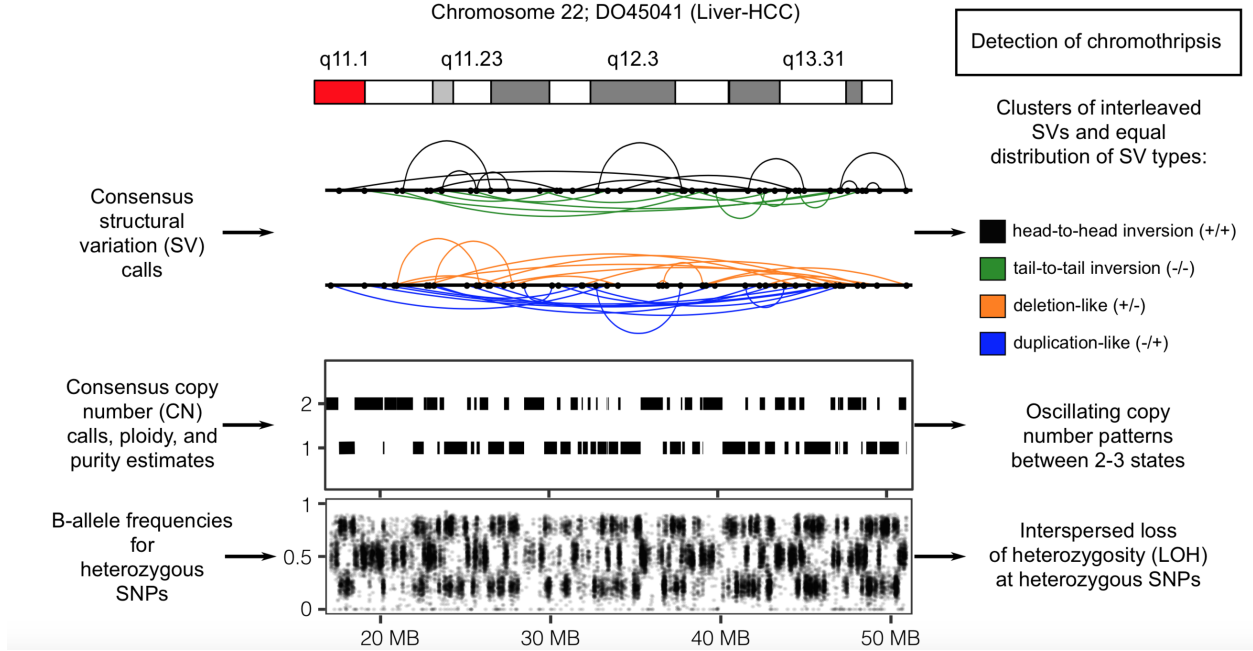


Figure 1: ssdfT

## 2.1 Detection of clusters of interleaved SVs

Given that chromothripsis events generate clusters of interleaved rearrangements, we firstly scanned the cancer genomes for the presence of clusters of interleaved SVs. We consider that two SVs are interleaved if the genomic regions bridged by their breakpoints overlap but are not nested. To find clusters within a given chromosome, Shatterseek constructs an undirected graph using intrachromosomal SVs whose nodes correspond to SVs and whose edges connect interleaved SVs. Thus, clusters of SVs can be detected by finding the connected components in the graph. The connected component in each chromosome with the highest number of SVs is considered for further analysis. By default, Shatterseek only considers chromosomes 1-22 and X. In the following, we use the term *chromothripsis region* to refer to the genomic regions affected by chromothripsis within one chromosome, whereas we term *chromothripsis event* the set of chromothripsis regions involved in a single catastrophic event. Thus, a chromothripsis event can affect a single chromosome, multiple regions within a chromosome, or comprise genomic regions from multiple chromosomes.

## 2.2 Statistical criteria implemented in Shatterseek

Once the clusters of interleaved SVs are detected, Shatterseek evaluates the following statistical criteria based on the work by Korbelt and Campbell [6].

### 2.2.1 Equal distribution of fragment joints

Once the SV clusters are detected, Shatterseek evaluates whether the distribution of DNA fragment joints, i.e., deletion-like (+/-; or "head"/"tail"), duplication-like (-/+), head-to-head (+/+), and tail-to-tail (-/-) inversions, diverges from a multinomial distribution with equal probabilities for each SV category using the goodness-of-fit test for the multinomial distribution using the `stats` package (`chisq.test` function). The types of SVs are determined by the orientation of the reads mapped to the breakpoints (see [9] for further details).

### 2.2.2 Chromosomal enrichment of breakpoints

Massive chromothripsis events, which involve hundreds of SVs, generally affect entire or multiple chromosomes against the background of quiescent genomes. Thus, in these cases the chromosomes harboring chromothripsis are highly enriched for breakpoints. An example of a chromothripsis event is depicted in Figure 1.

To test for the enrichment of breakpoints in each chromosome, Shatterseek uses the binomial test corrected for mappability. We note that this test might be misleading for focal chromothripsis events appearing in highly rearranged genomes, where other chromosomes not displaying chromothripsis harbor a number of SVs (interleaved or not) comparable to that detected in chromosomes displaying chromothripsis. Therefore, this criterion needs to be interpreted carefully on a per-case basis.

### 2.2.3 Random distribution of breakpoints

Shatterseek also evaluates whether the distribution of the breakpoints comprised the SV clusters differs from an exponential as described by Korbelt and Campbell [6].

### 2.2.4 Number of oscillating copy number segments

A major feature of chromothripsis that can be detected from copy number data is the presence of copy number profiles oscillating between two or three copy number states. This feature has been widely used before to detect chromothripsis from genome-wide copy number profiles using operational definitions, e.g., at least 10 contiguous segments oscillating between two copy number states [10].

To uncover oscillating patterns in the genomic regions delimited by the distal breakpoints composing the clusters of interleaved SVs, Shatterseek examines whether contiguous genomic segments oscillate between two or three CN states.

### 2.2.5 Interspersed loss of heterozygosity (LOH)

We decided not to use the criteria of loss of heterozygosity proposed by Campbell and Korbel [6], as we noticed that some bona fide chromothripsis cases identified in low-purity samples might not meet this criterion due to the infiltration of normal tissue. In such a case, the allelic ratios for heterozygous SNPs (i.e., B allele frequencies) would not divert significantly from 0.5, and thus, would hamper the observation of alternating LOH patterns associated to copy losses [11]. In addition, assessing LOH in aneuploid tumors is often difficult, where the BAF values for oscillating segments between high CN levels do not vary strongly.

However, given that the loss of heterozygosity profiles are very useful to determine the temporal profile of chromothripsis events, and to distinguish chromothripsis from chromoanasythesis in nearly-diploid cases, Shatterseek provides capabilities to visualize LOH profiles at chromothripsis regions.

## 2.3 Detection of multichromosomal chromothripsis events

Shatterseek considers that two or more chromothripsis regions belong to the same catastrophic event if these regions +/- 10Kb (default value) are linked by at least two interchromosomal SVs. Given that the sensitivity of SV detection algorithms is still limited, it is not always possible to detect rearrangements between all regions belonging to the same chromothripsis event. Thus, in cases where at least three chromosomes were involved, Shatterseek applies transitive reasoning to identify

the full extent of the events. For instance, if the chromothripsis regions detected in chromosomes 1 and 2 are linked, and those detected in chromosome 2 are also linked to a chromothripsis region in chromosome 3, Shatterseek considers that the chromothripsis patterns detected in these three chromosomes were generated as a result of the same catastrophic event.

### 3 Recommended cut-off values to interpret the output of Shatterseek

After manual curation of hundreds of massive and focal chromothripsis calls, we derived the following guidelines to detect chromothripsis using Shatterseek.

We assign two levels of confidence depending on the set of statistical tests satisfied by a candidate chromothripsis region.

- High confidence: at least 6 interleaved intrachromosomal SVs, 7 contiguous segments oscillating between 2 CN states, the fragment joints test, and either the chromosomal enrichment or the exponential distribution of breakpoints test.
- High confidence: at least 3 interleaved intrachromosomal SVs and 4 or more interchromosomal SVs, 7 contiguous segments oscillating between 2 CN states and the fragment joints test.
- Low confidence: at least 6 interleaved intrachromosomal SVs, 4, 5 or 6 adjacent segments oscillating between 2 CN states, the fragment joints test, and either the chromosomal enrichment or the exponential distribution of breakpoints test.

Application of these criteria to 2,600 whole genomes still led to a small set of false positives that were removed by visual inspection. This is mostly due to the fact that multiple layers of rearrangements other than chromothripsis might coexist, and the aggregate of these might satisfy the statistical tests described above. In addition, chromothripsis regions are highly heterogeneous, running the gamut from massive chromothripsis cases involving hundreds of SVs in an otherwise quiescent genome, to focal events in the context of genomes highly rearranged by other mutational processes. This heterogeneity makes it almost impossible to define a set of criteria with perfect discriminative power for chromothripsis. Therefore, we advocate, as far as possible, for the manual curation of candidate chromothripsis regions. To facilitate this process, Shatterseek provides functionalities to visually depict the candidate chromothripsis regions.

## 4 Shatterseek tutorial

In the following sections, we illustrate how to install and use Shatterseek to detect and visualize chromothripsis using SV and CN data. This tutorial assumes minimal knowledge of the R programming language.

### 4.1 Installation

Shatterseek is entirely written in R. To install Shatterseek type the following in R:

```
sessionInfo()  
require(devtools)  
install_github()
```

Alternatively, please download and install Shatterseek by running in a bash terminal:

```
wget XXXXXXXXXX  
R CMD INSTALL shatterSeek_0.4
```

### 4.2 Load SVs and CN data into R

XX data from any caller can be used XX ein PCAWG consensus, etc.. XX total copy number in integer format (output of e.g. ASCAT)

We first load Shatterseek and the test data that is provided with the package. This corresponds to the SV and CN data for a kidney renal cell carcinoma tumor (ICGC ID: DO17373).

```
library(shatterSeek)  
data(DO17373)
```

Running this command loads to the workspace two R dataframes, corresponding to the somatic SV (SV\_DO17373) and CNV (SCNA\_DO17373) calls.

Shatterseek requires the SV data to be stored in a data.frame with the following columns:



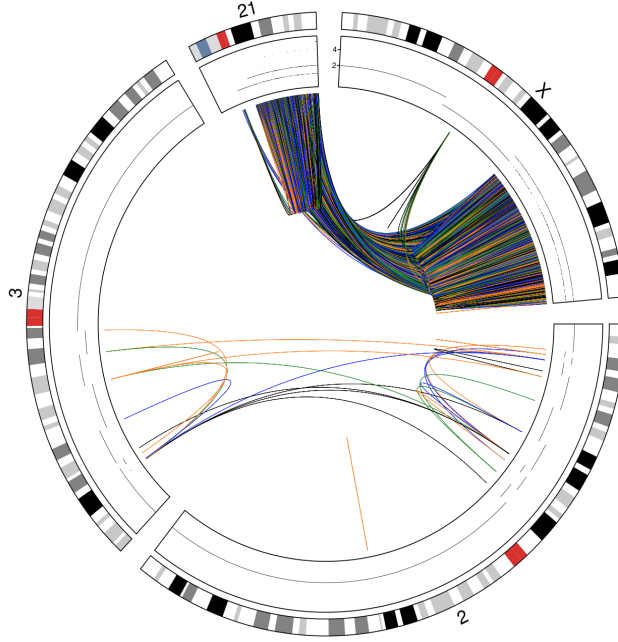


Figure 2: Evidence of two massive multichromosomal chromothripsis events involving chromosomes 21-X and 2-3 detected in a kidney renal cell carcinoma tumor in patient DO17373 (TCGA-CJ-5681).

- chrom1 (character): chromosome for the first of the two breakpoints composing each SV
- pos1 (character): position for the first breakpoint
- chrom2 (character): chromosome for the second breakpoint composing each SV
- pos2 (character): position for the second breakpoint
- SVtype (character): type of SV, encoded as: DEL (deletion-like; +/-), DUP (duplication-like; -/+), h2hINV (head-to-head inversion; +/+), and t2tINV (tail-to-tail inversion; -/-).

Chromosomes are expected to be in Ensembl chromosome notation, i.e. NOT contain the prefix "chr".

```
head(SV_D017373)
```

```
##   chrom1   start1   end1 chrom2   start2   end2   sv_id
## 1      1 142618685 142618686    21 21708092 21708093 SVMERGE1159
```

```
## 2      1 142623413 142623414      21 25233137 25233138  SVMERGE269
## 3      1 142639910 142639911      21 33703275 33703276  SVMERGE127
## 4      1 143536371 143536372      21 29698088 29698089  SVMERGE1343
## 5     11   8540384   8540385     11  8541772  8541773   SVMERGE55
## 6     14  61896353  61896354     14 62061595 62061596  SVMERGE1006
##      pe_support strand1 strand2 svclass          svmethod
## 1         18      +      -      TRA          SNOWMAN_DELLY
## 2         50      -      -      TRA      SNOWMAN_dRANGER_DELLY
## 3         34      +      -      TRA          SNOWMAN_dRANGER
## 4         39      -      +      TRA          SNOWMAN_dRANGER
## 5         37      +      -      DEL      SNOWMAN_BRASS_DELLY
## 6        111      -      +      DUP SNOWMAN_BRASS_dRANGER_DELLY
```

Please remember that Shatterseek only considers chromosomes 1-22 and X. Thus, make sure that the SVs comprised in your input data only correspond to these chromosomes. The SV data is loaded into an object of class "SV", using the function *SVs*:

```
SV_data <- SVs(chrom1=as.character(SV_D017373$chrom1),
               pos1=as.numeric(SV_D017373$start1),
               chrom2=as.character(SV_D017373$chrom2),
               pos2=as.numeric(SV_D017373$end2),
               SVtype=as.character(SV_D017373$svclass),
               strand1=as.character(SV_D017373$strand1),
               strand2=as.character(SV_D017373$strand2))
```

Shatterseek requires the CN to be in the following format: (i)

- chrom (character): chromosome (also in Ensembl notation)
- start (character): start position for the CN segment
- end (character): end position for the CN segment
- CN (character): absolute copy number value

```
head(SCNA_D017373)
```

##	chromosome	start	end	total_cn
## 1	1	1	249250620	2
## 3	2	20016	11969465	2
## 4	2	11969466	14420187	1
## 5	2	14420188	16916033	2
## 6	2	16916034	16937471	1
## 7	2	16937472	17054487	2

The CN data is loaded into an object of class "CNVsegs", using the function *CNVsegs*:

```
CN_data <- CNVsegs(chrom=as.character(SCNA_D017373$chromosome),  
                  start=SCNA_D017373$start,  
                  end=SCNA_D017373$end,  
                  log2copyratio=SCNA_D017373$total_cn)
```

Once the input data has been loaded, we proceed to run the main function of the package, namely *shatterseek*. This function runs the code to detect clusters of interleaved SVs, and subsequently evaluates each candidate chromothripsis region for the statistical criteria described above.

```
chromothripsis <- shatterseek(SV.sample=SV_data, seg.sample=CN_data)  
  
## Error in Rle(SV.df$chrom1[ind.chr]): could not find function "Rle"  
  
print(chromothripsis@chromSummary)  
  
## Error in print(chromothripsis@chromSummary): object 'chromothripsis' not found
```

The 'clusterSize' column indicates the size of the largest cluster of interleaved SVs detected per chromosome.

```
chromothripsis@chromSummary = statistical_criteria(chromothripsis)  
  
## Error in statistical_criteria(chromothripsis): object 'chromothripsis' not found
```

The function *shatterseek* returns a list containing two objects:

- `sd`
- `chromSummary`: a `data.frame` containing the results for all statistical criteria described above and additional information. Each row corresponds to a chromosome (i.e., 1-22, and X), whereas the columns contain the following information:
  - `sdf`
  - 
  - 
  - 
  - 
  -

## 5 Visualization of chromothripsis regions

```
library(ggplot2)
source("shatterSeek_0.4/R/plot_function.R")

## Error in file(filename, "r", encoding = encoding): cannot open the connection

plots = plot_chromothripsis(shatterSeek_output = chromothripsis, chr = 22)

## Error in plot_chromothripsis(shatterSeek_output = chromothripsis, chr = 22): object
'chromothripsis' not found

grid.arrange(plots[[1]], plots[[2]], plots[[3]], plots[[4]], ncol=1)

## Error in grid.arrange(plots[[1]], plots[[2]], plots[[3]], plots[[4]], : could not
find function "grid.arrange"
```

## 6 Visualization of multichromosomal chromothripsis events

## 7 Bibliography

### References

- [1] P. J. Stephens, C. D. Greenman, B. Fu, F. Yang, G. R. Bignell, L. J. Mudie, E. D. Pleasance, K. W. Lau, D. Beare, L. A. Stebbings, S. McLaren, M.-L. Lin, D. J. McBride, I. Varela, S. Nik-Zainal, C. Leroy, M. Jia, A. Menzies, A. P. Butler, J. W. Teague, M. A. Quail, J. Burton, H. Swerdlow, N. P. Carter, L. A. Morsberger, C. Iacobuzio-Donahue, G. A. Follows, A. R. Green, A. M. Flanagan, M. R. Stratton, P. A. Futreal, and P. J. Campbell, “Massive genomic rearrangement acquired in a single catastrophic event during cancer development.,” *Cell*, vol. 144, no. 1, pp. 27–40, 2011, ISSN: 1097-4172. DOI: 10.1016/j.cell.2010.11.055. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21215367><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3065307>.
- [2] T.-M. Kim, R. Xi, L. J. Luquette, R. W. Park, M. D. Johnson, and P. J. Park, “Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes.,” *Genome research*, vol. 23, no. 2, pp. 217–27, 2013, ISSN: 1549-5469. DOI: 10.1101/gr.140301.112. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/23132910><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3561863>.
- [3] H. Cai, N. Kumar, H. C. Bagheri, C. von Mering, M. D. Robinson, and M. Baudis, “Chromothripsis-like patterns are recurring but heterogeneously distributed features in a survey of 22,347 cancer genome screens,” *BMC Genomics*, vol. 15, no. 1, p. 82, 2014, ISSN: 1471-2164. DOI: 10.1186/1471-2164-15-82. [Online]. Available: <http://bmcbgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-82>.
- [4] F. Notta, M. Chan-Seng-Yue, M. Lemire, Y. Li, G. W. Wilson, A. A. Connor, R. E. Denroche, S.-B. Liang, A. M. K. Brown, J. C. Kim, T. Wang, J. T. Simpson, T. Beck, A. Borgida, N. Buchner, D. Chadwick, S. Hafezi-Bakhtiari, J. E. Dick, L. Heisler, M. A. Hollingsworth, E. Ibrahimov, G. H. Jang, J. Johns, L. G. T. Jorgensen, C. Law, O. Ludkovski, I. Lungu, K. Ng, D. Pasternack, G. M. Petersen, L. I. Shlush, L. Timms, M.-S. Tsao, J. M. Wilson, C. K. Yung, G. Zogopoulos, J. M. S. Bartlett, L. B. Alexandrov, F. X. Real, S. P. Cleary, M. H. Roehrl, J. D. McPherson, L. D. Stein, T. J. Hudson, P. J. Campbell, and S. Gallinger, “A renewed model of pancreatic cancer evolution based on genomic rearrangement patterns,” *Nature*, vol. 538, no. 7625, pp. 378–382, 2016, ISSN: 0028-0836. DOI: 10.1038/nature19823. [Online]. Available: <http://www.nature.com/doifinder/10.1038/nature19823>.

- [5] Y. Li, C. Schwab, S. L. Ryan, E. Papaemmanuil, H. M. Robinson, P. Jacobs, A. V. Moorman, S. Dyer, J. Borrow, M. Griffiths, N. A. Heerema, A. J. Carroll, P. Talley, N. Bown, N. Telford, F. M. Ross, L. Gaunt, R. J. Q. McNally, B. D. Young, P. Sinclair, V. Rand, M. R. Teixeira, O. Joseph, B. Robinson, M. Maddison, N. Dastugue, P. Vandenberghe, C. Haferlach, P. J. Stephens, J. Cheng, P. Van Loo, M. R. Stratton, P. J. Campbell, and C. J. Harrison, "Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia.," *Nature*, vol. 508, no. 7494, pp. 98–102, 2014, ISSN: 1476-4687. DOI: 10.1038/nature13115. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24670643><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3976272>.
- [6] J. O. Korbel and P. J. Campbell, "Criteria for inference of chromothripsis in cancer genomes.," *Cell*, vol. 152, no. 6, pp. 1226–36, 2013, ISSN: 1097-4172. DOI: 10.1016/j.cell.2013.02.023. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/23498933>.
- [7] S. K. Govind, A. Zia, P. H. Hennings-Yeomans, J. D. Watson, M. Fraser, C. Anghel, A. W. Wyatt, T. van der Kwast, C. C. Collins, J. D. McPherson, R. G. Bristow, and P. C. Boutros, "ShatterProof: operational detection and quantification of chromothripsis," *BMC Bioinformatics*, vol. 15, no. 1, p. 78, 2014, ISSN: 1471-2105. DOI: 10.1186/1471-2105-15-78. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24646301><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3999944><http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-15-78>.
- [8] J. Yang, J. Liu, L. Ouyang, Y. Chen, B. Liu, and H. Cai, "CTLPScanner: a web server for chromothripsis-like pattern detection," *Nucleic Acids Research*, vol. 44, no. W1, W252–W258, 2016, ISSN: 0305-1048. DOI: 10.1093/nar/gkw434. [Online]. Available: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw434>.
- [9] C.-Z. Zhang, M. L. Leibowitz, and D. Pellman, "Chromothripsis and beyond: rapid genome evolution from complex chromosomal rearrangements.," *Genes & development*, vol. 27, no. 23, pp. 2513–30, 2013, ISSN: 1549-5477. DOI: 10.1101/gad.229559.113. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/24298051><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3861665>.
- [10] T. Rausch, D. Jones, M. Zapatka, A. Stütz, T. Zichner, J. Weischenfeldt, N. Jäger, M. Remke, D. Shih, P. Northcott, E. Pfaff, J. Tica, Q. Wang, L. Massimi, H. Witt, S. Bender, S. Pleier, H. Cin, C. Hawkins, C. Beck, A. vonDeimling, V. Hans, B. Brors, R. Eils, W. Scheurlen, J. Blake, V. Benes, A. Kulozik, O. Witt, D. Martin, C. Zhang, R. Porat, D. M. Merino, J. Wasserman, N. Jabado, A. Fontebasso, L. Bullinger, F. G. Rücker, K. Döhner, H. Döhner, J. Koster, J. Molenaar, R. Versteeg, M. Kool, U. Tabori, D. Malkin, A. Korshunov, M. Taylor, P. Lichter, S. Pfister, and J. Korbel, "Genome Sequencing of Pediatric Medulloblastoma Links Catastrophic DNA Rearrangements with TP53 Mutations," *Cell*, vol. 148, no. 1-2,

pp. 59–71, 2012, ISSN: 00928674. DOI: 10.1016/j.cell.2011.12.013. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0092867411015169>.

- [11] S. Song, K. Nones, D. Miller, I. Harliwong, K. S. Kassahn, M. Pinese, M. Pajic, A. J. Gill, A. L. Johns, M. Anderson, O. Holmes, C. Leonard, D. Taylor, S. Wood, Q. Xu, F. Newell, M. J. Cowley, J. Wu, P. Wilson, L. Fink, A. V. Biankin, N. Waddell, S. M. Grimmond, and J. V. Pearson, “qpure: A Tool to Estimate Tumor Cellularity from Genome-Wide Single-Nucleotide Polymorphism Profiles,” *PLoS ONE*, vol. 7, no. 9, A. H. Ting, Ed., e45835, 2012, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0045835. [Online]. Available: <http://dx.plos.org/10.1371/journal.pone.0045835>.