Clinical validation of Whole Genome Sequencing for routine cancer diagnostics

Paul Roepman[1*], Ewart de Bruijn[1], Stef van Lieshout[1], Lieke Schoenmaker[1], Mirjam C Boelens[2], Hendrikus J Dubbink[3], Willemina RR Geurts-Giele[3], Floris H Groenendijk[3], Manon MH Huibers[4#], Mariëtte EG Kranendonk[4$], Margaretha GM Roemer[5] Marloes Steehouwer[6], Wendy WJ de Leng[4], Alexander Hoischen[6,7], Bauke Ylstra[5], Kim Monkhorst[2], Jacobus JM van der Hoeven[1] and Edwin Cuppen[1,8]

1. Hartwig Medical Foundation, Science Park 408, 1098 XH Amsterdam, the Netherlands
2. Department of Pathology, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands
3. Department of Pathology, Erasmus MC Cancer Institute, Dr. Molewaterplein 40, 3015 GD Rotterdam, the Netherlands
4. Department of Pathology, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands
5. Department of Pathology, Amsterdam University Medical Center, de Boelelaan 1117, 1081 HV Amsterdam, the Netherlands
6. Department of Human Genetics, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Geert Grooteplein Zuid 10, 6525 GA Nijmegen, the Netherlands
7. Department of Internal Medicine, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Geert Grooteplein Zuid 10, 6525 GA Nijmegen, the Netherlands
8. Center for Molecular Medicine and Oncode Institute, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands.

[#] Current employment: Department of Genetics, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands

[$] Current employment: Princess Máxima Center for Pediatric Oncology, Heidelberglaan 25, 3584 CS Utrecht, The Netherlands

***** Corresponding author**:** Dr. Paul Roepman, Hartwig Medical Foundation, Science Park 408, 1098 XH Amsterdam, the Netherlands. Tel +31 (0)20 235 2640, p.roepman@hartwigmedicalfoundation.nl.

number of text pages:	22
number of tables:	4
number of figures:	5
short running head:	Orthogonal clinical validation of WGS

Author contribution

PR, EB and EC designed the study; PR, EB, SL, and LS performed the whole genome sequencing analysis; PR, MCB, HJD, WRRG, FHG, AH, MMHH, MEGK, WWJL, KM, MGMR, MS and BY performed the validation experiments and provided support in the analysis; PR, EB, JJMH and EC wrote the manuscript; all authors approved submission and publication of the manuscript.

Abstract

Whole genome sequencing (WGS) using tissue and matched blood samples from cancer patients is becoming in reach as the most complete genetic tumor diagnostic test. With a trend towards the availability of only small biopsies, and at the same time the need to screen for an increasing number of (complex) biomarkers, the use of a single all-inclusive test is preferred over multiple consecutive assays. To meet the high-quality diagnostics standards, we have optimized and validated the performance of a clinical grade WGS workflow, resulting in a technical success rate of 95.6% for samples with sufficient (≥20%) tumor cell percentage.

Independent validation of identified biomarkers against commonly used diagnostic assays showed a high sensitivity (98.5%) and specificity (98.4%) for detection of somatic SNV and indels, and high concordance (93.3%) for gene amplification detection. Gene fusion analysis showed a concordance of 91.3% between DNA-based WGS and an orthogonal RNA-based gene fusion assay.

Microsatellite (in)stability assessment showed a sensitivity of 100% with a specificity of 97%, and high-risk human papillomavirus detection showed an accuracy of 95.8% compared to standard pathological tests.

In conclusion, whole genome sequencing has a >95% sensitivity and specificity compared to routinely used DNA techniques in diagnostics and all relevant mutation types can be detected reliably in a single assay.

Introduction

Needs and complexity in molecular cancer diagnostics are rapidly increasing, driven by a growing number of targeted drugs and developments towards more personalized treatments [1,2]. Simultaneously, advances in next-generation DNA sequencing technology have greatly enhanced the capability of cancer genome analyses, thereby rapidly progressing diagnostic approaches from small targeted panels to large panels and exome sequencing. Currently, whole genome sequencing (WGS) using tissue and matched blood samples from patients with (metastatic) cancer [3] is getting in reach as the most complete genetic tumor diagnostics test. In the context of the Dutch national CPCT-02 clinical study (NCT01855477) Hartwig Medical Foundation has established a national WGS facility including robust sampling procedure and logistics in more than 45 (of the 87) hospitals located across the Netherlands for the centralized analysis of tumor biopsies by WGS. Since the start in 2016, more than 5,000 tumors and matched control samples have been analyzed by WGS, of which the first cohort of 2500 patients has been extensively characterized and described [4]. Originally, this clinical study aimed to analyse data for biomarker discovery, but with growing clinical demands for more extensive and broader DNA analysis for patient stratification towards targeted treatments [5], the scope of WGS is now entering routine diagnostic usage. As part of this development, the required amount of tumor tissue for as well as the turn-around-time of the WGS procedure was decreased, together with implementation of more extensive quality control metrics and independent validation required for accreditation. Currently, there is an ongoing trend towards the availability of only small biopsies, especially for advanced stage cancer where metastatic lesions are sampled using core needle biopsies, with at the same time a growing need to screen for an increasing number of (complex) biomarkers. For future-proof and efficient molecular diagnostics, the use of a single all-inclusive test is preferred over multiple consecutive assays that, together, often take more time, require more tissue and provide a far less complete profile of the molecular characteristics. For complex molecular diagnostic indications (e.g. non-small cell lung cancer) the expected cost of WGS is now in the same order as the combined multiple individual tests, especially when also taking into account (technical) personnel cost and costs of maintaining and updating multiple test setups [6,7].

To meet the high-quality diagnostics standards, we have optimized and clinically validated the performance of the WGS workflow, both technically as well as bioinformatically, as these are highly interconnected in determining the specificity and sensitivity of the test. The validation efforts include current *standard-of-care* biomarkers (oncogenic hotspots, inactivating mutations in tumor suppressor genes), but also broader analyses of gene fusions and other genomic rearrangements as well as emerging genome-wide or complex biomarkers like tumor mutational burden estimation, microsatellite instability (MSI) [8], and homologous repair deficiency (HRD) signatures [9,10]. Importantly, an *open-source* and data-driven filtering and reporting strategy has been put into place to reduce the wealth of information into a diagnostically manageable size and to provide an overview of all clinically relevant DNA aberrations.

Here we show that WGS has an overall >95% sensitivity and specificity as compared to other targeted detection techniques that are routinely used in cancer diagnostics and that all relevant mutation types can be readily and reliably detected in a single assay. Although WGS requires minimal quantity of input material and can be applied pan-cancer, the tumor purity can be a limiting factor below 20% tumor cells as well as the availability of fresh frozen tumor material, which is a prerequisite for high-quality results as described here. Together, WGS has now matured from a research technology into an ISO accredited test that is ready to be used for clinical decision making in routine cancer care.

Methods

Patient selection

For this study, samples were used from patients that were included as part of the CPCT-02 (NCT01855477), DRUP (NCT02925234) and WIDE (NL68609.031.18) clinical studies, which were approved by the medical ethical committees (METC) of the University Medical Center Utrecht and the Netherlands Cancer Institute.

Whole Genome Sequencing

Whole Genome Sequencing (WGS) was performed under ISO-17025 accreditation at the Hartwig Medical Foundation laboratory (Amsterdam). The WGS test uses high quality DNA extracted from tumor tissue and blood samples. Input tissue type includes fresh-frozen or frozen archived samples from solid metastatic or primary tumor samples. In addition to tissue samples, frozen cell pellets from pleural fluid samples and ascites can be used.

DNA extraction is performed on the QiaSymphony following standard reagents and protocols. 50-200 ng gDNA is fragmented by sonication on the Covaris LE220 Focused ultrasonicator (median fragment size 450 bp) for NGS Truseq nano library preparation including PCR amplification (8 cycles). All procedures are automated on the Beckman Coulter Biomek4000 and i7 liquid handling robots. The Illumina® HiSeqX and NovaSeq6000 platforms are used for sequencing >90x and >30x average read coverage of tumor and normal genomes, respectively. To improve cost effectiveness, shallow whole-genome sequencing (8-15x coverage depth) has been used to estimate the tumor purity of the received tumor sample, before continuing to "deep" sequencing (90 – 100x in total) in case of sufficient tumor cell content (≥20%).

Sequencing data is analyzed with an in-house developed bioinformatic open source software-based pipeline (https://github.com/hartwigmedical/). The pipeline is designed to detect all types of somatic alterations, including single and multiple nucleotide substitutions (SNV and MNV), small insertions and deletions (indels), copy number alterations (aneuploidies, amplifications and gene copy losses), genomic rearrangements, and structural variants (e.g. gene fusions, mobile element insertions) [11].

The blood sample is used to filter out the ubiquitously present germline polymorphisms to be able to report somatic variants only.

Additionally, genome-wide mutational characteristics are determined and reported including microsatellite instability (MSI), tumor mutational load (ML), mutational burden (TMB), and Homologous Recombination DNA repair-deficiency (HRD). Further, viral integrations in the tumor genome are detected and reported. All code and scripts used for analysis of the WGS data are available at GitHub (https://github.com/hartwigmedical/). The raw and analyzed WGS data used in this manuscript are available for validation and cancer research purposes through a standardized controlled access procedure (see https://www.hartwigmedicalfoundation.nl/applying-for-data/ for details).

Orthogonal validation tests

Independent validation was performed for all to-be-reported types of clinically relevant DNA aberrations, including mutations (SNV, MNV and indels) with specific focus on *BRAF*, gene copy number (*ERBB2* as example), microsatellite (in)stability, gene fusions, and viral infection (HPV as example). WGS results were retrospectively compared against (as far as possible) routine diagnostic assays performed independently in ISO15189 accredited pathology laboratories. If a clinical assay was not available for the validation purpose, a custom research-use-only test was performed. The following independently performed validation experiments were performed. An overview of the used tumor samples per validation assay is available as **Suppl Data 1**.

Validation of SNV, MNV and indel detection

A custom designed (research-use-only based) single molecule Molecular Inversion Probe (smMIP) sequencing panel was designed for independent confirmation of variants detected by WGS in an independent lab (Radboudumc). The smMIP panel sequencing was designed and processed similar to previous reports [12,13]. In total 415 smMIPs were designed to test 192 randomly selected variants (including driver and passenger variants) that were detected by WGS across 29 tumor samples. smMIP validation was performed using the same isolated DNA as was used for WGS, and SeqNext (JSI medical systems) was used for analysis.

Orthogonal clinical validation of variant detection was performed using 48 samples and compared against a custom-made Oncomine NGS gene-panel (ThermoFisher), processed independently (double blind) in a routine pathology laboratory under ISO15189 accreditation (Erasmus MC). The custom Oncomine assay covered 25.2 kb exonic regions across 40 genes (design (v5.1) available in supplementary data[14]) and was performed using the same isolated DNA as was used for WGS, thereby ruling out potential tumor heterogeneity. Analysis was done using SeqNext (JSI medical systems) and a formal clinical report was generated. Additionally, for 10 samples a comparison was made between the WGS based mutational load (ML) assessment and the Oncomine Tumor Mutational Load (TML) assay (Thermofisher).

Validation of copy number assessment

WGS based copy number assessment was validated against fluorescent in situ hybridization (FISH) using COLO829 and a cohort of diagnostic tumor samples. For COLO829, a comparison was made for the ploidy of chromosomes 9, 13, 16, 18, 9p24 (CD274/*PDCD1LG2*), and 2q23 (*ALK*) at the Amsterdam UMC. Chromosome Enumeration Probes (CEP) for the centromeric region of chromosome 9, 13, 16 and (CEP9, CEP13, CEP16, CEP18) were used, as well as locus specific break-apart probes for 2p23 (*ALK*) fusion (Abbott Vysis) and 9p24 (*CD274/PDCD1LG2*) fusion (Leica Biosystems). Slides were visualized on a Leica DM5500 fluorescence microscope and for each marker, 100 cells/slide were scored for the percentages of cells with respective numbers of chromosomes (signals) counted.

Diagnostic *ERBB2* copy number readout was validated using 16 tumor samples and using HER2/neu FISH analysis at an independent routine pathology laboratory under ISO15189 accreditation (University Medical Center Utrecht). FISH scoring was performed according to guidelines [15]. New tumor sections (fresh-frozen) were used for probe hybridization (Cytocell LPS001), scanned using the Leica DM6000 scanner and analyzed with Cytovision software (Leica Biosystems). A formal clinical report was generated that was compared with the WGS results.

Validation of fusion gene detection

Validation of gene fusion detection by WGS was performed against RNA-based Anchored Multiplex PCR NGS assay (Archer FusionPlex Solid Tumor, ArcherDx). Twenty-four samples were selected based on the WGS results to include multiple fusion genes. Matching RNA (200 ng), isolated from the same tissue as the DNA that was used for WGS, was analyzed according to routine pathological procedures (ISO15189 certified) (Erasmus MC). A formal clinical report was generated and was compared with the WGS results.

Validation of microsatellite (in)stability readout

For a set of 48 tumor samples, the microsatellite status was validated using the MSI analysis system (Promega) and performed at a routine pathology laboratory (Erasmus MC) [16]. This fluorescent multiplex PCR assays analyzed five nearly monomorphic mononucleotide microsatellite loci (BAT-25, BAT-26, NR-21, NR-24, and MONO-27). Matching tumor and blood samples were analyzed for accurate detection. Both the number of positive loci as well as binary classification of microsatellite instable (MSI) and stable (MSS) were reported.

Validation of tumor associated virus detection

WGS based detection of genomically integrated high-risk Human Papillomavirus (HPV) DNA was validated against routine pathological testing (Netherlands Cancer Institute) using the QIAscreen HPV PCR Test (Qiagen). If available, results of routine testing were used for comparison with WGS. If not, HPV status was determined retrospectively using an aliquot of the DNA (20 ng) that was used for WGS.

Results

Analytical performance and reproducibility of clinical-grade WGS

In addition to the orthogonal clinical validation experiments that are described in the next paragraphs, the analytical performance and consistency of our WGS setup is continuously monitored using a Genome-in-a-bottle (GIAB) mix-in sample (tumor 30% NA12878: normal 100% NA24385) for which all DNA aberrations are known. The accuracy of GIAB genome-wide variant detection (SNV and short indels) by WGS was very high and stable across different runs and using multiple sequencers (in a time period of eight months) with a precision of 0.998 (range 0.994-0.998) and a sensitivity of 0.989 (range 0.973-0.990) (**Table 1**). Most importantly, all F-scores for variant detection exceeded the pre-set 0.98 lower limit for high-quality sequencing data (median 0.993, range 0.985-0.994). WGS coverage analysis across a set of 25 randomly selected tumor samples indicated stable and high coverage across the entire genome (median coverage 106x, range 84-130) (**Table 1**).

Robustness and reproducibility of the bioinformatic data analysis pipeline was assessed by re-analysis of 18 samples (selected across a period of six months) starting from raw sequencer output files. Compared to the initial output, results from the reanalysis show near identical results with a percentage positive agreement (PPA) of 99.98 for SNVs, 99.96 for MNVs and 99.88 for indels (**Table 1**). The observed small differences are partially caused due to random feeds by the algorithms but mainly due to (periodic) improvements in the bioinformatics pipeline. The reproducibility of the complete workflow was furthermore confirmed on two diagnostic cases (non-small cell lung cancer and an undifferentiated pleomorphic sarcoma) in which the replicated tests starting from independent biopsy/blood isolation provided highly similar molecular profiles with identical diagnostic reports (**Figure 1**).

Sample quality and overall WGS success rate

Samples used for WGS analysis currently comprise predominantly of fine needle biopsies taken from a metastatic lesion from patients with stage IV cancer. To determine whether WGS quality is

dependent on the (primary) tumor type, a large-scale analysis was performed on samples that were processed as part of the Dutch CPCT-02 trial. Eighty-six percent of the analyzed samples (n=2,520)[4] passed all quality criteria, with a lowest success rate for kidney (72.3%), liver (77.3%), and lung (79.1%) cancer patients (**Figure 2A**). An insufficient amount of tumor cell (<20% based on WGS-derived tumor purity) was the most prevalent failure rate: 6.4% of samples showed a tumor DNA purity between 5-20% and for 2.9% of the cases a seemingly absence (<5%) of tumor DNA was observed despite prior pathological assessment. As a consequence of the restricted use of only fresh frozen biopsies as input material, insufficient sequencing data quality was only observed for 4.4% of the samples, indicating a high technical success rate of 95.6% for samples with sufficient (≥20%) tumor purity (Figure 2A). Of note, technical success rate has further increased to >98% in a currently ongoing prospective clinical study [17].

Although the physical damage to the DNA is expected to be much lower for fresh-frozen samples as compared to formalin-fixed paraffin-embedded (FFPE) samples, we used the previously described Global Imbalance Value (GIV) score as a measure indicative for DNA damage [18]. The analyzed set of 2,520 samples showed very low $GIV_{G>T}$ scores with a median of only 1.02 (range 0.495 - 2.495) indicating only 3 samples (0.11%) were considered as damaged samples with a GIV score >1.5 (**Figure 2B)**. In comparison, 41% of the 1000 Genomes Project samples had a $GIV_{G>T}$ score of at least 1.5, while 73% of the TCGA samples showed a $GIV_{G>T}$ score >2 [18].

For accurate determination of absolute tumor-specific allele frequencies and copy number status, it is crucial to correctly assess the tumor cell contribution to a sample (tumor purity). Traditionally, pathological tumor cell percentages (pTCP) are used as representation for tumor DNA purity. However, the tumor purity can be determined more accurately from WGS data by genome-wide determination of the ratio of normal and aberrant genomic segments or nucleotides (mTCP). While the pTCP scores show a modest but significant correlation with the tumor DNA purity for samples with higher tumor content (r=0.40 p=0.002), this association was absent for samples with lower (<30%) tumor purity (r=0.08, p=0.76) (**Figure 2C**). Instead of using pTCP, molecular tumor cell purities (mTCP) that are based on analysis of shallow sequencing data (~8-15x average coverage) of the tumor's genome were found to be a more reliable measurement. Validation of the mTCP assessments by shallow sequencing showed a very good correlation with the mTCP of deep WGS (~90-110x) (R^2

of 0.931, n=43, **Figure 2D**), with an average deviation between both purities of only 3.2% (range 0% to 35% caused by an outlying non-small cell lung cancer case). This data confirms that shallow sequencing data is sufficient for reliable initial tumor purity estimation detection and can be a valuable and cost-effective approach for upfront selection of suitable samples for deep whole genome sequencing.

Specificity and sensitivity of SNV, MNV and indel detection

Specificity of the variants detected by WGS was assessed by a tailored single molecule Molecular Inversion Probe (smMIP) panel sequencing [12,13]. Across 29 samples, 192 randomly selected variants were sequenced and analyzed by a custom designed smMIP panel (no reliable panel design was possible for 17.6% of the initial selected WGS variants). Nearly all (98.4%) of the variants were confirmed by smMIP sequencing indicating a very high specificity of WGS for small variant detection. In addition, the observed variant allele frequencies showed a high correlation (R^2=0.733 between both assays (**Figure 3A**).

Orthogonal clinical validation of mutations in a specific oncogene, *BRAF*, was performed using 48 selected samples and compared against the custom-made Oncomine gene-panel NGS assay (ThermoFisher). Twenty-five samples showed a *BRAF* exon 15 or exon 11 mutation by WGS that were all confirmed by panel NGS (**Figure 3B**). Vice-versa, 26 *BRAF* mutations that were detected using panel-based sequencing were also identified using WGS. A single *BRAF* p.Gly469Ala mutation identified by panel NGS was not confirmed using the WGS analysis due to low mutation frequency (~2%). On the other hand, WGS identified two less common *BRAF* variants (p.Ala762Val and p.Pro403fs) that were not found by the custom Oncomine assay as the panel design does not include the corresponding exons. Both variants are unlikely to result in BRAF activation and are predicted passenger variants, especially because both tumors were MSI with a high TMB. All other 20 *BRAF* wild-type samples by WGS were confirmed by panel sequencing.

Next, all somatic non-synonymous mutations across the NGS panel design were evaluated (25.2 kb covering hotspot exons of 40 genes). Combined with the *BRAF* results, in total 139 mutations were detected by at least one of the tests of which 137 (98.6%) were reported by WGS and 133 using panel

sequencing (**Figure 3B**) resulting in an overall 98.5% sensitivity for WGS compared to panel based NGS and 95.6% for panel compared to WGS. A *PTEN* p.Lys327Arg mutation that was identified using the panel, was not reported by the WGS test. Re-analysis of the WGS read data confirmed the presence of this variant at a low VAF (7% with a coverage of 8 out of 116 reads). On the contrary, the panel assay did not report a pathogenic *PTEN* variant (p.Tyr27Ser), which was identified by WGS (VAF of 12%) using the same input DNA. The variant was present in the NGS panel data (VAF 6%), but did not meet the criteria for clinical reporting. The panel also missed identification of the *APC* p.Thr1556fs inactivating mutation in three samples. This *APC* codon lies within a homopolymeric DNA region and the IonTorrent sequencing technology used for the panel sequencing is known to face more difficulties in repetitive DNA regions.

Although the performance of tumor mutational load (ML) estimations are directly following the performance of accurate non-synonymous variant calling (analytically, ML is only a simple summation of the observed variants), mutational burden readout was compared on 10 additional samples between WGS and Oncomine Tumor Mutational Load (TML) assay (Thermofisher). Both readouts showed a high correlation (R^2=0.94) but this was mainly caused by a single high ML sample (ML > 1200) (**Figure 1C**). Binary classification based on both tests (WGS based ML cutoff of 140 mut vs. TML based TMB cutoff of 10 mut/Mb) indicated a concordance for 7 out of 9 samples (1 sample was not evaluable by Oncomine TML), but also indicated a lower correlation in the cutoff region (R^2=0.16 when excluding 2 highest ML/TMB samples). This result illustrates the challenges of accurate mutational burden readout using a more limited gene panel as compared to exome or genome-wide measurements, as discussed elsewhere [19,20].

Taken together, the Oncomine and smMIP NGS validation results indicate both a high sensitivity (98.5%) and a high specificity (98.4%) of detection of SNV, MNV and indels using WGS and biopsies with ≥20% tumor purity, which is similar as compared to commonly used panel-based approaches.

Copy number alterations

WGS-based chromosomal ploidy and local genomic copy number analytical performance were initially assessed by independent FISH analysis on 6 genomic locations of the COLO829 tumor cell line (centromeric region of chromosomes 9, 13, 16, and 18, and 2q23 *ALK* and 9p24 *CD274/PDCD1LG2* (PD-L1/PD-L2) using diagnostic 'break-apart' probes). WGS and FISH analysis showed highly similar purity and ploidy calculations with Chr9 showing 4x in ~55% of cells, Chr13 3x in ~55%, Chr18 3x in ~60%, 2q23 locus 3x in 70-80% and complete diploid Chr 16 and 9q24 locus for all cells **(Figure 4**).

Further orthogonal clinical validation focussed on accurate detection of *ERBB2* (*Her2/neu*) amplification. Sixteen samples from various tumor types were used (11 mamma, 2 colorectal, 1 stomach, 1 bladder and 1 melanoma). Importantly, samples were representative of the full spectrum of *ERBB2* amplifications also including samples with only marginal amplification and samples with increased ploidy of the complete chromosome 17. New tissue sections from the same biopsy or a second biopsy obtained at the same moment as the samples used for WGS were analyzed by FISH at an independent routine pathology laboratory (**Table 2**). For one sample (#5) FISH analysis failed due to insufficient tumor cells (confirmed by immunohistochemistry). All other FISH results were considered representative. All samples with a WGS-based *ERBB2* copy-number greater than 6x were confirmed by FISH to harbor substantial *ERBB2* amplified signals (defined as *ERBB2* >6). For *ERBB2* WGS copy-numbers between 2-6, at best an *ERBB2* gain was observed by FISH but considered insufficient for classification as *ERBB2* amplified (classified as *ERBB2* gain or equivocal). A borderline discordant *ERBB2* status was observed for a single case (sample #16, FISH 2-4x in 83% compared to WGS 6x). No technical explanation could be identified, but this might be caused due to tumor heterogeneity between the sections used for WGS and FISH. Of note, this specific case involved a colorectal tumor for which the FISH assay is not used in routine practice.

The copy number validation data showed a high concordance (93%, 14 of the 15 cases) of WGS and FISH analysis indicating that WGS can reliably be used for detection of sufficiently high gene amplifications. For lower copy numbers (range 2 to 6) the concordance showed more variability but the question remains whether such low gains are biologically and/or clinically relevant [21].

Detection of fusion genes

Detection of gene fusions by WGS was compared with results obtained with an RNA-based Anchored Multiplex PCR NGS assay (Archer FusionPlex, ArcherDx) and was performed independently on 24 samples using matching DNA and RNA from the same biopsy. Samples were selected based on the WGS results to include one or more clinically relevant fusion genes. The Archer NGS assay confirmed the WGS findings for 21 of the 23 samples (91.3%), including fusion of *ALK*, *NRG1* and *ROS1* (**Table 3**). For one sample no comparison could be made, as the *TMPRSS2-ERG* fusion is not covered by the used Archer NGS assay.

A *NTRK1* fusion detected by Archer NGS (*MEF2D-NTRK1*: (22 reads, 60% VAF) could not be identified using WGS, possibly due to a complex structural variation pattern involving multiple break-junctions in the intronic regions and thus more difficult to call using WGS data compared to analysis of RNA. Vica versa, one fusion (*SPAG17-ALK*) detected by WGS showed no evidence in the tumor RNA. Although based on fusion at DNA level a viable in-frame fusion protein was predicted, it can very well be that the corresponding RNA was expressed at low levels (e.g. due to temporal or spatial expression variation) that are insufficient for reliable detection by the Archer assay.

Quantification of microsatellite instability (MSI)

WGS-based MSI classification was validated independently using 48 selected samples including multiple tumor types (32 colorectal, 5 prostate, 3 esophagus, 2 pancreatic and 6 other) and using the routinely used 5-marker PCR MSI panel [16,22]. Assessment of microsatellite (in)stability by WGS, defined as the number of small indels per million bases occurring in ≥5-mer homopolymers and in di-, tri- and tetranucleotide repeats [8], showed an average microsatellite instability (MSI) score of 1.11 with the vast majority of samples having a low score and a long tail towards higher MSI scores (range 0.004 to 93, n=2520, **Figure 5A**). 2.7 percent of the samples were classified as MSI using a cutoff of 4 (cutoff was based on the apparent bi-nominal distribution of the MSI scores). On the validation set (n=48) the sensitivity of WGS MSI classification was 100% (95%CI 82.6-100%) with a specificity of

97% (95%CI 88.2-96.9%) and a Cohen's kappa score of 0.954 (95%CI 0-696-0.954). In addition to the binary MSI/MSS concordance, the MSI score correlated with the number of positive PCR markers, in which samples with only 1 or 2 positive PCR markers showed a marginal MSI score (**Figure 5B**). The only discordant results were from a lymphoma sample with a complex pathology showing 1/5 positive PCR markers (classified as MSS) but a WGS MSI score of 5.9 (classified MSI). IHC analysis showed no substantial loss of mismatch repair (MMR) proteins although WGS analysis indicated a somatic *PMS2* p.Ile193Met variant in combination with a likely inactivating *PMS2* structural variant. The p.Ile193Met mutation is classified with a high prior in de Leiden Open Variant Database (LOVD, https://databases.lovd.nl/shared/variants/PMS2) and thus likely represents a pathogenic variant. Both the MSI PCR test as well as the MMR IHC had not been validated for use in lymphoma cases so a definitive conclusion remained difficult.

Tumor-genome integrated virus detection

Recently it has been shown that the presence of viruses can be detected with great accuracy using WGS [23]. Assessment of the presence of integrated viral DNA was validated against standard routine pathological assessment, typically a PCR test. We focused on Human papillomavirus (HPV) due to the prevalence and clinical importance and the availability of HPV routine testing (e.g. QIAscreen HPV PCR assay, Qiagen). Twenty-four tumor samples (including 10 GI-tract, 6 female reproductive, 3 head-neck, 2 male reproductive and 3 other cancer types) were used for independent validation between WGS and PCR assay. The concordance of WGS and standard pathology was very high with an accuracy of 95.8% and a sensitivity of 90.9% (95%CI 67.6-90.9) and specificity of 100% (80.3-100%). Cohen's kappa score of 0.915 (95%CI 0.48-0.92) indicates an 'almost perfect agreement', in which also the HPV high-risk types were concordant between both tests (**Table 4**).

A single sample showed a discordant result in which the PCR assay indicated HPV type 16 while no such evidence was found by WGS. A follow-up PCR test on the same DNA that was used for WGS analysis showed the same result, thereby ruling out sample heterogeneity. This result can most likely be explained due to a non-integrated HPV infection, as the WGS analysis pipeline only considered

viral DNA fragments that were integrated in the host genome (shared viral-human read pairs), or due to integration into a non-sequenceable part of the genome.

Discussion

During the past few years, whole genome sequencing (WGS) and the associated bioinformatic data processing and interpretation has matured from a research-use-only tool to a diagnostic-level technology [24]. Together with the clinical need to screen for an increasing number of (complex) biomarkers in an increased number of tumor types (or even pan-cancer) [1,25] and the availability of limited amounts of biopsy material, the use of a single all-inclusive DNA test is a more than welcome development for efficient molecular diagnostics. While costs are currently still relatively high, sequencing technology continues to evolve including decreasing costs. Here we report on (retrospective) orthogonal validation efforts of WGS and show, to our knowledge for the first time, that the performance of WGS is equal to the range of routinely used diagnostic tests with technical concordances of >95%. More specifically, we show that a single WGS-based tumor-normal test can replace separate test for 1) actionable small variant (SNV, indel) driver mutations (previously detected by targeted PCR-based or NGS panel-based tests), 2) gene amplifications (FISH), 3) fusion genes (FISH or RNA panels), 4) microsatellite instability (amplicon fragment analysis), 5) HPV infection, and 6) tumor mutational load determination (NGS-panels). Prospective clinical validation and integration into routine workflow is currently being evaluated by a direct comparison of simultaneously obtained routine diagnostics and WGS-based test results [17]. To make the WGS test suitable for diagnostic use, the turn-around-time has already been reduced towards a clinically acceptable 10 working-days.

The good performance of WGS for diagnostic use is primarily the result of two important aspects of the workflow that are fundamentally different from most existing molecular diagnostics procedures for cancer: 1) the use of only fresh frozen tumor material yielding consistent high quality DNA and sequencing results, and 2) parallel processing of the patient's fresh blood sample to serve as a control/baseline for the matching tumor sample. Hereby, all germline variants can be subtracted automatically from the tumor data thereby allowing for precise pinpointing of all tumor specific

changes. Even with focus on a set of ~500 cancer related (driver) genes [4], the bulk of all missense variants observed in the tumor are in fact inherited germline polymorphisms without clinical significance, making comprehensive (manual) tumor-only interpretation and filtering a daunting task. This challenge is not unique for WGS but in principle also applies for all large NGS panels [26,27]. Filtering out germline variants using population database information is challenging due to various reasons (e.g. biases in such databases toward Caucasian population, rare patient or sub-population specific variants) and although known driver mutations are readily detected by panel-based tests, the impact on tumor mutational load measurements is likely severely impacted when germline or somatic status of a variant cannot be discriminated accurately.

With the increase in (technical) sequencing capabilities, the bioinformatics part ('dry-lab') has become essential for a good analysis and interpretation of the sequencing data of WGS but also for the emerging larger comprehensive panels. Traditionally, (hospital) laboratories have focused most on the wet-lab performance and automatization but it has become clear that the downstream bioinformatics, and the ICT infrastructure to handle (and store) all data, pose the greatest challenge. Complex bioinformatics and high-end reporting tools are essential for an understandable communication of the results to the (clinical) end-users. An example of our current WGS report which is multilayered to serve the different end-users (oncologist, pathologist), is provided (**Suppl Data 2**).

Currently, WGS still requires a tumor content that is somewhat higher than focussed panel based approaches (minimal 20% for WGS versus 5-10% for panel NGS). This limitation is caused due to a lower sequencing depth by WGS, but with ongoing price reductions we anticipate that WGS with ~250x coverage will be feasible in the next coming years and thus will also be able to analyse samples with lower tumor content and to detect minor tumor subclones. A more challenging limitation is the need of fresh-frozen (or freshly lysed) samples for WGS analysis as this will, for most hospitals, require an adaptation in the pathology laboratories that are currently mostly FFPE orientated. The adaption of WGS in a routine pathology workflow is currently being evaluated and optimized in a prospective clinical study [17].

DNA sequencing tests are often performed as laboratory-developed tests (LDTs) and the technical parameters, validation requirements and quality assurance are typically governed by national regulation and legislation that can differ. Various expert groups have drafted guidelines and

recommendations for the standardization of multigene panel testing [2,28] and for our validation efforts we have followed the guidelines for setup and validation of (new) sequencing tests in ISO-accredited pathological laboratories in the Netherlands. However, with the ongoing approval of NGS panel assays by the FDA [29] and the upcoming new European Regulations for in-vitro diagnostic medical devices IVDR (2017/746) in 2022 [30], it is anticipated that (whole) genome sequencing tests will become regulated following international guidelines, standardization and quality schemes. Clinical validation by comparison with common standards, as described here, will be a key component of such regulations.

With the rapid development of more targeted drugs and their associated biomarkers, it is next to standardization of the (complex) test results, important to be able to efficiently and quickly add new biomarkers/genes to the clinical reports (e.g. *NRG1* and *NTRK* fusions and *PIK3CA* activating mutations). WGS will allow such a rapid and efficient co-development of (all) future diagnostic DNA markers, because it 'only' requires an update of the bioinformatics and reporting aspects, without the need of laborious and costly new test developments or adaptations of panel designs including the required laboratory analytical validation experiments. In addition, the data from previously tested patients can, in principle and upon request from the treating physician, be reanalyzed for the presence of the (all) new biomarkers and recontacting of the patient can be considered [31].

Setting aside the direct impact WGS can have for routine clinical use and comprehensive screening for clinical study eligibility, a whole-genome view of the tumor will yield a wealth of valuable research data and provide the opportunity to increase our insights in oncogenic processes and to better explain or predict the response to targeted or immunotherapy. Such a learning-health-care system, where we learn from today's patients will greatly enhance our understanding of this complex disease and facilitate the discovery of newly identified (complex) biomarkers, targeted therapies, and improved treatment decision making for future patients.

Acknowledgements

The authors would like to thank Peggy Atmodimedjo, Isabelle Meijssen, Ronald van Marion and Hanna Schoep for (technical) assistance with collecting the data, Sandra van den Broek for data analysis support and Immy Riethorst for sample logistics. This publication and the underlying study have been made possible partly on the basis of the data that Hartwig Medical Foundation and the Center of Personalised Cancer Treatment (CPCT) have made available to the study.

References

1. Hyman DM, Taylor BS, Baselga J. Implementing Genome-Driven Oncology. Cell, 2017, 168:584–99

2. Mosele F, Remon J, Mateo J, Westphalen CB, Barlesi F, Lolkema MP, Normanno N, Scarpa A, Robson M, Meric-Bernstam F, Wagle N, Stenzinger A, Bonastre J, Bayle A, Michiels S, Bièche I, Rouleau E, Jezdic S, Douillard J-Y, Reis-Filho J, Dienstmann R, André F. Recommendations for the use of next-generation sequencing (NGS) for patients with metastatic cancers: A report from the ESMO Precision Medicine Working Group. Annals of Oncology, 2020. https://doi.org/10.1016/j.annonc.2020.07.014

3. Weinstein JN, The Cancer Genome Atlas Research Network, Collisson EA, Mills GB, Mills Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. The Cancer Genome Atlas Pan-Cancer analysis project. Nature Genetics, 2013, 45:1113–20

4. Priestley P, Baber J, Lolkema MP, Steeghs N, de Bruijn E, Shale C, Duyvesteyn K, Haidari S, van Hoeck A, Onstenk W, Roepman P, Voda M, Bloemendal HJ, Tjan-Heijnen VCG, van Herpen CML, Labots M, Witteveen PO, Smit EF, Sleijfer S, Voest EE, Cuppen E. Pan-cancer whole-genome analyses of metastatic solid tumours. Nature, 2019, 575:210–6

5. Manolio TA, Rowley R, Williams MS, Roden D, Ginsburg GS, Bult C, Chisholm RL, Deverka PA, McLeod HL, Mensah GA, Relling MV, Rodriguez LL, Tamburro C, Green ED. Opportunities, resources, and techniques for implementing genomics in clinical care. Lancet, 2019, 394:511–20

6. Layfield LJ, Hammer RD, White SK, Furtado LV, Schmidt RL. Molecular Testing Strategies for Pulmonary Adenocarcinoma: An Optimal Approach With Cost Analysis. Arch Pathol Lab Med, 2019, 143:628–33

7. Pasmans CTB, Tops BBJ, Steegs EMP, Coupe V, Grunberg K, de Jong EK, Schuuring EMD, Willems SM, Ligtenberg M, Retel VP, van Snellenberg H, de Bruijn E, Cuppen E, Frederix GWJ. Micro-costing Diagnostics in Oncology: From Single-Gene Testing to Whole Genome Sequencing, n.d. https://doi.org/10.1101/19009969

8. Huang MN, McPherson JR, Cutcutache I, Teh BT, Tan P, Rozen SG. MSIseq: Software for Assessing Microsatellite Instability from Catalogs of Somatic Mutations. Sci Rep, 2015, 5:13321

9. Davies H, Glodzik D, Morganella S, Yates LR, Staaf J, Zou X, Ramakrishna M, Martin S, Boyault S, Sieuwerts AM, Simpson PT, King TA, Raine K, Eyfjord JE, Kong G, Borg Å, Birney E, Stunnenberg HG, van de Vijver MJ, Børresen-Dale A-L, Martens JWM, Span PN, Lakhani SR, Vincent-Salomon A, Sotiriou C, Tutt A, Thompson AM, Van Laere S, Richardson AL, Viari A, Campbell PJ, Stratton MR, Nik-Zainal S. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. Nat Med, 2017, 23:517–25

10. Nguyen L, Martens J, Van Hoeck A, Cuppen E. Pan-cancer landscape of homologous recombination deficiency, n.d. https://doi.org/10.1101/2020.01.13.905026

11. Cameron DL, Baber J, Shale C, Valle-Inclan JE, Besselink N, Cuppen E, Priestley P, Papenfuss AT. GRIDSS2: harnessing the power of phasing and single breakends in somatic structural variant detection, n.d. https://doi.org/10.1101/2020.07.09.196527

12. Eijkelenboom A, Kamping EJ, Kastner-van Raaij AW, Hendriks-Cornelissen SJ, Neveling K, Kuiper RP, Hoischen A, Nelen MR, Ligtenberg MJL, Tops BBJ. Reliable Next-Generation Sequencing of Formalin-Fixed, Paraffin-Embedded Tissue Using Single Molecule Tags. J Mol Diagn, 2016,

18:851–63

13. Acuna-Hidalgo R, Sengul H, Steehouwer M, van de Vorst M, Vermeulen SH, Kiemeney LALM, Veltman JA, Gilissen C, Hoischen A. Ultra-sensitive Sequencing Identifies High Prevalence of Clonal Hematopoiesis-Associated Mutations throughout Adult Life. Am J Hum Genet, 2017, 101:50–64

14. Pruis MA, Geurts-Giele WRR, Von der TJH, Meijssen IC, Dinjens WNM, Aerts JGJV, Dingemans AMC, Lolkema MP, Paats MS, Dubbink HJ. Highly accurate DNA-based detection and treatment results of MET exon 14 skipping mutations in lung cancer. Lung Cancer, 2020, 140:46–54

15. Wolff AC, Elizabeth Hale Hammond M, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, Bilous M, Ellis IO, Fitzgibbons P, Hanna W, Jenkins RB, Michael F. Press, Spears PA, Vance GH, Viale G, McShane LM, Dowsett M. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. Journal of Clinical Oncology, 2018, 36:2105–22

16. van Lier MGF, Wagner A, van Leerdam ME, Biermann K, Kuipers EJ, Steyerberg EW, Dubbink HJ, Dinjens WNM. A review on the molecular diagnostics of Lynch syndrome: a central role for the pathology laboratory. J Cell Mol Med, 2010, 14:181–97

17. Roepman P, Bosch L, Samsom K, Schipper L, de Bruijn E, Hoes L, Riethorst I, Schoenmaker L, van der Kolk L, van Snellenberg H, Voest EE, Cuppen E, Monkhorst K, Meijer G. WGS implementation in standard cancer diagnostics for every cancer patient (WIDE). Annals of Oncology, 2019, 30:v575

18. Chen L, Liu P, Evans TC Jr, Ettwiller LM. DNA damage is a pervasive cause of sequencing errors, directly confounding variant identification. Science, 2017, 355:752–6

19. Mankor JM, Paats MS, Groenendijk FH, Roepman P, Dinjens WNM, Dubbink HJ, Sleijfer S, CPCT Consortium, Cuppen E, Lolkema MPJK. Impact of panel design and cut-off on tumour mutational burden assessment in metastatic solid tumour samples. Br J Cancer, 2020, 122:953–6

20. Budczies J, Allgäuer M, Litchfield K, Rempel E, Christopoulos P, Kazdal D, Endris V, Thomas M, Fröhling S, Peters S, Swanton C, Schirmacher P, Stenzinger A. Optimizing panel-based tumor mutational burden (TMB) measurement. Ann Oncol, 2019, 30:1496–506

21. Eijkelenboom A, Tops BBJ, van den Berg A, van den Brule AJC, Dinjens WNM, Dubbink HJ, Ter Elst A, Geurts-Giele WRR, Groenen PJTA, Groenendijk FH, Heideman DAM, Huibers MMH, Huijsmans CJJ, Jeuken JWM, van Kempen LC, Korpershoek E, Kroeze LI, de Leng WWJ, van Noesel CJM, Speel E-JM, Vogel MJ, van Wezel T, Nederlof PM, Schuuring E, Ligtenberg MJL. Recommendations for the clinical interpretation and reporting of copy number gains using gene panel NGS analysis in routine diagnostics. Virchows Arch, 2019, 474:673–80

22. Patil DT, Bronner MP, Portier BP, Fraser CR, Plesec TP, Liu X. A five-marker panel in a multiplex PCR accurately detects microsatellite instability-high colorectal tumors without control DNA. Diagn Mol Pathol, 2012, 21:127–33

23. Zapatka M, Pathogens P, Borozan I, Brewer DS, Iskar M, Grundhoff A, Alawi M, Desai N, Sültmann H, Moch H, Cooper CS, Eils R, Ferretti V, Lichter P, PCAWG Consortium. The landscape of viral associations in human cancers. Nature Genetics, 2020, 52:320–30

24. Wrzeszczynski KO, Felice V, Abhyankar A, Kozon L, Geiger H, Manaa D, London F, Robinson D, Fang X, Lin D, Lamendola-Essel MF, Khaira D, Dikoglu E, Emde A-K, Robine N, Shah M, Arora K, Basturk O, Bhanot U, Kentsis A, Mansukhani MM, Bhagat G, Jobanputra V. Analytical Validation of

Clinical Whole-Genome and Transcriptome Sequencing of Patient-Derived Tumors for Reporting Targetable Variants in Cancer. J Mol Diagn, 2018, 20:822–35

25. van der Velden DL, Hoes LR, van der Wijngaart H, van Berge Henegouwen JM, van Werkhoven E, Roepman P, Schilsky RL, de Leng WWJ, Huitema ADR, Nuijen B, Nederlof PM, van Herpen CML, de Groot DJA, Devriese LA, Hoeben A, de Jonge MJA, Chalabi M, Smit EF, de Langen AJ, Mehra N, Labots M, Kapiteijn E, Sleijfer S, Cuppen E, Verheul HMW, Gelderblom H, Voest EE. The Drug Rediscovery protocol facilitates the expanded use of existing anticancer drugs. Nature, 2019, 574:127–31

26. Jones S, Anagnostou V, Lytle K, Parpart-Li S, Nesselbush M, Riley DR, Shukla M, Chesnick B, Kadan M, Papp E, Galens KG, Murphy D, Zhang T, Kann L, Sausen M, Angiuoli SV, Diaz LA Jr, Velculescu VE. Personalized genomic analyses for cancer mutation discovery and interpretation. Sci Transl Med, 2015, 7:283ra53

27. Cheng DT, Prasad M, Chekaluk Y, Benayed R, Sadowska J, Zehir A, Syed A, Wang YE, Somar J, Li Y, Yelskaya Z, Wong D, Robson ME, Offit K, Berger MF, Nafa K, Ladanyi M, Zhang L. Comprehensive detection of germline variants by MSK-IMPACT, a clinical diagnostic platform for solid tumor molecular oncology and concurrent cancer predisposition testing. BMC Med Genomics, 2017, 10:33

28. Deans ZC, On behalf of IQN Path ASBL, Costa JL, Cree I, Dequeker E, Edsjö A, Henderson S, Hummel M, Ligtenberg MJL, Loddo M, Machado JC, Marchetti A, Marquis K, Mason J, Normanno N, Rouleau E, Schuuring E, Snelson K-M, Thunnissen E, Tops B, Williams G, van Krieken H, Hall JA. Integration of next-generation sequencing in clinical diagnostic molecular pathology laboratories for analysis of solid tumours; an expert opinion on behalf of IQN Path ASBL. Virchows Archiv, 2017, 470:5–20

29. US Food and Drug Administration. List of Cleared or Approved Companion Diagnostic Devices, 2020. https://www.fda.gov/medical-devices/vitro-diagnostics/list-cleared-or-approved-companion-diagnostic-devices-vitro-and-imaging-tools. (accessed September 6, 2020)

30. European Parliament and the council of the European Union. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU. Official Journal of the European Union L, 2017, 117:176–332

31. Sirchia F, Carrieri D, Dheensa S, Benjamin C, Kayserili H, Cordier C, van El CG, Turnpenny PD, Melegh B, Mendes Á, Halbersma-Konings TF, van Langen IM, Lucassen AM, Clarke AJ, Forzano F, Kelly SE. Recontacting or not recontacting? A survey of current practices in clinical genetics centres in Europe. Eur J Hum Genet, 2018, 26:946–54

Tables

Table 1

Performance characteristics for clinical-grade WGS using Genome-in-a-bottle (GIAB) and tumor biopsy samples. The GIAB sample has been analyzed in duplicate runs using multiple sequencers and across a time period of eight months. Data from 25 randomly selected tumor samples were used for coverage performance and a bioinformatics reanalysis of another set of 18 tumor samples (selected across a period of six months) was used to determine the pipeline reproducibility.

Quality Metric	Sample type (n)	Median Value	Range Value
Total Read count	GIAB (23)	872M	644M-1429M
Percentage Mapped	GIAB (23)	0.97	0.788-0.988
Precision SNVs	GIAB (23)	0.998	0.994-0.998
Sensitivity SNVs	GIAB (23)	0.989	0.973-0.990
F-score SNVs	GIAB (23)	0.993	0.985-0.994
Coverage hotspots regions	tumor (25)	107x	84-139
Coverage protein coding regions	tumor (25)	105x	82-133
Coverage whole-genome	tumor (25)	106x	84-130
Reproducibility bioinformatics pipeline - SNV	tumor (18)	99.98	99.98-99.99
Reproducibility bioinformatics pipeline - MNV	tumor (18)	99.96	99.89.99.98
Reproducibility bioinformatics pipeline - indels	tumor (18)	99.88	99.80-99.92
Reproducibility bioinformatics pipeline - all	tumor (18)	99.97	99.97-99.98

Table 2

ERBB2 copy number analysis by WGS and FISH. *ERBB2* FISH results were scored solely on tumor cells and categorized as; normal signals, 2-4 signals, 4-6 signals and more than 6 *ERBB2* signals (according to guidelines [15]). For WGS, the *ERBB2* copy number as well as the median ploidy of the complete chr17 is shown.

Case	WGS ERBB2	WGS chr17 ploidy	FISH HER2 normal	FISH HER2 2-4	FISH HER2 4-6	FISH HER2 >6	FISH classification
1	9	5	11%	53%	7%	30%	amplification
2	9	6	8%	34%	10%	48%	amplification
3	8	2	12%	31%	7%	51%	amplification
4	8	4	2%	32%	15%	50%	amplification
6	71	3	2%	34%	10%	54%	amplification
7	45	2	5%	10%	8%	76%	amplification
8	43	5	1%	18%	4%	77%	amplification
9	25	2	6%	41%	8%	45%	amplification
10	8	6	3%	22%	10%	65%	amplification
11	6	5	10%	82%	7%	1%	no amplification (equivocal)
12	5	2	4%	59%	37%	0%	no amplification (gain)
13	4	3	10%	20%	68%	2%	no amplification (equivocal)
14	2	1	11%	63%	25%	0%	no amplification (gain)
15	4	4	28%	44%	24%	4%	no amplification (gain)
16	6	6	52%	46%	3%	0%	no amplification

Table 3

Fusion genes detected by WGS and the Archer FusionPlex on matching DNA and RNA samples of 24 tumor biopsies.

Fusion gene details	WGS	Archer NGS	nr of samples
no fusion	none	none	7
EIF2AK2 ex12 - **ALK** ex3	yes	yes	1
EML4 ex13 - **ALK** ex20	yes	yes	5
EML4 ex2 - **ALK** ex18	yes	yes	1
EML4 ex6 - **ALK** ex20	yes	yes	2
SPAG17 ex20 - **ALK** ex9	yes	none	1
EZR ex10 - **ROS1** ex34	yes	yes	1
GOPC ex8 - **ROS1** ex35	yes	yes	2
MEF2D ex1 - **NTRK1** ex2	none	yes	1
PTPRF ex11 - **NRG1** ex6	yes	yes	1
TRPS1 ex1 - **NRG1** ex2	yes	yes	1
TMPRSS2 ex2 - **ERG** ex3	yes	n/a*	1
Total			24

* TMPRSS2-ERG fusions are not included in the used Archer FusionPlex assay.

Table 4

Detection and typing of HPV in tumor biopsies using WGS and PCR analysis.

PCR HPV result	WGS HPV result	nr samples
no high-risk HPV	no HPV detected	10
HPV high-risk type 16	HPV high-risk type 16	9
HPV high-risk type 16	no HPV detected	1
HPV high-risk type 18	HPV high-risk type 18	3
HPV high-risk (other)	HPV high-risk type 45	1
	Total	24

Figures

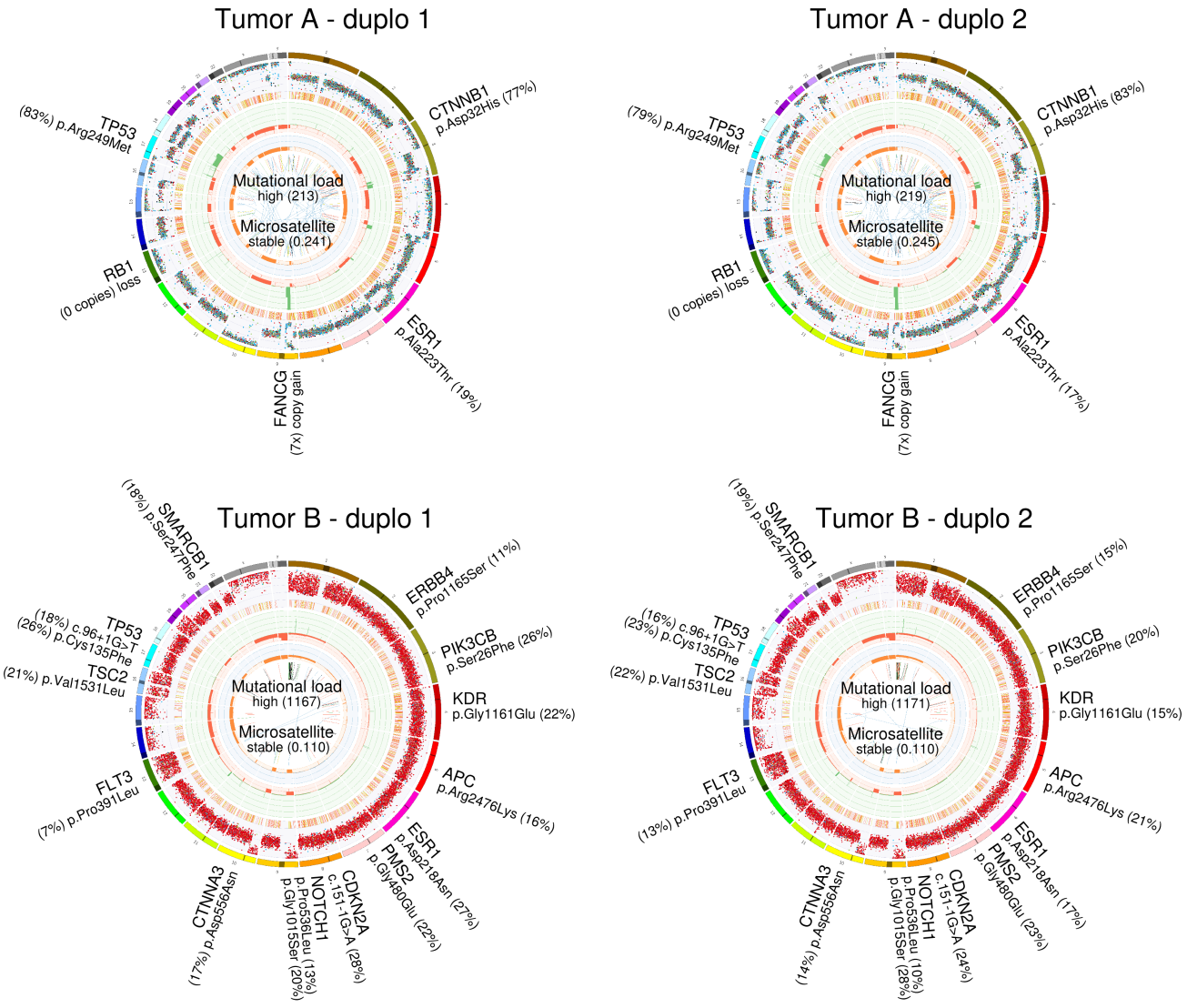


Figure 1

Representation of all tumor specific DNA aberrations as detected using WGS. For each case the complete CIRCOS is shown as well as the reported genomics events, including the mutational burden and microsatellite readout. WGS is performed in duplicate (starting with DNA isolation) for 2 tumor samples (A, non-small cell lung cancer; B, undifferentiated pleomorphic sarcoma).

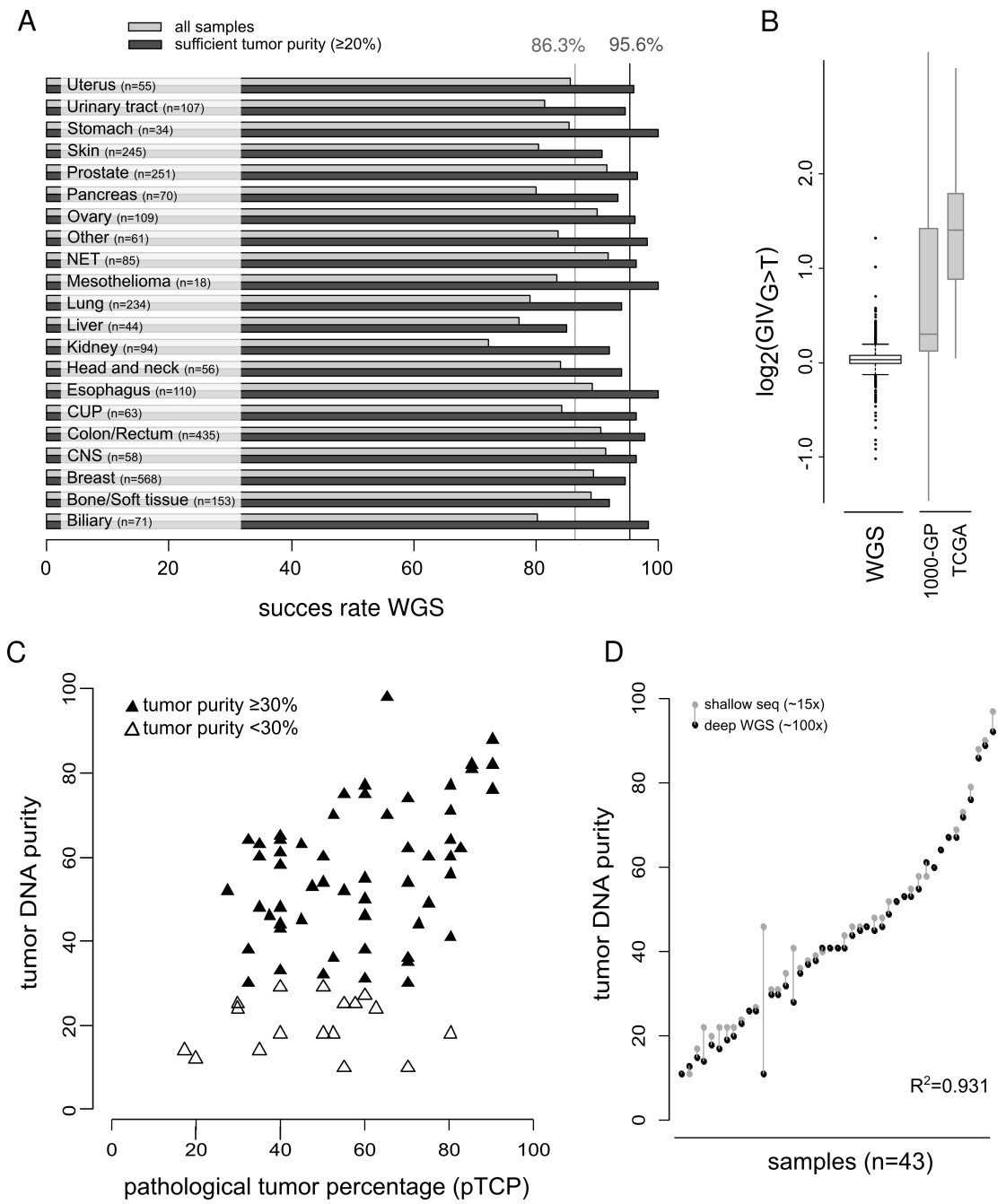


Figure 2

(**A**) WGS success rates for different primary tumor types. Success rates are shown for all samples and for samples that have sufficient tumor content. The average overall success rate across all tumor types is indicated by the vertical lines. (**B**) Global Imbalance Value G to T scores ($GIV_{G>T}$) (n=2520). As a reference the $GIV_{G>T}$ score range is depicted for the 1000 Genomes Project (1000-GP) and a TCGA subset that are described previously [18]. (**C**) Comparison of pathological tumor percentage scoring (pTCP) with sequencing based tumor DNA purity. (**D**) Comparison of tumor purity assessment using shallow sequencing (grey) (~15x) and based on deep whole genome sequencing (black) (~100x) (n=43).

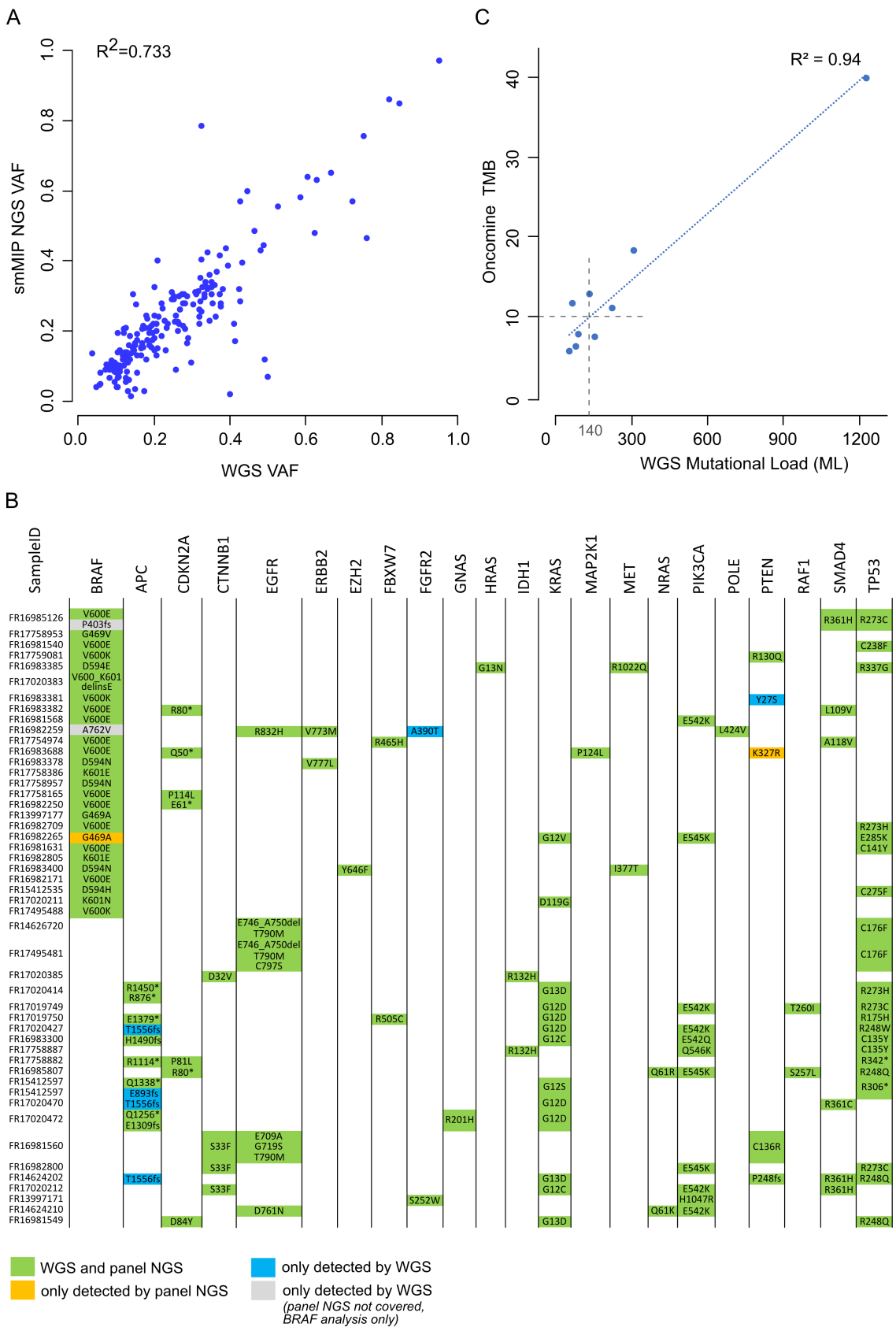


Figure 3

(**A**) Variant allele frequencies (VAF) for SNV, MNV and short indel variants that are detected using WGS and confirmed by smMIP NGS panels sequencing. (**B**) Overview of all protein-changing

mutations that are detected by WGS and or the custom-made Oncomine NGS assay. Mutations reported by both assays are marked in green, variants only reported by WGS in blue and only using the panel NGS assay in orange. For *BRAF*, also mutations detected by WGS but which are not included in the panel assay design are shown (in grey). For all other genes, only mutations included in the panel design are considered. (**C**) Comparison of WGS based mutational load (ML) readout with NGS panel based tumor mutational burden (TMB).

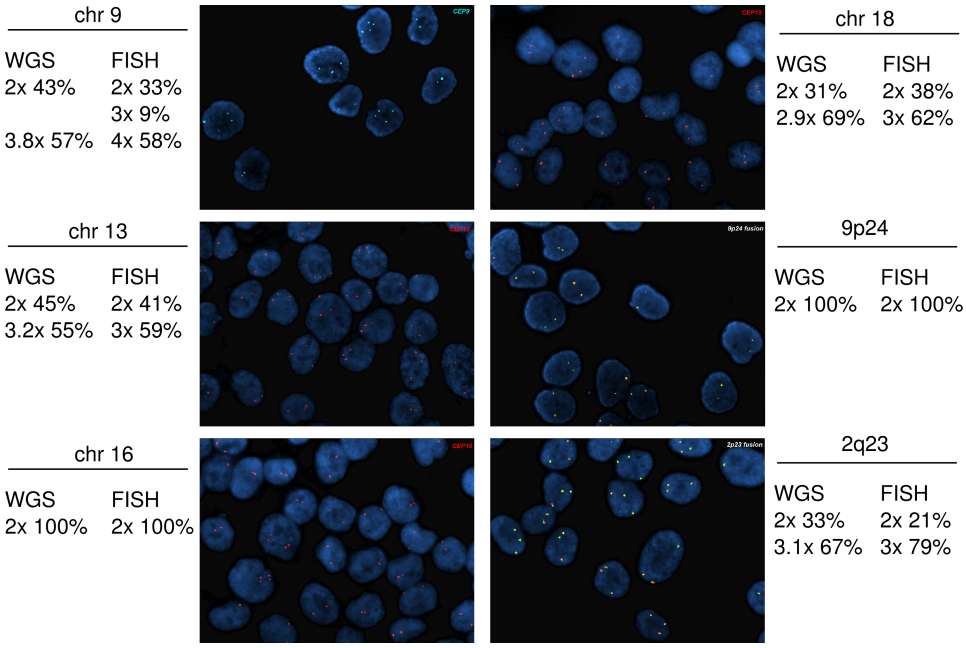


Figure 4

Comparison of COLO829 copy number analysis based on WGS and using FISH probes for copy number assessment of chromosomes 9, 13, 16 and 18, and for 9p24 (*CD274/PDCD1LG2*) and 2q23 (*ALK*). For both tests the copy number as well as the percentage of tumor cells is determined.

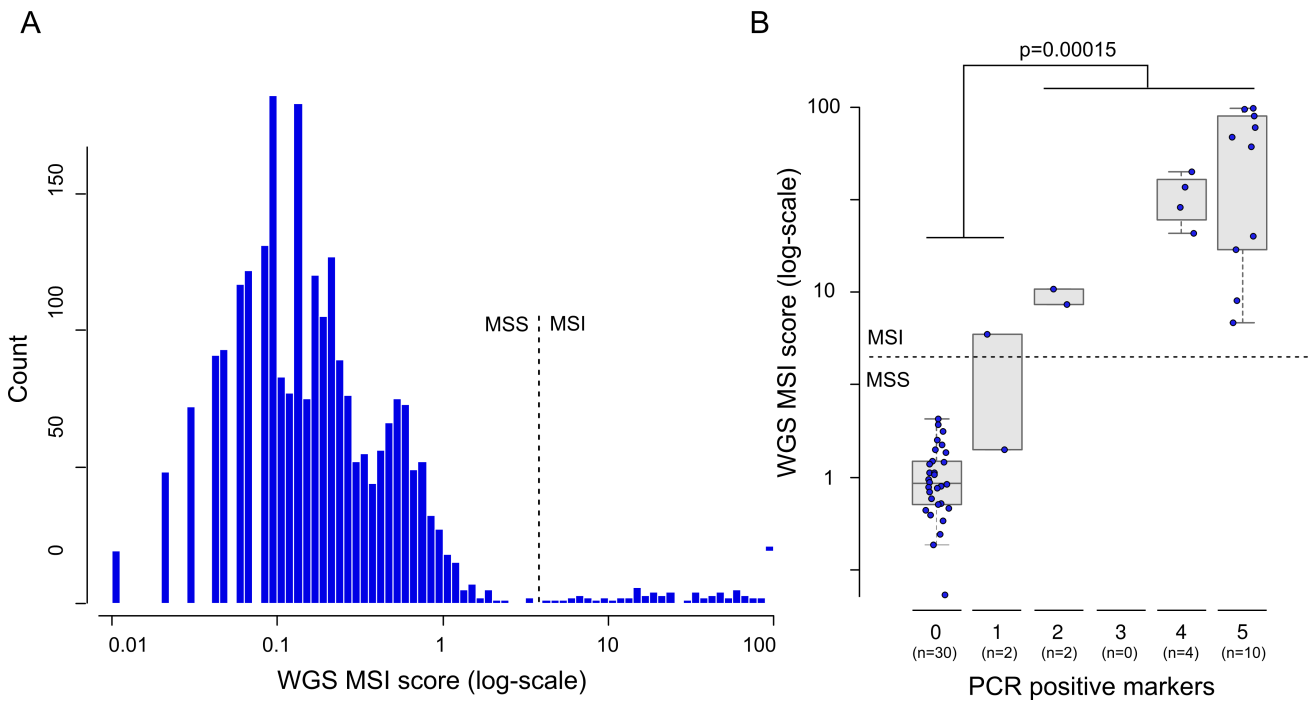


Figure 5

WGS based microsatellite instability (MSI) quantification across a cohort of 2520 metastatic cancer samples (**A**), and compared to the 5-marker PCR based test using an independent set of 48 validation samples (**B**).

Supplementary data 1: Overview validation samples

SampleID	tumor type	validation test	external ISO laboratory
0010701	Prostate cancer	smMIP panel sequencing (RUO)	Radboudumc
0120103	Head and Neck cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
0180201	Melanoma	Oncomine TML assay	Erasmus MC
0200111	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
0210701	Esophageal cancer	MSI analysis system	Erasmus MC
0410701	Prostate cancer	MSI analysis system	Erasmus MC
0500121	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
0500831	Pancreatic cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
0530701	Prostate cancer	Archer FusionPlex Solid	Erasmus MC
0590101	Lung cancer	Archer FusionPlex Solid	Erasmus MC
0700111	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
0710103	Cervical cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
0750103	Cervical cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
0840201	Kidney cancer	smMIP panel sequencing (RUO)	Radboudumc
0870201	Prostate cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
0920401	Breast cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
1000412	Prostate cancer	MSI analysis system	Erasmus MC
1000431	Melanoma	smMIP panel sequencing (RUO)	Radboudumc
1000502	Colorectal Cancer	MSI analysis system	Erasmus MC
1000731	Colorectal Cancer	MSI analysis system	Erasmus MC
1100521	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
1100631	Breast cancer	Archer FusionPlex Solid	Erasmus MC
1140501	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
1200211	Colorectal Cancer	MSI analysis system	Erasmus MC
1200611	Colorectal Cancer	MSI analysis system	Erasmus MC
1210601	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
1280201	Melanoma	Oncomine TML assay	Erasmus MC
1280201	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
1300611	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
1310701	Prostate cancer	MSI analysis system	Erasmus MC
1320601	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
1380101	Cancer of unknown primary	Archer FusionPlex Solid	Erasmus MC
1400701	Bladder cancer	MSI analysis system	Erasmus MC
1410801	Esophageal cancer	Archer FusionPlex Solid	Erasmus MC
1460201	Mesothelioma	Archer FusionPlex Solid	Erasmus MC
1530701	Glioma	Oncomine NGS gene-panel (custom)	Erasmus MC
1570201	Melanoma	Oncomine TML assay	Erasmus MC
1810501	Melanoma	HER2/neu FISH	University Medical Center Utrecht
1900401	Colorectal Cancer	MSI analysis system	Erasmus MC
1920103	GI-tract	QIAscreen HPV PCR Test	Netherlands Cancer Institute
2000202	Breast cancer	smMIP panel sequencing (RUO)	Radboudumc
2000502	Colorectal Cancer	MSI analysis system	Erasmus MC
2010401	Colorectal Cancer	MSI analysis system	Erasmus MC
2010431	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
2080101	Lung cancer	Archer FusionPlex Solid	Erasmus MC
2080201	Colorectal cancer	Oncomine TML assay	Erasmus MC
2100321	Colorectal cancer	smMIP panel sequencing (RUO)	Radboudumc
2110103	Vaginal cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
2200102	Cervical cancer	MSI analysis system	Erasmus MC
2210701	Colorectal Cancer	MSI analysis system	Erasmus MC
2310011	Breast cancer	Archer FusionPlex Solid	Erasmus MC
2350301	Glioma	Oncomine NGS gene-panel (custom)	Erasmus MC
2420801	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
2510211	Lung cancer	Oncomine NGS gene-panel (custom)	Erasmus MC

SampleID	tumor type	validation test	external ISO laboratory
2510211	Lung cancer	Archer FusionPlex Solid	Erasmus MC
2580201	Liver cancer	Oncomine TML assay	Erasmus MC
2600401	Colorectal Cancer	MSI analysis system	Erasmus MC
2720103	Stomach cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
2800701	Prostate cancer	smMIP panel sequencing (RUO)	Radboudumc
2810103	Vulva cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
2900103	Neuroendocrine	QIAscreen HPV PCR Test	Netherlands Cancer Institute
3000421	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
3000502	Glioma	MSI analysis system	Erasmus MC
3010601	Colorectal Cancer	MSI analysis system	Erasmus MC
3120103	Stomach cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
3190101	Lung cancer	Archer FusionPlex Solid	Erasmus MC
3200111	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
3220601	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
3250101	Breast cancer	smMIP panel sequencing (RUO)	Radboudumc
3300211	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
3350101	Colorectal cancer	smMIP panel sequencing (RUO)	Radboudumc
3430201	Lung cancer	Archer FusionPlex Solid	Erasmus MC
3660101	Bladder cancer	HER2/neu FISH	University Medical Center Utrecht
3700111	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
3740201	Head and Neck cancer	smMIP panel sequencing (RUO)	Radboudumc
3800401	Pancreatic cancer	smMIP panel sequencing (RUO)	Radboudumc
3800431	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
3900501	Colorectal Cancer	MSI analysis system	Erasmus MC
3990101	Penile cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
4000901	Colorectal Cancer	MSI analysis system	Erasmus MC
4010501	Colorectal Cancer	MSI analysis system	Erasmus MC
4100102	Cancer of unknown primary	MSI analysis system	Erasmus MC
4100511	Breast cancer	Archer FusionPlex Solid	Erasmus MC
4100811	Cholangiocarcinoma	smMIP panel sequencing (RUO)	Radboudumc
4250101	Lung cancer	smMIP panel sequencing (RUO)	Radboudumc
4320801	Skin cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
4410311	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
4410501	Gallbladder cancer	MSI analysis system	Erasmus MC
4480201	Melanoma	Oncomine TML assay	Erasmus MC
4500401	Lung cancer	Archer FusionPlex Solid	Erasmus MC
4600103	Esophageal cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
4610701	Prostate cancer	MSI analysis system	Erasmus MC
4640103	Anus cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
4800801	Colorectal Cancer	MSI analysis system	Erasmus MC
4900321	Breast cancer	Archer FusionPlex Solid	Erasmus MC
4920401	Prostate cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
4980101	Lung cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
4980101	Lung cancer	Archer FusionPlex Solid	Erasmus MC
4980101	Lung cancer	Archer FusionPlex Solid	Erasmus MC
5000702	Gastrointestinal stromal tumor	smMIP panel sequencing (RUO)	Radboudumc
5100211	Colorectal Cancer	MSI analysis system	Erasmus MC
5100702	Prostate cancer	MSI analysis system	Erasmus MC
5110501	Colorectal Cancer	MSI analysis system	Erasmus MC
5120401	Lung cancer	Archer FusionPlex Solid	Erasmus MC
5120401	Lung cancer	Archer FusionPlex Solid	Erasmus MC
5180201	Cancer of unknown primary	Oncomine TML assay	Erasmus MC
5200102	Colorectal Cancer	MSI analysis system	Erasmus MC
5200221	Colorectal cancer	HER2/neu FISH	University Medical Center Utrecht

SampleID	tumor type	validation test	external ISO laboratory
5200331	Colorectal Cancer	MSI analysis system	Erasmus MC
5220103	Testis cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
5230301	Sarcoma	smMIP panel sequencing (RUO)	Radboudumc
5400121	Esophageal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
5530103	Esophageal cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
5550101	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
5600411	Prostate cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
5640201	Gastrointestinal stromal tumor	smMIP panel sequencing (RUO)	Radboudumc
5700111	Neuroendocrine	Oncomine NGS gene-panel (custom)	Erasmus MC
5720103	Anus cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
5730501	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
5740201	Breast cancer	smMIP panel sequencing (RUO)	Radboudumc
5870201	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
5940101	Melanoma	smMIP panel sequencing (RUO)	Radboudumc
6000321	Colorectal cancer	smMIP panel sequencing (RUO)	Radboudumc
6000321	Colorectal Cancer	MSI analysis system	Erasmus MC
6100121	Pancreatic cancer	MSI analysis system	Erasmus MC
6100131	Stomach cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
6100331	Lymphoma	smMIP panel sequencing (RUO)	Radboudumc
6110501	Head and Neck cancer	smMIP panel sequencing (RUO)	Radboudumc
6200102	Colorectal Cancer	MSI analysis system	Erasmus MC
6200511	Lymphoma	Oncomine NGS gene-panel (custom)	Erasmus MC
6220401	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
6250101	Melanoma	smMIP panel sequencing (RUO)	Radboudumc
6250201	Esophageal cancer	MSI analysis system	Erasmus MC
6280201	Melanoma	Oncomine TML assay	Erasmus MC
6370201	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
6420401	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
6520601	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
6600103	GI-tract	QIAscreen HPV PCR Test	Netherlands Cancer Institute
6690101	Lung cancer	Archer FusionPlex Solid	Erasmus MC
6770101	Lung cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
6770101	Lung cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
6840201	Colorectal Cancer	MSI analysis system	Erasmus MC
6900601	Colorectal Cancer	MSI analysis system	Erasmus MC
6920103	Cervical cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
6930103	Liver cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
6930501	Glioma	Oncomine NGS gene-panel (custom)	Erasmus MC
6960101	Skin cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
7000103	Cervical cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
7000321	Unknown	smMIP panel sequencing (RUO)	Radboudumc
7000711	Stomach cancer	HER2/neu FISH	University Medical Center Utrecht
7050201	Colorectal Cancer	MSI analysis system	Erasmus MC
7100031	Colorectal Cancer	MSI analysis system	Erasmus MC
7100103	Head-neck cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
7100131	Lymphoma	Oncomine NGS gene-panel (custom)	Erasmus MC
7100331	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
7190101	Lung cancer	Archer FusionPlex Solid	Erasmus MC
7280101	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
7280201	Prostate cancer	Oncomine TML assay	Erasmus MC
7310211	Breast cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
7330501	Breast cancer	Archer FusionPlex Solid	Erasmus MC
7410103	Head-neck	QIAscreen HPV PCR Test	Netherlands Cancer Institute
7480101	Lung cancer	Oncomine NGS gene-panel (custom)	Erasmus MC

SampleID	tumor type	validation test	external ISO laboratory
7600601	Kidney cancer	smMIP panel sequencing (RUO)	Radboudumc
7650101	Lung cancer	Archer FusionPlex Solid	Erasmus MC
7700801	Colorectal Cancer	MSI analysis system	Erasmus MC
7870201	Kidney cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
7930501	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
8000031	Breast cancer	smMIP panel sequencing (RUO)	Radboudumc
8000811	Colorectal Cancer	MSI analysis system	Erasmus MC
8100102	Colorectal Cancer	MSI analysis system	Erasmus MC
8100111	Esophageal cancer	smMIP panel sequencing (RUO)	Radboudumc
8110501	Colorectal Cancer	MSI analysis system	Erasmus MC
8210501	Colorectal Cancer	MSI analysis system	Erasmus MC
8230301	Colorectal Cancer	MSI analysis system	Erasmus MC
8250301	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
8340101	Neuroendocrine tumor	smMIP panel sequencing (RUO)	Radboudumc
8600431	Endometrial cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
8700401	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
8720501	Colorectal cancer	HER2/neu FISH	University Medical Center Utrecht
8740201	Breast cancer	smMIP panel sequencing (RUO)	Radboudumc
8740201	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
8980101	Colorectal cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute
9010211	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
9010211	Melanoma	Archer FusionPlex Solid	Erasmus MC
9100011	Colorectal Cancer	MSI analysis system	Erasmus MC
9100811	Unknown	smMIP panel sequencing (RUO)	Radboudumc
9110211	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
9180201	Melanoma	Oncomine TML assay	Erasmus MC
9180201	Melanoma	Oncomine NGS gene-panel (custom)	Erasmus MC
9200111	Breast cancer	HER2/neu FISH	University Medical Center Utrecht
9200721	Lung cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
9320501	Pancreatic cancer	Archer FusionPlex Solid	Erasmus MC
9620401	Colorectal cancer	Oncomine NGS gene-panel (custom)	Erasmus MC
9650101	Esophageal cancer	MSI analysis system	Erasmus MC
9700801	Melanoma	smMIP panel sequencing (RUO)	Radboudumc
9710501	Stomach cancer	MSI analysis system	Erasmus MC
9750101	Colorectal Cancer	MSI analysis system	Erasmus MC
9800501	Pancreatic cancer	MSI analysis system	Erasmus MC
9950103	Anus cancer	QIAscreen HPV PCR Test	Netherlands Cancer Institute

Supplementary data 2: Example WGS rapport

Hartwig MEDICAL FOUNDATION

Hartwig Medical OncoAct

HMF SAMPLE ID
PNT00012345T

REPORT DATE
12-Jun-2020

HOSPITAL
HMF Testing Center

DNA Analysis Report

Summary

PRIMARY TUMOR LOCATION
Skin

CANCER SUBTYPE
Melanoma

The information regarding 'primary tumor location' and 'cancer subtype' is based on information received from the originating hospital.

Conclusion

Melanoma sample showing:
- activating BRAF mutation that is associated with response to BRAF-inhibitors (in combination with a MEK-inhibitor)
- complete inactivation of CDKN2A, indicating potential benefit of CDK4/6 inhibitors
- complete inactivation/loss of PTEN likely resulting in an activation of the PI3K-AKT-mTOR pathway and indicating potential benefit of mTOR/PI3K inhibitors
- high mutational burden (mutational load (ML) of 180, tumor mutation burden (TMB) of 13.6) that is potentially associated with an increased response rate to checkpoint inhibitor immunotherapy

Treatment indications (tumor-type specific)

Number of alterations with therapy indication	1 | 7 treatment(s)
Number of alterations with clinical trial eligibility	2 | 5 trial(s)

Tumor characteristics

Tumor purity of biopsy	100%
Average tumor ploidy	3.1
Tumor mutational load	High
Microsatellite (in)stability	Stable
HR Status	Proficient

Genomic alterations

Driver genes with variant(s)	TERT, CDKN2A, BRAF
Number of reported variants	5
Genes with copy-gain	NONE
Genes with copy-loss	PTEN
Gene fusions	NONE

v7.12

Therapy details (Tumor type specific)

Tumor type specific evidence

VARIANT	MATCH	TREATMENT	LEVEL OF EVIDENCE	RESPONSE	SOURCE
BRAF p.Val600Glu	Specific	Binimetinib + Encorafenib	A	Responsive	OncoKb
BRAF p.Val600Glu	Specific	Cobimetinib + Vemurafenib	A	Responsive	OncoKb
BRAF p.Val600Glu	Specific	Dabrafenib	A	Responsive	OncoKb
BRAF p.Val600Glu	Specific	Dabrafenib + Trametinib	A	Responsive	OncoKb
BRAF p.Val600Glu	Specific	Trametinib	A	Responsive	OncoKb
BRAF p.Val600Glu	Specific	Vemurafenib	A	Responsive	OncoKb
BRAF p.Val600Glu	Gene-level	RO4987655	B	Responsive	CiViC

Tumor type specific clinical trials (NL)

VARIANT	MATCH	TRIAL	CCMO	SOURCE
BRAF p.Val600Glu	Gene-level	CLXH254X2101	NL55506.078.15	iClusion
BRAF p.Val600Glu	Specific	COWBOY	NL71732.091.19	iClusion
BRAF p.Val600Glu	Gene-level	DRUP	NL54757.031.16	iClusion
BRAF p.Val600Glu	Gene-level	EBIN (EORTC-1612-MG)	NL67202.031.18	iClusion
BRAF p.Val600Glu	Gene-level	POLARIS	NL69569.028.19	iClusion
CDKN2A p.Ala68fs	Gene-level	DRUP	NL54757.031.16	iClusion

Potential eligibility for DRUP is dependent on tumor type details therefore certain tumor types may not be eligible for the DRUP. Mutational signatures (e.g. MSI, TMB) are not yet automatically matched witch clinical studies. If applicable however, matches are reported in the conclusion of the report.

The Cancer Genome Interpreter (CGI), OncoKB and CiViC knowledge bases are used to annotate variants of all types with clinical evidence. Only treatment associated evidence with a high level of evidence (A validated association; B strong clinical evidence) are reported here. Potential evidence items with a lower level of evidence (C case study, limited clinical evidence; D pre-clinical) are not reported.

Therapy details (Other tumor types)

Evidence on other tumor types

VARIANT	MATCH	TREATMENT	LEVEL OF EVIDENCE	RESPONSE	SOURCE
BRAF p.Val600Glu	Gene-level	Alpelisib + Cetuximab + Encorafenib	B	Responsive	CiViC
BRAF p.Val600Glu	Specific	Bevacizumab	B	Resistant	CiViC
BRAF p.Val600Glu	Specific	CI-1040	B	Responsive	CiViC
BRAF p.Val600Glu	Specific	Cetuximab	B	Resistant	CGI
BRAF p.Val600Glu	Gene-level	Cetuximab + Encorafenib	B	Responsive	CiViC
BRAF p.Val600Glu	Specific	Cetuximab + Irinotecan + Vemurafenib	B	Responsive	CiViC
BRAF p.Val600Glu	Specific	Dabrafenib + Panitumumab + Trametinib	B	Responsive	CiViC
BRAF p.Val600Glu	Specific	Irinotecan	B	Resistant	CiViC
BRAF p.Val600Glu	Specific	Oxaliplatin	B	Resistant	CiViC
BRAF p.Val600Glu	Specific	Panitumumab	B	Resistant	CGI
BRAF p.Val600Glu	Gene-level	Vemurafenib	B	Resistant	CiViC
PTEN Deletion	Specific	EGFR mAB inhibitor	B	Resistant	CGI
PTEN Deletion	Specific	Everolimus	B	Responsive	CiViC

The Cancer Genome Interpreter (CGI), OncoKB and CiViC knowledge bases are used to annotate variants of all types with clinical evidence. Only treatment associated evidence with a high level of evidence (A validated association; B strong clinical evidence) are reported here. Potential evidence items with a lower level of evidence (C case study, limited clinical evidence; D pre-clinical) are not reported.

Genomic alteration details

Tumor specific variants

GENE	POSITION	VARIANT	PROTEIN	READ DEPTH	COPIES	TVAF	BIALLELIC	HOTSPOT	DRIVER
BRAF	7:140453136	c.1799T>A	p.Val600Glu	**154** / 227	6	68%		Yes	High
CDKN2A	9:21971153	c.203_204delCG	p.Ala68fs	**95** / 103	2	93%	Yes	Near	High
TERT	5:1295228	c.-125_-124delCCinsTT		**49** / 64	2	77%		Yes	High
SF3B1	2:198266779	c.2153C>T	p.Pro718Leu	**76** / 115	3	67%			Medium
TP63	3:189604330	c.1497G>T	p.Met499Ile	**52** / 119	4	44%	No		Low

Tumor specific gains & losses

CHROMOSOME	REGION	GENE	TYPE	COPIES
10	q23.31	PTEN	partial loss	0

Tumor specific gene fusions

NONE

Tumor specific homozygous disruptions

NONE

Tumor specific gene disruptions

LOCATION	GENE	DISRUPTED RANGE	TYPE	DISRUPTED COPIES	UNDISRUPTED COPIES
10q23.31	PTEN	Intron 5 -> Intron 6	DEL	2	0

Tumor specific viral insertions

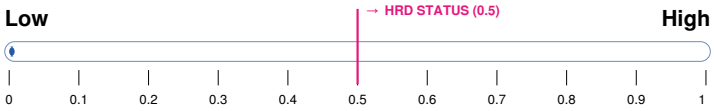
NONE

Tumor characteristics

HR-Deficiency score

Proficient 0

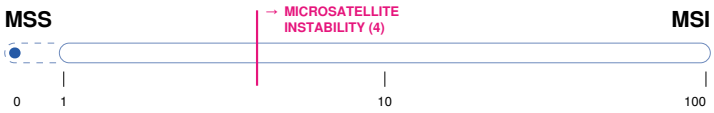
The HR-deficiency score is determined by CHORD, a WGS signature-based classifier comparing the signature of this sample with signatures found across samples with known BRCA1/BRCA2 inactivation.
Tumors with a score greater or equal than 0.5 are considered HR deficient by complete BRCA inactivation.



Microsatellite status

Stable 0.11

The microsatellite stability score represents the number of somatic inserts and deletes in (short) repeat sections across the whole genome of the tumor per Mb. This metric can be considered as a good marker for instability in microsatellite repeat regions. Tumors with a score greater than 4.0 are considered microsatellite unstable (MSI).



Tumor mutational load

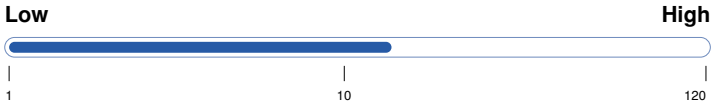
High 180

The tumor mutational load represents the total number of somatic missense variants across the whole genome of the tumor. Patients with a mutational load over 140 could be eligible for immunotherapy within the DRUP study.

Tumor mutational burden

13.6 variants per Mb

The tumor mutational burden score represents the number of all somatic variants across the whole genome of the tumor per Mb.



CIRCOS plot

The outer first circle shows the chromosomes. The darker shaded areas represent large gaps in the human reference genome: i.e. regions of centromeres, heterochromatin & missing short arms.

The second circle shows all tumor specific variants (incl. exon, intron and intergenic regions) and are divided into an outer ring of single nucleotide polymorphism (SNP) allele frequencies and an inner ring of short insertion/deletion (INDEL) locations. Variant allele frequencies have been corrected for tumor purity and scale from 0 to 100%. Each dot represents a single variant and are colored according to the type of base change (e.g. C>T/G>A in red) and are in concordance with the coloring used in Alexandrov et al. 2013 Nature paper that describes the use of mutational signatures. INDELs are colored yellow and red for insertions and deletions respectively.

The third circle shows all observed tumor purity adjusted copy number changes, including both focal and chromosomal events. Copy number losses are indicated in red, green shows regions of copy number gain. The scale ranges from 0 (complete loss) to 6 (high level gains). If the absolute copy number is > 6 it is shown as 6 with a green dot on the diagram.

The fourth circle represents the observed 'minor allele copy numbers' across the chromosome. The range of the chart is from 0 to 3. The expected normal minor allele copy number is 1, and anything below 1 is shown as a loss and represents a LOH event (orange). Minor allele copy numbers above 1 indicate amplification events of both A and B alleles at the indicated locations (blue).

The innermost circle displays the observed structural variants within or between the chromosomes. Translocations are indicated in blue, deletions in red, insertions in yellow, tandem duplications in green and inversions in black.

Report explanation

Details on the report in general

The analysis is based on reference genome version GRCh37.

Transcripts used for reporting can be found on https://resources.hartwigmedicalfoundation.nl in directory Patient-Reporting and are generally the canonical transcripts as defined by Ensembl.

Variant detection in samples with lower tumor content is less sensitive. In case of a low tumor purity (below 20%) likelihood of failing to detect potential variants increases.

The (implied) tumor purity is the percentage of tumor cells in the biopsy based on analysis of whole genome data.

Details on the reported clinical evidence

The CGI, OncoKB and CIViC knowledgebases are used to annotate variants of all types with clinical evidence, with a hyperlink to the specific evidence items. NOTE: If a certain evidence item or drug-biomarker is missing from the knowledgebases it will also not be included in this report.

More information on (CGI) biomarkers can be found on https://www.cancergenomeinterpreter.org/biomarkers

Clinical trials are matched against the iClusion database (https://iclusion.org) including a link to the specific trial.

Details on reported somatic variants

The 'Read Depth' displays the raw number of reads supporting the variant versus the total number of reads on the mutated position.

The 'Copies' field indicates the number of alleles present in the tumor on this particular mutated position.

The 'tVAF' field displays the variant allele frequency corrected for tumor purity

The 'Biallelic' field indicates whether the variant is present across all alleles in the tumor (and is including variants with loss-of-heterozygosity).

The 'Driver' field is based on the driver probability calculated based on the HMF database. A variant in a gene with High driver likelihood is likely to be positively selected for during the oncogenic process.

Details on reported gene copy numbers

The lowest copy number value along the exonic regions of the canonical transcript is determined as a measure for the gene's copy number.

Copy numbers are corrected for the implied tumor purity and represent the number of copies in the tumor DNA.

Any gene with less than 0.5 copies along the entire canonical transcript is reported as a full loss.

Any gene where only a part along the canonical transcript has less than 0.5 copies is reported as a partial loss.

Any relevant gene with more copies than 3 times the average tumor ploidy is reported as a gain.

Details on reported gene fusions

The canonical, or otherwise longest transcript validly fused is reported.

Fusions are restricted to those in a known fusion list based on CIViC, OncoKB, CGI and COSMIC

We additionally select fusions where one partner is promiscuous in either 5' or 3' position.

Details on reported gene disruptions

Genes are reported as being disrupted if their canonical transcript has been disrupted

The range of the disruption is indicated by the intron/exon/promoter region of the break point and the direction the disruption faces.

The type of disruption can be INV (inversion), DEL (deletion), DUP (duplication), INS (insertion), SGL (single) or BND (translocation).

A gene for which no wild type exists anymore in the tumor DNA due to disruption(s) is reported in a separate section called 'homozygous disruptions'

Details on reported viral insertions

Currently only integrated virus DNA can be detected.

The list of viruses that are considered can be found on https://resources.hartwigmedicalfoundation.nl.

Sample details & disclaimers

Sample details

The samples have been sequenced at **Hartwig Medical Foundation, Science Park 408, 1098XH Amsterdam**

The samples have been analyzed by Next Generation Sequencing

The results in this report have been obtained between **01-Jan-2020** and **12-Jun-2020**

The HMF sample ID is: **PNT00012345T** and the tissue ID of pathology is: **1234**

This experiment is performed on the tumor sample which arrived on **05-Jan-2020** with internal tumor barcode **FR12345678**

This experiment is performed on the blood sample which arrived on **01-Jan-2020** with internal blood barcode **FR12123488**

This experiment is performed according to lab procedures: **PREP013V23-QC037V20-SEQ008V25**

This report is generated and verified by: liekeschoenmaker

This report is addressed to: **AB, HMF Testing Center, Zip City**

Comments: This is a test report and is based off COLO829

Disclaimer

This report is generated by patient reporter **version 7.12**

The data on which this report is based is generated from tests that are performed under ISO/ICE-17025:2005 accreditation and have passed all internal quality controls.

The results stated in this report are based on the tested tumor and blood sample.

The 'primary tumor location' and 'cancer subtype' have influence on the clinical evidence/study matching. No check is performed to verify the received information.

The conclusion of this report is based solely on the results of the DNA sequencing of the tumor and the received tumor type. All other patient/tumor characteristics that might influence the interpretation of these results, are not considered. Final interpretation of the clinical consequence of this report should therefore always be performed by the treating physician.

Based on a tumor purity of at least 30%, the test has a sensitivity of >95% for detection of somatic variants and >95% for detection of translocations and gene copy number changes. For samples with a purity above 20%, the test has a sensitivity of >90%.

For feedback or complaints please contact **qualitysystem@hartwigmedicalfoundation.nl** and for general questions, please contact **info@hartwigmedicalfoundation.nl**

RVA

Edwin Cuppen,
Director Hartwig Medical Foundation

— End of report —