Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Pharmacogenomics

Pan-cancer pharmacogenetics: targeted sequencing panels or exome sequencing?

Laurentijn Tilleman[1], Björn Heindryckx[2], Dieter Deforce[1] & Filip Van Nieuwerburgh*,[1]

[1]Laboratory of Pharmaceutical Biotechnology, Ghent University, Ottergemsesteenweg 460, Ghent 9000, Belgium

[2]Department for Reproductive Medicine, Ghent-Fertility & Stem Cell Team (G-FaST), Ghent University Hospital, Corneel Heymanslaan 10, Ghent 9000, Belgium

*Author for correspondence: Filip.VanNieuwerburgh@ugent.be

Aim: This study provides clinicians and researchers with an informed choice between current commercially available targeted sequencing panels and exome sequencing panels in the context of pan-cancer pharmacogenetics. **Materials & methods:** Nine contemporary commercially available targeted pan-cancer panels and the xGen Exome Research Panel v2 were investigated to determine to what extent they cover the pharmacogenetic variant–drug interactions in five available cancer knowledgebases, and the driver mutations and fusion genes in the Cancer Genome Atlas. **Results:** xGen Exome Research Panel v2 and TrueSight Oncology 500 target 71.0 and 68.9% of the pharmacogenetic interactions in the available knowledgebases; and 93.7 and 86.0% of the driver mutations in the Cancer Genome Atlas, respectively. All other studied panels target lower percentages. **Conclusion:** Exome sequencing outperforms pan-cancer targeted sequencing panels in terms of covered cancer pharmacogenetic variant–drug interactions and pharmacogenetic cancer variants.

First draft submitted: 12 March 2020; Accepted for publication: 10 July 2020; Published online: 6 October 2020

Keywords: exome sequencing • pan-cancer • pharmacogenetic knowledgebase • targeted sequencing • TCGA

In recent years, the number of published tumor-related sequences has increased remarkably, leading to an accelerated understanding of the genomic profile of different cancer tissues [1–3]. The exponentially expanding literature about cancer pharmacogenetics has been aggregated into accessible knowledgebases [4–12]. However, tumor sequencing is required to implement this knowledge in clinical practice and to allow a more personalized therapy for the patient [13,14]. Additionally, sequencing of circulating tumor cells and circulating cell-free tumor DNA (ctDNA) in liquid biopsies is used as a screening test or as a less invasive alternative for solid tumor biopsy [15–18]. Today, several panels for targeted sequencing of tumor tissue and liquid biopsy samples are commercially available. This diversity of available panels makes it difficult to choose the appropriate panel. The question arises if several targeted sequencing panels should be adopted to analyze the different cancer samples, or if it makes more sense to sequence these cancer samples using the exome sequencing pipeline that is used in many medical genetics sequencing centers. This study aims to provide insight into this matter. Comparing these panels based on the number of sequenced bases, genes or targeted regions would result in the conclusion that the exome panel covers the highest number. Therefore, this study examines to what extent these panels determine the known variants in the existing cancer pharmacogenetics knowledgebases and the Cancer Genome Atlas.

This study focuses on recently developed or updated pan-cancer targeted sequencing panels of the leading commercial players in the field. Integrated DNA Technologies (IDT; IA, USA) has a hybridization-based targeted sequencing panel: xGen Pan-Cancer Panel v2.4, which includes 532 genes. Illumina (CA, USA) has a PCR-based targeted sequencing panel: TruSight Oncology 500, this assay targets 523 genes and covers most of the variants in the National Comprehensive Cancer Network (NCCN) guidelines [19]. Thermo Fischer Scientific's (MA, USA) broadest cancer panel is the Ion AmpliSeq Comprehensive Cancer Panel, which is a PCR-based targeted sequencing panel containing 409 genes. These genes are selected from the Sanger Institute Cancer Gene Census [20]. Qiagen (Hilden, Germany) provides a set of targeted DNA panels for cancer research. The broadest is the QIAseq

Future Medicine

targeted Human Comprehensive Cancer Panel, which contains 275 genes covering the most commonly occurring mutations in cancers [21]. Roche (Basel, Switzerland) has focused on the analysis of ctDNA and has developed three specialized AVENIO ctDNA panels for liquid biopsy. The AVENIO ctDNA Targeted Kit contains 17 genes and may be used for pan-cancer research applications, but is specially optimized for lung cancer and colorectal cancer targeted treatments. The AVENIO ctDNA Expanded Kit has an extended number of 77 genes to include a broader spectrum of therapies. The AVENIO ctDNA Surveillance Kit contains 197 genes optimized for longitudinal tumor burden monitoring in lung cancer and colorectal cancer [22]. Foundation Medicine (MA, USA) is the first company that developed a US FDA approved comprehensive genomic profiling test and has two main panels. FoundationOne CDx for all solid tumors which incorporates multiple companion diagnostics, covering 324 genes and FoundationOne Liquid, Foundation Medicine's liquid biopsy pan-cancer panel, which covers 70 genes [23].

In addition to the cancer-specific targeted sequencing panels, whole-exome sequencing (WES) is often used in cancer research [24–27]. Different commercial WES panels are on the market. The differences between these WES panels are described in the literature [28–30] and are out of the scope of this study. IDT's xGen Exome Research Panel v2 was selected as a representative for commonly used WES designs [30]. The xGen Exome Research Panel v2 spans a 34 Mb target region of the human genome and includes 19,433 genes [31].

These targeted sequencing panels are diverse and differ in many aspects. This study examines to what extent these panels genotype-relevant pharmacogenetic variants described in the literature. As mentioned above, this literature has been aggregated in accessible knowledgebases by various research institutes [4–12]. Unfortunately, this resulted in various cancer pharmacogenetics knowledgebases with a low level of overlap of information, each emphasizing other aspects of cancer pharmacogenetic knowledge. Therefore, the Variant Interpretation for Cancer Consortium (VICC) has recently started to aggregate this information in a consistent meta-knowledgebase [32]. This meta-knowledgebase aggregates the cancer-pharmacogenetic data of six different knowledgebases: the Cancer Biomarkers database of the Cancer Genome Interpreter (CGI) [10,33], Clinical Interpretations of Variants in Cancer (CIViC) [9,34], the Jackson Laboratory Clinical Knowledgebase (JAX-CKB) [8,35], Molecular Match [36], the Precision Oncology Knowledgebase (OncoKB) [6,37] and the Precision Medicine Knowledgebases (PMKB) [5,38]. At the moment of writing, this was still an ongoing project, and only a prototype of this meta-knowledgebase was available [39].

The Cancer Biomarkers database of CGI is a collection of genomic biomarkers of anticancer drug response and includes information of 488 genomic biomarkers interacting with 183 drugs in three response categories: sensitivity, resistance and toxicity. This database was last updated on 17 January 2018 [10,33]. CIViC is an expert-crowdsourced knowledgebase. This knowledgebase creates a platform for international experts to collaborate to gather all cancer information systematically. CIViC currently contains 5958 curated interpretations of clinical relevance for 2141 variants affecting 377 genes and 455 drugs [9,34]. JAX-CKB is a database of variant annotations, therapy knowledge, diagnostic information and clinical trials. The data are checked and updated daily by an automated algorithm, and experts then further curate the data. JAX-CKB database consists of two databases: The CKB CORE database, freely accessible via the web, contains 85 commonly known driver genes. The CKB BOOST database is only available when a yearly fee is paid and provides over 1000 genes [8,35]. Molecular Match is a data-as-a-service company and asks a monthly fee to access their database [36]. OncoKB is an expert-guided precision oncology knowledgebase. It contains biological, clinical and therapeutical information curated from the literature and recommendations derived from the FDA labeling and the NCCN guidelines, among others. This knowledgebase was last updated on 28 February 2019 and contained 4472 alterations in 595 cancer-associated genes and 79 drugs [6,37]. PMKB was initially developed for the interpretation of the AmpliSeq 50-gene panel and later expanded to support a wide range of features for signing out WES reports in the Weill-Cornell Medicine's Institute for Precision Medicine. The knowledgebase is open for new contributions, which are reviewed by approved users. PMKB currently contains 2233 variant descriptions with 1750 interpretations. Information about therapies is described in full text and not automatically extractable [5,38]. Another knowledgebase, not aggregated in the VICC, is DEPO, which is a curated database. The primary information comes from the Cancer Biomarkers oncology database within CGI. Additional variant–drug interactions are manually curated by reviewing peer-reviewed literature and conference abstracts, the NCCN Biomarkers Compendium, the Personalized Cancer Therapy website, and others, which are cross-referenced against primary sources. DEPO was last updated on 28 March 2018 [12,40].

In addition to the aggregated knowledge in the diverse knowledgebases, the Cancer Genome Atlas has recently published the Pan-Cancer Atlas, which gives an overview of the molecular landscape of more than 10,000 tumor specimens originating from 33 tumor types. These specimens were subjected to genomic, epigenomic, transcrip-

tomic, proteomic and histological evaluation to gather a broad overview of the possible molecular alterations in these tumor types [1,3]. The findings of the different molecular techniques were extensively published in three main categories: the cell-of-origin patterns, in which the genetic background of different tumor types is defined, and a new classification is described based on the genetic background instead of the cell of origin [41]; oncogenic processes, in which germline genetic variants and somatic mutations are described including their influence in cancer progression [42]; signaling pathways, in which tumor signaling pathways are described, possibly leading to better-personalized treatments [43]. The Pan-Cancer Atlas studies have led to the discovery and further characterization of new and already known genomic alterations [42]. The sequencing data of the more than 10,000 tumor samples were processed by the Multi-Center Mutation Calling in Multiple Cancers (MC3) working group, resulting in the discovery of around 3.5 million somatic mutations [44]. After filtering to reduce the false-positive rate, 9079 tumor samples and approximately 1.5 million mutations remained, where from the cancer driver genes and cancer driver mutations were selected. Two hundred and fifty-eighth driver genes were selected using a systematic approach, which was supplemented with 41 manual curated genes, resulting in a total of 299 driver genes. In addition to the driver genes, driver mutations were selected using three different approaches. Mutations that were determined by two different approaches were considered as somatic driver mutations (3437). The mutations determined by all three approaches were considered as the consensus mutations (579) [45]. Another group of genomic alterations studied in the Pan-Cancer Atlas is gene fusion. A total of 25,664 fusions were identified after extensive filtering using several pool-of-normals databases [46].

Materials & methods

In this study, a custom meta-knowledgebase was built using the tools provided by the VICC. Only the noncommercial and machine-readable knowledgebases were used. Based on these selection criteria, the paid JAX-CKB BOOST and 'Molecular Match' databases, and the nonmachine readable PMKB were excluded from the custom meta-knowledgebase. DEPO, a curated database, not aggregated in the VICC, was included in the custom meta-knowledgebase.

The custom meta-knowledgebase was built as follows. First, the harvesting tools were downloaded from the GitHub of the Computational Biology department of the Oregon Health and Science University [47]. Next, data from OncoKB, CGI and COSMIC were downloaded from their respective websites [48–50]. The data of CIViC were downloaded from their API [51], and the data of JAK-CKB and DEPO were scraped from their respective websites [35,40]. Evidence levels were harmonized according to the evidence levels in Supplementary Table 1. All the data from the knowledgebases were harvested with an adapted version of the harvesting tool described in Wagner *et al.* [32] and outputted in JSON-files. These files were loaded in a local MongoDB database for further use. Only pharmacogenetic variant–drug interactions that contain values for the gene, drug, response type and evidence label fields, and those that had genomic coordinates for the genetic variant were used. Overlapping pharmacogenetic variant–drug interactions between knowledgebases were merged using a custom python script to remove duplicated pharmacogenetic variant–drug interactions in the analysis. The highest evidence level of overlapping variant–drug interactions was used as the evidence level for the merged interaction.

The Pan-Cancer Atlas driver mutations were downloaded from the Supplementary data of Bailey *et al.* [45]. Genomic coordinates of GRCh37 were retrieved via the COSMIC translation table. Genomic coordinates from mutations that were not present in COSMIC were retrieved from Ensembl. One hundred and four mutations were not present in either COSMIC or Ensembl. These mutations were excluded from the analysis. The Pan-Cancer Atlas fusion genes were downloaded from the Supplementary data of Gao *et al.* [46]. Genomic coordinates were transformed from the GRCh38 reference to the GRCh37 reference to compare with the targeted sequencing panels, in which targeted positions are in the GRCh37 reference coordinates.

The five knowledge bases were compared based on the number of pharmacogenetic variant–drug interactions and the number of pharmacogenetic variants in these interactions.

Nine contemporary cancer-specific targeted sequencing panels were compared with the custom meta-knowledgebase built in this study based on their targeted regions. Therefore, their target regions were downloaded from different sources. The targeted regions from the Ion AmpliSeq Comprehensive Cancer Panel (Thermo Fischer Scientific) were downloaded from the supplier's website [52]. The targeted regions from the xGen Pan-Cancer Panel v2.4 (Integrated DNA Technologies), the TrueSight Oncology 500 (Illumina), the QIAseq Targeted Human Comprehensive Cancer Panel (Qiagen), the AVENIO ctDNA Targeted Kit (Roche), the AVENIO ctDNA Expanded Kit (Roche) and the AVENIO ctDNA Surveillance Kit (Roche) were provided at our request by the

suppliers. All genomic coordinates from the targeted regions were present in the GRCh37 reference coordinates, except for the AVENIO panels. These coordinates were from the GRCh38 reference and were transformed to the GRCh37 reference coordinates with CrossMap [53]. Gene lists with coding exons and selected introns of the FoundationOne CDx and FoundationOne Liquid assay were downloaded from the FoundationMedicine websites [54,55]. Genomic coordinates of the targeted exons and introns of the targeted gene were received via the UCSC Table Browser [56]. Along with the cancer-specific targeted sequencing panels, the xGen Exome Research Panel v2 (Integrated DNA Technologies) was also included in the comparison. The exome design was downloaded from the supplier's website [31].

The nine cancer-specific targeted sequencing panels were first compared with each other based on the number of genes they have in common. Second, the nine cancer-specific targeted sequencing panels and the xGen Exome Research Panel v2 were compared with the pharmacogenetic variants in the custom meta-knowledgebase and the driver mutations and fusion genes in the Pan-Cancer Atlas. A panel covers a variant if all the positions of that variant were covered, except for fusion, deletion or duplication of a gene. A fusion of two genes was covered if the panel includes the position of the fusion between the two genes. A deletion or a duplication of a gene was covered if the panel includes a part of that gene. For each panel, the numbers of variants of the Pan-Cancer Atlas covered by that panel were categorized by the different cancer tissues derived from the Pan-Cancer Atlas. Finally, the panels were compared based on the number of pharmacogenetic variant–drug interactions in the custom meta-knowledgebase they cover. A panel covered a pharmacogenetic variant–drug interaction in the meta-knowledgebase if the panel covers the pharmacogenetic variant in that pharmacogenetic variant–drug interaction.

All python scripts used for building the meta-knowledgebase and used for the different comparisons are available on GitHub [57] .

Results

Custom meta-knowledgebase

The custom meta-knowledgebase contains 8326 unique and well-annotated pharmacogenetic variant–drug interactions describing the influence of 3132 pharmacogenetic variants on the response of 860 drugs. These pharmacogenetic variants include SNPs, frameshift mutations, haplotypes and expression variants. The wild-type variants are also present in the meta-knowledgebase. Two hundred and seventy-one pharmacogenetic variant–drug interactions in the meta-knowledgebase had no genomic coordinates and were removed for further analysis. The pharmacogenetic variant–drug interactions in the meta-knowledgebase are categorized into three response types: sensitivity: the pharmacogenetic variant is a target for a specific treatment (5241 pharmacogenetic variant–drug interactions). Resistance: the pharmacogenetic variant reduces therapy efficiency (3061 pharmacogenetic variant–drug interactions). Toxicity: the pharmacogenetic variant causes a toxic effect of the drug (24 pharmacogenetic variant–drug interactions). All the pharmacogenetic variant–drug interactions in the meta-knowledgebase are listed in Supplementary Table 2.

The pharmacogenetic variant–drug interactions in the custom meta-knowledgebase originate from five different sources: CGI (791 pharmacogenetic variant–drug interactions), DEPO (665 pharmacogenetic variant–drug interactions), JAX-CKB (5149 pharmacogenetic variant–drug interactions), OncoKB (145 pharmacogenetic variant–drug interactions) and CIViC (1921 pharmacogenetic variant–drug interactions). Figure 1 shows the overlap of pharmacogenetic variant–drug interactions between the sources. Only 3.5% of the pharmacogenetic variant–drug interactions were present in more than one source. All other pharmacogenetic variant–drug interactions in the meta-knowledgebase originate from only one source. In accordance with the low level of overlap in pharmacogenetic variant–drug interactions, there is also a low level of overlap in the number of pharmacogenetic variants described in these pharmacogenetic variant–drug interactions. About 86.1% of these pharmacogenetic variants are only described in one of the five knowledgebases. About 9.0, 3.2, 1.3 and 0.4% of the pharmacogenetic variants were represented in two, three, four and all the five knowledgebases, respectively.

Cancer-specific targeted sequencing panels

The total number of genes, targeted by each of the nine cancer-specific targeted sequencing panels, were compared. Table 1 shows the total number of genes targeted by each panel. The number of targeted genes in common between two panels is also shown. The xGen Pan-Cancer Panel v2.4 from IDT targets the most genes. The TrueSight Oncology 500 Panel from Illumina targets the second most genes of all the studied panels. This panel has a high overlap with the xGen Pan-Cancer Panel v2.4, around 85% of the genes are targeted in both panels. This high

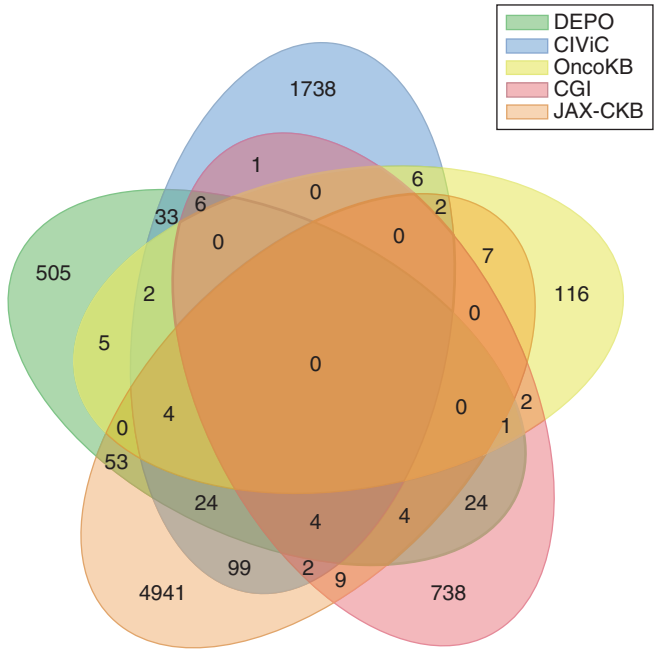

Figure 1. Venn diagram of the pharmacogenetic variant–drug interactions of the five knowledgebases in the meta-knowledgebase.
CGI: Cancer Genome Interpreter; CIViC: Clinical Interpretations of Variants in Cancer; DEPO: Database of Evidence for Precision Oncology; JAX-CKB: Jackson Laboratory Clinical Knowledgebase; OncoKB: Precision Oncology Knowledgebase.

Table 1. Comparison of the number of genes targeted by the cancer-specific sequencing panels.

Panels	Ion AmpliSeq Comprehensive Cancer Panel	AVENIO ctDNA Targeted Kit	AVENIO ctDNA Expanded Kit	AVENIO ctDNA Surveillance Kit	xGen Pan-Cancer Panel v2.4	TrueSight Oncology 500	QIAseq Targeted Human Comprehensive Cancer Panel	FoundationOne Liquid	FoundationOne CDx
FoundationOne CDx	191	74	25	15	318	291	209	70	324
FoundationOne Liquid	62	51	19	15	70	70	66	70	
QIAseq Targeted Human Comprehensive Cancer Panel	189	72	28	15	243	256	275		
TrueSight Oncology 500	247	75	31	15	449	525			
xGen Pan-Cancer Panel v2.4	227	75	29	15	532				
AVENIO ctDNA Surveillance Kit	15	17	17	17					
AVENIO ctDNA Expanded Kit	30	25	198						
AVENIO ctDNA Targeted Kit	69	77							
Ion AmpliSeq Comprehensive Cancer Panel	409								

The total number of genes targeted by each panel can be found on the diagonal and are under-lined.
The number of genes targeted by two panels can be found above the diagonal.

Table 2. Number of cancer variants targeted by each panel.

Variants	Pan-Cancer Atlas driver mutations	Pan-Cancer Atlas fusion	Pharmacogenetic variants in the meta-knowledgebase
Total variants in database	3442	25,664	3132
xGen Exome Research Panel v2	3227	16,149	2339
xGen Pan-Cancer Panel v2.4	2882	324	2290
TrueSight Oncology 500	2961	290	2300
Ion AmpliSeq Comprehensive Cancer Panel	2657	82	2028
QIAseq Targeted Human Comprehensive Cancer Panel	2684	44	2244
AVENIO ctDNA Targeted Kit	710	0	1160
AVENIO ctDNA Expanded Kit	1564	0	1820
AVENIO ctDNA Surveillance Kit	940	0	1243
FoundationOne CDx	2602	91	2218
FoundationOne Liquid	1517	4	1928

overlap suggests that there is a consensus about genes that need to be sequenced in cancer research. However, the Ion AmpliSeq Comprehensive Cancer Panel from Thermo Fischer Scientific, which targets 409 genes, targets only around 50% of the genes targeted in the xGen Pan-Cancer Panel v2.4 and the TrueSight Oncology 500 Panel. The QIAseq Targeted Human Comprehensive Cancer Panel is a smaller panel targeting 275 genes. Around 90% of its genes are also targeted by the xGen Pan-Cancer Panel v2.4 and the TrueSight Oncology 500 Panel, and about 70% of its genes are targeted by the Ion AmpliSeq Comprehensive Cancer Panel. Most of the targeted genes in the FoundationOne CDx Panel, 318 and 291 out of the 324 genes, are also targeted by the xGen Pan-Cancer Panel v2.4 and TrueSight Oncology 500 Panel, respectively. The FoundationOne Liquid panel targets 70 of the genes in the FoundationOne CDx Panel, which are also targeted by both the xGen Pan-Cancer Panel v2.4 and TrueSight Oncology 500 Panel. All panels have a low overlap with the AVENIO Expanded Panel, which targets 198 genes. The other two AVENIO panels target a smaller subset of the genes in the AVENIO Expanded Panel.

Sequencing panels versus genetic variants in meta-knowledgebase & the Pan-Cancer Atlas dataset

The targeted regions of the nine cancer-specific targeted sequencing panels and the xGen Exome Research Panel v2 were compared with the genetic variants in the custom meta-knowledgebase and the Pan-Cancer Atlas dataset. Table 2 & Figure 2 show the numbers of targeted pharmacogenetic variants in the meta-knowledgebase by the sequencing panels. The xGen Exome Research Panel v2 targets the most pharmacogenetic variants (74.7%) in the meta-knowledgebase. The xGen Pan-Cancer Panel v2.4 and the Truesight Oncology 500 Panel target both 73% of the pharmacogenetic variants. The QIAseq Targeted Human Comprehensive Cancer Panel and the FoundationOne CDx Panel target both around 71%. These four panels are also the bigger panels in this study. The Ion AmpliSeq Comprehensive Cancer Panel and the FoundationOne Liquid Panel target 64.8 and 61.6% of the pharmacogenetic variants, respectively. All the AVENIO panels target less than 60% of the pharmacogenetic variants in the meta-knowledgebase.

In Table 2, the numbers of driver mutations targeted by each sequencing panel are shown. The xGen Exome Research Panel v2 targets the highest number (93.7%) of driver mutations. The two most extensive targeted sequencing panels, the xGen Pan-Cancer Panel v2.4 and the TrueSight Oncology 500 Panel, target 83.7 and 86.0% of the driver mutations, respectively. The FoundationOne CDx Panel, the Ion AmpliSeq Comprehensive Cancer Panel, and the QIAseq Targeted Human Comprehensive Cancer Panel target around 75%. The other sequencing panels, optimized for liquid biopsy, target less than 50% of the driver mutations. Only the xGen Exome Research Panel v2 targets a substantial part (62.9%) of the fusion genes discovered in the Pan-Cancer Atlas. All other panels target less than 1% of the fusion genes described in the Pan-Cancer Atlas (Table 2).

The driver mutations and fusion genes of the Pan-Cancer Atlas were discovered in 33 cancer types. An overview of the numbers of variants discovered in each cancer type and targeted by each panel are shown in Table 3. The xGen Exome Research Panel v2, the xGen Pan-Cancer Panel v2.4, the Truesight Oncology 500 Panel, the Ion AmpliSeq Comprehensive Cancer Panel, the QIAseq Targeted Human Comprehensive Cancer Panel and the FoundationOne CDx Panel have a uniform distribution of targeted mutations over the different cancer types. The AVENIO Expanded and the FoundationOne Liquid Panel have a slightly different distribution. The

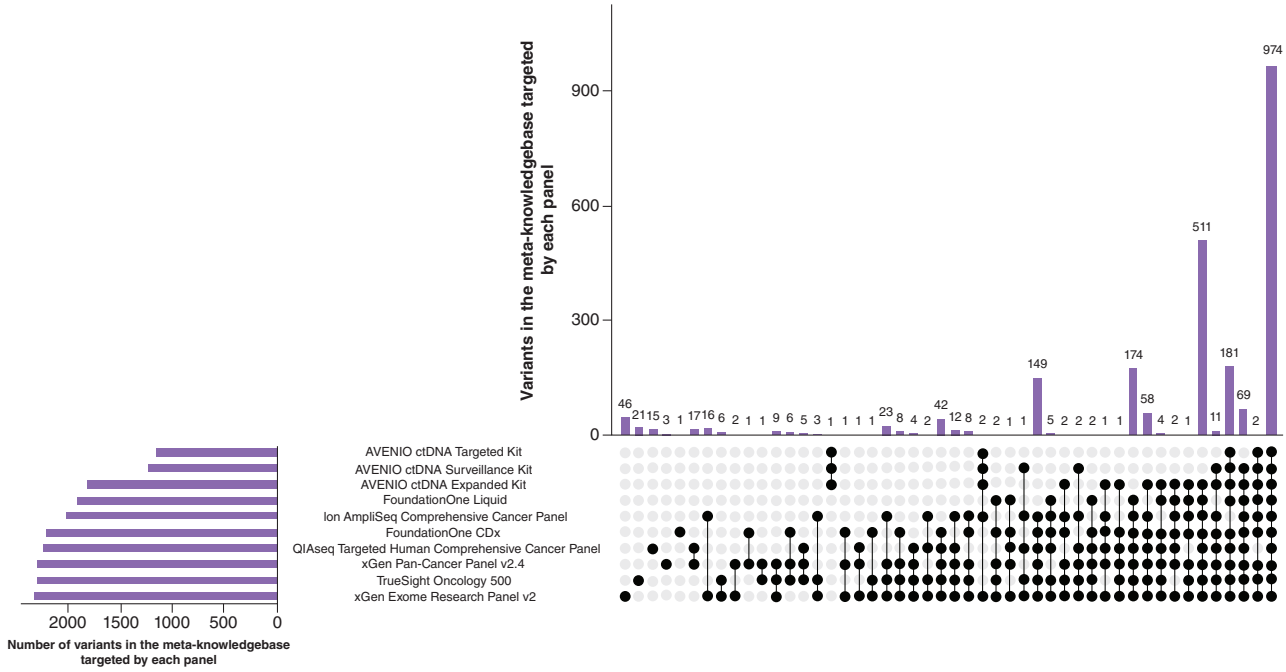

Figure 2. UpSet plot of pharmacogenetic variants in the meta-knowledgebase of the ten targeted sequencing panels.

AVENIO Targeted and Surveillance panels proportionally target about twice as many mutations concerning Chromophobe Kidney Cancer than the other panels. The proportions of targeted mutations concerning Brain Lower Grade Glioma, Ovarian Serous Cystadenocarcinoma and Pancreatic Adenocarcinoma are also increased in the AVENIO Targeted and Surveillance panels. The proportions of targeted mutations concerning Kidney Renal Clear Cell Carcinoma, Pheochromocytoma and Paraganglioma, Uterine Corpus Endometrial Carcinoma, and Uveal Melanoma are decreased in these panels.

Sequencing panels versus pharmacogenetic variant–drug interactions in meta-knowledgebase

In the previous paragraphs, the panels were compared based on the number of targeted pharmacogenetic variants. In this paragraph, the panels are compared based on the covered number of pharmacogenetic variant–drug interactions in the meta-knowledgebase. Table 3 gives an overview of the number of pharmacogenetic variant–drug interactions covered by each panel. The pharmacogenetic variant–drug interactions are categorized by the different response types and their level of evidence. The xGen Exome Research Panel v2 targets not only the most driver mutations and fusion genes described in the Pan-Cancer Atlas and the most pharmacogenetic variants in the meta-knowledgebase, but it also covers the most pharmacogenetic variant–drug interactions (71.0%) in the meta-knowledgebase. The TruSight Oncology 500 Panel is the cancer-specific panel that covers the most pharmacogenetic variant–drug interactions (68.9%) in the meta-knowledgebase. The xGen Pan-Cancer Panel v2.4, the QIAseq Targeted Human Comprehensive Cancer Panel, the AmpliSeq Comprehensive Cancer Panel and the FoundationOne CDx Panel cover around 68% of the pharmacogenetic variant–drug interactions in the meta-knowledgebase. These five panels cover 5530 pharmacogenetic variant–drug interactions in common, which are approximately 94% of the pharmacogenetic variant–drug interactions covered by these panels. This illustrates an extensive overlap in the covered pharmacogenetic variant–drug interactions and shows that there is a consensus as to which pharmacogenetic variant–drug interactions are assayed.

Discussion

In this study, nine contemporary pan-cancer targeted sequencing panels and WES were compared to determine what extent they cover the available knowledge in cancer pharmacogenetics. Therefore, a meta-knowledgebase was built using several available knowledgebases. The used knowledgebases have several shortcomings, and we acknowledge

Table 3. Number of pharmacogenetic variant–drug interactions covered by each panel, categorized by the different response types and evidence levels.

Interactions	Pharmacogenetic interactions in the meta-knowledgebase	xGen Exome Research Panel v2	xGen Pan-Cancer Panel v2.4	TrueSight Oncology 500	Ion AmpliSeq Comprehensive Cancer Panel	QIAseq Targeted Human Comprehensive Cancer Panel	AVENIO ctDNA Targeted Kit	AVENIO ctDNA Expanded Kit	AVENIO ctDNA Surveillance Kit	FoundationOne CDx	FoundationOne Liquid
Total pharmacogenetic interactions in the meta-knowledgebase	8326	5912	5730	5736	5533	5717	3593	5057	3735	5666	5320
Sensitive A	283	214	219	216	189	207	157	192	157	219	193
Sensitive B	516	247	231	230	224	222	159	216	176	227	220
Sensitive C	1131	786	778	777	718	765	517	696	538	771	714
Sensitive D	3311	2393	2313	2319	2234	2326	1387	1965	1449	2282	2090
Resistance A	66	49	51	49	49	51	49	49	49	49	49
Resistance B	308	165	167	167	164	161	113	154	121	162	157
Resistance C	621	503	495	497	487	492	291	455	302	489	474
Resistance D	2066	1542	1475	1480	1457	1492	915	1324	938	1466	1422
Toxic A	19	9	0	0	9	0	5	5	5	0	0
Toxic B	5	4	1	1	2	1	0	1	0	1	1
Toxic C	0	0	0	0	0	0	0	0	0	0	0
Toxic D	0	0	0	0	0	0	0	0	0	0	0

that the meta-knowledgebase is not curated. Curating the meta-knowledgebase would be a tremendous effort and could be the effort of an international consortium of experts. This curation falls out of the scope of this study. There does not seem to be a freely available knowledgebase covering most cancer pharmacogenetics knowledge, which is the reason for initiatives such as the VICC initiative to build a meta-knowledgebase. The meta-knowledgebase build in this study is based on the VICC initiative in a best effort to gather as much information as possible.

The five knowledgebases used in this study have a low level of overlap between the pharmacogenetic variant–drug interactions and pharmacogenetic variants they describe. Therefore, it seems necessary to consult the different available knowledgebases when consulting pharmacogenetic information about cancer treatment, for example, by building a meta-knowledgebase. Each knowledgebase has different resources and strategies to collect pharmacogenetic information. CGI and DEPO are no longer updated and contain older data. However, they need to be included in the meta-knowledgebase because there is nearly no overlap with other knowledgebases. The remaining knowledgebases are still being updated. CIViC uses the cancer research community to aggregate as much data as possible about cancer mutations and their effect on therapy. This knowledgebase is continually updated and reviewed by the cancer research community and by qualified experts. JAX-CKB and OncoKB are both knowledgebases maintained and curated by experts. A difference between these two knowledgebases is the way they update their content and their resources. JAX-CKB updates its knowledgebase automatically, while OncoKB works with releases. Another difference is that OncoKB mainly focuses on guidelines from the FDA and the NCCN, while JAX-CKB focuses on variant–drug interactions in published articles.

Overall the xGen Exome Research Panel v2 outperforms the cancer-specific panels in terms of the number of sequenced genes, driver mutations and fusion genes in the Pan-Cancer Atlas dataset, pharmacogenetic variants and pharmacogenetic variant–drug interactions in the meta-knowledgebase. It is evident that an exome panel, which

targets a larger part of the genome, covers more variants. However, not all cancer mutations are located in the exonic regions of the human genome [58]. Thus, an exome panel might not be the most appropriate choice. Nevertheless, most exome panels are also covering many nonexonic regions such as promoter regions and splice-site mutations. Considering this, the question arises if one or more targeted sequencing panels should be adopted to analyze the different cancer samples, or if it makes more sense to sequence cancer samples using the exome sequencing pipeline that is in place at many medical genetics sequencing centers. This study provides insight into this matter. Comparing the number of sequenced variants would obviously result in the conclusion that the exome panel covers the highest number of variants. Therefore, we first constructed a meta-knowledgebase, gathering the most freely available knowledge on cancer-related pharmacogenetic variants and pharmacogenetic variant–drug interactions, and used this as a basis to see how much of this knowledge is covered by the different panels.

The nine cancer-specific panels target different regions, but the TruSight Oncology 500 Panel, the xGen Pan-Cancer Panel v2.4, the QIAseq Targeted Human Comprehensive Cancer Panel, the AmpliSeq Comprehensive Cancer Panel and the FoundationOne CDx Panel show some consensus, especially regarding coverage of known pharmacogenetic variant–drug interactions. The xGen Exome Research Panel targets only a slightly higher number of pharmacogenetic variant–drug interactions, making these cancer-specific panels more appropriate under certain circumstances. The other cancer-specific panels, optimized for liquid biopsies, target only a subset of the studied pharmacogenetic variant–drug interactions. Some of these pharmacogenetic variant–drug interactions are only targeted by these panels, making them necessary in specific situations.

In addition to the technical aspects of a targeted sequencing panel, the total cost is also an important parameter. In the last years, the cost of whole-genome sequencing (WGS) has been reduced [59]. This reduction, in turn, reduces the cost-efficiency of a targeted sequencing approach compared with WGS. However, not every research institution or clinical laboratory has the throughput to achieve a low sequencing cost. The exact cost of a method is highly dependent on the specific setting and laboratory setup. It is dependent on the throughput and sequencing capacity of the laboratory since suppliers offer substantial order volume-based discounts. In addition to the reagents' cost, there are substantial other costs such as the cost for the instruments, the laboratory space, lab technicians' hands-on time and data analysis. Laboratories that have a high sequencing capacity and that already have routine exome sequencing in place, might want to avoid the cost of additionally implementing, validating and maintaining several separate cancer sequencing methods, even when these panels have a lower reagent list price.

Around 33% of the pharmacogenomic variants in the meta-knowledgebase are covered by neither of the studied panels. Half of these variants are SNPs located outside the exon regions of the genome. These SNP positions could easily be added to new versions of the panels. Another 20% of the variants that are not covered are gene expression level variants, which cannot be detected by targeted DNA sequencing. RNA sequencing or proteome analysis could detect these expression variants. The remaining variants are haplotypes.

Conclusion

In this study, nine commercially available pan-cancer targeted sequencing panels and WES were compared with determine to what extent they cover the available cancer pharmacogenetic knowledge. Overall, WES outperforms the cancer-specific targeted sequencing panels in terms of covered genes, driver-mutations and fusion genes in the Pan-Cancer Atlas, and pharmacogenetic variants and pharmacogenetic variant–drug interactions in the studied pharmacogenetic cancer knowledgebases. This indicates that WES, WGS not taken into consideration, is the most comprehensive pan-cancer genomic diagnostic sequencing approach. The TruSight Oncology 500 Panel, the xGen Pan-Cancer Panel v2.4, the QIAseq Targeted Human Comprehensive Cancer Panel, the AmpliSeq Comprehensive Cancer Panel and the FoundationOne CDx Panel target only a slightly lower number of pharmacogenetic variant–drug interactions than the xGen Exome Research Panel and show some consensus about the pharmacogenetic variant–drug interactions that need to be targeted. Moreover, it is not necessarily the larger sequencing panels that target the most cancer variants or pharmacogenetic variant–drug interactions.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/pgs-2020-0035

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Summary points

Knowledgebases

- We have built a meta-knowledgebase from five freely accessible knowledgebases, as a reference to check how many of the currently known cancer pharmacogenetic variant–drug interactions can be determined using the contemporary pan-cancer targeted sequencing panels or whole-exome sequencing (WES). This meta-knowledgebase contains 8326 pharmacogenetic variant–drug interactions between 3132 pharmacogenetic variants and 860 drugs.
- The contemporary pan-cancer targeted sequencing panels and WES were also compared with the 3437 driver mutations and 25,664 fusion genes in the Pan-Cancer Atlas.

Pan-cancer targeted sequencing panels

- In this study, nine contemporary pan-cancer targeted sequencing panels and WES were compared with the knowledgebases mentioned above. These panels are the Ion AmpliSeq Comprehensive Cancer Panel, the xGen Pan-Cancer Panel v2.4, the TrueSight Oncology 500 Panel, the QIAseq Targeted Human Comprehensive Cancer Panel, the AVENIO ctDNA Targeted Kit, the AVENIO ctDNA Expanded Kit, the AVENIO ctDNA Surveillance Kit, the FoundationOne CDx Panel and FoundationOne Liquid Panel, and the xGen Exome Research Panel v2 representative for WES.

Results

- The xGen Exome Research Panel v2 outperforms cancer-specific targeted sequencing panels in terms of the covered pharmacogenetic variant–drug interactions and cancer variants.
- From the cancer-specific targeted sequencing panels, the xGen Pan-Cancer Panel v2.4, the TrueSight Oncology 500 Panel, the Ion AmpliSeq Comprehensive Cancer Panel, the QIAseq Targeted Human Comprehensive Cancer Panel and the FoundationOne CDx Panel cover about the same number of pharmacogenetic variants and determine a consensus set of pharmacogenetic variant–drug interactions. The other panels, optimized for liquid biopsies, target only a subset of pharmacogenetic variant–drug interactions.
- The cancer-specific targeted sequencing panels determine virtually none of the known cancer gene fusions.

References

1. Weinstein JN, Collisson EA, Mills GB *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* 45(10), 1113–1120 (2013).
2. Hudson TJ, Anderson W, Artez A *et al.* International network of cancer genome projects. *Nature* 464(7291), 993–998 (2010).
3. Hutter C, Zenklusen JC. The Cancer Genome Atlas: creating lasting value beyond its data. *Cell* 173(2), 283–285 (2018).
4. Yeh P, Chen H, Andrews J, Naser R, Pao W, Horn L. DNA-Mutation Inventory to Refine and Enhance Cancer Treatment (DIRECT): a catalog of clinically relevant cancer mutations to enable genome-directed anticancer therapy. *Clin. Cancer Res.* 19(7), 1894–1901 (2013).
5. Huang L, Fernandes H, Zia H *et al.* The cancer precision medicine knowledge base for structured clinical-grade mutations and interpretations. *J. Am. Med. Informatics Assoc.* 24(3), 148 (2016).
6. Chakravarty D, Gao J, Phillips SM *et al.* OncoKB: a precision oncology knowledge base. *JCO Precis. Oncol.* 1(1), 1–16 (2017).
7. Dienstmann R, Jang IS, Bot B, Friend S, Guinney J. Database of genomic biomarkers for cancer drugs and clinical targetability in solid tumors. *Cancer Discov.* 5(2), 118–123 (2015).
8. Patterson SE, Liu R, Statz CM, Durkin D, Lakshminarayana A, Mockus SM. The clinical trial landscape in oncology and connectivity of somatic mutational profiles to targeted therapies. *Hum. Genomics* 10(1), 4 (2016).
9. Griffith M, Spies NC, Krysiak K *et al.* CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat. Genet.* 49(2), 170–174 (2017).
10. Tamborero D, Rubio-Perez C, Deu-Pons J *et al.* Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* 10(1), 25 (2018).

11. Damodaran S, Miya J, Kautto E *et al.* Cancer driver log (CanDL): catalog of potentially actionable cancer mutations. *J. Mol. Diagnostics* 17(5), 554–559 (2015).
12. Sun SQ, Mashl RJ, Sengupta S *et al.* Database of evidence for precision oncology portal. *Bioinformatics* 34(24), 4315–4317 (2018).
13. Hyman DM, Taylor BS, Baselga J. Implementing genome-driven oncology. *Cell* 168(4), 584–599 (2017).
14. Beltran H, Eng K, Mosquera JM *et al.* Whole-exome sequencing of metastatic cancer and biomarkers of treatment response. *JAMA Oncol.* 1(4), 466 (2015).
15. Lai J, Du B, Wang Y, Wu R, Yu Z. Next-generation sequencing of circulating tumor DNA for detection of gene mutations in lung cancer: implications for precision treatment. *Onco. Targets Ther.* 11, 9111–9116 (2018).
16. Birkenkamp-Demtröder K, Christensen E, Nordentoft I *et al.* Monitoring treatment response and metastatic relapse in advanced bladder cancer by liquid biopsy analysis. *Eur. Urol.* 73(4), 535–540 (2018).
17. Zhu Z, Qiu S, Shao K, Hou Y. Progress and challenges of sequencing and analyzing circulating tumor cells. *Cell Biol. Toxicol.* 34(5), 405–415 (2018).
18. Manier S, Park J, Capelletti M *et al.* Whole-exome sequencing of cell-free DNA and circulating tumor cells in multiple myeloma. *Nat. Commun.* 9(1), 1691 (2018).
19. TruSight Oncology 500. (2019). https://www.illumina.com/tso500
20. Ion AmpliSeq Comprehensive Cancer Panel – Thermo Fisher Scientific. (2019). https://www.thermofisher.com/order/catalog/product/4477685
21. QIAseq Targeted Human Comprehensive Cancer Panel. (2019). https://www.qiagen.com/us/shop/sample-technologies/dna/genomic-dna/qiaseq-targeted-dna-panels/?catno=DHS-3501Z#orderinginformation
22. AVENIO ctDNA Analysis Kits. (2019). https://sequencing.roche.com/en/products-solutions/by-category/assays/ctdna-analysis-kits.html
23. Foundation Medicine Panels. (2019). https://www.rochefoundationmedicine.com/home/services.html
24. McKeown C, Connors S, Stapleton R *et al.* A pilot study of exome sequencing in a diverse New Zealand cohort with undiagnosed disorders and cancer. *J. R. Soc. New Zeal.* 48(4), 262–279 (2018).
25. Erinjeri NJ, Nicolson NG, Deyholos C, Korah R, Carling T. Whole-exome sequencing identifies two discrete druggable signaling pathways in follicular thyroid cancer. *J. Am. Coll. Surg.* 226(6), 950–959 (2018).
26. Riahi A, Radmanesh H, Schürmann P *et al.* Exome sequencing and case-control analyses identify *RCC1* as a candidate breast cancer susceptibility gene. *Int. J. Cancer* 142(12), 2512–2517 (2018).
27. Das R, Kundu S, Laskar S, Choudhury Y, Ghosh SK. Assessment of DNA repair susceptibility genes identified by whole exome sequencing in head and neck cancer. *DNA Repair (Amst).* 66–67 , 50–63 (2018).
28. Chilamakuri CS, Lorenz S, Madoui M-A *et al.* Performance comparison of four exome capture systems for deep sequencing. *BMC Genomics* 15(1), 449 (2014).
29. Zhang G, Wang J, Yang J *et al.* Comparison and evaluation of two exome capture kits and sequencing platforms for variant calling.*BMC Genomics* 16, 581 (2015).
30. Wang VG, Kim H, Chuang JH. Whole-exome sequencing capture kit biases yield false negative mutation calls in TCGA cohorts. *PLoS ONE* 13(10), e0204912 (2018).
31. xGen Exome Research Panel v2 targets. (2020). https://sfvideo.blob.core.windows.net/sitefinity/docs/default-source/supplementary-product-info/xgen-exome-research-panel-v2-targets-hg1902a5791532796e2eaa53ff00001c1b3c.bed?sfvrsn=6dd1707_10&download=true
32. Wagner AH, Walsh B, Mayfield G *et al.* A harmonized meta-knowledgebase of clinical interpretations of cancer genomic variants. *Nature Genetics* 52, 448–457 (2020).
33. Cancer Genome Interpreter. (2019). https://www.cancergenomeinterpreter.org
34. Clinical Interpretations of Variants in Cancer. (2019). https://civicdb.org
35. The Jackson Laboratory Clinical Knowledgebase. (2019). https://ckb.jax.org
36. Molecular Match. (2019). https://www.molecularmatch.com/
37. Precision Oncology Knowledge Base. (2019). https://oncokb.org/
38. The Precision Medicine Knowledgebase. (2019). https://pmkb.weill.cornell.edu/
39. The Variant Interpretation for Cancer Consortium Meta-Knowledgebase. (2019). search.cancervariants.org
40. Database of Evidence for Precision Oncology. (2019). http://depo-dinglab.ddns.net/
41. Hoadley KA, Yau C, Hinoue T *et al.* Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* 173(2), 291–304 (2018).

42. Ding L, Bailey MH, Porta-Pardo E *et al.* Perspective on oncogenic processes at the end of the beginning of cancer genomics. *Cell* 173(2), 305–320 (2018).
43. Sanchez-Vega F, Mina M, Armenia J *et al.* Oncogenic signaling pathways in the Cancer Genome Atlas. *Cell* 173(2), 321–337 (2018).
44. Ellrott K, Bailey MH, Saksena G *et al.* Scalable open science approach for mutation calling of tumor exomes using multiple genomic pipelines. *Cell Syst.* 6(3), 271–281 (2018).
45. Bailey MH, Tokheim C, Porta-Pardo E *et al.* Comprehensive characterization of cancer driver genes and mutations. *Cell* 173(2), 371–385 (2018).
46. Gao Q, Liang WW, Foltz SM *et al.* Driver fusions and their implications in the development and treatment of human cancers. *Cell Rep.* 23(1), 227–238 (2018).
47. g2p-aggregator. (2019). https://github.com/ohsu-comp-bio/g2p-aggregator
48. The Precision Oncology Knowledge Base API. (2019). https://oncokb.org/api/v1/
49. Cancer Genome Interpreter Cancer Biomarkers database. (2019). https://www.cancergenomeinterpreter.org/data/cgi_biomarkers_latest.zip
50. Catalogue Of Somatic Mutations In Cancer Mutation Data. (2019). https://cancer.sanger.ac.uk/cosmic/file_download/GRCh37/cosmic/v88/CosmicMutantExport.tsv.gz
51. Clinical Interpretations of Variants in Cancer API. (2019). https://civicdb.org/api
52. Ion AmpliSeq Comprehensive Cancer Panel Data Files. (2019). https://ampliseq.com/protected/tmpl/download.action?tmplDesignId=50986615
53. Zhao H, Sun Z, Wang J, Huang H, Kocher J-P, Wang L. CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* 30(7), 1006–1007 (2014).
54. FoundationOne CDx. (2019). https://www.foundationmedicine.com/genomic-testing/foundation-one-cdx
55. FoundationOne Liquid. (2019). https://www.foundationmedicine.com/genomic-testing/foundation-one-liquid
56. Karolchik D, Hinrichs AS, Furey TS *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* 32(90001), 493D–496 (2004).
57. PanCancerTools. (2020). https://github.com/laurentijntilleman/PanCancerTools
58. Hornshøj H, Nielsen MM, Sinnott-Armstrong NA *et al.* Pan-cancer screen for mutations in non-coding elements with conservation and cancer specificity reveals correlations with expression and survival. *NPJ Genom. Med.* 3(1), (2018).
59. Wetterstrand KA. The cost of sequencing a human genome. *National Human Genome Research Institute* (2020). https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost