Tumor Genomic Profiling Reports from Different Vendors: A Comparison with Respect to Clinical Action Ability of the Provided Data

Yuriko Mori', Victor Levenson and Jeffrey Otto

Corresponding author: Yuriko Mori, CHI Institute for Research and Innovation, Center for Translational Research, Baltimore, USA

Background: Next generation sequencing (NGS) of selected genes is an expanding field of solid tumor characterization. Multiple vendors offer this service, but panel design and policies on interpretive reporting are variable. This study compared reports from selected vendors, with emphasis on clinical Action ability reporting.

Methodology: DNA aliquots of five breast and five colorectal cancers were sent to providers offering the following solid tumor NGS tests: Foundation oneTM, StrandAdvantageTM, CANCP, and GeneTrailsTM. The interpretive reports were compared for the reporting of clinically actionable variants.

Results: Detection of mutants was mostly consistent (17 of 21, 81% concordant), while detection of copy number variants was highly discordant (6 of 23, 26.1% concordant). Discordance of copy number variation appeared to be primarily due to difference in analytical sensitivity. Actionable variants with approved targeted therapies for the same tumor type were detected and interpreted uniformly by all tests. Reporting of variants with other action ability was variable, largely due to difference in panel design and analytical sensitivity, and, to a lesser degree, to different annotation policies. Actionable variant reporting by Foundation oneTM provided better coverage of clinical trials for targeted therapies while Strand advantage uniquely covered chemotherapeutic response modulation. GeneTrailsTM provided moderate coverage of clinical trials despite its small gene panel size. Utility of rearrangement targets in Foundation oneTM and StrandAdvantageTM panel was not evident, due to the tumor type assessed.

Conclusion: All enrolled tests reported actionable variants with FDA-approved target therapy for the specific tumor type with consistent clinical interpretations. In contrast, reporting of variants with other Action ability was heavily impacted by panel design, reporting policy, and annotation policy. Larger studies will be needed to quantify the qualitative differences outlined above and evaluate the benefit to the patient in order to improve the practice of precision medicine in the future.

Introduction

Genetic profiling of cancer aims to identify molecular variants in the tumor that might help guide patient management. For instance, it is recommended that KRAS and NRAS genes be genotyped in metastatic colon cancer tissue and patients with mutations in codons 12, 13 of exon 2, codons 59, 61 of exon 3, and codons 117 and 146 of exon 4 should not be treated with either cetuximab or panitumumab [1-3]. Similarly, amplification of the MET gene is also associated with poor response to cetuximab and panitumumab therapy [4-10]. Large clinical studies have found that a mutation in codon 600 of BRAF is a negative prognostic factor in metastatic colorectal cancer [11,12].

Increasing utility of multi-gene tumor profiling for patient management has led to an increase in the utilization of Next-Generation Sequencing (NGS) assays in order to identify clinically informative variants in multiple genes. These complex assays require careful planning on multiple design parameters; panel design (i.e. selection of genes and number of targets to include), assay type, NGS platforms, variant calling algorithms, thresholds to call Copy Number Variation (CNV) and rearrangements. Additionally, multiple interpretive considerations need to be included in this planning (e.g. how to interpret the oncogenicity of a variant, or its ability to activate a particular pathway, and what clinical utility to report for the variant based on literature assessment).

Different vendors make different choices along each of these dimensions. For instance, some vendors use small sequencing panels of up to approximately 50 genes, often choosing to sequence only mutation hotspot regions in these genes. Yet others use larger panels of hundreds of genes, choosing to sequence whole genes. Some vendors use the Illumina platform, others the Life Technologies platform, and yet others-the Roche 454 platform. Selection of algorithms for calling variants, amplifications, deletions and translocations to a large degree depends on the workflow upstream of sequencing and is different for amplicon sequencing and hybridization-based capture methods. The next step is the rather difficult task of curating literature and various databases to determine if a variant is indeed oncogenic, and if so, whether studies indicate that it modulates response to any of the several applicable therapies in the tumor type of interest. There are several choices in this process as well: what threshold of study size is deemed to provide sufficient evidence of clinical utility, whether to include studies on other cancer types as the supporting evidence, whether to use an inclusive or a restrictive approach in interpreting the clinical utility of variants whose significance is not fully clear, how to
handle "off-label" use of Food and Drug Administration (FDA)-approved drugs, etc.

Given this plethora of choices, it is important to understand how these choices reflect in the final report that an oncologist receives for the case at hand. Do the clinical recommendations in these reports vary widely based on these choices or are they largely consistent irrespective of these choices? To the best of our knowledge, there hasn’t been a formal study addressing this question. Here we present an analysis of clinical reports for 10 cases, sent to 4 different vendors, who have made different choices in designing their cancer profiling assays and the subsequent interpretation and reporting protocols.

Materials and Methods

Clinical samples and DNA

De-identified formalin-fixed, paraffin-embedded (FFPE) tumor tissues from 5 CRC (Colorectal Carcinoma) and 5 hormone receptor (HR)-negative breast cancer (BC) from the databank at the CHI-Center for translational research were chosen for the study. All but one BC samples were triple negative breast cancer (TNBC). All specimens demonstrated tumor cell percentages greater than 50%. This study was conducted under a protocol approved by Western Institutional Review Board (Puyallup, WA). Written consent was obtained from all patients. DNA was extracted from the samples at the CTR Molecular Diagnostic Laboratory using the Qubit 2.0 DNA concentration by using Qubit 2.0 analysis of clinical reports for 10 cases, sent to 4 different vendors, who have made different choices in designing their cancer profiling assays and the subsequent interpretation and reporting protocols.

NGS assays

The Foundation one gene panel from Foundation Medicine Inc. (Cambridge, MA) comprises 315 genes as well as introns of 28 genes involved in rearrangements. This is a hybridization-based capture panel, with sequencing performed on an Illumina platform [13].

The CANCP panel from the Mayo Medical Laboratories at Mayo Clinic (Rochester, MN) has 50 genes. It is a PCR-based amplicon hotspot panel, with sequencing performed on the Life Technologies Ion PGM platform. Unlike other tests, CANCP doesn’t assess Copy Number Variation (CNV).

The GeneTrails solid tumor panel from the Knight Diagnostic Laboratory, at the Oregon Health and Science University (Portland, OR) has 37 genes. This is a PCR-based amplicon sequencing panel, sequenced on the Life Technologies Ion PGM platform.

The StrandAdvantage panel from Strand Genomics Inc (Aurora, CO) comprises 151 genes and rearrangements in the ALK, ROS and RET genes. This is also a hybridization-based capture panel, with sequencing performed on an Illumina platform.

Targets assessed by each test are listed in supplementary Table 1. A frozen DNA aliquot per sample was shipped out to each sequencing vendor, with de-identified clinicopathological data, according to each vendor’s specification. Although FFPE tissue was the preferred submission format for all vendors, DNA submission was selected for this study in order to eliminate data variability arising from FFPE section unevenness. Additionally, a representative HE-stained tissue section for each block was provided to each vendor, each vendor performed its own quality control of the received aliquots prior to NGS assay. The NGS assay and reporting was conducted from December, 2014 to April, 2015. Thus, NGS assay and reporting described in this study reflect each vendor’s practice during this period. The list of targets assessed by each test is available in supplementary Table 1. Specific genetic regions assessed were unavailable and were not considered in our analyses. Interpretive NGS reports received from these vendors including the limit of detection data, when permission for publication is granted by the vendor, are available as part of supplementary material S1.

Analysis of reports

These reports were compared across the four tests, with primary focus on clinically actionable variants. For the purposes of comparison, such variants were defined as those that fall in one of the following 4 categories.

- Variants modulating targeted therapy approved for use in the same tumor type
- Variants modulating targeted therapy approved for use in another tumor type
- Variants modulating targeted therapy currently in trials
- Variants modulating efficacy or toxicity of chemotherapy and/or radiotherapy

Reported variants were excluded from our analysis if all reporting vendors designated them as: a) variants of unknown significance (VUS), or b) variants with clinical significance that didn’t confer any of the four forms of actionability.

The variants and their action ability annotations were then compared, with follow up literature review wherever required to assess the accuracy of the Action ability ascribed. For test reports that provided summary and detailed narrative separately (Foundation one and Strand Advantage), annotations given in the summary were used for the analysis. Discrepancies between summary and narrative were noted if existed. For test reports that did not always provide explicit statement on FDA-approved tumor type (GeneTrails and CANCP), we assigned the category according to the current approval status in order to enable the comparison. Erroneous FDA-approval status assignment in reports was also corrected based on our literature review for the purpose of comparison.

Results

Characteristics of reports

Not all of the 10 chosen samples ran successfully on all tests; 1-2 samples for each vendor were dropped due to the factors beyond the vendors’ control, such as sample exhaustion or compromised sample during shipment. Table 1 shows sample description and report availability.
Foundation one provided reports for all 10 samples, but inspection of initial reports revealed that samples #358 (TNBC) and #38 (CRC) were interchanged by the vendor at a preanalytical stage. Reanalysis and acquisition of corrected report were done for sample #358 but not for sample #38 due to sample exhaustion. Thus, the Foundation One report for sample #38 was dropped from the analysis. Of note, DNA submission was a non-preferred sample format for all vendors, thus preanalytical workflow in this study was not their standard practice.

All tests included actionable variants, as well as VUS and clinically significant variants without Action ability (defined in the Method section), in their reports. Some differences in reporting were observed; a) CANCP didn’t assess CNVs, b) Strand Advantage reported germline polymorphisms, based on the population allele frequency and variant allele prevalence in the tumor, if they were deemed actionable, c) Foundation One stated that the Action ability annotations were “not evaluated for source or level of published evidences,” d) GeneTrails and CANCP did not always explicitly state the FDA-approval status of the referred agents (i.e. approved for the patient’s or other tumor types), e) CANCP mentioned potential targeted therapies under research but did not indicate whether pertinent clinical trial existed or not, f) Foundation One and Strand Advantage provided detailed clinical trial information (e.g. trial ID, title, and phase), while Gene Trails provided only trial ID, g) Gene Trails did not provide supporting literature for clinical significance, and h) Foundation One and Strand Advantage provided summary of detected variants and their action ability separately from narrative interpretation. Median report length in page number greatly differed among vendors; 16 for Foundation one, 9.5 for StrandAdvantage, three for GeneTrail, and four for CANCP. The difference in report length was attributed to factors e)- h) listed above. Table 2 lists the clinically actionable variants reported for each sample by each vendor, along with their Action ability annotation according to the four criteria defined in the Method section.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Cancer Type</th>
<th>Type</th>
<th>Foundation one</th>
<th>Strand Advantage</th>
<th>GeneTrails</th>
<th>CANCP</th>
</tr>
</thead>
</table>
Sixty variants were reported as clinically actionable by at least one test in this study. Thirty-nine of these variants were assessed, at least in one sample, by multiple tests that were capable of reporting them based on the panel design and vendor's reporting policy. A total of 44 variant detection events met this description, and could be evaluated for the subsequent analyses, we corrected FDA-approval status error/ambiguity in the reports to the actual approval status. Intra-vendor annotation discrepancies were left unrevised, since it is impossible to surmise which of the conflicting statements represents the vendor's standing at the time.

**Test-to-test comparison in variant detection**

Sixty variants were reported as clinically actionable by at least one test in this study. Thirty-nine of these variants were assessed, at least in one sample, by multiple tests that were capable of reporting them based on the panel design and vendor's reporting policy. A total of 44 variant detection events met this description, and could be evaluated discrepant action ability annotation of loss of function mutations in TP53 for stage IIA TNBCs (i.e. category III-actionable in sample #84 and #241, versus non-actionable in sample #275).
for test-to-test consistency in variant detection. Discordance among tests was significantly more frequent for detection of CNV (17 of 23, 73.9%) than mutations (4 of 21, 19.0%; p < 0.01, Fisher’s exact test). There was no significant difference in discrepancy rate based on the type of mutation (frameshift vs. missense; data not shown) or CNV (amplification vs. deletion; data not shown). Of note, some discrepancies in mutation detection may have stemmed from the difference in specific genetic regions covered by each vendor due to the difference in target capture methods in hot spot amplicon designs. We could not assess this possibility because the region information was not available for some vendors offering hotspot panels.

There were differences in reporting when discrepancy in variant detection was reported (Table 3).

Table 3: Discrepantly detected variants.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Sample</th>
<th>Foundation One</th>
<th>Strand Advantage</th>
<th>GeneTrails</th>
<th>CANCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC p.Arg232X</td>
<td>80</td>
<td>y</td>
<td>y</td>
<td>NA</td>
<td>n</td>
</tr>
<tr>
<td>SMO p.Ala601Val</td>
<td>240</td>
<td>y</td>
<td>y</td>
<td>NA</td>
<td>n</td>
</tr>
<tr>
<td>TP53 p.Trp53X</td>
<td>241</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>n</td>
</tr>
<tr>
<td>TP53 p.Val122fsX26</td>
<td>84</td>
<td>y</td>
<td>NA</td>
<td>y</td>
<td>n</td>
</tr>
<tr>
<td>AURKA amplification</td>
<td>240</td>
<td>n</td>
<td>y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>EGFR amplification</td>
<td>240</td>
<td>n</td>
<td>y</td>
<td>n</td>
<td>NA</td>
</tr>
<tr>
<td>FGFR1 amplification</td>
<td>358</td>
<td>n</td>
<td>y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>FGFR2 amplification</td>
<td>358</td>
<td>n</td>
<td>y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>GNAS amplification</td>
<td>26</td>
<td>y</td>
<td>NA</td>
<td>n</td>
<td>NA</td>
</tr>
<tr>
<td>JAK2 amplification</td>
<td>241</td>
<td>n</td>
<td>y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>KRAS amplification</td>
<td>241</td>
<td>n</td>
<td>y</td>
<td>n</td>
<td>NA</td>
</tr>
<tr>
<td>MET amplification</td>
<td>202</td>
<td>y (eq.)</td>
<td>y</td>
<td>n</td>
<td>NA</td>
</tr>
<tr>
<td>MET amplification</td>
<td>358</td>
<td>n</td>
<td>y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>MYC amplification</td>
<td>241</td>
<td>n</td>
<td>y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>NTRK1 amplification</td>
<td>358</td>
<td>n</td>
<td>y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SMO amplification</td>
<td>358</td>
<td>n</td>
<td>y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>TP53 deletion</td>
<td>80</td>
<td>n</td>
<td>n</td>
<td>y</td>
<td>NA</td>
</tr>
<tr>
<td>TP53 deletion</td>
<td>202</td>
<td>n</td>
<td>n</td>
<td>y</td>
<td>NA</td>
</tr>
<tr>
<td>TP53 deletion</td>
<td>289</td>
<td>n</td>
<td>n</td>
<td>y</td>
<td>NA</td>
</tr>
<tr>
<td>TSC2 deletion</td>
<td>289</td>
<td>n</td>
<td>n</td>
<td>y</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 3: Discreptantly detected variants.

StrandAdvantage didn’t test sample #26 and #84. GeneTrails didn’t test sample #358. CANCP didn’t perform CNV analyses. Other events marked as NA were not assessed because the specific target was not in the panel. This table includes variants reported as non-actionable or VUS.

These involved 4 mutations (APC p.Arg232X, SMO p.Ala601Val, TP53 p.Trp53X, and TP53 p.Val122fsX26). In all four cases, CANCP was the only vendor that didn’t report the mutations. For CNV, discrepancy involved amplification of 11 loci (AURKA, EGFR, FGFR1, FGFR2, GNAS, JAK2, KRAS, MET, MYC, NTRK1, and SMO) and deletion of 3 loci (TP53, BRCA2, and TSC2). GeneTrails was most often the sole vendor for reporting deletions (4 events), followed by StrandAdvantage (one event). StrandAdvantage most frequently was the sole vendor reporting amplifications (11 events), followed by Foundation one (three events). Interestingly, one amplification event was reported as “amplification” by StrandAdvantage, “equivocal amplification” by Foundation one, and not reported by GeneTrails. These findings strongly suggest that detection discrepancy depended on difference in definition of reportable CNV and analytical sensitivity setting, rather than technical error. An independent evaluation of discrepant variants on the studied samples was not performed by us.

Test-to-test comparison of actionable variant annotation and detection

The number of clinically actionable variants reported greatly varied among the vendors. Figure 1A summarizes this number for each vendor.

StrandAdvantage didn’t test sample #26 and #84. GeneTrails didn’t test sample #358. CANCP didn’t perform CNV analyses. Other events marked as NA were not assessed because the specific target was not in the panel. This table includes variants reported as non-actionable or VUS.

These involved 4 mutations (APC p.Arg232X, SMO p.Ala601Val, TP53 p.Trp53X, and TP53 p.Val122fsX26). In all four cases, CANCP was the only vendor that didn’t report the mutations. For CNV, discrepancy involved amplification of 11 loci (AURKA, EGFR, FGFR1, FGFR2, GNAS, JAK2, KRAS, MET, MYC, NTRK1, and SMO) and deletion of 3 loci (TP53, BRCA2, and TSC2). GeneTrails was most often the sole vendor for reporting deletions (4 events), followed by StrandAdvantage (one event). StrandAdvantage most frequently was the sole vendor reporting amplifications (11 events), followed by Foundation one (three events). Interestingly, one amplification event was reported as “amplification” by StrandAdvantage, “equivocal amplification” by Foundation one, and not reported by GeneTrails. These findings strongly suggest that detection discrepancy depended on difference in definition of reportable CNV and analytical sensitivity setting, rather than technical error. An independent evaluation of discrepant variants on the studied samples was not performed by us.

Test-to-test comparison of actionable variant annotation and detection

The number of clinically actionable variants reported greatly varied among the vendors. Figure 1A summarizes this number for each vendor.

Figure 1A: Overall trends of Action ability reporting by each vendor this bar graph shows the number of variants reported as clinically actionable by each vendor according to the 4 Action ability categories. Category description is as follows: I- targeted therapy approved for the same tumor type, II- targeted therapy approved for another tumor type, III- targeted therapy in clinical trials, IV- modulation of efficacy or toxicity of chemotherapy or radiotherapy.

However, simple comparison of reported variants may not be appropriate, since multiple factors may impact Action ability; a) error in detection or annotation, b) variant detection discrepancy due to analytical characteristics of assays, which were not independently validated in this study, c) different panel design and variant types.
assayed, and d) different clinical Action ability annotation. Furthermore, as was mentioned earlier, not all vendors could assess all samples due to sample exhaustion or incidents during shipment, which impacted numbers of the detected variants.

In order to eliminate the impact of samples that were not assessed by a specific vendor from this analysis, we calculated the Action ability detection rate of each vendor relative to total Action ability detected by all vendors in the samples that the vendor could report (Figure 1B).

Figure 1B: This bar graph shows the ratio of actionable variants reported by a vendor to the total actionable variants reported by any of the four vendors for all samples. Sample that wasn't assessed by a vendor were excluded from calculation of the total number applied for the vendor's ratio calculation.

At first glance, category I Action ability (FDA-approved targeted therapy for the same tumor type) reporting was comparable for all vendors regardless of the panel size. CANCP reported actionable variants only in this category. Category II Action ability reporting (FDA-approved targeted therapy for other tumor types) was markedly higher for Foundation one and StrandAdvantage than GeneTrails. Foundation one had the highest report rate for category III Action ability (targeted therapy in clinical trial), followed by both StrandAdvantage and GeneTrails. Category IV Action ability (modulation of efficacy or toxicity of chemotherapy and radiotherapy) was reported only by StrandAdvantage. Figure 1C shows that this trend held true in the analysis excluding discrepantly detected variants (17 CNVs and 4 mutants described earlier), which should most directly reflect panel design and annotation/reporting policy difference.

Figure 1C: This bar graph principally shows the same data as B, except, only the data for variants without any detection discordance.

To better visualize differences in Action ability reporting between vendors and elucidate their potential causes, all clinically actionable variants reported in this study are listed according to the vendors and their status for each vendor (Figures 2-5 for Action ability categories I-IV, respectively).

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Tumor type</th>
<th>EGFR amp.</th>
<th>BRAF p.Val600Glu</th>
<th>KRAS p.Gly13Asp</th>
<th>ERBB2 amp.</th>
<th>Foundation One</th>
<th>Strand Advantage</th>
<th>GeneTrails</th>
<th>CANCP</th>
<th>Discrepancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>240</td>
<td>CRC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
</tr>
<tr>
<td>289</td>
<td>CRC</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>CRC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>HR- HER2+ BC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Foundation One annotated Everolimus as FDA-approved in the same tumor type (Category 1), which actually is approved for only for HR+ HER2- breast cancers (Category 2) in sample #275 (PTEN p.Gly132Asp) and #385 (AKT amplification). These possibly erroneous annotations were excluded from this figure.

Figure 3: Variants annotated as category II-actionable. This figure shows variants annotated as actionable by at least one vendor, in the same fashion as Figure 2. White asterisk in dark grey cell indicates variants that were reported as actionable in summary table yet described as VUS in narrative of the report. CANCP didn’t report category II-IV Action ability explicitly, and was not considered in the evaluation of vendor-to-vendor discrepancies for these categories.

Figure 4: Variants annotated as category III-actionable. This figure shows variants designated as with category III Action ability by at least one vendor, in the same fashion as Figures 2 and 3.

Figure 5: Variants annotated as category IV-actionable. This figure shows variants designated as with category IV Action ability by at least one vendor, in the same fashion as Figures 2-4.

As was noted earlier in Results, StrandAdvantage reports several amplifications which are not reported by other vendors, while GeneTrails reports a few deletions which are not reported by other vendors.

Larger panel design provided more actionable variants in category III (Foundation one and StrandAdvantage; Figure 4) or category IV (StrandAdvantage; Figure 5). Six variants were reported in targets assessed only by Foundation one (mutation of PTCH1 and BARD1,
CNV of CDK6, MYCL1, and CCND2, as well as PTEN intron 1 rearrangement). The report also indicated possible targeted therapy clinical trials (except for PTCH1, where targeted therapy was approved for another tumor type). Six variants were reported for targets assessed only by StrandAdvantage (germline mutation of ERCC2. ERCC5, XRCC1, ABCC1, DPYD, and MLH1) which indicated either sensitivity or toxicity to chemotherapy.

In addition to the already described factors (i.e. difference in panel design and CNV detection sensitivity, or absence of category II-IV Action ability reporting by CANCPr), discrepancies in Action ability reporting among vendors was likely dependent on differences in annotation. Major difference in annotation was observed for a total of 21 variants (including two occasions of double mutations of a gene in the same sample) corresponding to 12 genes (APC, DDR2, ERBB2, FBXW7, FGFR1, KRAS, NOTCH2, PIK3CA, PIK3R1, SMO, TP53).

Distribution of different Action ability annotation was as follows: seven for category II (Figure 3), 13 for category III (Figure 4, including one occasion of double mutations in a single gene), and eight for category IV (Figure 5, including one occasion of double mutations in a single gene). No differences in action ability annotation were observed for category I.

Most differences in annotation appeared to depend on annotation focus and principles. This topic has been summarized in introduction and will be discussed further in Discussion. However, there were several occasions where either format of the report or an error seemed to contribute to reporting differences. GeneTrails free-text format reports didn’t always state specific drug name or FDA-approved tumor types, and often referred to off-label use of FDA-approved drugs under clinical trials. This reporting format may have attributed to the lack of category II annotation by GeneTrails (Figure 3) for KRAS p.Gly13Asp (sample #20), PIK3CA p.Cys420Arg (sample #26), ERBB2 amplification (sample #26), and FGFR1 amplification (sample #84). For variants SMO p.Ala601Val, PIK3CA p.Asn345Ser, and PIK3CA p.Asn1068fsX3+, Foundation one reported clinical trial availability in the summary, yet in the narrative stated that “mutation reported here has not been characterised.” This inconsistency within the report, possibly error, is likely the source of annotation discordance for these variants, since other vendors labelled these variants as VUS (Figure 4).

Additionally, Foundation one annotated Everolimus as FDA-approved in the same tumor type (Category I) for two TNBC cases (#275, with PTEN p.Gly132Asp, and #385, with AKT amplification; Table 2), when it was approved only for HR+ HER2- breast cancers (Category II). This potential error was corrected prior to Action ability annotation comparison by us, as was stated under the Characteristics of reports, thus it is not reflected in Figures.

Discussion

The challenges in identifying actionable variants have been highlighted by various groups [14,15]. It is known that only certain ‘driver’ mutations are of consequence in tumorigenesis. Such driver mutations are usually shown to be functional in in vitro assays where oncogenic potential or activation of downstream pathways is observed. In the absence of this information, frequent occurrence in tumors may be taken as a surrogate. In the absence of either line of evidence, this inference may be speculative. Further, even if a variant is indeed confirmed to be oncogenic, its connection to therapy may be indirect and reliant on activation of a certain pathway. Even further, such pathway activation may be confounded by the presence of additional variants that modulate the pathway, which must be included in the assessment in vivo. Finally, reporting of Action ability requires considerations of practical matters, such as the handling of unapproved drugs and “off-label” use of approved drugs. Off-label use of drug is allowed when appropriate, but patient may encounter issues with reimbursement or access to pertinent clinical trials [16,17]. Similarly, recommendation of clinical trials where preliminary safety data have yet to be established (i.e. phase I) may require extra caution.

In the current study, there were multiple variants which are found by at least two vendors, but reported with differing clinical Action ability. As was mentioned in Results, CANCPr did not explicitly report any Action ability other than FDA-approved targeted therapy in same tumor type. Some variants involved major discordance (i.e. presence versus absence of actionable variants in a category), and the others involved minor difference in recommendation within the category.

Major discordance

Major discordance involves Action ability annotation at category level. This type of discordance was observed for samples with variants of the TP53 and APC genes, NOTCH2 p.Arg2400X, FBXW7 p.Arg465His, PIK3CA p.Asn345Ser, PIK3R1 p.Lys676Glu, PTEN p.Gly132Asp, and DDR2 amplification. Our review revealed a few key contributing factors for this discordance: a) different functional significance annotation, due to different choices made at multiple decision points described above (in general, GeneTrails was the most aggressive, and CANCPr was the most conservative), b) subtle difference in vendor’s field of focus (i.e. StrandAdvantage was almost the only vendor reporting Action ability on chemotherapy modulation), and c) handling of unapproved drugs and off-label use of approved drugs in the absence of clinical trials at phase II or later (Foundation one and GeneTrails made recommendation, while StrandAdvantage did not). Also, Foundation one noted impact of co-founding mutation/variant (e.g. HR status for Everolimus, oncogene mutation for FBX7 mutant targeting mTOR inhibitor use) in targeted therapy but did not mention it in the summary recommendation. Each of these variants is discussed in depth below with emphasis on differences in interpretation between vendors.

TP53 variants

TP53 somatic mutations were frequently detected by all vendors, but reported clinical Action ability markedly differed. Foundation one and GeneTrails referred to clinical trials, while StrandAdvantage only mentioned chemotherapy sensitivity/toxicity specific to the loss-of-function variant, p.Arg282Trp, with some reservation (sample #240, TNBC). All clinical trials reported by Foundation one and GeneTrails were phase I. StrandAdvantage cited indirect evidences for it’s claim: TP53 p.Arg282Trp mutation was associated with shorter overall survival in a mixed cancer patient population, as well as with possible resistance to cytochrome 450 metabolized drugs in cancer cell lines [18].

APC variants

APC somatic mutations were frequently detected by all vendors (Foundation one, StrandAdvantage, and CANCPr) that tested APC in their panel [19]. However, StrandAdvantage is the only vendor that attributes these variants (truncation mutations in exon 14 or 16) to poor response to chemotherapy. Supportive evidence for this claim included a clinical study of CRC patients that suggested an association between nonsense or frameshift mutations in exon 16 of APC and poor
response to 5-fluorouracil [20], with additional preclinical evidence to this effect.

**Other variants with major discordance**

The NOTCH2 truncation variant p.Arg2400X was detected in sample #241 by Foundation one and StrandAdvantage. It is a truncating mutation in a gene that behaves both as a tumor suppressor and as an oncogene [21], and has been reported as an indicator of potential response to Notch inhibitors (in clinical trial) by Foundation one and as non-actionable by StrandAdvantage. StrandAdvantage didn't provide explanation for this designation, indicating that it annotated this variant as VUS. Foundation one cited a study of this variant showing its in vitro NOTCH2 activating effect as well as its co-occurrence only with NOTCH2 copy number gain in lymphoma patients [22]. All clinical trials mentioned by Foundation one were at phase I, and no clinical study data are available at present for NOTCH2 inhibitor effectiveness in breast cancer patients.

The missense variant FBXW7 p.Arg465His in sample #289 was detected by Foundation one, StrandAdvantage, and CANCP. It was deemed as actionable with clinical trials by Foundation one and non-actionable by StrandAdvantage. The referenced clinical trials were at phase I or II. Both vendors mentioned preclinical effectiveness of mTOR inhibitor therapy against FBXW7 variants [23]. Foundation one further supported this claim with a clinical case report [24], while StrandAdvantage dismissed the Activity ability in this particular patient based on a clinical study showing that concomitant KRAS mutation mitigates the effectiveness of mTOR inhibitor therapy targeting FBXW7 variants [25,26]. CANCP didn't report any clinical trials, but cited the same study and mentioned the negative interaction by co-existing mutations in oncogenes, such as KRAS.

The PIK3CA p.Asn345Ser variant was detected in sample #80 by all vendors that assessed these genes (CANCP, only analysed PIK3R1). However, they were listed as functional only by GeneTrails (Foundation one listed this PIK3CA variant as actionable in summary but as VUS in narrative—an error mentioned in Results). StrandAdvantage didn't provide literature evidence for this annotation. CANCP stated the research on the PIK3CA variant was insufficient to ascertain its functionality. GeneTrails cited no reference for PIK3R1 variant, but stated that another mutant at the PIK3CA codon 345 demonstrated in vitro tumorigenicity [27]. This study was cited by Foundation one, too. Foundation one also stated that the detected PIK3CA and PIK3R1 missense mutations were located within domains of interaction between PIK3R1 and PIK3CA to negatively regulate the Akt pathway [28]. Foundation one also acknowledged that PIK3R1 p.Lys567Glu heterozygous cells was shown to be non-tumorigenic in vitro [29] and that direct evidence was missing for oncogenicity of PIK3CA p.Asn345Ser. The clinical trials recommended by GeneTrails were at phase I or II.

The PTEN p.Gly132Asp variant was detected in sample #275 by all vendors. Foundation one and CANCP regarded it non-functional (Foundation one report had summary-narrative inconsistency for this variant; while summary stated this variant to be actionable narrative referred to it as functionally unclear). GeneTrails and StrandAdvantage designated this variant as actionable with approved drug for another tumor type in clinical trials. Foundation one, CANCP, and GeneTrails didn't provide evidence for the functional significance of this missense variant. StrandAdvantage provided indirect support: in vitro data of PTEN functional loss caused by missense mutations at neighbouring codons (codons 121-131, and 134; [30,31]), and in-silico functional prediction results (Strand's internal data). Referenced trials were phase I or II. StrandAdvantage cited clinical studies that showed mTOR inhibitor effectiveness in human cancers both dependent [32] and independent [33] of tumor PTEN-status, in addition to a few preclinical studies.

DDR2 amplification was detected in sample #358 by Foundation one and StrandAdvantage, but only StrandAdvantage reported it as clinically actionable. Foundation one designated this variant as VUS with no further description. StrandAdvantage stated that, while direct evidence for DDR2 amplification was lacking, DDR2 activation was shown to be oncogenic in a few preclinical studies of DDR2 activating mutation [34-36]. Similarly, the response to tyrosine kinase inhibitor therapy in patients with DDR2 activating mutations was referenced as a supportive evidence [34,37,38]. The recommended clinical trial was at phase II.

**Minor discordance (intra-category discrepancy)**

An example of minor discrepancy in recommendation between vendors was seen for KRAS p.Gly13Asp mutation in a CRC sample #80. All vendors report resistance to cetuximab and panitumumab (both category I) for this sample, but only StrandAdvantage noted additional detail that this variant may be more responsive to these drugs than other KRAS codon 12/13 mutants, based both clinical and preclinical studies [39-41]. Also, only StrandAdvantage reported resistance to VEGF inhibitor Bevacizumab (category I) for this variant based on meta-analysis of 12 clinical trial data [42]. Further, a category II recommendation for this variant differed between Foundation one (Trametinib) versus StrandAdvantage (Sorafenib). Sorafenib was reported by StrandAdvantage to have clinical evidence specifically in CRC as a second or later-line treatment [43], while clinical effectiveness of Trametinib was reported in melanoma [44] and other solid tumors [45,46], but not yet specifically in CRC.

**Conclusions**

Our comparison between the 4 vendors shows substantial commonality as well as differences of commercial NGS-based gene panel analysis of solid tumor in panel design, reporting policies, and clinically actionable information provided. While the study was too small to derive statistically significant inferences, some qualitative trends were observed. All enrolled tests, regardless of the panel design, were capable of reporting well-established actionable variants (i.e. those with FDA-approved target therapy for the specific tumor type) with consistent clinical interpretations. In contrast, reporting of other Action ability was heavily impacted by panel design, reporting policy, and annotation policy.

In general, the larger panels (Foundation one and StrandAdvantage) detected and reported more actionable variants besides targeted therapies approved for the specific tumor type than the remaining smaller panels. Of the latter, GeneTrails covered clinical trial Action ability well, while CANCP massively limited its actionable variant detection capability by not reporting CNVs and not referring to specific clinical trials. Foundation one reported more genes and rearrangements with targeted therapy in trials, particularly in early phase trials with novel investigational drugs, compared the other panels. StrandAdvantage provided information on variant modulation of chemotherapy efficacy and toxicity by germline and somatic variants, which the other vendors didn't provide. StrandAdvantage
reported relatively less clinical trial Action ability for its expansive panel size, in part due to its annotation policy that is further described in the next paragraph. Of note, the assessed tumor types were not those in which gene rearrangements have large impact on Action ability (e.g. non-small cell lung cancer). Therefore, the current study is unsuited for full evaluation of the utility of the Foundation one and StrandAdvantage panels’ rearrangement targets.

Based on the examination of variants with different interpretations several important distinctions were noted in policies for clinical trials reporting: decision to report (Foundation one and GeneTrail) or not (StrandAdvantage) variants when only phase I trials are available; decision to report off-label recommendations in the absence of a trial relevant to the tumor type (Foundation one) or not (StrandAdvantage); decision to report Action ability supported only by preclinical data (Foundation one) or not (StrandAdvantage and GeneTrails). CANCP didn’t report availability of specific clinical trial at all. Accordingly, StrandAdvantage was generally more conservative than other vendors in reporting available clinical trials, thus their coverage of clinical trials was relatively small for its panel size.

Analytical sensitivity/specificity of mutation detection appeared comparable among vendors, while CNV reporting was highly variable. The difference in reporting accounted for significant differences in provided clinically actionable information. StrandAdvantage detected more clinically actionable amplifications, while GeneTrails detected more clinically actionable deletions than the other vendors. It should be noted that the veracity of CNV calls was not independently verified in our study. Vendors are encouraged to perform this validation on this own. It may also be desirable to develop a consensus in the definition of clinically impactful CNV.

Report formats of Foundation one and StrandAdvantage were the most comprehensive with statement of drug FDA-approval status, clinical trial description, and references to supporting literature. Both vendors provided concise summary for convenience, but caution should nonetheless be exercised for occasional discrepancy between summary and narrative (observed in Foundation one). GeneTrails and CANCP had shorter report formats. GeneTrails didn’t provide literature information, thus verification of the recommendation required independent literature search. CANCP mentioned ongoing research with literature but without information on availability of a specific drug or a clinical trial, requiring active literature/database search to take full advantage of the variant data.

All the vendors continue to improve their reports and gene panel contents, so the trends reported in our study, may change. It is clear though that sequencing methods, choice of genes, bioinformatics methods, and variant curation and interpretation policies will have a major bearing on the accuracy and nature of reporting. Larger studies will be needed to quantify the qualitative differences outlined above and evaluate the benefit to the patient. Such rigorous analysis will no doubt improve the practice of precision medicine in the future.

**References**


