The Journal of Molecular Diagnostics, Vol. 23, No. 12, December 2021

ELSEVIER

the Journal of Molecular Diagnostics
jmdjournal.org

Automated Detection of Arm-Level Alterations for Individual Cancer Patients in the Clinical Setting

Yann Christinat,* Prasad Chaskar,†‡ Sophie Clément,* Liza Ho,* Mélinda Charrier,‡ Thomas McKee,* and Petros Tsantoulis†‡

From the Division of Pathology, Department of Diagnosis, and the Department of Oncology,‡ Geneva University Hospitals; and the Faculty of Medicine,† Department of Medicine, University of Geneva, Geneva, Switzerland*

Accepted for publication August 18, 2021.

Address correspondence to Petros Tsantoulis, M.D., Ph.D., Department of Oncology, Hôpitaux Universitaires de Genève, Rue Gabrielle Perret-Gentil 4, Geneva 1204, Switzerland; or Yann Christinat, Ph.D., Division of Pathology, Diagnostic Department, Hôpitaux Universitaires de Genève, Cluse-Roseraie Centre Médical Universitaire (CMU), Niveau 5, 9I.E5.2949, Geneva, Switzerland. E-mail: petros.tsantoulis@hcuge.ch or yann.christinat@hcuge.ch.

Copy number alterations are genetic events that promote tumor initiation and progression and are used in clinical care as diagnostic, prognostic, and predictive biomarkers. Based on the length of the alteration, they are roughly classified as focal and arm-level alterations. Although genome-wide techniques to detect arm-level alterations are gaining momentum in hospital laboratories, the high precision and novelty of these techniques pose new challenges: there is no consensus on the definition of an arm-level alteration and there is a lack of tools to compute them for individual patients. Based on 376 clinical samples analyzed with the OncoScan formalin-fixed, paraffin-embedded assay, a bimodal distribution of the percentage of bases with copy number alterations within a chromosomal arm was observed, with the second peak starting at 90% of arm length. Two approaches were tested for the definition of arm-level alterations: sum of altered segments (SoS) >90%, or the longest segment (LS) >90%. These approaches were validated against expert annotation of 25 clinical cases. The SoS method outperformed the LS method as indicated by a higher concordance (SoS, 95.2%; LS, 79.9%). Some of the discordances ultimately were attributed to human error, highlighting the advantages of automation. The increase in reliability led to the development of publicly available software and its inclusion into routine clinical practice at Geneva University Hospitals. *(J Mol Diagn 2021, 23: 1722–1731; https://doi.org/10.1016/j.jmoldx.2021.08.003)*

Copy number alterations (CNAs) can be gains, losses, or loss of heterozygosity (LOH) of a chromosome segment. Based on the length of the altered segment, they are crudely classified as focal alterations and arm-level alterations. Focal alterations are shorter and harbor losses of tumor-suppressor genes (eg, *PTEN*, *ARID1A*, *CDKN2A*, and *ATM*) and gains or amplifications of oncogenes (eg, *MYC*, *EGFR*, *CCND1*, and *ERBB2*).[1,2] Arm-level alterations are broader and contain hundreds of genes, and although their contribution to the cancer phenotype is clear, it has been complicated and sometimes impossible to ascribe this effect to individual genes.[3]

Improvements in genomics technology and the availability of large amounts of data have helped provide insights into the potential prognostic value of arm-level alterations in cancer.[4–7] For example, the co-loss of chromosome arms 1p and 19q is one of the essential diagnostic criteria for oligodendrogliomas (World Health Organization 2016 classification of tumors of the central nervous system[8]) and is associated with a favorable prognosis in patients receiving radiotherapy, chemotherapy, or both.[9] In colorectal and cervical cancers, the status of chromosome arm CNAs has been used as a marker of the disease stage.[10,11] Loss of chromosomal arm 3p is frequent in squamous cell carcinomas from several sites,[12] hyperdiploidy defines a diagnostic entity among B-lymphoblastic leukemia/lymphomas,[13] and the ploidy category is a prognostic group in multiple myelomas.[13] A recent review article by Ben-David and Amon[4] redefined the concept of aneuploidy (alterations of

Supported by the Ligue Genevoise Contre le Cancer (P.C. and P.T.), and the Philanthropia Foundation (M.C.).

Y.C. and P.C. contributed equally to this work.

Disclosures: None declared.

https://doi.org/10.1016/j.jmoldx.2021.08.003

chromosome arms) and summarized its prognostic impact in several cancer types. Failure to convert many of these important observations into clinical practice probably is related to the fact that they help refine diagnostic and prognostic categories but are not always integrated into treatment decisions. In addition, the current commercially available sequencing approaches use integrated methods that identify focal amplifications and deletions, but not arm-length alterations.

The classic clinical tools to assess copy number variation of specific genes are fluorescence *in situ* hybridization and chromogenic *in situ* hybridization assays. However, single-nucleotide polymorphism arrays and, more recently, next-generation sequencing, are commonly used for genome-wide copy number profiling of tumor samples.[3,14–16] The tumor biopsy specimens obtained in a clinical setting are often in the form of formalin-fixed, paraffin-embedded (FFPE) samples, and genomic DNA extracted from such samples is often of lower yield and can be degraded. The OncoScan assay (ThermoFisher Scientific, Waltham, Massachusetts, United States), in contrast to the more research-oriented sixth generation of Affymetrix SNP (SNP6) array, is designed and optimized for the prediction of CNAs from FFPE samples and is a popular choice in clinical settings.

One of the widely used methods for the identification and classification of CNAs from single-nucleotide polymorphism array data is Genomic Identification of Significant Targets in Cancer (GISTIC)[17] from the Broad Institute (*https://www.broadinstitute.org*, last accessed March 23, 2021). The initial version of GISTIC used a statistical scoring approach based on the amplitude and frequency of CNAs at each position in the genome to separate the CNAs into focal and arm-level alterations. In GISTIC, the amplitude is a free parameter and hence subject to change across different research studies. The current version of GISTIC (GISTIC2.0) uses a probabilistic scoring approach that performs a data-driven estimation of the background rates of CNA. One limitation of this method is that the trends produced by GISTIC2.0 are dependent on the characteristics of the measuring platform, which may change over time.[18]

Furthermore, the current definitions of arm-level alterations are based on batch-level population studies[3,18] and are not necessarily robust in a single-sample clinical use case. For example, in a study by Roy et al,[19] an arm-level alteration of the arm 9p was defined as an alteration of 70% of the chromosome arm. In another study by Mermel et al,[18] a threshold of 98% was used to separate focal alterations from arm-level alterations. Several other articles have used a context-dependent proportion to separate focal events from the arm-level events and there is no standard agreement on the optimal cut-off value. The present study attempts to provide a precise definition of arm-level alterations that can be applied in a per-sample clinical setting.

Materials and Methods

Study Cohort

The Clinical Pathology Department of the Geneva University Hospitals (Geneva, Switzerland) performed OncoScan FFPE assays for more than 400 patients as part of their routine laboratory analyses from 2016 to 2018. The re-use of data from 376 patients who were not opposed to the re-use of their data for research was authorized by the ethics committee of the Geneva canton (CCER-2020-02377).

Among these 376 samples, 25 were selected manually to validate our method against the expert annotations from the clinical report. Based on the content of the clinical reports, the selection was made to represent a diverse range of arm-level alterations in terms of chromosomal distribution, CNA type (gain versus loss of copies), tumor ploidy, and the number of arms altered.

DNA Extraction and Copy Number Profiling

Genomic DNA was purified from FFPE tumor tissues using the QIAamp DNA FFPE Tissue Kit (cat. 56404; QIAGEN, Hilden, Germany) and quantified using the Quant-iT dsDNA HS Assay Kit (cat. Q32854; Life Technologies, Carlsbad, CA). Copy number variation analysis was performed with 80 ng genomic DNA using the OncoScan FFPE Assay Kit (cat. 902695; ThermoFisher Scientific) following the manufacturer's instructions. The arrays were stained in GeneChip Fluidics Station (ThermoFischer Scientific) and scanned using the Gene Chip scanner (ThermoFisher Scientific, Waltham, Massachusetts, United States). CEL files (Affymetrix DNA microarray image analysis software output) generated from the scanned array image were converted to OncoScan array data (OSCHP files) and analyzed using Chromosome Analysis Suite (ChAS) software (version 4.0; ThermoFisher). Reference files for the OncoScan copy number variation array (NetAffxGenomicAnnotations.Homo_sapiens, hg19; NA33.r1) were downloaded from the ThermoFisher website. As part of the standard clinical procedure, the copy number variation segments were evaluated manually in ChAS and exported in a text file format. Parameters in ChAS were adapted from the default settings to exclude segments with fewer than 50 markers or smaller than 50 kbp (kilo base pairs) (Figure 1A).

Estimation of Percentage Arm Alteration

The segmented copy number information provided by ChAS software was used to detect arm-level alterations, and then was processed further with a custom script (*https://github.com/yannchristinat/oncoscanR-public*, last accessed March 23, 2021) in the R programming language (*https://www.r-project.org*, last accessed December 11, 2019).

CNA segments were further classified into four categories: gain, when there are one or two extra copies with respect to the diploid state; amplification, in case of a gain of three or more copies; loss, when the number of copies is lower than the normal number (two in a human genome); and loss of heterozygosity (LOH), when there is a loss of the maternal or paternal allele without any loss of copies.

Along with this information, the position and length of each chromosomal arm covered by the OncoScan array was extracted from the reference files mentioned in the above section (DNA Extraction and Copy Number Profiling). All arms are represented in a text file contains the following columns: name of the arm, starting position, and end position.

Because any technique has a detection limit and biases, two filters were applied to the set of segments: trimming of small segments to remove artifacts and smoothing to account for the imprecision of the segment positions (Figure 1, B and C). The trimming filter consists of removing segments whose length is below a certain threshold. The smoothing filter was performed for each CNA category (gain, loss, LOH, and amplification) by merging segments closer than a defined threshold into a single segment (Figure 1, B and C and Supplemental Figure S1).

A commonly accepted but imprecise definition of an arm-level alteration is that the altered segment should be a continuous and broad region of the chromosome arm.[19] To refine this definition, two different approaches were tested (Figure 1C). In the first approach, called sum of all segments (SoS), the percentage of arm altered (PAA) was obtained by summing the length of all the segments in a chromosome arm for a given CNA category and normalizing by the length of the chromosomal arm. Note the assumption of nonoverlapping segments. In the second approach, called longest continuous segment (LS), the PAA was calculated using only the length of the single longest continuous segment within the chromosome arm with the same normalization method. The division by the length of the arm is required because chromosomal arms have quite different lengths (from 15 to 148 Mbp (mega base pairs)).

The two methods can be described mathematically as follows: where A is the set of all chromosomal arms, T is the set of all CNA types (gain, loss, amplification, and LOH), and $S_{a,t}$ is the set of all CNA segments of type t within arm a.

$$SoS:\ \forall a \in A,\ \forall t \in T,\ PAA_{a,t} = \sum_{s \in S_{a,t}} length_s \Big/ length_a \quad (1)$$

$$LS:\ \forall a \in A, \forall t \in T,\ PAA_{a,t} = \max_{s \in S_{a,t}} (length_s) \Big/ length_a \quad (2)$$

The methods were applied on the set of segments restricted to each CNA category and after trimming and smoothing of data. Note that amplified segments are included in the computation of the gain category. A

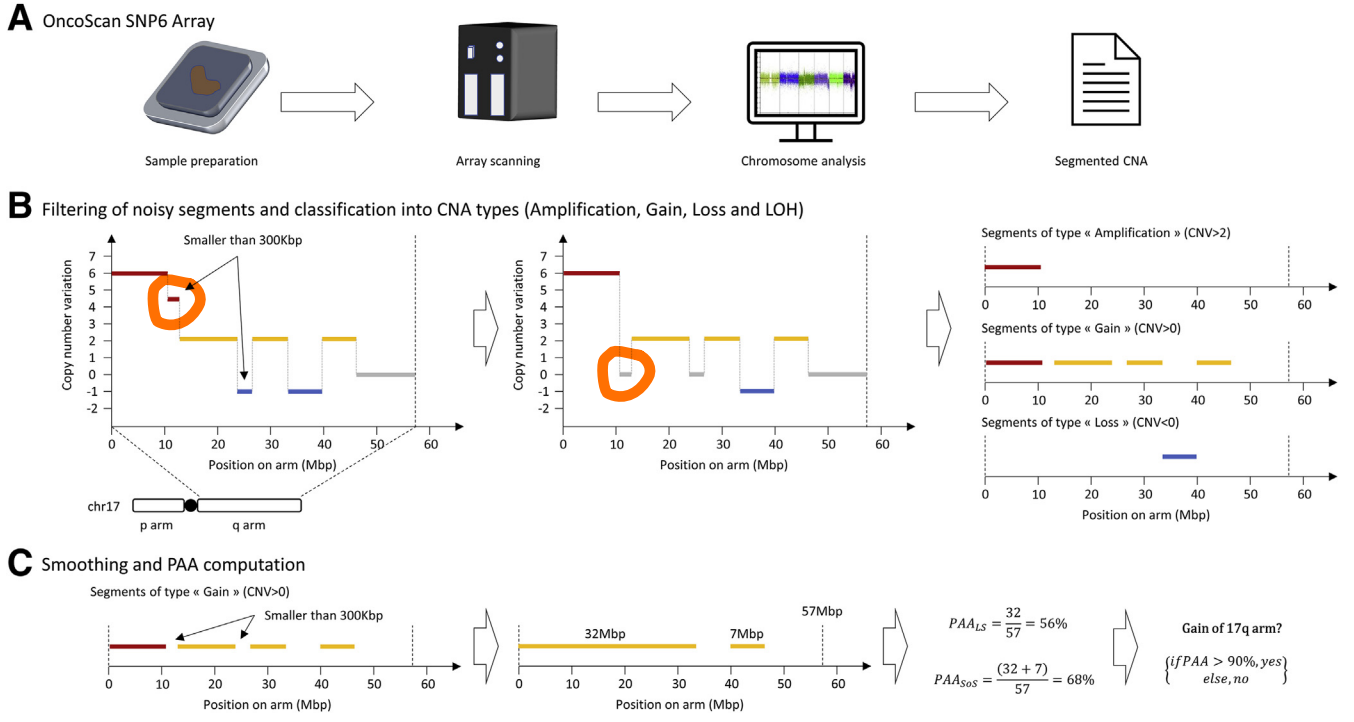

Figure 1 Flow chart for arm-level copy number alteration detection in the clinical setting. **A:** OncoScan (ThermoFisher Scientific) assay sample preparation and data processing. **B:** Trimming filter. **C:** Smoothing filter and estimate of the percentage of arm altered using both the longest segment and the sum of all segments approaches.

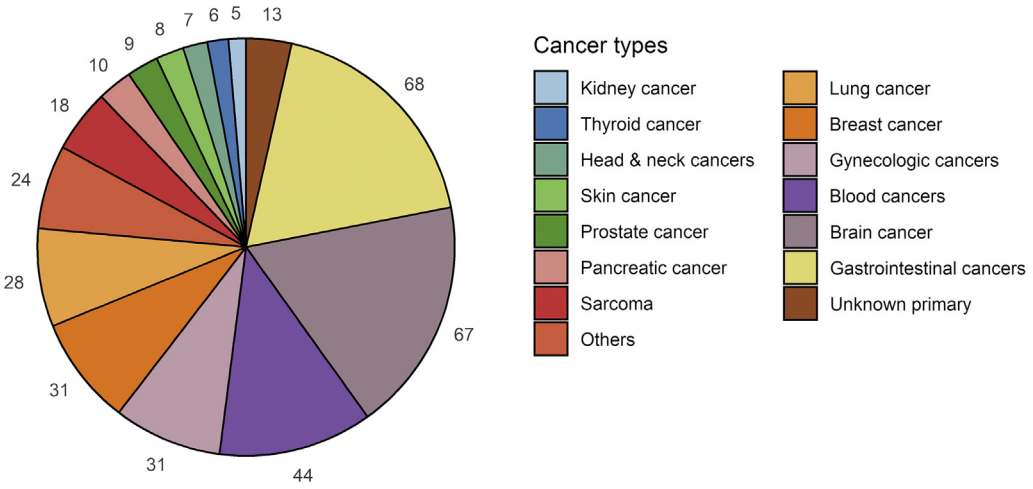

Figure 2 Cancer types covered in the present study.

chromosomal arm is defined as altered if the PAA is higher than a given threshold.

The filtering, smoothing, and PAA computation methods are contained in an R package available on Github (*https://github.com/yannchristinat/oncoscanR-public*; v0.1.1). In practice, any text file input with one line per segment and three tab-separated columns named CN State, Type, and Full Location can be processed. The three columns describe, respectively, the absolute number of copies, the type of CNA (gain, loss, or LOH), and the chromosomal position of the segment (ie, *chr7:55219200-55234200*).

Results

Cohort Description

The cohort consisted of 376 samples from Geneva University Hospitals in Switzerland. There was no bias in sex distribution (49.7% male). The median age at the time of analysis was 59 years for women (interquartile range, 19 y) and 60 years for men (interquartile range, 20 y). Ethnicity was not recorded, but given the lack of exclusion criteria, should reflect the ethnically diverse population of cancer patients in Geneva. The primary cancer types covered were gastrointestinal cancers (68), brain cancer (67), blood cancers (44), breast cancer (31), gynecologic cancers (31), lung cancer (28), sarcoma (18), pancreatic cancer (10), prostate cancer (9), skin cancer (8), head and neck cancers (7), thyroid cancer (6), and kidney cancer (5). This data set also included 24 other minor cancer types, 13 cases of cancer of unknown primary origin, and 7 samples from non-tumoral tissue (Figure 2).

Effect of Data Preprocessing

Array-based techniques such as comparative genomic hybridization or next-generation sequencing are popular methods for the study of CNAs in cancer. The resolution is the minimum size of the segment detectable by a given assay. Segments smaller than this value are most likely artifacts and can be removed. The OncoScan assay has a genome-wide resolution of 300 kbp and an even finer resolution of 50 to 100 kbp on approximately 900 cancer genes. Herein, the effect of different filtering thresholds on the PAA was tested. Four different thresholds were tested: 0 bp (no trimming or smoothing), 300 kbp (the OncoScan resolution), and 1 Mbp (three times the OncoScan resolution) and 3 Mbp (10 times the OncoScan resolution). The same threshold was applied for the trimming and smoothing procedures because it accounted for the imprecision level.

The current results show a bimodal distribution of the PAA for both gain and loss, with a first peak ending at approximately 10% and a second starting after 90%, irrespective of the method used for PAA computation (Figures 3 and 4). Herein, the distribution of PAA (gain/loss) for the SoS method remained unchanged after filtering of likely artifacts that occurred below assay resolution (0 to 300 kbp), or at 1 Mbp but not at 3 Mbp. For the LS method (Figure 4), the trimming and smoothing at 300 kbp drastically changed the distribution. Many chromosomal arms that were between 10% and 50% PAA were pushed toward higher values. The most striking was the histogram peak near 100 PAA that increased from 500 arms to 1500 arms. Unlike with the SoS method, an increase from 300 kbp to 1 Mbp showed a noticeable change in the high end of the PAA distribution. Therefore, for further analysis, segments smaller than the resolution of the OncoScan method (300 kbp) were eliminated and segments less than 300 kbp apart were joined because this seemed to provide a more robust procedure, irrespective of the method used for the PAA computation.

Next, the PAA based on the use of the LS method was compared against the SoS method. As seen in Figure 5, the LS approach yielded a lower PAA and this behavior was exacerbated in arms with a high PAA, independently of the alteration type. This was expected because the longest segment is always included in the sum of all segments (SoS), hence $PAA_{LS} \leq PAA_{SoS}$. Nonetheless, if 90% PAA is used as a threshold to classify an arm as altered, both approaches are

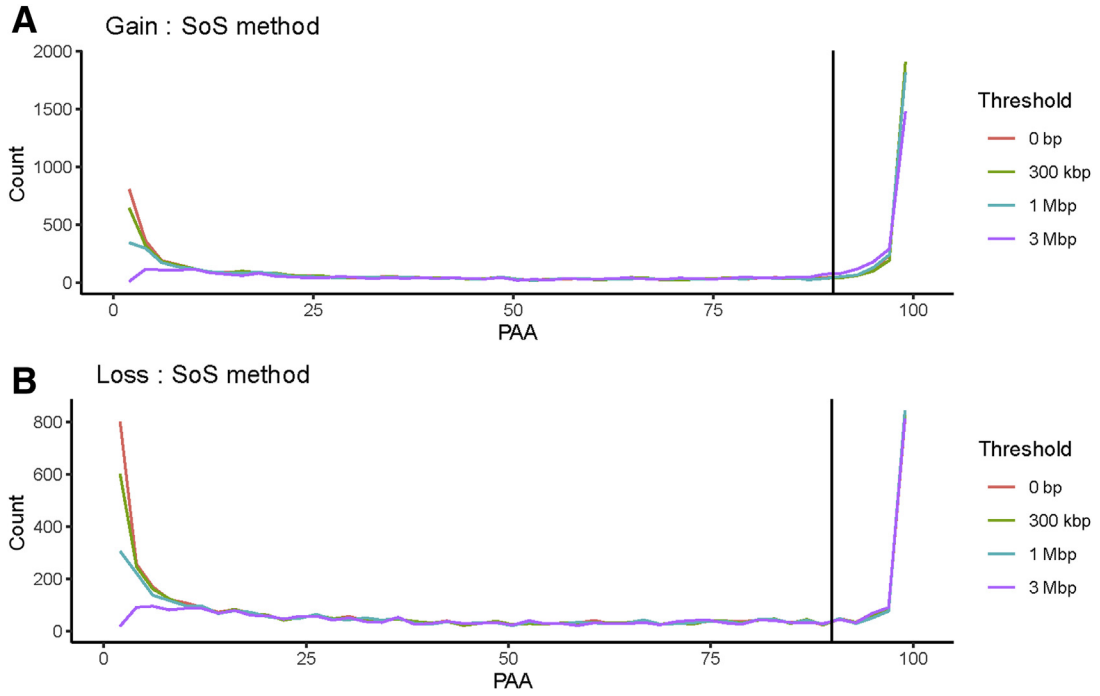

Figure 3 Effect of different trimming and smoothing thresholds on the percentage of arm altered (PAA) distribution for the sum of all segments (SoS) method. **A:** Segments with gains of extra copies. **B:** Segments with copy losses. Arms with a PAA of 0 (61% of cases for gain and 71% of cases for loss) were removed for clarity.

highly consistent with very few discordant observations (Table 1) (97.1% agreement; Cohen's κ[20] = 0.912). However, in case of the altered cases, the SoS method classified roughly 16% more arms as altered (gain, loss, or amplification) than the LS method (3371 versus 2913).

Validation against Human Annotations

In the present study, the PAA had a bimodal distribution with a first peak at approximately 10% and a second peak starting at approximately 90% (Figures 3 and 4). This

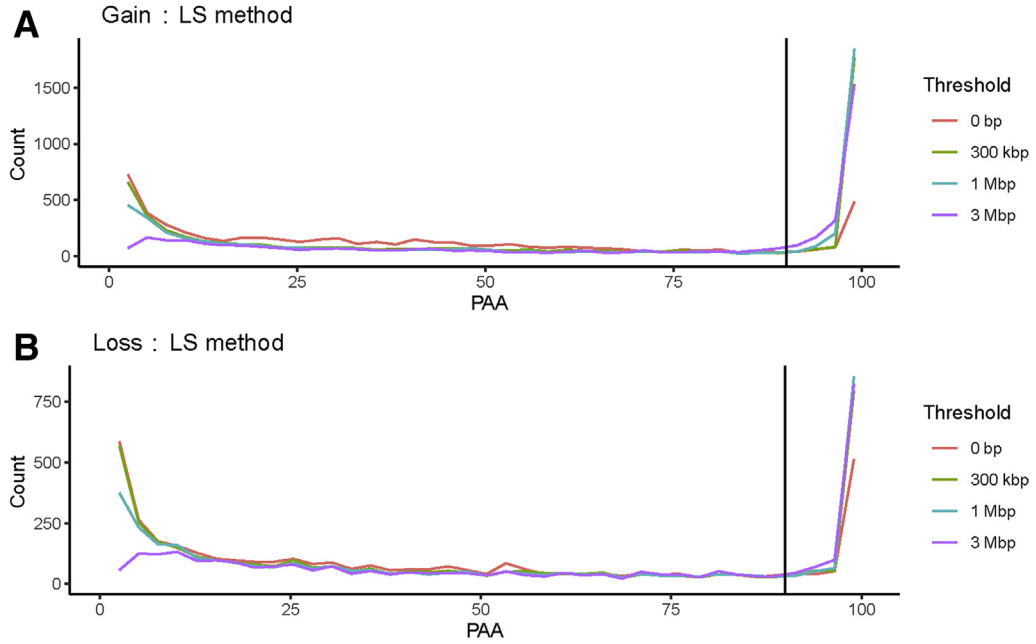

Figure 4 Effect of different trimming and smoothing thresholds on the percentage of arm altered (PAA) distribution for the longest segment (LS) method. **A:** Segments with gains of extra copies. **B:** Segments with copy losses. Arms with a PAA of 0 (61% of cases for gain and 71% of cases for loss) were removed for clarity.

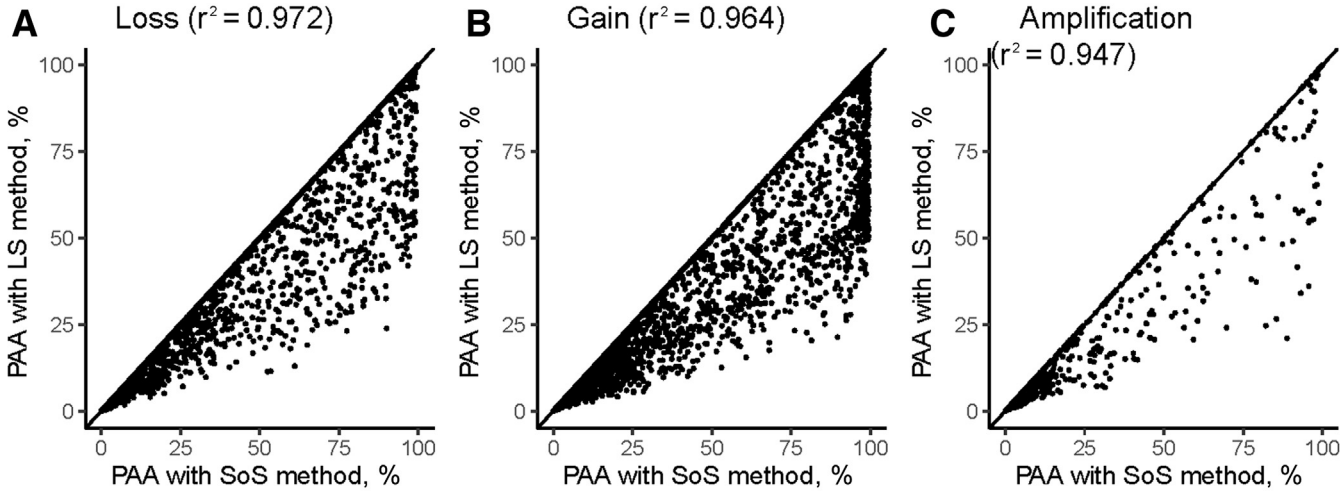

Figure 5 **A**–**C:** Discordant arms between the sum of all segments (SoS) and the longest segment (LS) approach for the different copy number alteration categories: percentage arm lost (**A**), percentage arm gained (**B**), and percentage arm amplified (**C**). PAA, percentage of arm altered.

observation was used to separate smaller focal events from the broader arm-level events by setting up a PAA threshold of 90%, above which the whole arm was considered to be altered. Algorithmic estimates of arm-level alterations were compared against manual annotation in a cohort of 25 patients (Figure 6) for both the LS and SoS methods. Compared with the LS approach, the SoS method correlated better with the expert annotations.

The SoS approach correctly predicted 122 of 126 (96.8%) arm-level gains, 96 of 101 (95.0%) arm-level losses, 16 of 19 (84.2%) arm-level LOH events, and 3 of 3 (100%) arm-level amplifications (Supplemental Figure S2A). On the other hand, the LS approach correctly predicted 99 of 126 (78.6%) arm-level gains, 82 of 101 (81.2%) arm-level losses, 15 of 19 (78.9%) LOH events, and 3 of 3 (100%) arm-level amplifications (Supplemental Figure S2B).

In general, both computational approaches (SoS and LS) showed a high number of arm-level alterations (gain, 27; loss, 14; and LOH, 8), which were missed by the manual annotation but detected by the current approach. An in-depth verification showed that two of the 12 false-negative results were caused by human mistakes in the original reports. For example, in one of the test cases, the analyst (human annotator) reported a loss of the "p" arm of chromosome 1 (1p); however, the karyotype indicated no such deletion of 1p (Supplemental Figure S3). The current approach predicted a PAA of approximately 85% in the remaining 10 false-negative results, which was just below the 90% cut-off point of arm-level alteration. Cases with true-positive results also were verified independently, and no discordant results were found.

Table 1 Classification of Chromosomal Arms When Using LS or SoS with a Threshold of 90% PAA

LS or SoS method				
	Not altered	Loss	Gain	Amplification
Not altered	79.28%	0.00%	0.00%	0.00%
Loss	0.66%	5.18%	0.00%	0.00%
Gain	2.12%	0.00%	11.64%	0.00%
Amplification	0.06%	0.00%	0.08%	0.38%

The values represent the percentage of chromosomal arms.

LS, longest segment; PAA, percentage of arm altered; SoS, sum of all segments.

To evaluate the performance in highly altered cases, four cases with alterations reported on almost every chromosome and a high aneuploidy were selected (H19008810, H19007081, H19005930, and H19009958). The performance in calling arm-level alterations was comparable with that of the 25 cases (recall, 96%; precision, 80%).

Arm-level alteration classification accuracy was evaluated on a series of PAA thresholds to re-assess the choice of the current PAA cut-off value (90%). These results indicated that in comparison with the smaller PAA cut-off values of 80% and 85%, a PAA of 90% yields a balanced accuracy (Figure 7). Here, the PAA of 95% slightly increased precision, but the gain from 90% to 95% was weak and resulted in a higher rate of false-negative results. These results justify the choice of a PAA cut-off value of 90%.

Arm-Level Alterations in Cancer

The arm-level alteration frequency measured across 376 copy number profiles from different cancer types highlighted the frequent arm-level copy number gains in 8q (30.5%), 7p (29.2%), 1q (28.5%), and 20q (27.1%); losses in 17p (19.4%), 10p (14.1%), 10q (14.1%), 18q (13.3%), and 9p (12.5%); and LOH in 17p (12.8%) (Supplemental Figure S4).

Arm-level alterations have a context-specific prognostic impact in several cancers.[4] To investigate the cancer

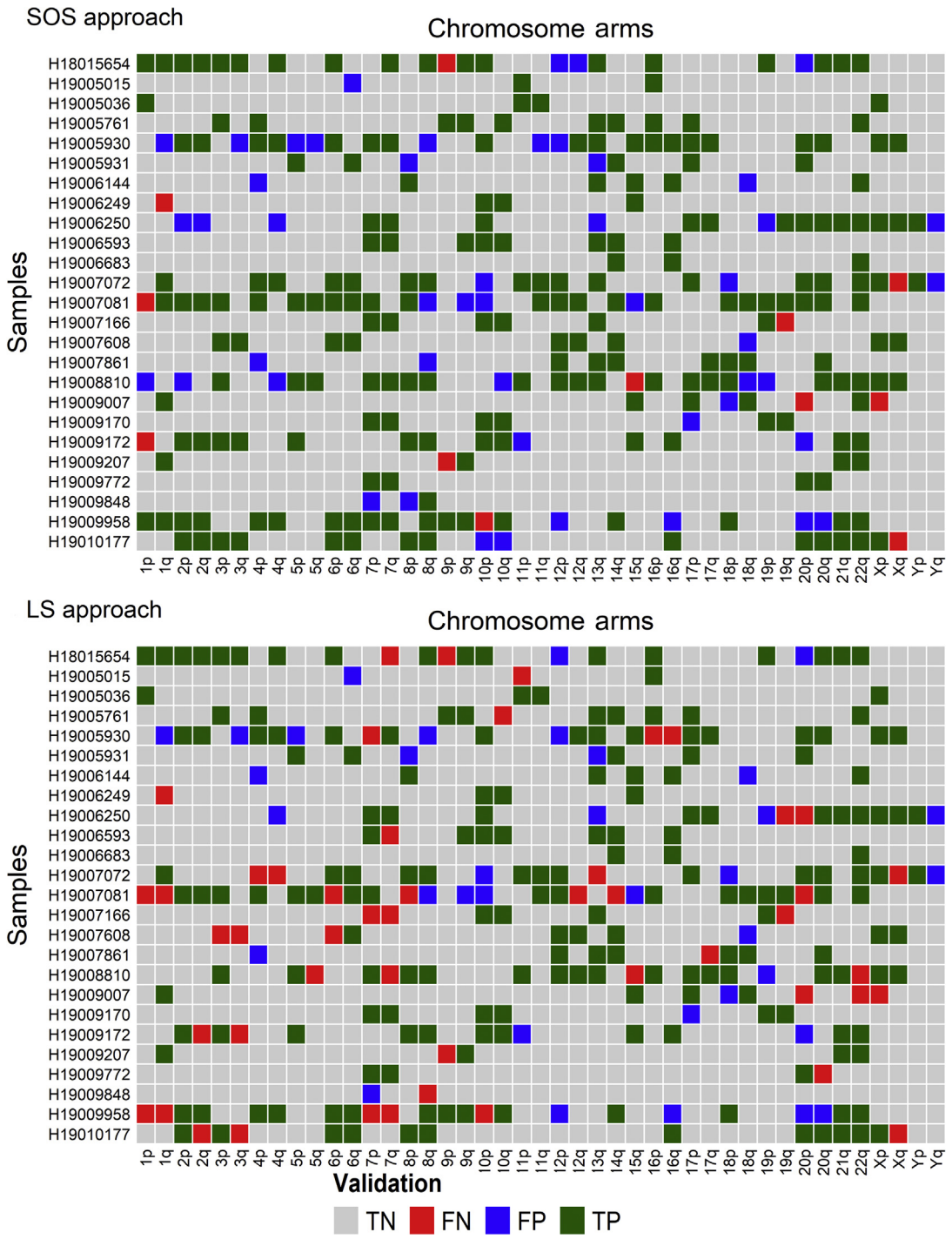

Figure 6 Validation of arm-level alteration prediction against human interpretation. Sum of all segments (SoS) approach (**top**) and longest segment (LS) approach (**bottom**). True negatives (TN) are shown in gray, false negatives (FN) are shown in red, false positives (FP) are shown in blue, and true positives (TP) are shown in green. CNA, copy number alteration; LOH, loss of heterozygosity.

type−specific arm-level alterations in the current data set, a subset of brain cancers (67 samples) was chosen. The subtypes of brain cancers covered in this study were as follows: astrocytoma (14), glioblastoma (28), glioma (20), medulloblastoma (2), and neuroblastoma (3). In most CNA profiles from patients with glioblastoma and astrocytoma, arm-level gains were observed in 7q and 7p and arm-level losses were observed in 10p and 10q (Figure 8). This is corroborated by Stichel et al,[21] who proposed the potential use of chromosome 7 gain/chromosome 10 loss (7+/10−) as a diagnostic biomarker for reclassifying *IDHwt* astrocytoma to *IDHwt* glioblastoma.

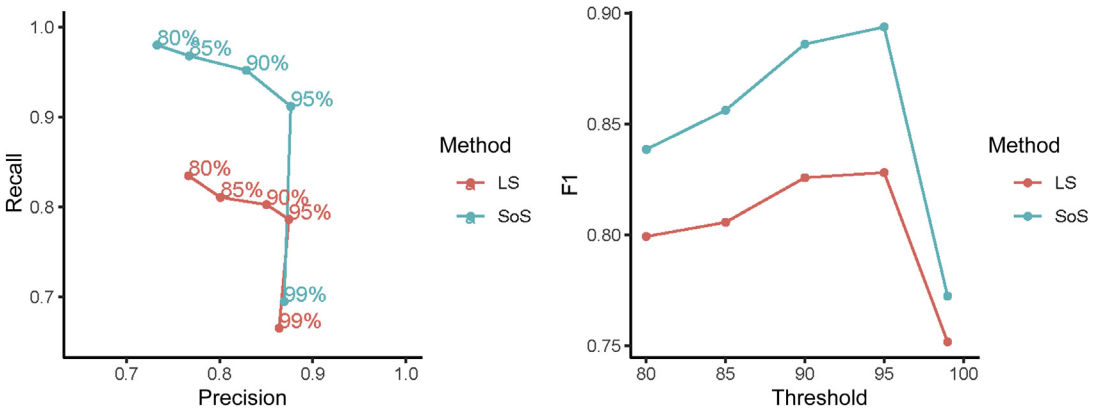

Figure 7 Arm-level alteration classification accuracy based on the different percentage of arm altered thresholds. Performance metric recall (**left**) and F1 score (**right**) are shown. LS, longest segment; SoS, sum of all segments.

In-depth analysis showed a frequent co-loss (or LOH) of chromosome arms 1p and 19q in 15 brain cancer cases, which is a known genetic signature of oligodendrogliomas and is associated with a favorable prognosis for patients receiving radiotherapy or chemotherapy[22] (Figure 8). We could confirm that our method correctly identified the 1p/19q co-loss in all cases diagnosed as adult oligodendrogliomas (15 of 15) and that no other brain cancer sample

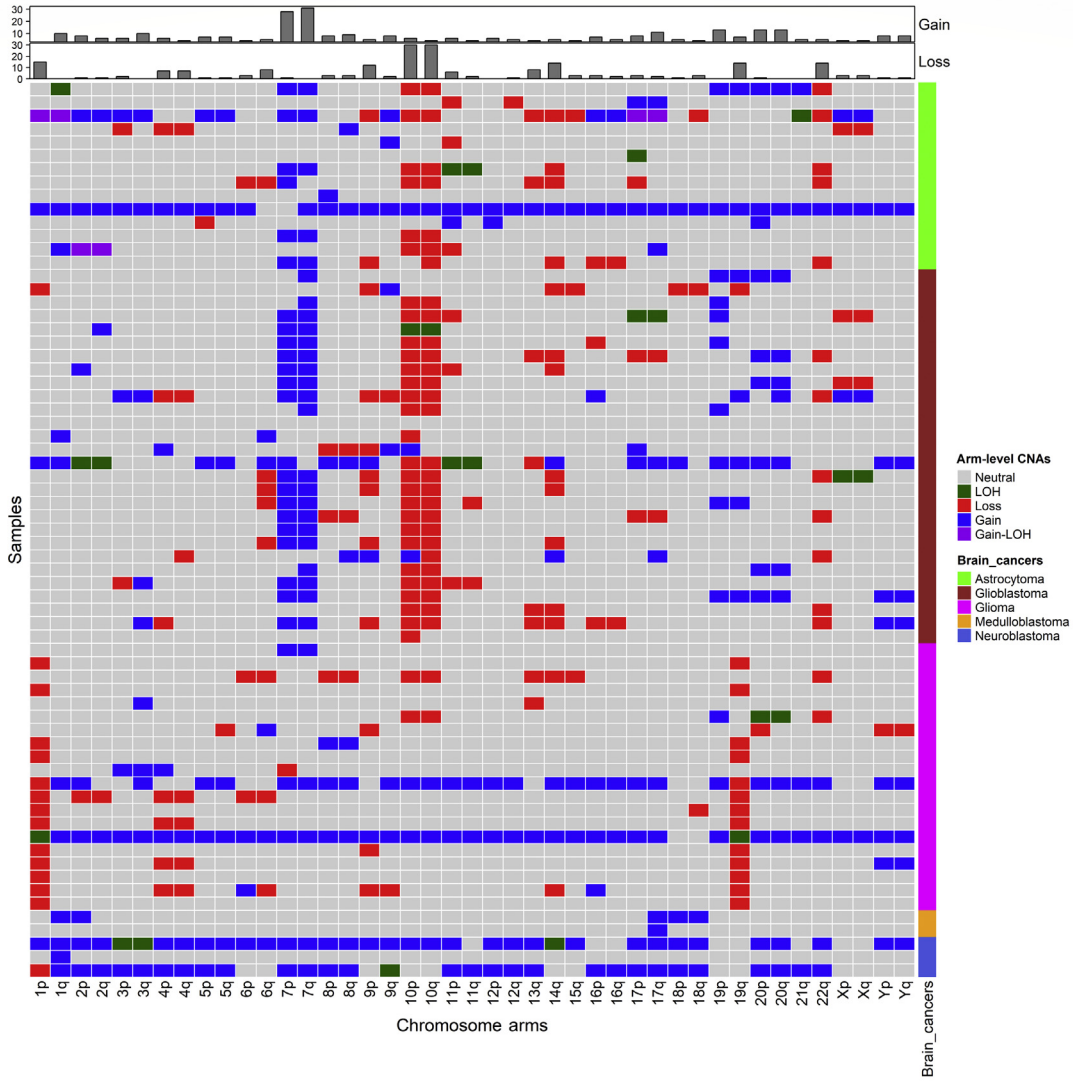

Figure 8 Landscape of chromosome arm-level copy number alterations in brain cancers. CNA, copy number alteration; LOH, loss of heterozygosity.

with a 1p/19q co-deletion had a diagnosis other than adult oligodendroglioma.

Discussion

The current definitions of arm-level alterations are based on batch-level population studies[3,18] and are not applicable in a single-sample clinical use case. The present study attempted to provide a precise definition of arm-level alterations that could be applied in a per-sample clinical setting. During validation, our algorithm successfully identified the vast majority of manually annotated arm-level alterations. Computational classification also uncovered alterations otherwise missed by manual reporting.[23]

In an extensive study of copy number profiles (8227 profiles from 107 cancer studies) performed by Kim et al,[23] the investigators presented a global view of CNAs in cancer. Some of the most recurrent alterations proposed by the above study[23], such as a frequent gain of chromosome arms 8q, 20q, and 1q, and a loss of 17p, were picked up successfully by our approach (Supplemental Figure S4). Most of these chromosome arms harbor important oncogenes such as *MYC* (8q), *BCL2L1* (20q), and *MCL1* (1q), and tumor-suppressor genes such as *TP53* (17p). The current study assessed the tumor-specific, arm-level CNA in a subset of brain cancers (67 samples). This approach identified the combined gain of chromosome 7 and loss of chromosome 10, which is characteristic of glioblastomas without isocitrate dehydrogenase mutations.[21,24] Furthermore, a focused analysis on chromosome arms 1p and 19q confirmed the validity of our method as a diagnostic tool for oligodendrogliomas.

Tumors, however, can contain more than a few chromosomal arm alterations. Whole-genome doubling events or aneuploidy can lead to highly altered tumors, as seen in the hyperdiploid entity of B-lymphoblastic leukemias. More generally, high-level aneuploidy has been observed in advanced cancer of various types, such as gastroesophageal and colorectal cancer.[4] Although these are only a minority of tumors, the presence of a highly altered genome could have performance implications for the detection of arm-level alterations. Our software was specifically tested on four selected, highly altered, polyploid samples with good results. Nevertheless, our method relied on good quality copy number variation calling, such as that provided by the ChAS software and the OncoScan assay upstream. Polyploid tumors can be extremely challenging in this respect. Naturally, when copy number variation calls are difficult or uncertain, the downstream results of our method should be interpreted with care.

Conclusion

Overall, our computational method is highly accurate and robust for detecting CNAs across diverse cancer types in a clinical setting. This method performs as accurately as human experts but at a fraction of the time. A software tool also increases the reliability because typographic and annotation errors were observed in some manually curated cases. The method and tool described herein are now routinely used in the Department of Clinical Pathology at Geneva University Hospitals and are available to the community.

Acknowledgments

We thank the Department of Oncology at University of Geneva Hospital and the University of Geneva (UNIGE) for permitting us to conduct this research.

Supplemental Data

Supplemental material for this article can be found at *http://doi.org/10.1016/j.jmoldx.2021.08.003*.

References

1. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, Lawrence MS, Zhsng C-Z, Wala J, Mermel CH, Sougnez C, Gabriel SB, Hernandez B, Shen H, Laird PW, Getz G, Meyerson M, Beroukhim R: Pan-cancer patterns of somatic copy number alteration. Nat Genet 2013, 45:1134−1140
2. Myllykangas S, Böhling T, Knuutila S: Specificity, selection and significance of gene amplifications in cancer. Semin Cancer Biol 2007, 17:42−55
3. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al: The landscape of somatic copy-number alteration across human cancers. Nature 2010, 463:899−905
4. Ben-David U, Amon A: Context is everything: aneuploidy in cancer. Nat Rev Genet 2020, 21:44−62
5. Gross AM, Orosco RK, Shen JP, Egloff AM, Carter H, Hofree M, Choueiri M, Coffey CS, Lippman SM, Hayes DN, Cohen EE, Grandis JR, Nguyen QT, Ideker T: Multi-tiered genomic analysis of head and neck cancer ties TP53 mutation to 3p loss. Nat Genet 2014, 46:939−943
6. Bown N, Cotterill S, Łastowska M, O'Neill S, Pearson ADJ, Plantaz D, Meddeb M, Danglot G, Brinkschmidt C, Christiansen H, Laureys G, Nicholson J, Bernheim A, Betts DR, Vandesompele J, Van Roy N, Speleman F: Gain of chromosome arm 17q and adverse outcome in patients with neuroblastoma. N Engl J Med 1999, 340: 1954−1961
7. Caron H, van Sluis P, de Kraker J, Bökkerink J, Egeler M, Laureys G, Slater R, Westerveld A, Voûte PAA, Versteeg R: Allelic loss of chromosome 1p as a predictor of unfavorable outcome in patients with neuroblastoma. N Engl J Med 1996, 334:225−230
8. Louis D, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee W, Ohgaki H, Wiestler O, Kleihues P, Ellison D: The 2016 World Health Organization classification of tumors of the central nervous system: a summary. Acta Neuropathol 2016, 131:803−820
9. Weller M, Stupp R, Hegi ME, van den Bent M, Tonn JC, Sanson M, Wick W, Reifenberger G: Personalized care in neuro-oncology coming of age: why we need MGMT and 1p/19q testing for malignant glioma patients in clinical practice. Neuro Oncol 2012, 14 Suppl 4:iv100−iv108
10. Laubert T, Freitag-Wolf S, Linnebacher M, König A, Vollmar B, Habermann JK: Stage-specific frequency and prognostic significance of

aneuploidy in patients with sporadic colorectal cancer—a meta-analysis and current overview. Int J Colorectal Dis 2015, 30:1015–1028
11. Heselmeyer K, Schröck E, Du Manoir S, Blegen H, Shah K, Steinbeck R, Auer G, Ried T: Gain of chromosome 3q defines the transition from severe dysplasia to invasive carcinoma of the uterine cervix. Proc Natl Acad Sci U S A 1996, 93:479–484
12. Taylor AM, Shih J, Ha G, Gao GF, Zhang X, Berger AC, Schumacher SE, Wang C, Hu H, Liu J, Lazar AJ; Cancer Genome Atlas Research Network, Cherniack AD, Beroukhim R, Meyerson M: Genomic and functional approaches to understanding cancer aneuploidy. Cancer Cell 2018, 33:676–689.e3
13. Swerdlow SH, Campo E, Pileri SA, Lee Harris N, Stein H, Siebert R, Advani R, Ghielmini M, Salles GA, Zelenetz AD, Jaffe ES: The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood 2016, 127:2375–2390
14. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, Girtman K, Mathew S, Ma J, Pounds SB, Su X, Pui CH, Relling MV, Evans WE, Shurtleff SA, Downing JR: Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature 2007, 446:758–764
15. Van Loo P, Nordgard SH, Lingjærde OC, Russnes HG, Rye IH, Sun W, Weigman VJ, Marynen P, Zetterberg A, Naume B, Perou CM, Børresen-Dale A-L, Kristensen VN: Allele-specific copy number analysis of tumors. Proc Natl Acad Sci U S A 2010, 107:16910–16915
16. Matsuzaki H, Dong S, Loi H, Di X, Liu G, Hubbell E, Law J, Berntsen T, Chadha M, Hui H, Yang G, Kennedy GC, Webster TA, Cawley S, Walsh PS, Jones KW, Fodor SPA, Mei R: Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. Nat Methods 2004, 1:109–111
17. Beroukhim R, Getz G, Nghiemphu L, Barretina J, Hsueh T, Linhart D, Vivanco I, Lee JC, Huang JH, Alexander S, Du J, Kau T, Thomas RK, Shah K, Soto H, Perner S, Prensner J, Debiasi RM, Demichelis F, Hatton C, Rubin MA, Garraway LA, Nelson SF, Liau L, Mischel PS, Cloughesy TF, Meyerson M, Golub TA, Lander ES, Mellinghoff IK, Sellers WR: Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. Proc Natl Acad Sci U S A 2007, 104:20007–20012
18. 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:R41
19. Roy DM, Walsh LA, Desrichard A, Huse JT, Wu W, Gao JJ, Bose P, Lee W, Chan TA: Integrated genomics for pinpointing survival loci within arm-level somatic copy number alterations. Cancer Cell 2016, 29:737–750
20. Cohen J: A coefficient of agreement for nominal scales. Educ Psychol Meas 1960, 20:37–46
21. Stichel D, Ebrahimi A, Reuss D, Schrimpf D, Ono T, Shirahata M, Reifenberger G, Weller M, Hänggi D, Wick W, Herold-Mende C, Westphal M, Brandner S, Pfister SM, Capper D, Sahm F, von Deimling A: Distribution of EGFR amplification, combined chromosome 7 gain and chromosome 10 loss, and TERT promoter mutation in brain tumors and their potential for the reclassification of IDHwt astrocytoma to glioblastoma. Acta Neuropathol 2018, 136:793–803
22. Reifenberger J, Reifenberger G, Liu L, James CD, Wechsler W, Collins VP: Molecular genetic analysis of oligodendroglial tumors shows preferential allelic deletions on 19q and 1p. Am J Pathol 1994, 145:1175–1190
23. Kim T, Xi R, Luquette LJ, Park RW, Johnson MD, Park PJ: Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes. Genome Res 2012, 23:217–227
24. Körber V, Yang J, Barah P, Wu Y, Stichel D, Gu Z, Fletcher MNC, Jones D, Hentschel B, Lamszus K, Tonn JC, Schackert G, Sabel M, Felsberg J, Zacher A, Kaulich K, Hübschmann D, Herold-Mende C, von Deimling A, Weller M, Radlwimmer B, Schlesner M, Reifenberger G, Höfer T, Lichter P: Evolutionary trajectories of IDH WT glioblastomas reveal a common path of early tumorigenesis instigated years ahead of initial diagnosis. Cancer Cell 2019, 35:692–704.e12