Biomedicine & Pharmacotherapy 159 (2023) 114210

Contents lists available at ScienceDirect

Biomedicine & Pharmacotherapy

journal homepage: www.elsevier.com/locate/biopha

Identifying somatic changes in drug transporters using whole genome and transcriptome sequencing data of advanced tumors

Wesley S. van de Geer [a,b,c], Ron H.J. Mathijssen [a], Job van Riet [a,b,c], Neeltje Steeghs [d], Mariette Labots [e], Carla van Herpen [f], Lot A. Devriese [g], Vivianne C.G. Tjan-Heijnen [h], Emile E. Voest [d,i,j], Stefan Sleijfer [a,i], John W.M. Martens [a], Edwin Cuppen [j,k], Harmen J.G. van de Werken [b,c,l,*,1], Sander Bins [a,**,1]

[a] Dept. of Medical Oncology, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, the Netherlands
[b] Cancer Computational Biology Center, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, the Netherlands
[c] Dept. of Urology, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, the Netherlands
[d] Dept. of Medical Oncology, Netherlands Cancer Institute, Antoni van Leeuwenhoek, Amsterdam, the Netherlands
[e] Amsterdam UMC location Vrije Universiteit Amsterdam, Department of Medical Oncology, Amsterdam, the Netherlands
[f] Dept. of Medical Oncology, Radboud University Medical Center, Nijmegen, the Netherlands
[g] Dept. of Medical Oncology, Cancer Center, University Medical Center Utrecht, Utrecht, the Netherlands
[h] Dept. of Medical Oncology, GROW, Maastricht University Medical Center, Maastricht, the Netherlands
[i] Center for Personalized Cancer Treatment, Rotterdam, the Netherlands
[j] Center for Molecular Medicine and Oncode Institute, University Medical Center Utrecht, Utrecht, the Netherlands
[k] Hartwig Medical Foundation, Amsterdam, the Netherlands
[l] Dept. of Immunology, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, the Netherlands

ARTICLE INFO

Keywords:
Pharmacogenomics
Pharmacogenetics
Drug transporters
Metastatic cancer
Whole-genome sequencing
RNA sequencing
Precision medicine

ABSTRACT

Drug resistance is a perpetual problem in cancer therapy with many underlying mechanisms. Alterations in drug transport over the cancer cell membrane can severely alter intratumoral drug exposure, contributing to resistance. Here, we present the somatic mutational landscape of 48 ATP-binding cassette and 416 solute carrier transporter genes in a cohort (CPCT-02; NCT01855477) of 3290 patients with different types of advanced and metastasized cancer through analysis of whole genome and transcriptome sequencing. In order to identify potential stressor mechanisms, we stratified patients based on previous systemic therapies and subsequently investigated the enrichment of mutations and copy-number alterations of transporter genes. In tumors from patients pretreated with protein kinase inhibitors (PKIs), genes encoding for specific copper (*SLC31A1* and *SLC31A2*, χ^2-test adjusted *p*-values: 6.9e-09 and 2.5e-09) and nucleoside transporters (*SLC28A2* and *SLC28A3*, χ^2-test adjusted *p*-values: 3.5e-06 and 6.8e-07) were deleted significantly more frequently than in patients pretreated with chemotherapy. Moreover, we detected 16 transporters that were differentially expressed at RNA level between these treatment groups. These findings contradict mechanisms of selective pressure, as they would be expected to originate during treatment with chemotherapy rather than with PKIs. Hence, they might constitute primary drug resistance mechanisms and, therefore, warrant further study.

1. Introduction

Genomic instability and mutations are fundamental characteristics of cancer cells [1], e.g. by gain and change of function of proto-oncogenes [2]. Signaling from tumor-driving factors can be inhibited by an increasing number of small molecules [3]. Upon

* Correspondence to: Cancer Computational Biology Center & Department of Urology & Department of Immunology, Erasmus MC Cancer Institute, University Medical Center, PO Box 2040,, NA-1218, 3000 CA Rotterdam, the Netherlands.
** Correspondence to: Department of Medical Oncology, Erasmus MC Cancer Institute, University Medical Center, PO Box 2040, Nt-558, 3000 CA Rotterdam, the Netherlands.
E-mail addresses: h.vandewerken@erasmusmc.nl (H.J.G. van de Werken), s.bins@erasmusmc.nl (S. Bins).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.biopha.2022.114210
Received 14 October 2022; Received in revised form 23 December 2022; Accepted 30 December 2022
Available online 6 January 2023

treatment with targeted therapies, however, drug resistance eventually occurs due to outgrowth of resistant subclones, e.g. *EGFR* T790M in non-small cell lung cancer treated with an EGFR inhibitor such as gefitinib [4–6]. Factors impacting intracellular drug concentrations might also contribute to resistance. For most small molecules and conventional chemotherapies, influx and efflux over the cell membrane is largely facilitated by membrane transporters [7], the majority from the ATP-binding cassette (ABC) or solute carrier (SLC) families [8]. Germline variation in the genes encoding drug transporters, the study of which being referred to as pharmacogenetics, is one of the factors known to influence transporter activity and thus intracellular drug concentrations [9]. Moreover, these 'pharmacogenes' become genetically and epigenetically affected during cancer progression by somatic genetic and epigenetic events that could, additionally, alter intratumoral drug exposure [10]. We hypothesized that treatment with mutagenic therapies such as chemotherapy or with small molecules that are substrates to a specific drug transporter might lead to disease progression through the outgrowth of tumor subclones that harbor somatic pharmacogenetic aberrations conferring resistance to treatments given [11]. Here, we explored the pan-cancer landscape of somatic pharmacogenetic aberrations based on whole genome sequencing (WGS) of paired tumor biopsies and germline DNA from 3290 metastatic cancer patients (Fig. S1, Table S1) with solid tumors who had received one or multiple treatment regimens ($N = 1947$, types specified in data file S1) or were treatment naive ($N = 1343$). In the 464 known transporter genes (48 ABC-transporters and 416 SLC-transporters), we analyzed small mutations, which are single-nucleotide variants (SNVs), multi-nucleotide variants (MNVs) and insertions/deletions (InDels), as well as large somatic events, such as copy-number alterations (CNAs) and structural variants (SVs). With access to 2059 matched RNA-Seq samples, we extended our WGS analyses with examinations of differential expression of transporters.

2. Materials and methods

2.1. Patient cohort and study procedures

Patients with metastatic cancer were recruited as part of the CPCT-02 (NCT01855477) clinical study, part of the Center for Personalized Cancer Treatment [12]. Briefly, main inclusion criteria were: (1) age $\geq$ 18 years; (2) locally advanced or metastatic solid tumor; (3) indication for new line of systemic treatment with registered anti-cancer agents; (4) safely biopsiable tumor lesion according to the intervening physician. This study and its protocol were approved by the medical ethical committee (METC) of the University Medical Center Utrecht, as well as the Netherlands Cancer Institute. In accordance with the Declaration of Helsinki, written informed consent was acquired before study procedures. Core needle biopsies from metastatic lesion and peripheral whole blood samples, along with clinical data were collected in the same manner across 41 Dutch hospitals. Obtained tumor biopsies from patients included in this paper were whole-genome sequenced between the 3rd of May 2016 and the 15th of April 2019.

2.2. Sample collection and sequencing

Sampling, sequencing and initial computational steps have been previously described in-depth by Priestley et al. [13]. Briefly, fresh frozen tumor biopsies and whole blood samples were sent to the central sequencing facility at the Hartwig Medical Foundation for whole-genome sequencing on a HiSeqX system generating 2 × 150 base read pairs using standard settings (Illumina, San Diego, CA, USA). Read mapping to reference GRCh37 and subsequent (somatic, structural and copy-number) variant calling, along with several quality control or correction steps were performed systematically. Samples with an estimated tumor cell percentage of > 20%, both based on pathology and sequencing data, were deemed eligible for further analysis. Only the first biopsy of each patient was selected for analysis.

2.3. Variant annotation and filtering

Somatic VCF files were annotated with gene symbols, HGVS notations and gnomAD [14] frequencies, using VEP [15] (database release 95, merged cache) for GRCh37, with setting "–per_gene". Exclusively (somatic SNV, InDel, and MNV) variants with sufficient coverage (>= 3 alternative read observations) and quality (variant caller quality control), were included in the analyses. Furthermore, population variants were removed to prevent germline leakage, based on the gnomAD database (v2.0.2): gnomAD exome (ALL) allele frequency $\geq$ 0.001; and (3) gnomAD genome (ALL) $\geq$ 0.005. Population variants specific to the CPCT cohort were removed based on a panel-of-normals from 1 762 whole blood samples representing the Dutch population. The most deleterious mutation was used to annotate the overlapping gene for each sample. Variants were selected for analysis based on the hg19 (GRCh37) genomic coordinates retrieved from Ensembl (version 87 [16]) of transporter genes from the ABC and SLC gene families (gene lists from HUGO Gene Nomenclature Committee [17]), excluding pseudogenes. We classified drug transporters from which extensive functional or clinical literature is available as 'established' (data file S1), based on the position paper on membrane transporters in drug development by the International Transporter Consortium [7].

2.4. Germline clinical significance

Germline variants were imported from the VCF files, which were annotated by Hartwig Medical Foundation. Variants with at least two reads and passing variant caller quality control were included. Annotated variant data (GRCh37) was downloaded from ClinVar [18] and filtered for transporter genes and a review status of "practice guideline", "expert panel", "multiple submitters" or "multiple submitters, no conflicts". Clinical significance of transporter germline variants was determined by overlapping with these 7108 variants. Clinical significance was grouped into Pathogenic/Likely pathogenic, Drug response, Benign/Likely benign and Variant of Unknown Significance (VUS).

2.5. Sample grouping for enrichment analyses

Each treatment in the clinical data was assessed separately for enrichment, meaning that patients were grouped based on each drug or combination of drugs (depending on how this was annotated). Monoclonal antibodies were excluded from this analysis as their pharmacokinetics do not depend on the investigated drug transporters. Patients who received more than one treatment were present in multiple groups. For several visualizations, only groups larger than 30 patients are shown, as indicated.

2.6. Ploidy and copy-number analysis

Computational ploidy estimation and copy-number (CN) analysis was performed using the custom pipeline PURPLE (PURity & PLoidy Estimator) as described by Priestley et al. [13]. Briefly, this approach combines B-allele frequency (BAF), read depth, and structural variants to estimate the purity and CN profile of a tumor sample. Broad and focal somatic CN alterations were identified by GISTIC2.0 [19] (v2.0.23), using the following parameters: genegistic 1, gcm extreme, maxseg 4000, broad 1, brlen 0.98, conf 0.95, rx 0, cap 3, saveseg 0, armpeel 1, smallmem 0, res 0.01, ta 0.1, td 0.1, savedata 0, savegene 1, qvt 0.1. Distinction between shallow and deep CN events per region was based on thresholding performed by GISTIC2.0. The thresholding can be divided into five categories, two for deletions (−2 = deep, possibly homozygous loss, −1 = shallow, possibly heterozygous loss), one for diploid (0 = diploid) and two for amplifications (1 = few additional copies, often broad gain, 2 = more copies, often focal gain). Annotation

of GISTIC2.0 peaks, split into several groups, was performed as follows: A) Wide peaks were annotated with all overlapping canonical UCSC genes within its limits. B) Focal peaks were annotated based on overlapping genomic coordinates, using custom R scripts and UCSC gene annotations. Focal peaks were only considered for the enrichment analysis if a single gene overlapped the peak. The GISTIC2.0 analysis was performed on each treatment group individually, with separate internal multiple-testing corrections (Benjamini-Hochberg false discovery rate method).

2.7. Structural variants analysis

Calling of structural variants (SVs) was performed using the Genome Rearrangement IDentification Software Suite [20] (GRIDSS) at Hartwig Medical Foundation. SVs were imported using custom R scripts, overlapping transporter genes on at least one breakpoint, using hg19 (GRCh37) genomic coordinates retrieved from Ensembl (version 87 [16]). Structural variants with an upstream or downstream Lesser Allele Frequency (LAF) below 0.1 were filtered. Elements that could not be annotated with an Ensembl gene ID were removed.

2.8. Hotspot detection

Coding variants were extracted from the dataset and counted based on position (i.e., not exact substitution). All positions that counted > 2 somatic mutations (any substitution) were fit with a generalized linear model using a Poisson distribution. Any position with an adjusted p-value < 0.05 (Benjamini-Hochberg false discovery rate method) based on the generalized linear model was viewed as a hotspot mutation. Tissue of origin was extracted from the metadata of our cohort and compared with those described in COSMIC for the hotspots to determine overlap [21].

2.9. Detection of genes under selective pressure

To identify genes undergoing mutational selection, we utilized the dN/dS [22] model (192 Poisson rate parameters; under the full trinucleotide model) with the R package dndscv (v0.0.0.9). Briefly, this model tests the ratio between nonsynonymous (missense, nonsense and essential splice site) and background (synonymous) mutations, with an additional InDel model, while correcting for sequence composition, gene length and mutational signatures. To identify genes that drive selection, a p-value < 0.01 (both including and excluding the InDel model) was used. The dN/dS analysis was performed on each treatment group individually. To determine a q-value, the p-value underwent multiple testing correction by Benjamini–Hochberg procedure [23] internally, for each group.

2.10. Coding region length correction

To calculate the absolute number of coding mutations (per megabase) overlapping each transporter gene, we corrected for the respective coding lengths (in bases, retrieved from Ensembl) and the number of samples ($N = 3290$). Therefore, we multiplied by a megabase (Mb, 1E6). This equation can be represented as:

$$mutations\ per\ Mb\ for\ each\ transporter = \left(\frac{1E6 * nMuts}{(length * 3290)}\right)$$

2.11. Correlation of the number of somatic variants in selected genes versus coding length

For each gene, the sum of gene-respective somatic nonsynonymous SNVs, InDels and MNVs were correlated to its respective coding region length (in bp) using Pearson's correlation coefficient.

2.12. Statistical analysis with Kruskal-Wallis and χ^2 tests

Frequencies of nonsynonymous SNVs and copy-number alterations (gain/loss of at least one allele) of the ABC and SLC transporter genes were tested for prior treatment independence using the Kruskal-Wallis test. Subsequently, χ^2-tests were performed as "group" versus untreated, with the groups being: chemotherapy, hormonal therapy, targeted therapy, and protein kinase inhibitors (PKIs, being a subset of targeted therapy and removed from the targeted therapy group). We assumed statistical significance with an adjusted p-value < 0.05, rejecting the null hypothesis and concluding that the two groups were independent. Since chemotherapeutics and PKIs are best characterized as substrates of the investigated membrane transporters, we decided to perform the χ^2-test for these two groups and utilize this contrast for differential expression analyses as well.

2.13. RNA Sequencing

RNA was isolated from biopsy using the QIAsymphony RNA Kit (Qiagen, Hilden, Germany) for tissue and quantified on the Qubit. Between 50 and 100 nanograms of RNA was used as input for the KAPA RNA HyperPrep Kit with RiboErase (Human/Mouse/Rat) library preparation (Roche) on an automated liquid handling platform (Beckman Coulter). RNA was fragmented (high temperature in the presence of magnesium) to a target length of 300 bp. Barcoded libraries were sequenced as pools on either a NextSeq 500 (V2.5 reagents) generating 2×75 base read pairs or on a NovaSeq 6000 generating 2×150 base read pairs using standard settings (Illumina, San Diego, CA, USA). BCL output from the sequencing platform was converted to FASTQ using Illumina's bcl2fastq tool (versions 2.17–2.20) using default parameters. RNA-Seq data was aligned using STAR [24] to GRCh37 resulting in unsorted BAMs including chimeric reads as output. Gene counts and transcript per million (TPM) counts were generated and used for subsequent fusion detection using Isofox at Hartwig Medical Foundation (https://github.com/hartwigmedical/hmftools/tree/master/isofox). Samples with less than 5 million read pairs mapped were excluded from the analyses.

2.14. Association of copy-number alterations and expression

Copy-number alterations overlapping transporter genes were divided into three ordered levels: Deleted (loss of at least one copy), Neutral, or Amplified (gain of at least one copy). These alterations were fit in an ordinal logistic regression for each transporter gene with uncorrected transcripts per million as predictor. Statistically significant findings (p-value < 0.05) with odds ratios above 1 were viewed as having a positive effect.

2.15. Differential expression analyses

Raw read counts were imported into R and filtered on protein coding genes based on Ensembl GTF file (Homo sapiens GRCh37, version 87). Differential expression analysis between patients treated with chemotherapy and those treated with PKIs was performed on these counts using DESeq2 [25] and the Likelihood ratio test (LRT). Full model was "~primaryTumorLocation + group", with the reduced model "~primaryTumorLocation" for the LRT, resulting in significant genes due to "group" (chemotherapy/PKI). For visualization purposes, transcripts per million (TPM) values per transporter were corrected by the median value (+1 for nonnegative values) of the TPM of that transporter gene per tumor type as:

$$Tumor\ Corrected\ TPM = \left(\frac{Uncorrected\ TPM}{(tumor\ type\ median\ TPM + 1)}\right)$$

Results from the DESeq2 analyses were filtered on transporter genes

and adjusted p-value < 0.001 to be strict, as $\log_2$ fold change cannot be used directly due to the LRT approach.

2.16. Matched proteome and transcriptome abundances

We downloaded the 'Supporting Information' Table EV6 from Wang et al. [26]. This table contained genes that were detected in at least ten tissues, which we filtered for ABC and SLC transporters genes. P-values < 0.05 were labeled as significant. The Kolmogorov-Smirnov test was applied to the correlation distribution, comparing to a normal distribution.

3. Results

3.1. Somatic coding landscape of (drug) transporters

We investigated somatic mutations in the coding regions of transporter genes ($N = 4659$ in ABC transporters and $N = 16{,}249$ in SLC transporters, full overview in data file S1). These coding variants were predominantly nonsynonymous ($N = 14{,}890$, 71.22%) of which 85.84% were missense, and were most frequent in two genes encoding the largest transporter proteins, i.e. *ABCA13* ($N = 344$) and *ABCA12* ($N = 149$) (data file S1). Linear regression analysis illustrated that there is indeed a clear linear relation (R^2 of 0.76 for all transporter genes and 0.66 for established pharmacogenes, Fig. S2) between gene length and number of nonsynonymous coding mutations. From the established pharmacogenes (Fig. 1 A and B), *ABCB1* ($N = 133$, of which 85 missense variants) and *SLCO1B3* ($N = 111$, of which 74 missense variants) were most frequently affected. After correcting for the length of the gene-respective coding region, the established pharmacogenes (Fig. 1 C) with the highest number of somatic coding mutations per Mb were *SLC10A2* (16.0 per Mb), *SLCO1B3* (15.7 per Mb) and *SLCO1A2* (12.1 InDels per Mb). Conjointly, amongst the rest of the transporter gene families these were *SLC35A4* (22.6 per Mb), *SLC17A6* (17.4 per Mb) and *SLCO6A1* (16.8 per Mb). The established drug transporters most frequently affected by structural variants were *ABCB1*, *ABCC4* and *SLCO1B3*, which had 27–30 structural variants. In the rest of the transporter gene families, structural variants occurred 2616 times, overlapping 399 transporter genes (Fig. 1D), most frequently in zinc transporter *SLC39A11* ($N = 57$). In addition, we found eight statistically significant hotspot positions (Table S2) when comparing the number of hotspots against Poisson distributions. Seven of these variants were known variants and described in COSMIC [21]. These hotspot variants were detected in the same tissue as in COSMIC in 30.8–70%.

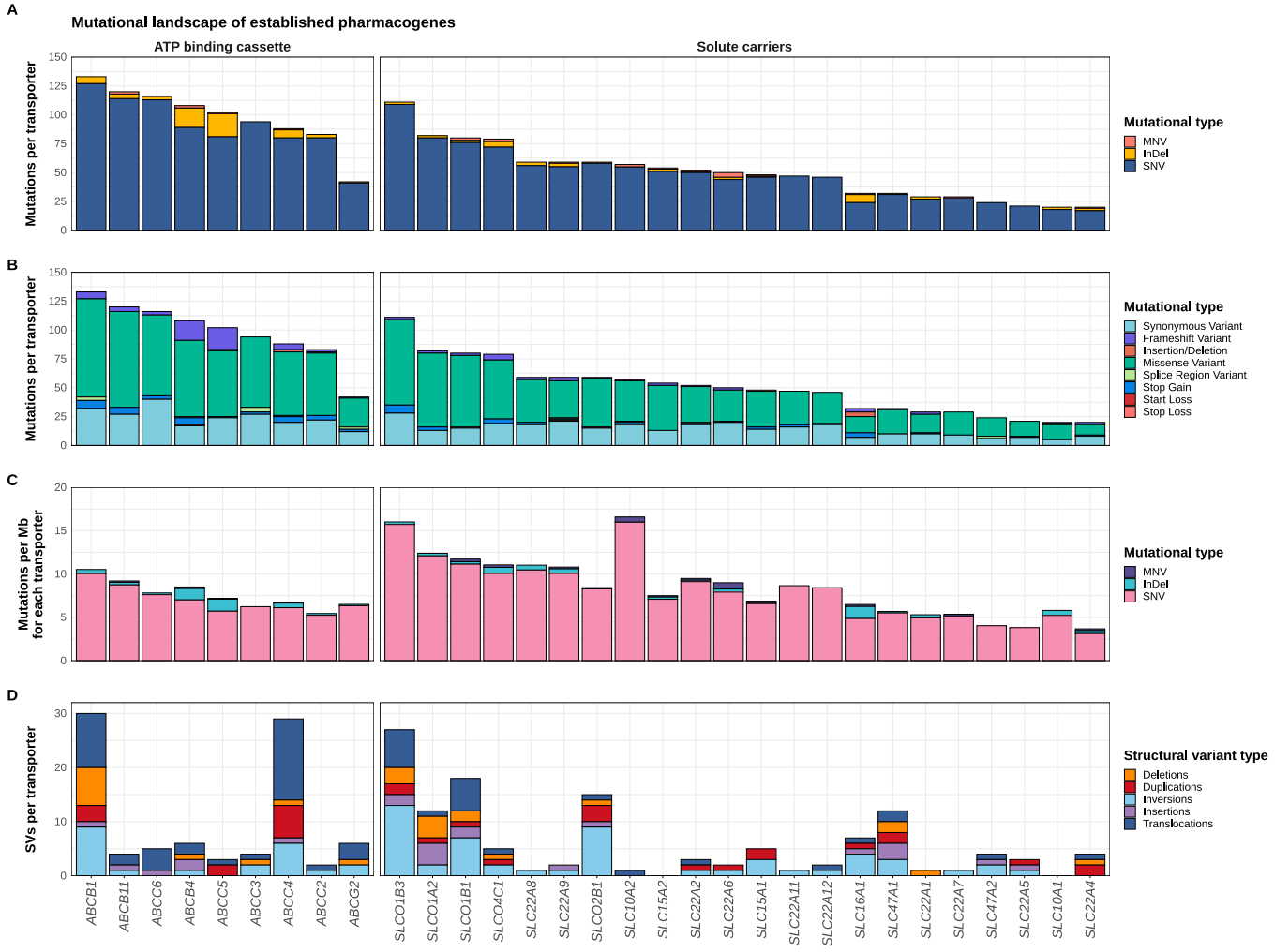

Fig. 1. General somatic mutational characteristics of established pharmacogenes ($N = 31$) within the metastatic pan-cancer whole-genome sequencing cohort ($N = 3290$), ordered descendingly by the number of mutations per transporter family. Each track is split in ATP binding cassette transporters and solute carrier genes, visualized as barcharts. Tracks A and B show the absolute number of coding somatic mutations, both illustrating the general category of the mutation through the color scheme, with B showing the results of the mutation at protein level. Track C shows the absolute number of mutations from tracks A and B, corrected by the coding length of each transporter gene. Lastly, track D presents the absolute number of structural variants overlapping with the transporter genes, with the color illustrating the type of structural variation. SNV = Single Nucleotide Variant, MNV = Multi-Nucleotide Variant, InDel = Insertion/Deletion.

3.2. Germline variants

In our cohort, we found 153,090,642 germline variants (data file S1) in the full gene bodies of transporter genes, of which 125,435,284 were SNVs, 27,635,620 InDels, and 19,738 MNVs (without discriminating between common variants and rare mutations). We determined the clinical significance of germline variants from our cohort through a comparison of 7108 unique, annotated variants from ClinVar [18]. We labeled 1661,621 variants based on ClinVar, 1.31% were classified as 'VUS', 1.52% as 'pathogenic/likely pathogenic', followed by 0.40% as 'drug response'. The bulk of these variants were 'benign/likely benign' (96.8%). Additionally, 214,261 germline variants overlap with their exact change on the variants selected in the DMET™ [27] plus panel ($N = 747$ unique variants). The mean number of identified DMET™ germline variants per transporter was 462, occurring in 48 of the 51 transporter genes on the panel. The gene with the most DMET™ germline variants was *SLCO1B7* ($N = 18,682$, 8.71% of total identified DMET variants, data file S1; sheet 'Germline').

3.3. Enrichment of somatic mutations within pretreatment groups

We used the dN/dS algorithm [22] to study selective effects by prior systemic treatments (groups specified in Fig. S3, Table S3). Somatic variants from patients who had received systemic treatment prior to biopsy ($N = 1947$) were categorized per treatment and subjected to dN/dS analysis, of which results with p-value < 0.01 are shown in Fig. S4. When correcting the p-values with Benjamini and Hochberg's method [23], none of the treatment groups showed a significant (q-value < 0.05) enrichment of mutations within transporter genes. Nutrient transporters, such as *SLC16A1* and *SLC5A8*, were relatively highly enriched with somatic variants and are known to have certain metabolic effects when affected by mutations [28]. In the list of established drug transporters, only *SLCO1B1* had a p-value < 0.01 (5-fluorouracil and gemcitabine groups, p-values of 0.0035 and 0.0011, respectively).

3.4. Focal copy-number alterations

Overexpression of several efflux transporters has been linked to drug resistance in several cancer types, in vitro [29]. Therefore, we studied somatic copy-number alterations (CNAs) in the 1947 patients who had received systemic therapy prior to the tumor biopsy. We identified focal (highly specific regions of a limited number of genes) and recurring copy-number alterations using GISTIC2.0 [19]. Deletions of the efflux transporter (*ABCD1*) and six influx transporters were enriched significantly (q-value < 0.1; Fig. S5). No statistically significant amplifications or deletions (q-value < 0.05) were found in established drug transporters, which likely illustrates the rarity of this phenomenon as a pharmacokinetic drug resistance mechanism in tumors.

3.5. WGS – Prevalence of somatic aberrations in patients with prior systemic treatment

We tested for differences in the prevalence of nonsynonymous SNVs, amplifications and deletions within transporter genes between patients with different prior systemic treatments, i.e. chemotherapy, hormonal therapy, targeted therapies and untreated patients using the Kruskall-Wallis test. Since Protein Kinase Inhibitors (PKIs) comprise a specific subset of targeted therapies, we included patients pretreated with PKIs as a separate subgroup and removed them from the targeted therapies group (Fig. 2). Since these three Kruskal-Wallis tests passed statistical significance at p-value < 0.05, we subsequently performed χ^2-tests, comparing each treatment group to the untreated patients for the three types of genetic changes. All comparisons showed statistically significant results (p-value < 0.05), except for targeted therapy in the SNV group ($P = 0.0707$).

Chemotherapeutic agents and PKIs are the best characterized substrates of the investigated membrane transporters. Therefore, we compared these two groups, after binarization of the genetic changes, with the χ^2-test and found that nonsynonymous coding SNVs ($P = <$ 2.2e-16), amplifications ($P = 4.35e-04$), deletions ($P = <$ 2.2e-16) were significantly different between patients pretreated with chemotherapy ($N = 1\ 553$) and PKIs ($N = 87$). In individual genes, there was no significant difference in SNV frequency between both groups. Statistically significant findings (adjusted p-value < 0.05) for amplifications and deletions are depicted in Table S4.

In Fig. 3 A and C, the percentage of affected samples for the two groups is visualized, with the absolute frequencies in Fig. 3B and Fig. 3D.

Four genes with significant differences have been clearly related with pharmacokinetics. *SLC31A1* and *SLC31A2* (χ^2-test adjusted p-values: 6.9e-09 and 2.5e-09) encode the copper transporters CTR1 and CTR2, respectively, which are known to mediate cellular influx of platinum compounds [30,31]. Although relatively more deletions were observed of *SLC31A1* in patients treated with PKIs, 30% of patients with *SLC31A1* deletions were pretreated with platinum containing chemotherapy. *SLC31A1/2* are located on the long arm of chromosome 9 and all 81 cases of *SLC31A2* deletions co-occurred with *SLC31A1* ($N = 82$).

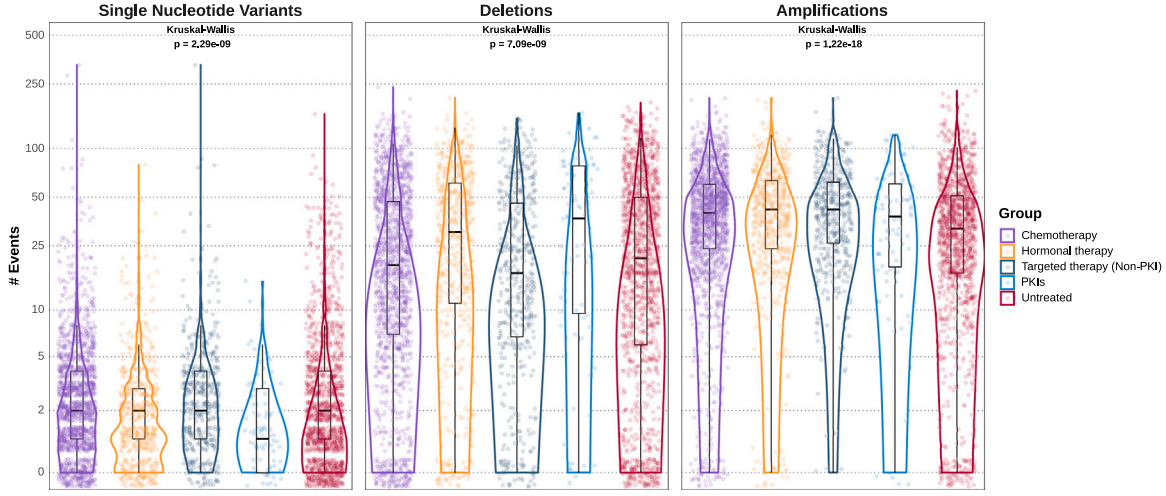

Fig. 2. The number of Single Nucleotide Variants, Deletions and Amplifications in transporter genes displayed as violin- and boxplots on a $\log_{10}$ scale from the CPCT-02 pan-cancer whole-genome sequencing cohort. Colors are indicative of pretreatment. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

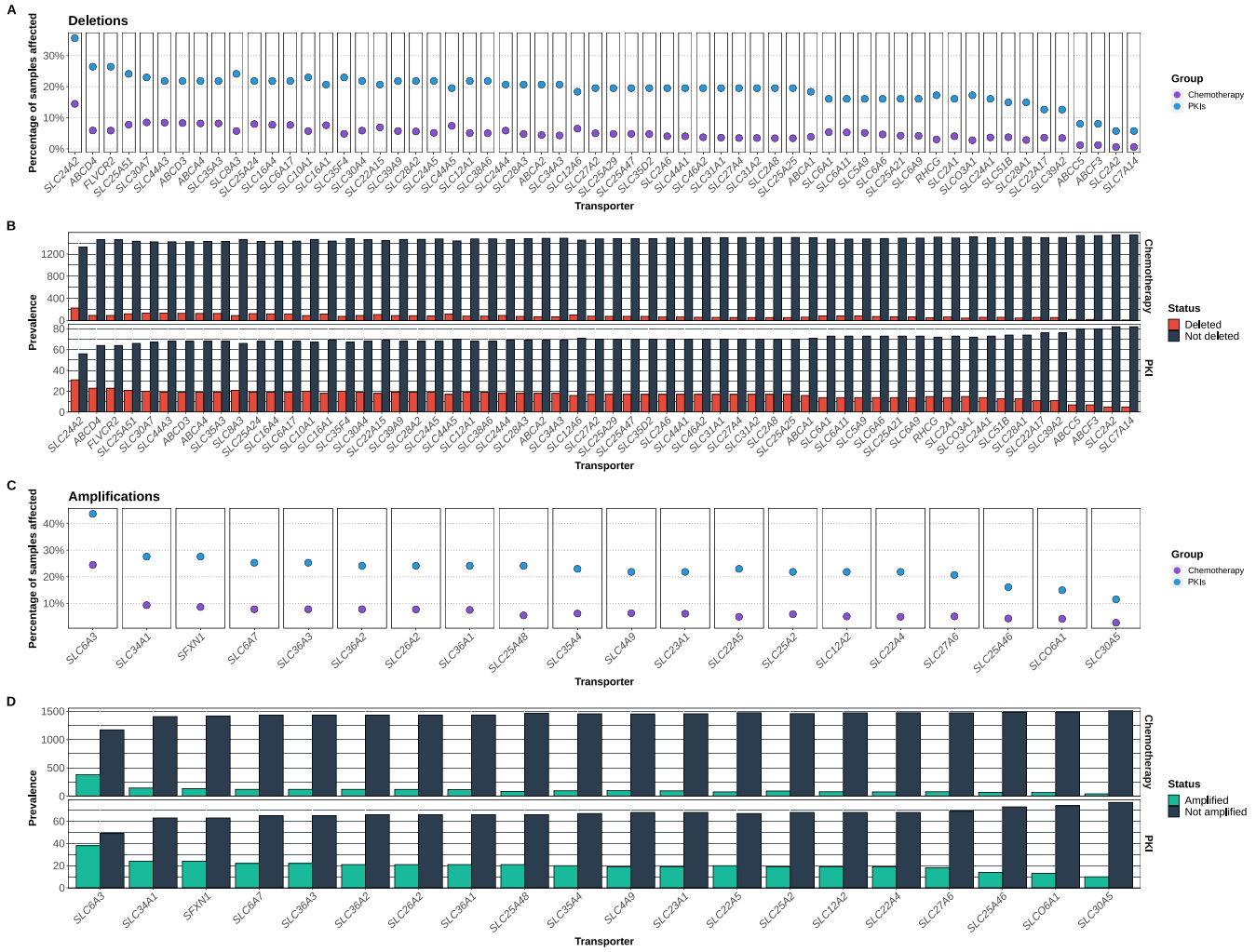

Fig. 3. Representation of differences in events within the transporter genes between chemotherapy and PKI pretreated patients in the CPCT-02 pan-cancer whole-genome sequencing cohort. Tracks A (deletions) and C (amplifications) show the percentage of patients affected by amplifications or deletions in transporter genes that had a significantly different prevalence based on the χ^2-test. Absolute frequencies of deletions (Track B) and amplifications (Track D) in patients pretreated with chemotherapy and protein kinase inhibitors, visualized as barcharts. Red bars indicate the number of patients with deleted allele(s) in the tumor, green bars show the gene was amplified, with the navy blue bars indicating unaffected samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Deletions in *SLC28A2* and *SLC28A3,* the encoding genes of the concentrative nucleoside transporters 2 (CNT2) and 3 (CNT3) were also relatively more frequent in the PKI-pretreated group (χ^2-test adjusted *p*-values: 3.5e-06 and 6.8e-07). These nucleoside transporters have been characterized extensively and are known to facilitate cellular uptake of nucleoside analogues such as gemcitabine [32,33] and cytarabine [34], and differential functioning of these transporters have been associated with marked survival differences after gemcitabine treatment [35]. Amplifications in known drug transporters were only significantly different in *SLC22A4* and *SLC22A5*, encoding the uptake transporters OCTN1 and OCTN2, respectively.

3.6. RNA-Seq – Differential expression analyses Chemotherapy versus PKI

Since 58.1% of transporter genes showed a positive association between deletion/amplifications and their mRNA expression (ordinal logistic regression, *p*-value < 0.05, Fig. S6), we continued the comparisons of chemotherapy versus PKI. In healthy human tissues, a correlation of transporter genes was detected between RNA expression and protein abundance [26]. Of the 195 of the transporter genes that were present in at least 10 tissue types in that analysis, there was a significant correlation ($P < 0.05$) between RNA expression and protein abundance was significantly correlated in 93 transporters (Fig. S7). We investigated differential expression of transporter genes between patients treated with chemotherapy ($N = 899$) versus those treated with PKIs ($N = 51$); taking tumor type into account. Statistically significant genes (adjusted $P < 0.001$, Likelihood Ratio Test) can be seen in Table 1. Several of these genes already showed a clear trend (Fig. 4), of which only *ABCC4* is known for drug transport.

In addition to showing an increased incidence of amplifications observed in WGS for *SLC27A6, SLC35A4* and *SLC12A2*, these genes now reveal a higher median expression in PKI-pretreated patients. This is in line with the positive correlation between RNA expression and protein abundance for *SLC12A2* and *SLC27A6* that has previously been found in healthy tissues [26].

4. Discussion

In this study, we describe the genomic and transcriptomic landscape of genes encoding SLC and ABC membrane transporters. Intratumoral drug exposure is a prerequisite for the effects of nearly all non-antibody

Table 1
Statistically significant differentially expressed genes corrected for tumor type (chemotherapy versus PKIs, *p*-adjusted < 0.001). The group with the highest tumor corrected median transcripts per million is listed (chemotherapy or PKI), as well as the difference in log likelihood deviance (stat) between the reduced model and the full model.

Transporter	Higher median	stat	Adjusted P
ABCA3	PKI	27.65	1.26E-05
ABCC4	PKI	18.29	5.20E-04
SLC2A1	Chemotherapy	17.70	6.61E-04
SLC4A4	PKI	29.63	5.71E-06
SLC7A1	Chemotherapy	37.31	2.43E-07
SLC7A5	Chemotherapy	17.47	7.27E-04
SLC12A2	PKI	54.91	1.38E-10
SLC12A6	Chemotherapy	23.79	5.80E-05
SLC26A4	PKI	27.98	1.09E-05
SLC27A6	PKI	19.41	3.32E-04
SLC35A4	PKI	35.81	4.35E-07
SLC35F5	Chemotherapy	17.13	8.22E-04
SLC35G1	PKI	35.95	4.16E-07
SLC37A2	Chemotherapy	20.89	1.79E-04
SLC38A5	Chemotherapy	19.32	3.42E-04
SLC39A2	Chemotherapy	28.29	9.67E-06

based systemic therapies and many preclinical studies address altered drug transporter functioning in cancer. However, to our knowledge, somatic genetic variation in drug transporter genes has not yet been studied on a large scale in advanced cancer patients. We observed interesting differences comparing genomic copy number alterations and transcriptomic expression in tumors from patients pretreated with chemotherapy and from patients pretreated with protein kinase inhibitors (PKIs), which are in line with prior knowledge on the pharmacokinetics of these drug classes. In particular, patients pretreated with PKIs had significantly more deletions in *SLC31A1* and *SLC31A2* (encoding the copper transporters CTR1 and CTR2, respectively) and in *SLC28A2* and *SLC28A3* (encoding the nucleoside transporters CNT2 and CNT3, respectively), which are involved in cellular uptake of platinum compounds and nucleoside analogues, respectively. Reduced function or expression of these proteins has been associated with worse survival after treatment with the mentioned substrates [36–39]. Moreover, single nucleotide polymorphisms in *SLC28A3* have been associated with an altered risk to anthracycline-induced cardiotoxicity [40–42], which could hint at its role in anthracycline transport, especially since doxorubicin seems to be transported by *SLC28A3* [43]. While the LRT analysis of the RNA-Seq data was corrected for tumor type, no correction factors were applied to the prevalence of genomic alterations in the WGS data. It should be noted that we observed significantly more of these aberrations in patients treated with PKIs, whereas these aberrations have been described to mediate resistance to chemotherapy. Our findings might therefore represent a primary resistance mechanism that is present before systemic treatment, which resulted in successful registration trials of treatment with PKIs and not with chemotherapy for the underlying cancer types, e.g. in renal cancer [44]. Alternatively, a substantial proportion of the tumors harboring a deletion in *SLC31A1* was previously treated with platinum agents, which could imply the presence of an underlying selective mechanism for these deletions [45]. It remains difficult, however, to pinpoint the drugs involved, since there is major overlap between deletions in these copper and nucleoside transporters. Although extensive, our dataset does not allow for differentiation between these two hypotheses, nor for testing of the functional consequences of the aberrations we found. Regardless the intrinsic or acquired nature of these deletions, these findings warrant further study as they may impair chemotherapy effects and should then potentially be assessed prior to the start of treatment.

In patients pretreated with protein kinase inhibitors, we also found significantly higher RNA expression of *ABCC4* and a higher prevalence of *SLC22A4* and *SLC22A5* amplifications. Both *ABCC4* and *SLC22A5* (encoding ergothioneine uptake transporter OCTN2) have been reported to facilitate imatinib transport [46]. While we have expression data from RNA-Seq, no quantitative proteomics data is available for these samples. Identifying protein abundance using proteomics approaches could be very helpful in understanding the relevance of these DNA amplifications in their biological behavior. Therefore, we alternatively used publicly available data to confirm that there is a correlation between RNA expression and protein abundance for transporters for transporters in healthy human tissue [26]. As both *SLC22A4* (OCTN1) and *SLC22A5* (OCTN2) are vital in (cancer) cell metabolism and show a different prevalence of amplifications, it is not unlikely that these differences are also caused by the selection of specific tumor types in the group of patients pretreated with PKIs [47].

By analyzing all membrane transporters from the ABC and SLC transporters, we also observed events in metabolic pathways, which may also be important in systemic treatment. For example, we found enrichment of *SLC16A1* (Monocarboxylate transporter 1; MCT1) in paclitaxel-pretreated patients, which seems to reflect metabolic effects due to altered functioning of MCT1 [48], especially since paclitaxel is not a MCT1 substrate [49]. Similarly, we found significantly more deletions in GABA transporters *SLC6A1* and *SLC6A11* in patients treated with PKIs, which could prove additional evidence to the theory that the GABA signaling system is a relevant component of the immunological

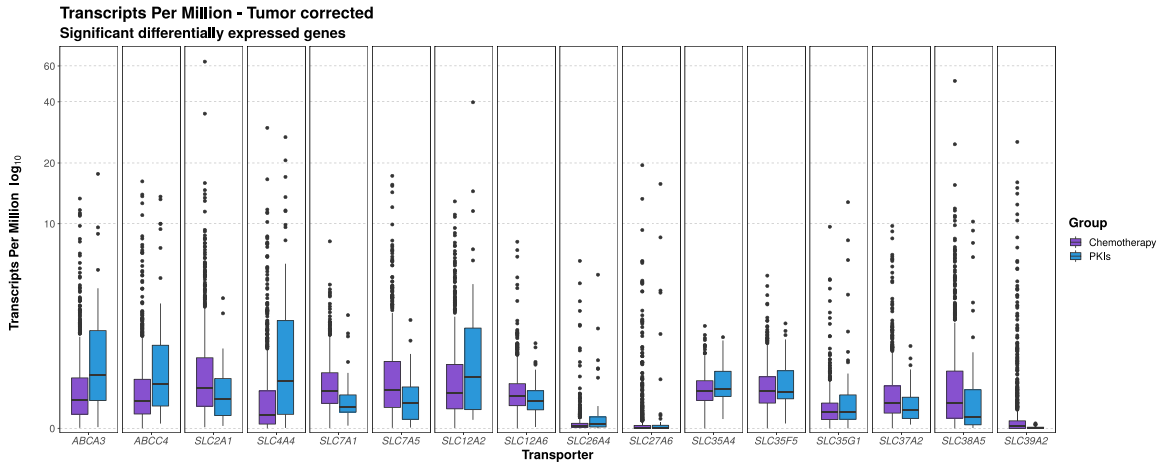

Fig. 4. Transcripts Per Million (tumor corrected) of statistically significant differentially expressed genes on a $\log_{10}$ scale of RNA-sequenced patients pretreated with chemotherapy or protein kinase inhibitors (PKIs). Colors are indicative of pretreatment group. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

anti-cancer response. However, these metabolic effects are outside the scope of this study and are therefore not further explored.

Additionally, we described the complete genomic landscape of genes encoding SLC and ABC membrane transporters, which predominantly comprises mutational events and copy-number alterations. The incidence of structural variants was marginal, especially in established drug transporters.

We are not aware of any previous effort to characterize the somatic pharmacogenetic landscape in cancer patients through large-platform sequencing data. This type of research can primarily serve for hypothesis generation. Transporters that are known to mediate transport of specific drugs, such as we describe for copper transporters and platinum compounds, can subsequently be studied in homogenous patient populations to estimate their contribution to drug resistance. For relatively unknown transporters, such as *SLC28A3*, more detailed determination of its substrates ideally precedes this step, for example in vitro, but possibly also by using data from The Cancer Genome Atlas, as has been done recently [50]. Major issues in analyzing these large-scale datasets are the heterogeneity in tumor type, disease stage and concomitant therapies between patients. The most informative analysis we performed was by dichotomizing patients into groups pretreated with either chemotherapy or protein kinase inhibitors.

In conclusion, we used whole genome sequencing and RNA sequencing data to scrutinize the intratumoral pharmacogenetic landscape. Comparing tumor specimens after chemotherapy and after PKI treatment, we observed a significantly different prevalence of deletions in genes that encode for copper and nucleoside transporters. Therefore, future pharmacologic research on drug resistance should focus on these membrane transporters. Lastly, this study illustrates that large scale sequencing data contain valuable information outside of the mainstream pharmacodynamic parameters.

Funding

No funding was received for this work.

CRediT authorship contribution statement

Wesley S. van de Geer: Writing – original draft, Conceptualization, Methodology, Formal analysis, Visualization. **Ron H.J. Mathijssen:** Writing – original draft, Conceptualization, Methodology, Supervision. **Job van Riet:** Software. **Neeltje Steeghs:** Investigation. **Mariette Labots:** Investigation. **Carla van Herpen:** Investigation. **Lot A. Devriese:** Investigation. **Vivianne C.G. Tjan-Heijnen:** Investigation. **Emile E. Voest:** Investigation. **Stefan Sleijfer:** Investigation. **John W.M. Martens:** Writing – review & editing. **Edwin Cuppen:** Software. **Harmen J.G. van de Werken:** Writing – original draft, Conceptualization, Methodology, Supervision. **Sander Bins:** Writing – original draft, Conceptualization, Methodology, Supervision.

Conflict of interest statement

The authors declare that they have no competing interests related to this study.

Data availability

The data that support the findings of this study are available from Hartwig Medical Foundation, which were used under data request number DR-027 for the current study. Both WGS and clinical data are freely available for academic use from the Hartwig Medical Foundation through standardized procedures and request forms can be found at https://www.hartwigmedicalfoundation.nl. All tools and scripts used for processing of the WGS data are available at https://github.com/hartwigmedical/ and/or can be provided by authors upon request.

Acknowledgements

This publication and the underlying study have been made possible partly on the basis of the data that Hartwig Medical Foundation and the Center of Personalized Cancer Treatment (CPCT) have made available to the study. We would like to thank J. Alberto Nakauma-Gonzalez for providing assistance with statistical testing with Poisson distributions.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.biopha.2022.114210.

References

[1] H.L. McLeod, Cancer pharmacogenomics: early promise, but concerted effort needed, Science 339 (2013) 1563–1566, 1563–6.
[2] M.H. Bailey, et al., Comprehensive characterization of cancer driver genes and mutations, Cell 173 (2018) 371–385.e18.
[3] M. Bar-Zeev, Y.D. Livney, Y.G. Assaraf, Targeted nanomedicine for cancer therapeutics: towards precision medicine overcoming drug resistance, Drug Resist. Updat. 31 (2017) 15–30.
[4] G. Housman, et al., Drug resistance in cancer: an overview, Cancers 6 (2014) 1769–1792.
[5] C. Holohan, S. Schaeybroeck, Van, D.B. Longley, P.G. Johnston, Cancer drug resistance: an evolving paradigm, Nat. Rev. Cancer 13 (2013) 714–726.
[6] S. Kobayashi, et al., EGFR mutation and resistance of non–small-cell lung cancer to gefitinib, Deaconess Med. Cent. (2005) (at <www.nejm.org>).
[7] K.M. Giacomini, et al., Membrane transporters in drug development, Nat. Rev. Drug Discov. 9 (2010) 215–236.
[8] S.K. Nigam, What do drug transporters really do? Nat. Rev. Drug Discov. 14 (2014) 29–44.
[9] M.V. Relling, et al., The clinical pharmacogenetics implementation consortium: 10 years later, Clin. Pharmacol. Ther. 107 (2020) 171–175.
[10] Walker, G. & Ahmad, T. , Drug Toxicity: Personalising IBD Therapeutics – The Use of Genetic Biomarkers to Reduce Drug Toxicity BT - Biomarkers in Inflammatory Bowel Diseases. (Sheng Ding, N. & Cruz, P. De) 257–269, Springer International Publishing, Cham, 2019.doi:10.1007/978–3-030–11446-6_21.
[11] J.E. Kucab, et al., A compendium of mutational signatures of environmental agents, Cell (2019), https://doi.org/10.1016/J.CELL.2019.03.001.
[12] S. Bins, et al., Implementation of a multicenter biobanking collaboration for next-generation sequencing-based biomarker discovery based on fresh frozen pretreatment tumor tissue biopsies, Oncologist 22 (2017) 33–40.
[13] P. Priestley, et al., Pan-cancer whole-genome analyses of metastatic solid tumours, Nature 575 (2019) 210–216.
[14] K.J. Karczewski, et al., The mutational constraint spectrum quantified from variation in 141,456 humans, Nature 581 (2020) 434–443.
[15] W. McLaren, et al., The ensembl variant effect predictor, Genome Biol. 17 (2016) 1–14.
[16] Yates, A.D. et al., Ensembl 2020. Nucleic Acids Res., 48, D682–D688 (2020).
[17] S. Povey, et al., The HUGO Gene Nomenclature Committee (HGNC), Hum. Genet. 109 (2001) 678–680.
[18] M.J. Landrum, et al., ClinVar: improvements to accessing data, Nucleic Acids Res. 48 (2020) D835–D844.
[19] C.H. Mermel, et al., GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers, Genome Biol. 12 (2011) 1–14.
[20] D.L. Cameron, et al., GRIDSS: Sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly, Genome Res. 27 (2017) 2050–2060.
[21] J.G. Tate, et al., COSMIC: the catalogue of somatic mutations in cancer, Nucleic Acids Res. 47 (2019) D941–D947.
[22] I. Martincorena, et al., Universal patterns of selection in cancer and somatic tissues, Cell 171 (1029–1041) (2017), e21.
[23] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. Ser. B 57 (1995) 289–300.
[24] A. Dobin, et al., STAR: ultrafast universal RNA-seq aligner, Bioinformatics 29 (2013) 15–21.
[25] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 550.
[26] D. Wang, et al., A deep proteome and transcriptome abundance atlas of 29 healthy human tissues, Mol. Syst. Biol. 15 (2019), e8503.
[27] J.K. Burmester, M. Sedova, M.H. Shapero, E. Mansfield, DMET™ microarray technology for pharmacogenomics-based personalized medicine, Methods Mol. Biol. 632 (2010) 99–124.
[28] V. Ganapathy, M. Thangaraju, P.D. Prasad, Nutrient transporters in cancer: Relevance to Warburg hypothesis and beyond, Pharmacol. Ther. (2009), https://doi.org/10.1016/j.pharmthera.2008.09.005.
[29] H. Xiao, Y. Zheng, L. Ma, L. Tian, Q. Sun, Clinically-relevant ABC transporter for anti-cancer drug resistance, Front. Pharmacol. 12 (2021), 648407.

[30] B.G. Blair, C.A. Larson, R. Safaei, S.B. Howell, Copper transporter 2 regulates the cellular accumulation and cytotoxicity of cisplatin and carboplatin, Clin. Cancer Res. 15 (2009) 4312–4321.
[31] C.A. Larson, B.G. Blair, R. Safaei, S.B. Howell, The Role of the Mammalian Copper Transporter 1 in the Cellular Accumulation of Platinum-Based Drugs, *Mol. Pharmacol.* 75, 324 LP – 330 (2009).
[32] R.J. Paproski, et al., Human concentrative nucleoside transporter 3 transfection with ultrasound and microbubbles in nucleoside transport deficient HEK293 cells greatly increases gemcitabine uptake, PLoS One 8 (2013), e56423.
[33] A. Khatri, et al., SLC28A3 genotype and gemcitabine rate of infusion affect dFdCTP metabolite disposition in patients with solid tumours, Br. J. Cancer 110 (2014) 304–312.
[34] A.H. Elsayed, et al., Comprehensive Ara-C SNP score predicts leukemic cell intracellular ara-CTP levels in pediatric acute myeloid leukemia patients, Pharmacogenomics 19 (2018) 1101–1110.
[35] C.J. Carter, A.H. Mekkawy, D.L. Morris, Role of human nucleoside transporters in pancreatic cancer and chemoresistance, World J. Gastroenterol. 27 (2021) 6844–6860.
[36] R. Maréchal, et al., Human Equilibrative nucleoside transporter 1 and human concentrative nucleoside transporter 3 predict survival after adjuvant gemcitabine therapy in resected pancreatic adenocarcinoma, Clin. Cancer Res. 15 (2009) 2913–2919.
[37] R.A. Soo, et al., Distribution of gemcitabine pathway genotypes in ethnic Asians and their association with outcome in non-small cell lung cancer patients, Lung Cancer 63 (2009) 121–127.
[38] X. Xu, et al., Genetic polymorphism of copper transporter protein 1 is related to platinum resistance in Chinese non-small cell lung carcinoma patients, Clin. Exp. Pharmacol. Physiol. 39 (2012) 786–792.
[39] Sun, C. et al., Genetic polymorphism of SLC31A1 is associated with clinical outcomes of platinum-based chemotherapy in non-small-cell lung cancer patients through modulating microRNA-mediated regulation. Oncotarget, 9 (35) (2018). <https://www.oncotarget.com/article/24794/text/>
[40] H. Visscher, et al., Pharmacogenomic prediction of anthracycline-induced cardiotoxicity in children, J. Clin. Oncol. 30 (2011) 1422–1428.
[41] H. Visscher, et al., Validation of variants in SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children, Pediatr. Blood Cancer 60 (2013) 1375–1381.
[42] K.M. Huang, S. Hu, A. Sparreboom, Drug transporters and anthracycline-induced cardiotoxicity, Pharmacogenomics 19 (2018) 883–888.
[43] K. Nagasawa, K. Nagai, N. Ohnishi, T. Yokoyama, S. Fujimoto, Contribution of specific transport systems to anthracycline transport in tumor and normal cells, Curr. Drug Metab. 2 (2001) 355–366.
[44] S. Buti, et al., Chemotherapy in metastatic renal cell carcinoma today? A systematic review, Anticancer. Drugs 24 (2013).
[45] M. Smid, S.M. Wilting, J.W.M. Martens, Lost by transcription: fork failures, elevated expression, and clinical consequences related to deletions in metastatic colorectal cancer, Int. J. Mol. Sci. 23 (2022) 5080.
[46] S. Hu, et al., Interaction of Imatinib with human organic ion carriers, Clin. Cancer Res. 14 (2008) 3141–3148.
[47] B. Juraszek, K.A. Nałęcz, SLC22A5 (OCTN2) Carnitine transporter—indispensable for cell metabolism, a jekyll and hyde of human cancer, Molecules 25 (2020).
[48] L. Hou, et al., Interfering cellular lactate homeostasis overcomes Taxol resistance of breast cancer cells through the microRNA-124-mediated lactate transporter (MCT1) inhibition, Cancer Cell Int. 19 (2019) 193.
[49] N. Vijay, M.E. Morris, Role of monocarboxylate transporters in drug delivery to the brain, Curr. Pharm. Des. 20 (2014) 1487–1498.
[50] M.T. Zimmermann, T.M. Therneau, J.-P.A. Kocher, The impact of pharmacokinetic gene profiles across human cancers, BMC Cancer 18 (2018) 577.