



Universiteit
Leiden
The Netherlands

Integration of germline pharmacogenomic burden to predict fluoropyrimidine-related toxicity: a secondary analysis of the PREPARE trial

Mattia, E. de; Park, Y.; Peruzzi, E.; Zhou, Y.T.; Roncato, R.; Polesel, J.; ... ; Cecchin, E.

Citation

Mattia, E. de, Park, Y., Peruzzi, E., Zhou, Y. T., Roncato, R., Polesel, J., ... Cecchin, E. (2025). Integration of germline pharmacogenomic burden to predict fluoropyrimidine-related toxicity: a secondary analysis of the PREPARE trial. *Oncogene*, 44(45), 4352-4362. doi:10.1038/s41388-025-03587-7

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/4298710>

Note: To cite this publication please use the final published version (if applicable).

ARTICLE



Integration of germline pharmacogenomic burden to predict fluoropyrimidine-related toxicity – A secondary analysis of the PREPARE trial

Elena De Mattia^{1,13}, Yoomi Park^{2,3,13}, Elena Peruzzi¹, Yitian Zhou², Rossana Roncato^{1,4}, Jerry Polesel⁵, Lucia Scarabel¹, Matthias Schwab^{6,7,8}, Henk-Jan Guchelaar⁹, Jesse Joachim Swen⁹, Michele Spina¹⁰, Fabio Puglisi^{4,10}, Giuseppe Toffoli¹, Volker M. Lauschke^{10,11,12,14} and Erika Cecchin^{1,14}

© The Author(s), under exclusive licence to Springer Nature Limited 2025

Testing for four dihydropyrimidine dehydrogenase (*DPYD*) variants (*DPYD**2 A, *DPYD**13, c.2846 A > T, *DPYD*-HapB3) is currently implemented in clinical practice to prevent fluoropyrimidines (FLs) related toxicity but with limited sensitivity. This study aimed to identify novel genetic factors in FL-related genes to enhance risk prediction using data from the PREPARE trial (NCT03093818). Two hundred seventy-four patients receiving FL-based chemotherapy with severe toxicity were sequenced for 60 candidate genes. Gene and pathway-level association analyses focusing mainly on rare variants were performed using dedicated statistical tests, including gene-wise variant burden (GVB) analysis. *DPYD* germline variant burden beyond the four routinely tested markers emerged to contribute to toxicity, indicating that rarer genetic variants could help in refining the optimal FL dosage ($p < 0.1$). Functional rare variant burden in *ABCB5*, *PARP1*, *ENOSF1*, *CYP3A4* and nuclear receptors pathway impacted on toxicity risk ($p < 0.05$ in at least one statistical test). GVB analysis confirmed *ABCB5* as a significant risk gene and highlighted *ABCC4*, *HNF4A*, and *XRCC3* as additional candidates. A predictive model combining genetic burden scores with clinical variables improved the identification of high-risk patients (sensitivity=0.71, specificity=0.74, accuracy=0.73). This study indicated a paradigm shift from population to individual-level arguing for an extension of testing beyond the four *DPYD* currently considered variants to predict FL-related toxicity.

Oncogene (2025) 44:4352–4362; <https://doi.org/10.1038/s41388-025-03587-7>

BACKGROUND

In recent years, human germline rare (minor allelic frequency [MAF] <1%) and novel variants, identified by next-generation sequencing (NGS) technologies, have been suggested to play an important role in the field of pharmacogenetics and to represent promising predictive markers for a personalized treatment [1–4]. Overall, it is estimated that approximately 30–40% of the overall functional variability in pharmacogenes (i.e., transporters, phase I and II enzymes, nuclear receptors) is caused by rare variants that are not commonly considered in conventional targeted genotyping tests [5].

For specific genes, a high contribution in the observed interindividual variability is hypothesized to be ascribable to rare variants, as for instance for the dihydropyrimidine dehydrogenase

(DPD, encoded by the *DPYD* gene) [6]. DPD is the first and rate-limiting enzyme in the detoxification pathway of fluoropyrimidines (FLs), which include 5-fluorouracil (5-FU) and its oral prodrug capecitabine. FL are widely prescribed for the treatment of many different solid tumors [7, 8] but are affected by high rates of severe toxicity, with a negative impact on treatment adherence, patients' quality of life, and medical costs [9–14]. At present, four prospectively validated *DPYD* variants (*DPYD**2 A, rs3918290; *DPYD**13, rs55886062; c.2846 A > T, rs67376798; c.1236 G > A-c.1129-5923 C > G [HapB3]; rs56038477-rs75017182) are routinely tested for pre-treatment drug adjustments based on specific clinical guidelines [7, 15–18]. However, they identify only a minority (about 17%) of patients experiencing severe FL-related toxicity [9, 13]. Among pharmacogenes, *DPYD* was found to have

¹Experimental and Clinical Pharmacology, Centro di Riferimento Oncologico di Aviano (CRO) IRCCS, via Franco Gallini n. 2, 33081 Aviano, PN, Italy. ²Department of Physiology and Pharmacology and Center for Molecular Medicine, Karolinska Institutet and University Hospital, 171 76 Stockholm, Sweden. ³Medical Research Center, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, South Korea. ⁴Department of Medicine, University of Udine, via Colugna 50, 33100 Udine, Italy. ⁵Unit of Cancer Epidemiology, Centro di Riferimento Oncologico di Aviano (CRO) IRCCS, via Franco Gallini n. 2, 33081 Aviano, PN, Italy. ⁶Dr Margarete Fischer-Bosch Institute of Clinical Pharmacology, Auerbachstraße 112, 70376 Stuttgart, Germany. ⁷Departments of Clinical Pharmacology, and of Biochemistry and Pharmacy, University Tuebingen, Geschwister-Scholl-Platz, 72074 Tübingen, Germania. ⁸Cluster of excellence iFIT (EXC2180) "Image-Guided and Functionally Instructed Tumor Therapies", University Tuebingen, Geschwister-Scholl-Platz, 72074 Tübingen, Germania. ⁹Department of Clinical Pharmacy and Toxicology, Leiden University Medical Center, Albinusdreef 2, 2300 RC Leiden, The Netherlands. ¹⁰Department of Medical Oncology, Centro di Riferimento Oncologico di Aviano (CRO), IRCCS, via Franco Gallini n. 2, 33081 Aviano, PN, Italy. ¹¹University of Tübingen, Geschwister-Scholl-Platz, 72074 Tübingen, Germany. ¹²Department of Pharmacy, the Second Xiangya Hospital, Central South University, 139 Renmin Middle Road, Changsha, Hunan 410011, China. ¹³These authors contributed equally: Elena De Mattia, Yoomi Park. ¹⁴These authors jointly supervised this work: Volker M. Lauschke, Erika Cecchin.

✉email: volker.lauschke@ki.s; ecechin@cro.it

Received: 10 June 2025 Revised: 29 August 2025 Accepted: 18 September 2025

Published online: 25 September 2025

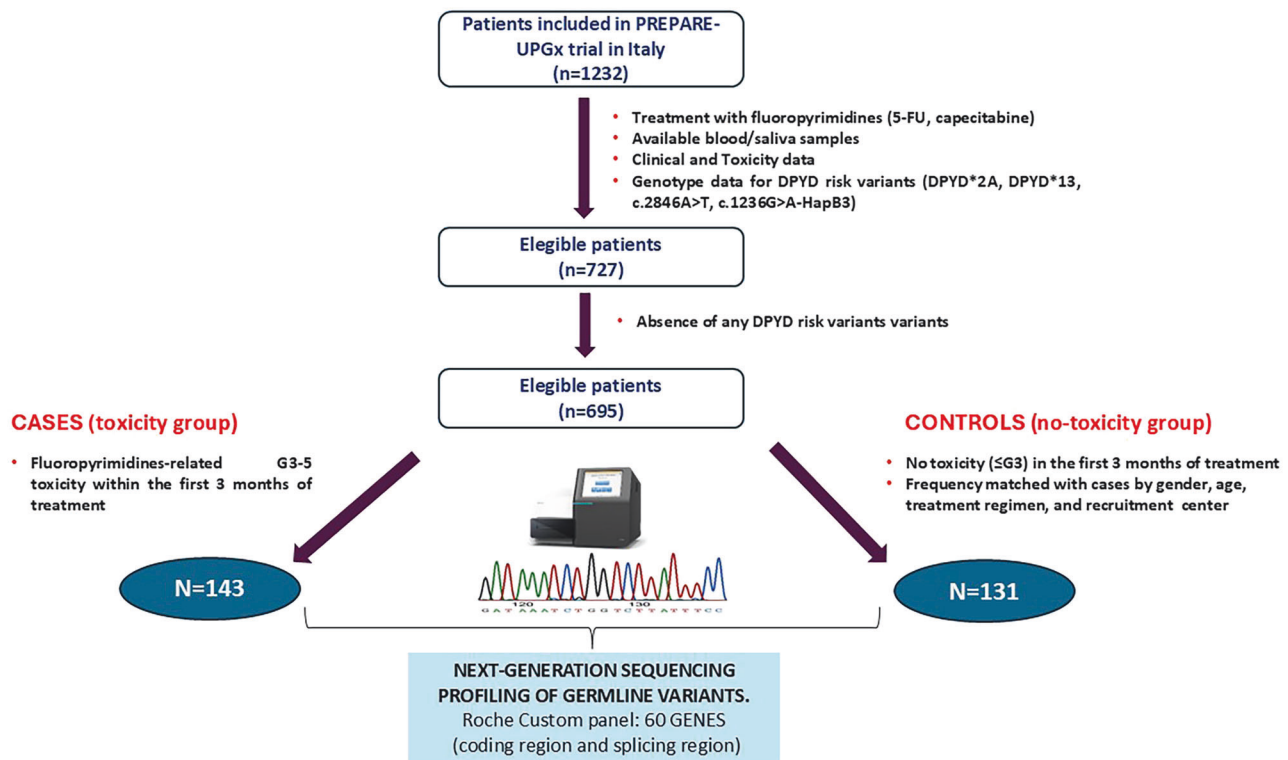


Fig. 1 Flowchart of the study strategy.

the highest number of relevant variants that are not captured by conventional targeted genotyping [6, 19] and several clinical case reports have shown that *DPYD* rare and novel variants could be causative for severe toxic reactions to FL treatment [1]. Previous studies suggested that *DPYD* rare and novel variants could negatively impact the DPD enzyme activity [20]. Testing *DPYD* sequence in a clinical setting failed to identify a role of rare *DPYD* variants in developing FL-related toxicity at a population level but pointed out a potential utility at the single patient level [21].

Our previous data suggested that *DPYD* rare and novel missense variants could collectively detect an appreciable additional fraction (about 13%) of patients at increased risk for developing severe FL-related toxicity, improving the preemptive identification of at-risk patients [22]. Successively our group further showed that the burden of rare and novel variants in the *DPYS*, encoding another member of the catabolic FL pathway, is an additional predictive marker for the risk of developing severe FL-related toxicity [23]. Based on these findings, it could be anticipated that rare variants also in other genes involved in FL disposition could play an important role in the pharmacogenetics of FLs.

The aim of the present study is to identify additional markers for the prediction of severe FL-related toxicity. To this end, data from the multicenter, international, controlled, open, block-randomized PREemptive Pharmacogenomic Testing for Preventing Adverse Drug REactions (PREPARE)- Ubiquitous Pharmacogenomics trial [24] were analyzed.

PATIENTS AND METHODS

Patient cohorts and clinical data collection

The study population for this single-country secondary analysis was selected from patients enrolled in the prospective PREPARE trial (NCT03093818) between March 7, 2017, and June 30, 2020 in three participating clinical centers in Italy (Centro di Riferimento Oncologico, Aviano; San Filippo Neri Hospital, Rome; Ca' Foncello Hospital, Treviso) [24]. The PREPARE study adhered to the principles outlined in the 1975

Declaration of Helsinki (with the 1983 revision) and was granted ethical approval by the local ethic committee. All patients provided written informed consent before entering the study to donate their anonymized blood samples for further studies including NGS analysis. While the analyses described here were not pre-specified among the secondary aims in the PREPARE protocol, a sub-study focusing on the role of rare and novel genetic variants as markers of extreme toxicity phenotypes was planned from the beginning [25]. All experiments were carried out in accordance with the relevant guidelines and regulations of Centro di Riferimento Oncologico di Aviano.

The flowchart of the study is shown in Fig. 1. Of the 1232 patients with a histologically confirmed diagnosis of solid cancer enrolled in the PREPARE trial, 727 patients were eligible for the study based on the following criteria: (1) assumption of treatment containing fluoropyrimidines (5-FU or capecitabine), given either as monotherapy or in combination with other chemotherapy drugs; (2) available biological sample from peripheral blood or saliva; (3) available detailed demographic and clinical data collected within PREPARE trial as well as additional information extracted from the patients' clinical records in accordance with the study objective (e.g., tumor location, treatment regimen and schedule, setting of treatment, detailed on toxicity)[12]; and (4) available genotype data for the four *DPYD* risk genetic variants (i.e., *DPYD**2 A, *DPYD**13, c.2846 A > T, c.1236 G > A-HapB3).

The study set was selected from the 727 eligible patients considering the absence of any of the four *DPYD* risk genetic variants. Patients were selected for the "toxicity group" if they developed at least one episode of grade ≥ 3 hematologic or nonhematologic toxicity according to the Common Terminology Criteria for Adverse Events (CTCAE) v.4.0, during the first three months of therapy. Only those toxicities that were most likely to be related to FL administration were considered (i.e., neutropenia, leukopenia, and thrombocytopenia as hematological toxicity; nausea, vomiting, diarrhea, hand-foot syndrome, mucositis, and cardiotoxicity as non-hematological toxicity). Patients were selected for the "no-toxicity group" if they did not develop any grade ≥ 3 hematologic or nonhematologic FL-related toxicities during the first three months of treatment at the standard dose conditions.

Candidate genes and polymorphism selection

Candidate genes were selected based on a literature search (PubMed-MEDLINE) based on their described involvement within FL-related

pathways (e.g., drug metabolism, folate pathway, drug membrane transporters, gene transcription regulators, DNA repair enzymes). A set of 60 genes were eventually selected for the pharmacogenetic analysis. The custom design of the panel was carried out by HyperDesign software, based on Genome Build hg19/GRCh37 (February 2009), and captured the genetic variability of all exons and their adjacent splice junctions (about 35 bases up- and downstream of the exon). The list of genes included in the panel ($n = 60$), the pathway ($n = 5$) to which they belong, and the number of bases covered by the custom design for each gene are shown in Supplementary Table S1.

Library preparation and sequencing

Genomic DNA was extracted from peripheral blood samples, using the BioRobot EZ1 automated system in association with the “EZ1 DNA Blood Kit 350 μ l” Kit (Qiagen), or from saliva samples, using the manual purification method with the GeneJET Whole Blood Genomic DNA Purification Mini Kit (Thermo Scientific). The extracted DNA was stored at +4°C. The quality and quantity of DNA samples were assessed using the Nanodrop (Thermo Scientific) spectrophotometric method (i.e., 260/280 and 260/230 ratio) and the Quantus™ Fluorometer (Promega), respectively.

Gene sequencing was carried out by a hybrid capture-based custom Roche assay. DNA library preparation for all samples was conducted on 100 ng of input DNA. The library preparation was performed according to the manufacturer's instructions in KAPA HyperCap Workflow v3.4 User's Guide with the KAPA EvoPlus Kit (Roche). Briefly, the genomic DNA was enzymatically fragmented, and after adapter ligation, a double-sided size selection was performed to obtain fragments of approximately 300–350 bp, which were then amplified. Each library was pooled prior to in-solution hybridization with custom KAPA HyperCap Target Enrichment Probes (Roche) containing complementary oligonucleotide DNA baits. Following washing, the captured libraries were amplified by PCR. The concentration of single and pooled libraries was measured using a Quantus™ Fluorometer (Promega), while quality and size distribution were assessed using the 2200 TapeStation system (Agilent). Pooled libraries were sequenced on a MiSeq platform (Illumina) according to the manufacturer's instructions, using a 300-cycle kit with a 2 × 150 paired-end read setup.

Bioinformatic analysis, variants annotation and functional prediction

Raw paired-end sequencing reads were aligned to the hg19/GRCh37 human reference genome using BWA-MEM (v0.7.18) with appropriate read group identifiers to ensure consistent tracking across samples [26]. The resulting SAM files were converted into BAM format, sorted, and indexed with SAMtools (v1.20) [27]. Duplicate reads were then identified using Picard (v2.2.1) and base quality score recalibration was conducted using the Genome Analysis Toolkit (GATK, v4.2.1.0) with known variant sites from dbSNP build 138 [28]. Variants were subsequently called using GATK's HaplotypeCaller in GVCF mode, restricting analysis to targeted genomic regions as defined by a BED file. Each variant type underwent stringent quality filtering based on depth, strand bias, mapping quality and rank sum metrics to eliminate low-confidence calls. Based on allele frequencies from non-Finnish Europeans (NFE) in gnomAD, variants were classified into very rare ($MAF \leq 0.1\%$), rare ($0.1\% < MAF < 1\%$), common ($MAF \geq 1\%$), or novel (not registered, $MAF = 0$). Variant classes were inferred using SnpEff [29]. For missense variants, predictions from seven established in silico tools (CADD phred, Mutation Assessor, PROVEAN, Polyphen2 HDIV, Polyphen2 HVAR, REVEL, and SIFT) were obtained using ANNOVAR [30]. Additionally, the ADME-optimized prediction framework (APF) and the updated APF2 score, were calculated [31, 32].

Statistical analysis

To investigate associations between rare variants and severe FL-related toxicity, three statistical methods were employed: the burden tests, variance component tests and Sequence Kernel Association Tests (SKAT-O) tests [33]. Analyses considered rare variants ($MAF < 1\%$ or novel) that were predicted as putatively functional (i.e., variants causing frameshifts, stop-gain, start-lost and splice-site variants, in-frame indels, as well as deleterious missense variants). Tests were performed at gene and pathway levels using logistic regression under an additive genetic model adjusted for FL type, administration of an anticancer co-treatment, age and sex. A gene or pathway was considered significantly associated with toxicity if at least one of the three statistical models yielded a p -value < 0.05 . To correct

for multiple testing, raw p -values were adjusted using the Benjamini–Hochberg false discovery rate procedure, with q -value threshold of 0.05 defining statistical significance. Jackknife sampling approach was used to identify influential variants contributing to associations. Gene-wise variant burden (GVB) scores that quantify the cumulative effects of genetic variation at the individual level were calculated for each of the aforementioned nine variant effect predictors [34]. Scores were normalized to a 0–1 scale, with lower values indicating higher functional impact (Supplementary Table S2). Logistic regression analyses for GVB score were performed with adjustment for FL type, co-treatment, age and sex. Genes or pathways achieving a significant logistic regression p -value (< 0.05) in at least five out of the nine tools were prioritized. Individual variant effects were subsequently tested within prioritized genes and pathways. Interactions between genetic variants and co-medications (oxaliplatin or irinotecan) on toxicity were evaluated using logistic regression models incorporating GVB-co-medication interaction terms, adjusted for age and sex, at both the gene and pathway levels.

Predictive modeling of toxicity outcomes involved clinical, genetic, and combined clinical-genetic models. Clinical variables included sex, age, FLs type, co-treatment with other anti-cancer drugs (yes/ no), and type of co-administered anticancer drugs (irinotecan or oxaliplatin). Genetic predictors incorporated GVB scores from prioritized genes. All predictive models were constructed using Python's scikit-learn [35], employing standardized numerical features, one-hot encoding for categorical features, and logistic regression classifiers. Model performance was evaluated using repeated stratified k -fold cross-validation (a five-fold scheme repeated 100 times, 500 total iterations per model), with the area under the receiver operating characteristic curve (AUC) as the primary evaluation metric. At the optimal threshold, binary predictions were generated and confusion matrices used to calculate sensitivity, specificity, accuracy, false positive rate (FPR), false negative rate (FNR), and positive predictive values (PPV). The number needed to surveil (NNS) was calculated as $1/PPV$ and the ratio of false positives to true positives (FP/TP) were derived to quantify clinical surveillance burden. Standardized logistic regression coefficients were extracted from the final combined model to assess the relative contributions of individual predictors.

RESULTS

Sequencing and patient characteristics

A total of 274 FL-treated patients ($n = 143$ in the “toxicity group”; $n = 131$ in the “no-toxicity group”) were selected and analyzed by targeted NGS covering 60 genetic loci. All the sequenced samples passed quality checks by FastQC. Overall, an average of 528,457.9 ($\pm 115,375.3$, sd) mapped reads per sample was generated, with 99.95% (± 0.02) of reads aligning and 95.5% (± 1.7) of bases achieving a recalibrated quality score of Q20 or higher. The mean depth of coverage across targeted regions was 151.2 (± 31.3).

The main clinical-demographic features of the two study populations are shown in Table 1. The two groups were balanced in terms of clinical-demographic characteristics. Almost all patients (272 [99.2%]) were of self-reported European ethnicity; one patient each was of self-reported African and Hispanic ancestry. Of the 143 patients in the “toxicity group”, 105 (73.4%) developed grade 3 toxicity, 36 (20.2%) grade 4 toxicity, and 2 (1.4%) grade 5 toxicity, as the maximum grade of hematologic or non-hematologic toxicity during the first three months of treatment. 110 of 143 patients (76.9%) developed grade ≥ 3 hematologic toxicity, with neutropenia being the most common adverse event (89/143; 62.2%). 52 of 143 patients (36.4%) developed grade ≥ 3 non-hematologic toxicity of which diarrhea was the most common (29/143; 20.3%).

Genetic landscape in 60 FL-related pharmacogenes

The genetic variability in 60 genes involved in the mechanism of action of FLs was analyzed. These include genes involved in drug metabolism and folate cycle ($n = 9$), phase I-II enzymes ($n = 4$), ABC/SLC transporters ($n = 15$), nuclear receptors ($n = 10$), and DNA repair pathways ($n = 22$, Fig. 2A). From 274 individuals a total of 908 exonic single-nucleotide variants and indels were identified (Fig. 2B). Of these, 520 variants (57.3%) were classified as missense,

Table 1. Socio-demographic and clinical characteristics of solid cancer patients enrolled in the study.

	Toxicity group (n = 143)	No-toxicity group (n = 131)	p-value
Sex, N (%)^a			
Female	80 (55.9)	68 (51.9)	0.5448
Male	63 (44.1)	63 (48.1)	
Age, years (median, range)^b	63 (29-81)	64 (37-81)	0.4460
Cancer Type, N (%)^c			
Head-neck cancer	44 (30.8)	18 (13.7)	0.718
Colorectal cancer	43 (30.1)	61 (46.6)	
Gastro-esophageal cancer	21 (14.7)	24 (18.3)	
Pancreatic cancer	15 (10.5)	11 (8.4)	
Breast cancer	12 (8.4)	8 (6.1)	
Other	8 (5.6)	9 (6.9)	
Fluoropyrimidines, N (%)^a			
5-Fluorouracil	109 (76.2)	73 (55.7)	0.0005
Capecitabine	34 (23.8)	58 (44.3)	
Co-treatment, N (%)^a			
Monotherapy	14 (9.8)	18 (13.7)	0.3494
Association with oxaliplatin	104 (72.7)	93 (71.0)	0.7887
Association with irinotecan	25 (17.5)	23 (17.6)	1
Max Toxicity Grade*, N (%)			
0	-	50 (38.2)	
1	-	78 (59.5)	
2	-	3 (2.3)	
3	105 (73.4)	-	
4	36 (25.2)	-	
5	2 (1.4)	-	
Early G3-5 toxicity (within 30 days), N (%)	96 (67.1)	-	

^aP values were calculated using the Fisher's exact test. ^bP values were calculated using Wilcoxon test.

^cP values were calculated using chi-squared test for trend in proportions.

*Maximum grade of hematological or non-hematological toxicity experienced by the patients.

start-lost, stop-gained, frameshifts, or affected canonical splice sites. The remaining variants were synonymous changes (307; 33.8%) or were localized in untranslated regions (2 in the 5' UTR [0.2%] and 27 in the 3' UTR [3%]). Among the exonic variants, several putative loss-of-function variants were identified, including 22 frameshift variants, 8 stop-gain variants, 2 start-loss variants, and 1 variant affecting a splice site (Fig. 2B). Notably, the majority of variants were rare (69.9%) or very rare (52.1%). Furthermore, nearly half (48.3%) of all variants were only observed in a single individual, highlighting the genetic diversity within these pharmacogenes (Fig. 2C). To evaluate the functional importance of rare pharmacogenetic variants, the frequencies of putatively functional variants were calculated and substantial differences across the 60 genes were observed. Overall, the majority of the genetically encoded functional variability was attributed to rare variants whereas in *ABCG2*, *PPARG*, *PMS1*, *OGG1*, *XRCC3* and *MSH2*, functional variability was driven primarily by common variants

(Fig. 2D). The number of putatively functional variants was particularly low in *STAT3*, and no functional variants were detected in *ERCC1*. Overall, each patient harbored an average of 48.4 variants that affected the amino acid sequence of the respective gene products across the 60 target genes, with rare variants accounting for 4.2 (8.6%) of these (Fig. 2E).

DPYD variants and FL-induced toxicity

Since functional *DPYD* variants are known to be strongly associated with grade ≥ 3 toxicity during FL chemotherapy, genetic variability in *DPYD* was evaluated. Carriers of the four clinically established *DPYD* variants were excluded during cohort selection and these variants were thus not present in the selected patients. However, ten other *DPYD* variants with putative functional impacts (9 missense and 1 frameshift; Supplementary Table S3) were identified. While these variants have been previously reported in the PharmVar database [36], they had not been implicated in FL toxicity. Rare variant association tests for rare functional *DPYD* variants did not yield statistically significant results (burden test $p = 0.1624$; variance component test $p = 0.3320$; SKAT-O test $p = 0.2492$), likely at least in part due to the low numbers of carriers. However, when the aggregated functional impact of all identified *DPYD* variants was evaluated using GVB analysis, marginally significant associations were identified ($p < 0.1$, Supplementary Fig. S1), suggesting a residual increase in risk among rare variant carriers to develop severe FL-induced toxicity. This finding indicates that *DPYD* variability beyond the four clinically validated variants may contribute to toxicity in this specific set of Italian patients of self-declared European ancestry.

Rare variants association analyses

Next, rare variant association tests for the remaining 59 target genes were performed. It was hypothesized that rare variants tend to be functional; therefore, the effects of putative functional variants with low minor allele frequencies within defined regions were aggregated and analyzed using three complementary tests (i.e., burden tests, variance component tests and SKAT-O tests). Specifically, the burden test collapses all rare variants within a gene into a single score, assuming that variants act in the same direction. The variance component test instead treats each variant's effect as a random variable, allowing for heterogeneity in both direction and magnitude. SKAT-O adaptively combines the two approaches to maximize power across a range of effect scenarios. Importantly, four genes showed significant associations with FL-induced toxicity at the gene level (Table 2A). *ABCB5* exhibited the strongest association, with significant results across all three tests (burden test $p = 0.0027$; variance component test $p = 0.0330$; SKAT-O test $p = 0.0052$). *PARP1* showed a strong signal in the burden ($p = 0.0058$) and SKAT-O ($p = 0.0087$) tests, suggesting that variants in *PARP1* likely have similar directional and comparable magnitudes effects on toxicity. In contrast, *ENOSF1* and *CYP3A4* yielded the strongest signals in the variance component test (*ENOSF1*: variance component test $p = 0.0105$ and SKAT-O test $p = 0.0178$; *CYP3A4*: variance component test, $p = 0.0488$; SKAT-O test, $p = 0.0776$), indicating that these genes may harbor both risk and protective variants or include many non-causal variants. At the pathway level, nuclear receptors emerged as significantly associated with severe toxicity (burden test, $p = 0.0060$; SKAT-O test, $p = 0.0113$, Table 2B). After Benjamini-Hochberg multiple testing correction, the nuclear receptor pathway remained significantly associated with toxicity (FDR $q = 0.0302$), whereas individual genes were only borderline significant (*ABCB5*, $q = 0.13$ and *PARP1*, $q = 0.14$).

When individuals with at least one nonsynonymous rare variant were recoded as carriers, the variant carriers in *ABCB5*, *PARP1*, and *CYP3A4* were associated with increased risk of toxicity, whereas carriers in *ENOSF1* and the nuclear receptor pathway exhibited a

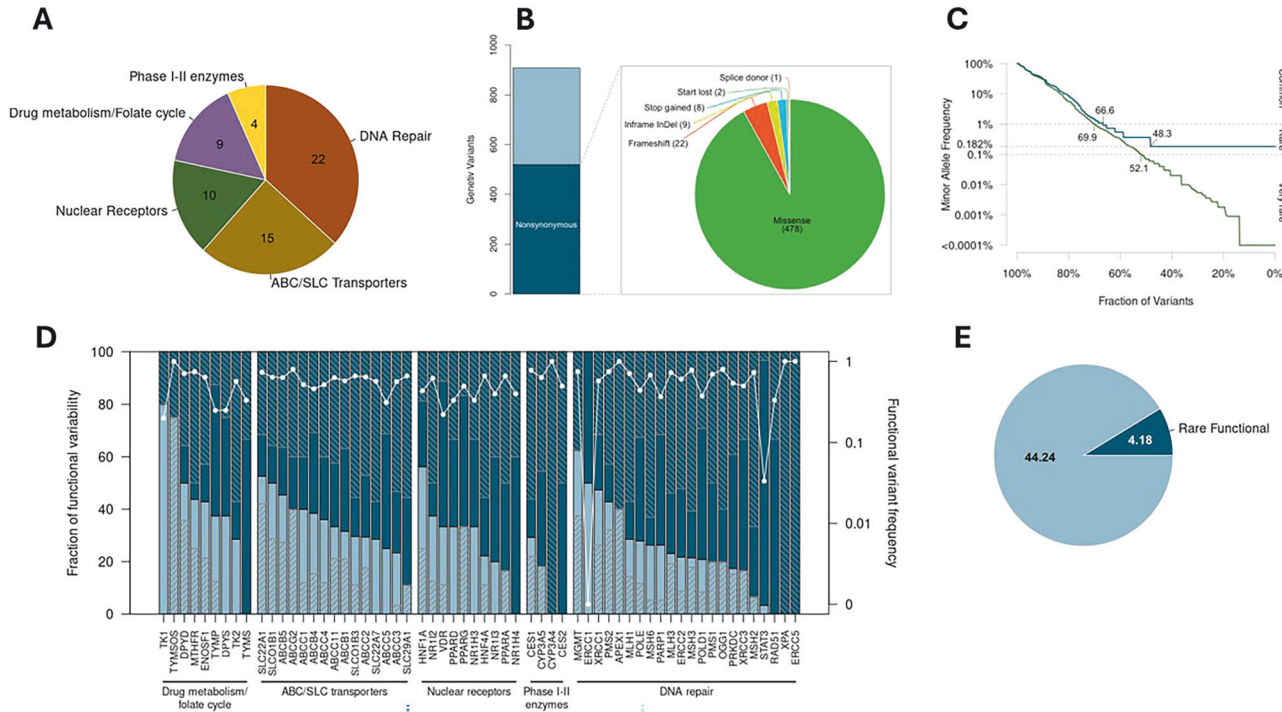


Fig. 2 The landscape of genetic variants in the study population across all tested target genes. Only variants annotated with target genes were included in the analysis. Multi-allelic variants were split and counted as separate variants and single variants with unclear classification (e.g. structural interaction variants) were merged into “missense” **A** Pie chart illustrating the distribution of the 60 target genes across various pathway. **B** 520 (57.3%) of the 908 identified variants were nonsynonymous. The pie chart details the distribution of each nonsynonymous variant type. **C** The minor allele frequency (MAF) distribution of all the identified variants is displayed for both the gnomAD Non-Finnish European (NFE, green line) and the PREPARE study population (blue line). Based on gnomAD NFE frequencies, 69.9% of variants were classified as rare (MAF < 1%) and 52.1% as very rare (MAF < 0.1%), whereas in PREPARE, 66.6% were rare and 48.3% were singletons. Variants found only in the PREPARE dataset (i.e., absent in gnomAD NFE) were assigned a pseudo-value of 1e-6 to ensure that all variants are represented on the plot. **D** The frequency of putative functional variants is plotted on a log scale, shown as dots connected by a white line for each of the 60 target genes (right y-axis). The fraction of overall genetic variability allotted to common (light blue) and rare (dark blue) variants is shown on the left y-axis. The proportion of putative functional variants is indicated by a hatched overlay. **E** Across the 60 target genes analyzed, each individual carried, on average, 4.18 rare and 44.24 common putative functional variants.

Table 2. Genes **A** and pathways **B** significantly ($p < 0.05$) associated with fluoropyrimidine-induced toxicity in at least two of the three rare-variant association tests.

	Burden test*	Variance component test*	SKAT-O*
(A) Gene			
<i>ABCB5</i>	0.0027	0.0330	0.0052
<i>PARP1</i>	0.0058	0.1480	0.0087
<i>ENOSF1</i>	0.5488	0.0105	0.0178
<i>CYP3A4</i>	0.3260	0.0488	0.0776
(B) Pathway			
Nuclear Receptors	0.0060	0.1131	0.0113

*p-values were estimated adjusting for sex, age, and fluoropyrimidine type and co-treatment (yes/no). Associations with $p < 0.05$ are evidenced in bold.

decreased risk. Specifically, *ABCB5* variant carriers have a 4.8-fold increased risk of developing toxicity compared to non-carriers (95%CI:1.7–13.3, $p = 0.0028$), while individuals with at least one nonsynonymous rare variant in the nuclear receptor pathway experience about a 58% reduction in the odds of developing toxicity (OR = 0.42, 95%CI:0.19–0.93, $p = 0.0327$). Using jackknife sampling to prioritize the main contributing variants, the aggregated effects of rare variants (9 of 12 in *ABCB5*, 6 of 6 in

PARP1, 3 of 5 in *CYP3A4*, 3 of 7 in *ENOSF1*, and 15 of 24 in the nuclear receptor pathway) that contributed to FL-induced toxicity was identified (Supplementary Fig. S2, Supplementary Table S4).

Gene wise variant burden (GVB) analysis

To evaluate the cumulative impact of all identified variants within defined regions, GVB analyses were performed. For each of the 274 samples, a GVB score was computed for each of the 60 target genes and logistic regression analyses were conducted. Four genes (*ABCB5*, *ABCC4*, *HNFA4*, *XRCC3*) exhibited significant associations (Table 3). Specifically, *ABCB5* and *ABCC4* were associated with an increased risk of severe FL-induced toxicity, as indicated by negative beta coefficients (lower GVB score corresponds to higher risk). In contrast, *HNFA4* and *XRCC3* exhibited protective effects (positive beta coefficients). Notably, two common variants (rs6461515 in *ABCB5* and rs861539 in *XRCC3*) showed statistically significant frequency differences between the “toxicity” and “no-toxicity” groups ($p < 0.05$) by themselves, suggesting that they substantially contributed to toxicity risk (Supplementary Fig. S3). Approximately 40% of the variants in each gene (*ABCB5*: 8/21, *ABCC4*: 5/12, *HNFA4*: 2/5, *XRCC3*: 2/3) were detected exclusively in the “toxicity group”. In the putative risk genes, over 60% of these rare variants (*ABCB5*: 5/8, *ABCC4*: 3/5) were predicted to be damaging, suggesting that while other individual variants may not reach statistical significance on their own, their aggregated effect constitutes a substantial risk signal. A high degree of concordance (84%) in functional predictions was observed across nine established in

Table 3. Gene-wise variant burden (GVB) analysis; **A** Gene lists associated with fluoropyrimidine-induced toxicity in at least five of the nine in silico prediction tools used. Y=The logistic regression test $p < 0.05$ N= The logistic regression test $p \geq 0.05$. **B** Details about the effect described by the GVB score and the p-value obtained by logistic regression test for the selected genes.

(A)												
Gene	Significant results	APF	APF2	CADD phred	Mutation Assessor	PROVEAN	Polyphen2 HDIV	Polyphen2 HVAR	REVEL	SIFT		
ABCB5	5	N	N	Y	Y	Y	N	N	Y	Y		
ABCC4	5	N	Y	N	Y	Y	N	N	Y	Y		
HNF4A	5	N	Y	Y	Y	Y	N	N	N	Y		
XRCC3	5	Y	N	N	Y	Y	Y	N	N	Y		
(B)												
Gene	APF	APF2	CADD phred	Mutation Assessor	PROVEAN	Polyphen2 HDIV	Polyphen2 HVAR	REVEL	SIFT			
ABCB5	Effect	P*	Effect	P*	Effect	P*	Effect	P*	Effect	P*	Effect	P*
ABCB5	NS	NS	risk	0.007	risk	0.023	NS	NS	risk	0.044	risk	0.002
ABCC4	NS	0.022	NS	0.029	risk	0.034	NS	NS	risk	0.016	risk	0.035
HNF4A	NS	0.027	protective	0.028	protective	0.031	NS	NS	NS	NS	protective	0.019
XRCC3	protective	0.039	NS	0.040	protective	0.041	protective	0.041	NS	NS	protective	0.046

*P-values were estimated adjusting for sex, age, and fluoropyrimidine type and co-treatment (yes/no). Associations with $p < 0.05$ are evidenced in bold.

Table 4. Interaction analysis by Gene-wise variant burden (GVB) test; **A** Gene lists associated with fluoropyrimidine-induced toxicity in at least five of the nine in silico prediction tools used in the oxaliplatin and irinotecan treated subgroup. Y=The logistic regression test $p < 0.05$ N= The logistic regression test $p \geq 0.05$. **B** Details about the effect described by the GVB score and the p-value obtained by logistic regression test for the selected genes.

(A)												
Gene	Significant results	APF	APF2	CADD phred	Mutation Assessor	PROVEAN	Polyphen2 HDIV	Polyphen2 HVAR	REVEL	SIFT		
Oxaliplatin												
MGMT	6	Y	Y	Y	N	Y	N	N	Y	Y		
SLCO1B1	5	N	Y	Y	N	N	N	Y	Y	Y		
ABCC3	5	N	N	Y	N	N	Y	Y	Y	Y		
Irinotecan												
<i>None</i>												
(B)												
Gene	APF	APF2	CADD phred	Mutation Assessor	PROVEAN	Polyphen2 HDIV	Polyphen2 HVAR	REVEL	SIFT			
Oxaliplatin	Effect	P*	Effect	P*	Effect	P*	Effect	P*	Effect	P*	Effect	P*
MGMT	risk	0.011	risk	0.025	risk	0.014	NA	NA	risk	0.012	risk	0.009
SLCO1B1	NS	NS	protective	0.012	protective	0.011	NS	NS	protective	0.009	protective	0.02
ABCC3	NS	NS	NS	0.038	NS	0.038	NS	NS	risk	0.03	risk	0.04

*P-values were estimated adjusting for sex, age, and fluoropyrimidine type and co-treatment (yes/no). Associations with $p < 0.05$ are evidenced in bold.

Table 5. Performance metrics for clinical, genetic, and combined predictive models of severe FL-induced toxicity.

Model	Sensitivity	Specificity	Accuracy	FPR	FNR	PPV	NNS	Unnecessary surveillance per TP	Optimal threshold
Clinical	0.76	0.46	0.62	0.54	0.24	0.61	1.65	0.65	0.472
Genetic	0.62	0.73	0.67	0.27	0.38	0.71	1.40	0.40	0.549
Combined	0.71	0.74	0.73	0.26	0.29	0.75	1.33	0.33	0.542

The key diagnostic measures for each model are evaluated at the optimal probability threshold determined by Youden's index. The number needed to surveil (NNS = 1/PPV), representing how many patients must be monitored to identify one true high-risk case, and the number of false positives per true positive (FP / TP), reflecting unnecessary surveillance burden.

FPR false positive rate, FNR false negative rate, PPV positive predicted value, NNS number needed to surveil, FP false positives, TP true positives.

Associations with $p < 0.05$ are evidenced in bold.

silico tools for the majority of rare variants identified in these key genes (Supplementary Table S5).

To investigate whether FL-toxicity was influenced by drugs co-administered with FL (i.e., oxaliplatin and irinotecan), an interaction analysis was performed. In the oxaliplatin treated group, three genes (*MGMT*, *SLCO1B1*, *ABCC3*) showed significant associations (Table 4). Specifically, the DNA methyltransferase *MGMT* and the ABC transporter *ABCC3* were shown to be associated with an increased risk of severe FL-induced toxicity, suggesting that variant carriers in these genes are more likely to develop toxicity when treated with oxaliplatin. Conversely, carriers of *SLCO1B1* variants had a lower risk of FL-toxicity under oxaliplatin co-treatment. Follow-up allele-specific interaction analyses of the reduced function alleles *SLCO1B1**5 and *SLCO1B1**15 abolished significance ($p > 0.35$) and recalculation of GVB score without these alleles abolished the protective interaction (all interaction p -values > 0.05). These results suggest that *5 and *15 are the primary drivers of the observed protective signal. No significant associations were identified in the subgroup analysis when stratifying by irinotecan treatment.

Integration of germline genetic burden into a predictive model for severe FL-induced toxicity

To enhance translational application, three predictive models were developed that considered clinical information, genetic data or both. For genetic data, the GVB scores for all significant results from the association analyses were used (i.e., *DPYD* along with *ABCB5*, *PARP1*, *ENOSF1*, *CYP3A4*, *HNF4A*, *ABCC4*, and *XRCC3* as well as ten genes from the Nuclear Receptors group). Performance of all models was evaluated using repeated stratified five-fold cross-validation. The genetic model exhibited comparable accuracy to the clinical model based solely on patient clinical characteristics (clinical model AUC = 0.6 ± 0.06 [mean $\pm$ SD] versus genetic model AUC = 0.62 ± 0.07). The combined clinical-genetic model, which integrated both clinical variables and genetic data, achieved the highest performance (AUC = 0.66 ± 0.066). At an optimal probability threshold of 0.542, the combined model achieved a sensitivity of 0.71, specificity of 0.74 and accuracy of 0.73 with false-positive and false-negative rates of 0.26 and 0.29, respectively (Table 5, Supplementary Fig. S4). The clinical model showed the highest sensitivity, reflecting that clinical risk factors can be identified in most individuals. In contrast, the genetic model, which is mostly composed of rare variants that are only identified in a subset of patients, exhibited lower sensitivity but higher specificity. By comparison, the clinical model had substantially lower specificity (0.46) and a much higher false-positive rate of 0.54. The positive predictive value of the combined model was 0.75, meaning that out of 100 patients flagged for increased genetic risk, 75 patients experienced severe FL-related toxicity. In contrast, only 61 patients would manifest toxicity per 100 identified patients when using clinical parameters only (PPV = 0.61).

To further evaluate the clinical utility of the combined model, we simulated a cohort of 10,000 patients receiving FL-based chemotherapy. Parameter estimates for toxicity prevalence, *DPYD*

variant carrier frequency, and the sensitivity and specificity of *DPYD*-only testing were derived from an aggregated literature-based dataset [37]. In the simulation, *DPYD* genotyping was applied to the full cohort, followed by application of our predictive combined model to *DPYD*-negative individuals. Predicted metrics were calculated for both *DPYD*-only testing and the combined strategy to enable direct comparison of performance and estimation of the additional number of at-risk patients identified by the integrated approach (Supplementary Table S6). Assuming a *DPYD* variant carrier prevalence of 1.45% and an overall toxicity prevalence of 25.2%, the simulation showed that *DPYD*-only testing would identify only 73 out of 2520 toxicity cases (2.9% sensitivity) but with high specificity (99.0%). NGS-based evaluation of non-carriers using the combined strategy identified an additional 1741 true positives (1814 toxicity cases in total; sensitivity 71%). However, specificity was reduced to 74%. This improvement may reflect the model's ability to integrate risk signals from a broader range of pharmacogenes, particularly rare variants not captured by standard *DPYD* genotyping. These results indicate that in patients without actionable *DPYD* variants who currently do not benefit from genetically guided dose adjustments, the consideration of genetic information can improve personalized risk stratification for severe FL-induced toxicity and that integration of clinical and genetic data yields the most efficient strategy to identify at-risk individuals.

DISCUSSION

Currently, genotyping for the four validated *DPYD* markers (*DPYD**2 A, *DPYD**13, c.2846 A > T, c.1236 G > A-HapB3) is the only validated test to preemptively identify patients at risk of developing severe FL-related toxicity. However, a significant number of severe toxicity events are missed by this targeted genotyping strategy. Evolving sequencing technologies now allow to obtain a comprehensive picture of a patient's genetic variability, facilitating the transition from candidate analyses to the interpretation of genetic signatures for the guidance FL-based therapy [1, 3, 22, 23].

The present study defines an analytical framework for the identification of genes that modulate FL toxicity risk based on germline variant burden and their integration in a clinical predictive model. The results suggest that other *DPYD* variants, beyond the four routinely tested in clinical practice, could contribute to refine the prediction of the risk of toxicity thus improving the selection of the optimal FL dosage. While statistical significance for individual rare and ultrarare variations is difficult to obtain using conventional single variant association tests, these results align with various clinical cases that indicate severe toxicity in rare variant carriers [38–44]. In addition, the analyses reveal that functional variants in *ABCB5*, DNA repair genes or in nuclear receptors may significantly modify the risk of developing FL-related toxicity at the patient level. While the effect sizes of these novel associations are lower than the four well-established *DPYD* variants (Supplementary Table S7) and the predictive performance

primarily driven by rare variants is moderate, the results suggest that the identified markers might be powerful complements to refine current *DPYD* genotyping strategies. However, larger, independent replication cohorts will be needed to refine effect size estimates. The predictive model, which combines genetic data with clinical information, significantly improves both sensitivity and specificity for the identification of at-risk patients, thereby opening possibilities for patients that do not benefit from conventional targeted *DPYD* genotyping. Importantly, this integrated clinical-genetic approach provides the most efficient surveillance strategy, requiring fewer patients to monitor per true at-risk case and minimizing unnecessary follow-up, highlighting its practical feasibility in the clinical setting. Overall, these findings demonstrated that integration of comprehensive sequencing and computational variant effect predictions can identify new FL risk genes and support the development of more personalized treatment strategies for severe FL-induced toxicity. Previous evidences demonstrated the value of developing predictive algorithms that incorporate the most relevant clinical factors [45] as well as the four routinely genotyped *DPYD* variants [46] to predict the high risk of life-threatening FL-related toxicity. However, at our knowledge this is the first study looking at the integration of clinical data with rare or novel variants markers in *DPYD* gene as well as in other FL-related pharmacogenes. Notably, for the presented analyses, only variations covered by our NGS panel were evaluated. As such, some regulatory variations, such as the *DPYD* rs4294451 enhancer variant, previously associated with an increased risk of severe FL-related toxicity [47], are not considered. The inclusion of such variants could potentially further enhance the predictive accuracy of the developed GVB model.

ABCB5 belongs to the most polymorphic members of the ABC transporter superfamily [48]. In the described cohort, patients carrying at least one nonsynonymous rare variant exhibited a significantly increased risk of developing FL toxicity compared to non-carriers. This association was also confirmed by considering both common and rare variants. The common missense *ABCB5* polymorphism rs6461515 (p.Glu970Lys) emerged as a significant predictor of severe FL-related toxicity in the GVB analysis. This finding aligns with previous preclinical studies in human colorectal cancer models that suggest that *ABCB5* mediates chemoresistance against 5-FU [49, 50]. Specifically, the *ABCB5* rs6461515 polymorphism impacts the substrates-interaction site at the protein level, potentially affecting the substrate specificity [51]. These results provide a first molecular underpinning to the identified association. Further in vitro studies, such as transporter activity assays, may be helpful to characterize the functional effect of the *ABCB5* rs6461515 polymorphism on FL efflux at the molecular level and thus elucidate the biological mechanism behind this genetic association. Similarly, the identification of *ABCC4* (MRP4), are in agreement with preclinical and clinical data showing that MRP4 is involved in modulating sensitivity to FLs and individual response to FL-based chemotherapy [52–54]. Thus, it can be hypothesized that deleterious variation in *ABCB5* or *ABCC4* may lead to reduced export of FL or its metabolites, leading to increased toxicity risk. Although further studies are needed to better understand the biological mechanism underlying this association, the burden of putatively functional variants in these genes may represent a good candidate marker for predicting FL-related toxicity and point out a major role of drug efflux in this clinical phenotype.

An additional gene related to the expression level of the major FL pharmacodynamic target thymidylate synthase (TS, *TYMS*) is enolase superfamily member 1 (*ENOSF1*), which encodes a mitochondrial enzyme also known as reverse TS (rTS). The *ENOSF1* gene locus is adjacent to *TYMS* and a regulatory interaction between the two genes has been suggested. High expression levels of *ENOSF1* have been found to suppress TS expression, either by inducing the production of an antisense RNA targeting *TYMS* or by inducing the expression of a protein that modulates TS protein

translation [55, 56]. Hence, alterations in *ENOSF1* might impact the pharmacodynamic effects of FLs either directly or indirectly.

DNA repair mechanisms have been demonstrated to play a crucial role in determining the sensitivity to FLs [57, 58]. In vitro, ex vivo, and in vivo experimental results have indicated that PARP1 may modulate the sensitivity to 5-FU through the regulation of the mismatch repair (MMR) pathway [59–61]. Additionally, GVB highlighted an enrichment in genetic polymorphisms in other genes involved in DNA damage repair, comprising both the mismatch repair and homologous recombination pathways. *XRCC3* is involved in the repair of 5-FU-induced DNA damages [57] and a specific *XRCC3* missense variant (rs861539, p.Thr241Met) [62, 63] emerged as protective marker in our analyses. While the functional effect of this polymorphism is still unclear, previous reports suggested associations with treatment outcomes and survival upon FLs therapies in various solid malignancies [64, 65].

Nuclear receptors encode for a family of transcriptional factors that can be induced by endo or xenobiotic molecules and modulate the expression of phase I and II drug metabolizing enzymes and membrane transporters [66–69]. Genetic variants in these genes were previously suggested to impact the bioavailability of anti-cancer drugs, including FLs [66–68], and thus represent potential candidates as pharmacogenetic marker. Among the nuclear receptors, the analysis presented here identified the strongest signal for HNF4A, which emerged as a significant protective marker against FL-related toxicity. This finding is supported by data from the literature, as multiple papers, including a meta-analysis combining the analysis of several microarray studies, have identified HNF4A as a hub gene that plays a key role in the regulation of responsiveness to FL-containing chemotherapy, probably because of its strong regulation of genes encoding the 5-FU membrane transporters [70, 71]. The results of the present work thus support the predictive value of rare *HNF4A* variations in the clinical setting. Besides HNF4A, experimental in vitro and ex vivo evidence also implicates other members of the nuclear receptor family, such as PXR (*NR1I2*), GR (*NR3C1*) and PPARG in the modulation of 5-FU sensitivity [72–74]. Mechanistically, these receptors can increase cellular drug efflux by inducing P-gp and MRP3, or they can act via controlling the expression of genes related to 5-FU pharmacoresistance including pro- and anti-apoptotic proteins.

Interaction analysis suggested that variants in *MGMT*, *ABCC3* and *SLCO1B1* have a more significant effect on the subgroup of patients receiving oxaliplatin cotreatment. MGMT, or O6-methylguanine-DNA methyltransferase, is a DNA repair enzyme that plays a crucial role in protecting cells from the damaging effects of chemotherapeutic drugs such as alkylating agents [75]. Although literature data have not directly associated MGMT with oxaliplatin cytotoxicity, DNA repair mechanisms play a crucial role in the activity of platinum agents [55, 56, 76], and thus support a potential interaction between MGMT and oxaliplatin [77]. Preliminary data have also suggested an involvement of the transporters encoded by *ABCC3* and *SLCO1B1* in membrane translocation of oxaliplatin, potentially modulating its bioavailability and toxicity profile [78–80]. Although preliminary, these data indicate that integrating FL-specific markers with additional genetic predictors related to the co-treatment may yield a more comprehensive assessment of an individual's risk of developing toxicity.

The disposition and cytotoxicity of FLs is the result of several interconnected biological mechanisms that interact and converge [56, 76, 81, 82]. It could be hypothesized that nuclear receptors, such as HNF4A, alter the expression of membrane transporters and phase I enzymes [66–68], thereby impacting FL pharmacokinetics and intracellular FL concentrations. The resulting altered exposure can impact FL effects on its target thymidylate synthase (TS) and the related enzyme system, thereby providing a plausible model that unifies effects in pharmacokinetic and pharmacodynamic genes [55, 81]. In addition, FL cytotoxicity can be

modulated by DNA damage repair proteins [58]. Notably, we identify FL sensitivity signatures in both pharmacokinetic (*ABCB5*, *ABCC4* and *CYP3A4*), pharmacodynamic (*ENOSF1*) and DNA damage repair pathways (*PARP1* and *XRCC3*), thus providing genetic evidence that corroborates the complexity and interaction of the different pathways in modulating FL toxicity risk.

The overall effect of these non-*DPYD* genes is promising but still limited as compared to current *DPYD* genotyping strategies, and further validation is needed to elucidate their real predictive value and the potential strategies for their prospective application to guide FLs precision dosing in the clinical practice. Furthermore, it will be important to determine the extent of dose adjustment required in patients with pharmacogenetic risk signatures beyond *DPYD* and whether the clinical benefit of upfront dose reductions in at-risk patients would outweigh unnecessary dose-reductions in false-positives. Although this analysis was focused on the identification of toxicity markers, it was already previously demonstrated how genetic signatures related to drug pharmacokinetics may affect treatment efficacy by impairing treatment exposure and adherence [12, 83–86]. It is worthwhile considering a future evaluation of what role the germline pharmacogenomic burden may have on the tumor response and survival in patients treated with FLs.

Combined, these data indicate that comprehensive sequencing of candidate genes, including but not limited to *DPYD*, can provide clinically useful additional information that is not captured by targeted genotyping approaches. However, challenges remain for the widespread implementation of such NGS-based strategies. These include higher turnaround times, higher costs, a lack of standardization with regards to the functional interpretation of rare and novel variations and a paucity of formal cost-effectiveness evaluations, which are required by most national health care providers to consider the reimbursement of such test costs [87]. Currently, efforts by several international working groups aim to address these limitations and pave the way for the integration of rare variant information into pharmacogenomic guidance documents [88]. These efforts are fertilized by the increasingly widespread use of NGS technology and reduced analytical costs, which jointly facilitate the adoption of such targeted sequencing approaches in routine clinical practice.

Some limitations of the present study need to be considered. First, the study included patients treated mainly with FL in combination with oxaliplatin and irinotecan. An interaction analysis was performed to account for comedication confounding, but the low patient number in the different strata did not allow for conclusions with sufficient statistical power. Similarly, the cohort size did not allow for an evaluation of the effect of genetic profiles on severe grade 4–5 toxicities [47]. An additional limitation is the lack of replication in an external cohort. While we aimed to mitigate this concern through rigorous and extensive internal cross-validation, validation in independent cohorts, ideally comprising also other ethnogeographic groups to increase the generalizability of our findings [89, 90], remains to be of high importance. With few exceptions, the rare variants identified in our study showed limited inter-ethnic variability (Supplementary Table S5); yet, external validation in multiethnic cohorts remains an important objective for future work. We anticipate that the predictive model could be further improved by incorporation of additional clinical parameters known to influence FL toxicity risk, such as kidney and liver function [45]. However, these data were unfortunately not available for the analyzed patient population.

In conclusion, the presented results underscore the potential of comprehensive pharmacogenomic strategies coupled to variant burden analyses to highlight genes that contribute to FL toxicity risk. Specifically, rare variants in ABC transporters (*ABCB5*, *ABCC4*), nuclear receptors (especially *HNF4A*), DNA repair enzymes (*PARP1*, *XRCC3*), regulators of TYMS activity (*ENOSF1*) and cytochrome P450s (*CYP3A4*), emerged as significant predictors that could improve patient stratification in addition to *DPYD*. The present

study thus paves the way for a paradigm shift from population-level screening of candidate variants towards a truly individualized pharmacogenetic approach of FL therapy that comprehensively considers genomic variability at the single patient level. Current clinical algorithms focus on four well-characterized *DPYD* variants - *DPYD**2 A, *DPYD**13, c.2846 A > T, c.1236 G > A-c.1129-5923 C > G [HapB3]. Our data, however, demonstrate that this population-based prevention concept leaves carriers of rare or novel deleterious alleles unaccounted for, exposing them to a persistent risk of severe toxicity. Additional studies are probably needed to validate a strategy for a prospective clinical application of novel and rare deleterious variants in FL risk genes, other than *DPYD*. On the other hand, due to the well-established association between DPD deficiency and FL safety, a comprehensive definition of the variant profiles in *DPYD* by NGS before treatment could be already envisioned as a promising strategy to increase the pool of at-risk patients that can be flagged for in-depth surveillance or dose adjustment, maximizing the sensitivity of risk prediction.

DATA AVAILABILITY

The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

REFERENCES

- De Mattia E, Milan N, Assaraf YG, Toffoli G, Cecchin E. Clinical Implementation of Rare and Novel *DPYD* Variants for Personalizing Fluoropyrimidine Treatment: Challenges and Opportunities. *Int J Biol Sci*. 2024;20:3742–59.
- Lauschke VM, Zhou Y, Ingelman-Sundberg M. Pharmacogenomics Beyond Single Common Genetic Variants: The Way Forward. *Annu Rev Pharm Toxicol*. 2024;64:33–51.
- Peruzzi E, Roncato R, De Mattia E, Bignucolo A, Swen JJ, Guchelaar HJ, et al. Implementation of pre-emptive testing of a pharmacogenomic panel in clinical practice: Where do we stand?. *Br J Clin Pharm*. 2025;91:270–82.
- Zhou Y, Tremmel R, Schaeffeler E, Schwab M, Lauschke VM. Challenges and opportunities associated with rare-variant pharmacogenomics. *Trends Pharm Sci*. 2022;43:852–65.
- Kozyra M, Ingelman-Sundberg M, Lauschke VM. Rare genetic variants in cellular transporters, metabolic enzymes, and nuclear receptors can be important determinants of interindividual differences in drug response. *Genet Med*. 2017;19:20–9.
- Zhou Y, Dagli Hernandez C, Lauschke VM. Population-scale predictions of DPD and TPMT phenotypes using a quantitative pharmacogene-specific ensemble classifier. *Br J Cancer*. 2020;123:1782–9.
- Knikman JE, Gelderblom H, Beijnen JH, Cats A, Guchelaar HJ, Henricks LM. Individualized Dosing of Fluoropyrimidine-Based Chemotherapy to Prevent Severe Fluoropyrimidine-Related Toxicity: What Are the Options?. *Clin Pharm Ther*. 2021;109:591–604.
- Lunenborg C, Henricks LM, Guchelaar HJ, Swen JJ, Deenen MJ, Schellens JHM, et al. Prospective *DPYD* genotyping to reduce the risk of fluoropyrimidine-induced severe toxicity: Ready for prime time. *Eur J Cancer*. 2016;54:40–8.
- Dalle Fratte C, Polesel J, Roncato R, De Mattia E, Ecça F, Bignucolo A, et al. *DPYD* Gene Activity Score Predicts Dose-Limiting Toxicity in Fluoropyrimidine-Treated Colorectal Cancer Patients. *J Mol Clin Med*. 2018;1:143–9.
- Fragoulakis V, Roncato R, Bignucolo A, Patrinos GP, Toffoli G, Cecchin E, et al. Cost-utility analysis and cross-country comparison of pharmacogenomics-guided treatment in colorectal cancer patients participating in the U-PGx PREPARE study. *Pharm Res*. 2023;197:106949.
- Fragoulakis V, Roncato R, Fratte C, Ecça F, Bartsakoulia M, Innocenti F, et al. Estimating the Effectiveness of *DPYD* Genotyping in Italian Individuals Suffering from Cancer Based on the Cost of Chemotherapy-Induced Toxicity. *Am J Hum Genet*. 2019;104:1158–68.
- Roncato R, Bignucolo A, Peruzzi E, Montico M, De Mattia E, Foltran L, et al. Clinical Benefits and Utility of Pretherapeutic *DPYD* and *UGT1A1* Testing in Gastrointestinal Cancer: A Secondary Analysis of the PREPARE Randomized Clinical Trial. *JAMA Netw Open*. 2024;7:e2449441.
- Toffoli G, Giodini L, Buonadonna A, Berretta M, De Paoli A, Scalone S, et al. Clinical validity of a *DPYD*-based pharmacogenetic test to predict severe toxicity to fluoropyrimidines. *Int J Cancer*. 2015;137:2971–80.
- Toffoli G, Innocenti F, Polesel J, De Mattia E, Sartor F, Dalle Fratte C, et al. The Genotype for *DPYD* Risk Variants in Patients With Colorectal Cancer and the Related Toxicity Management Costs in Clinical Practice. *Clin Pharm Ther*. 2019;105:994–1002.

15. Amstutz U, Henricks LM, Offer SM, Barbarino J, Schellens JHM, Swen JJ, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Dihydropyrimidine Dehydrogenase Genotype and Fluoropyrimidine Dosing: 2017 Update. *Clin Pharm Ther.* 2018;103:210–6.
16. Bignucolo A, De Mattia E, Roncato R, Peruzzi E, Scarabel L, D'Andrea M, et al. Ten-year experience with pharmacogenetic testing for DPYD in a national cancer center in Italy: Lessons learned on the path to implementation. *Front Pharm.* 2023;14:1199462.
17. de With M, Sadlon A, Cecchin E, Haufroid V, Thomas F, Joerger M, et al. Implementation of dihydropyrimidine dehydrogenase deficiency testing in Europe. *ESMO Open.* 2023;8:101197.
18. Lunenburg C, van der Wouden CH, Nijenhuis M, Crommentuijn-van Rhenen MH, de Boer-Veger NJ, Buunk AM, et al. Dutch Pharmacogenetics Working Group (DPWG) guideline for the gene-drug interaction of DPYD and fluoropyrimidines. *Eur J Hum Genet.* 2020;28:508–17.
19. Lopes JL, Harris K, Karow MB, Peterson SE, Kluge ML, Kotzer KE, et al. Targeted Genotyping in Clinical Pharmacogenomics: What Is Missing?. *J Mol Diagn.* 2022;24:253–61.
20. Larrue R, Fellah S, Hennart B, Sabaouni N, Boukrout N, Van der Hauwaert C, et al. Integrating rare genetic variants into DPYD pharmacogenetic testing may help preventing fluoropyrimidine-induced toxicity. *Pharmacogenomics J.* 2024;24:1.
21. Knikman JE, Zhai Q, Lunenburg C, Henricks LM, Bohringer S, van der Lee M, et al. Discovering novel germline genetic variants linked to severe fluoropyrimidine-related toxicity in- and outside DPYD. *Genome Med.* 2024;16:101.
22. De Mattia E, Silvestri M, Polesel J, Ecça F, Mezzalana S, Scarabel L, et al. Rare genetic variant burden in DPYD predicts severe fluoropyrimidine-related toxicity risk. *Biomed Pharmacother.* 2022;154:113644.
23. De Mattia E, Polesel J, Silvestri M, Roncato R, Scarabel L, Calza S, et al. The burden of rare variants in DPYD gene is a novel predictor of the risk of developing severe fluoropyrimidine-related toxicity. *Hum Genom.* 2023;17:99.
24. Swen JJ, van der Wouden CH, Manson LE, Abdullah-Koolmees H, Blagec K, Blagus T, et al. A 12-gene pharmacogenetic panel to prevent adverse drug reactions: an open-label, multicentre, controlled, cluster-randomised crossover implementation study. *Lancet.* 2023;401:347–56.
25. van der Wouden CH, Cambon-Thomsen A, Cecchin E, Cheung KC, Davila-Fajardo CL, Deneer VH, et al. Implementing Pharmacogenomics in Europe: Design and Implementation Strategy of the Ubiquitous Pharmacogenomics Consortium. *Clin Pharm Ther.* 2017;101:341–58.
26. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:13033997v2 [q-bio.GN].* 2013.
27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25:2078–9.
28. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20:1297–303.
29. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w11180; iso-2; iso-3. *Fly.* 2012;6:80–92.
30. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38:e164.
31. Zhou Y, Mkrchian S, Kumondai M, Hiratsuka M, Lauschke VM. An optimized prediction framework to assess the functional impact of pharmacogenetic variants. *Pharmacogenomics J.* 2019;19:115–26.
32. Zhou Y, Pirmann S, Lauschke VM. APF2: an improved ensemble method for pharmacogenomic variant effect prediction. *Pharmacogenomics J.* 2024;24:17.
33. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet.* 2014;95:5–23.
34. Park Y, Kim H, Choi JY, Yun S, Min BJ, Seo ME, et al. Star Allele-Based Haplotyping versus Gene-Wise Burden Scoring for Predicting 6-Mercaptopurine Intolerance in Pediatric Acute Lymphoblastic Leukemia Patients. *Front Pharm.* 2019;10:654.
35. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. *J Mach Learn Res.* 2011;12:2825–30.
36. Gaedigk A, Casey ST, Whirl-Carrillo M, Miller NA, Klein TE. Pharmacogene Variation Consortium: A Global Resource and Repository for Pharmacogene Variation. *Clin Pharm Ther.* 2021;110:542–5.
37. Meulendijks D, Henricks LM, Sonke GS, Deenen MJ, Froehlich TK, Amstutz U, et al. Clinical relevance of DPYD variants c.1679T>G, c.1236G>A/HapB3, and c.1601G>A as predictors of severe fluoropyrimidine-associated toxicity: a systematic review and meta-analysis of individual patient data. *Lancet Oncol.* 2015;16:1639–50.
38. Bukhari N, Azam F, Alfawaz M, Zahrani M. Identifying a Novel DPYD Polymorphism Associated with Severe Toxicity to 5-FU Chemotherapy in a Saudi Patient. *Case Rep Genet.* 2019;2019:5150725.
39. Del Re M, Michelucci A, Di Leo A, Cantore M, Bordonaro R, Simi P, et al. Discovery of novel mutations in the dihydropyrimidine dehydrogenase gene associated with toxicity of fluoropyrimidines and viewpoint on preemptive pharmacogenetic screening in patients. *EPMA J.* 2015;6:17.
40. Garcia-Gonzalez X, Lopez-Tarruella S, Garcia MI, Gonzalez-Haba E, Blanco C, Salvador-Martin S, et al. Severe toxicity to capecitabine due to a new variant at a donor splicing site in the dihydropyrimidine dehydrogenase (DPYD) gene. *Cancer Manag Res.* 2018;10:4517–22.
41. Henricks LM, Siemerink EJM, Rosing H, Meijer J, Goorden SMI, Polstra AM, et al. Capecitabine-based treatment of a patient with a novel DPYD genotype and complete dihydropyrimidine dehydrogenase deficiency. *Int J Cancer.* 2018;142:424–30.
42. Lazar A, Mau-Holzmann UA, Kolb H, Reichenmiller HE, Riess O, Schomig E. Multiple organ failure due to 5-fluorouracil chemotherapy in a patient with a rare dihydropyrimidine dehydrogenase gene variant. *Onkologie.* 2004;27:559–62.
43. Ly RC, Schmidt RE, Kiel PJ, Pratt VM, Schneider BP, Radovich M, et al. Severe Capecitabine Toxicity Associated With a Rare DPYD Variant Identified Through Whole-Genome Sequencing. *JCO Precis Oncol.* 2020;4:632–8.
44. Vilquin P, Medard Y, Thomas F, Goldwirth L, Teixeira L, Mourah S, et al. DPYD genotype should be extended to rare variants: report on two cases of phenotype / genotype discrepancy. *Cancer Chemother Pharm.* 2025;95:16.
45. Knikman JE, Lopez-Yurda M, Meulendijks D, Deenen MJ, Schellens JHM, Beijnen J, et al. A Nomogram to Predict Severe Toxicity in DPYD Wild-Type Patients Treated With Capecitabine-Based Anticancer Regimens. *Clin Pharm Ther.* 2024;115: 269–77.
46. Le Teuff G, Cozic N, Boyer JC, Boige V, Diasio RB, Taieb J, et al. Dihydropyrimidine dehydrogenase gene variants for predicting grade 4–5 fluoropyrimidine-induced toxicity: FUSAFE individual patient data meta-analysis. *Br J Cancer.* 2024;130:808–18.
47. De Mattia E, Polesel J, Scarabel L, Cecchin E. Clinical implications of a gain-of-function genetic polymorphism in DPYD (rs4294451) in colorectal cancer patients treated with fluoropyrimidines. *Front Pharm.* 2024;15:1516375.
48. Xiao Q, Zhou Y, Lauschke VM. Ethnogeographic and inter-individual variability of human ABC transporters. *Hum Genet.* 2020;139:623–46.
49. Kugimiya N, Nishimoto A, Hosoyama T, Ueno K, Enoki T, Li TS, et al. The c-MYC-ABCB5 axis plays a pivotal role in 5-fluorouracil resistance in human colon cancer cells. *J Cell Mol Med.* 2015;19:1569–81.
50. Wilson BJ, Schatton T, Zhan Q, Gasser M, Ma J, Saab KR, et al. ABCB5 identifies a therapy-refractory tumor cell population in colorectal cancer patients. *Cancer Res.* 2011;71:5307–16.
51. Moitra K, Scally M, McGee K, Lancaster G, Gold B, Dean M. Molecular evolutionary analysis of ABCB5: the ancestral gene is a full transporter with potentially deleterious single nucleotide polymorphisms. *PLoS ONE.* 2011;6:e16318.
52. Chen Q, Meng F, Wang L, Mao Y, Zhou H, Hua D, et al. A polymorphism in ABCC4 is related to efficacy of 5-FU/capecitabine-based chemotherapy in colorectal cancer patients. *Sci Rep.* 2017;7:7059.
53. Hagmann W, Jesnowski R, Faissner R, Guo C, Lohr JM. ATP-binding cassette C transporters in human pancreatic carcinoma cell lines. Upregulation 5-fluorouracil-resistant Cells *Pancreatol.* 2009;9:136–44.
54. Zhang G, Wang Z, Qian F, Zhao C, Sun C. Silencing of the ABCC4 gene by RNA interference reverses multidrug resistance in human gastric cancer. *Oncol Rep.* 2015;33:1147–54.
55. Cecchin E, De Mattia E, Ecça F, Toffoli G. Host genetic profiling to increase drug safety in colorectal cancer from discovery to implementation. *Drug Resist Updat.* 2018;39:18–40.
56. De Mattia E, Cecchin E, Toffoli G. Pharmacogenomics of intrinsic and acquired pharmacoresistance in colorectal cancer: Toward targeted personalized therapy. *Drug Resist Updat.* 2015;20:39–70.
57. Sethy C, Kundu CN. 5-Fluorouracil (5-FU) resistance and the new strategy to enhance the sensitivity against cancer: Implication of DNA repair inhibition. *Biomed Pharmacother.* 2021;137:111285.
58. Wyatt MD, Wilson DM. 3rd. Participation of DNA repair in the response to 5-fluorouracil. *Cell Mol Life Sci.* 2009;66:788–99.
59. Mini E, Landini I, Lucarini L, Lapucci A, Napoli C, Perrone G, et al. The Inhibitory Effects of HYDAMTIQ, a Novel PARP Inhibitor, on Growth in Human Tumor Cell Lines With Defective DNA Damage Response Pathways. *Oncol Res.* 2017;25:1441–51.
60. Paul S, Chatterjee S, Sinha S, Dash SR, Pradhan R, Das B, et al. Veliparib (ABT-888), a PARP inhibitor potentiates the cytotoxic activity of 5-fluorouracil by inhibiting MMR pathway through deregulation of MSH6 in colorectal cancer stem cells. *Expert Opin Ther Targets.* 2023;27:999–1015.
61. Wang F, Gouttiau OG, Wang L, Peng A. PARP1 Upregulation in Recurrent Oral Cancer and Treatment Resistance. *Front Cell Dev Biol.* 2021;9:804962.
62. Du L, Xiong T, He Q, Wang Y, Shen J, Peng Y, et al. The Thr241Met polymorphism in the XRCC3 gene is associated with increased risk of cancer in Chinese mainland populations. *Tumour Biol.* 2014;35:1371–6.
63. Qin LY, Chen X, Li P, Yang Z, Mo WN. Association between the XRCC3 Thr241Met polymorphism and cervical cancer risk: a meta-analysis. *Asian Pac J Cancer Prev.* 2014;14:6703–7.

64. Liu Y, Chen H, Chen L, Hu C. Prediction of genetic polymorphisms of DNA repair genes XRCC1 and XRCC3 in the survival of colorectal cancer receiving chemotherapy in the Chinese population. *Hepatogastroenterology*. 2012;59:977–80.
65. Ott K, Rachakonda PS, Panzram B, Keller G, Lordick F, Becker K, et al. DNA repair gene and MTHFR gene polymorphisms as prognostic markers in locally advanced adenocarcinoma of the esophagus or stomach treated with cisplatin and 5-fluorouracil-based neoadjuvant chemotherapy. *Ann Surg Oncol*. 2011;18:2688–98.
66. Cecchin E, De Mattia E, Toffoli G. Nuclear receptors and drug metabolism for the personalization of cancer therapy. *Expert Opin Drug Metab Toxicol*. 2016;12:291–306.
67. De Mattia E, Cecchin E, Roncato R, Toffoli G. Pregnane X receptor, constitutive androstane receptor and hepatocyte nuclear factors as emerging players in cancer precision medicine. *Pharmacogenomics*. 2016;17:1547–71.
68. De Mattia E, Dreussi E, Cecchin E, Toffoli G. Pharmacogenetics of the nuclear hormone receptors: the missing link between environment and drug effects?. *Pharmacogenomics*. 2013;14:2035–54.
69. Weikum ER, Liu X, Ortlund EA. The nuclear receptor superfamily: A structural perspective. *Protein Sci*. 2018;27:1876–92.
70. Abu N, Othman N, Nazarie KWH, Jamal WF. R. Integrative meta-analysis for the identification of hub genes in chemoresistant colorectal cancer. *Biomark Med*. 2020;14:525–37.
71. Hagos Y, Wegner W, Kuehne A, Floerl S, Marada VV, Burckhardt G, et al. HNF4alpha induced chemosensitivity to oxaliplatin and 5-FU mediated by OCT1 and CNT3 in renal cell carcinoma. *J Pharm Sci*. 2014;103:3326–34.
72. Jiang H, Chen K, He J, Pan F, Li J, Chen J, et al. Association of pregnane X receptor with multidrug resistance-related protein 3 and its role in human colon cancer chemoresistance. *J Gastrointest Surg*. 2009;13:1831–8.
73. Schockel L, Woischke C, Surendran SA, Michl M, Schiergens T, Holscher A, et al. PPARγ activation promotes the proliferation of colorectal cancer cell lines and enhances the antiproliferative effect of 5-fluorouracil. *BMC Cancer*. 2024;24:234.
74. Yu J, Chen M, Sang Q, Li F, Xu Z, Yu B, et al. Super-enhancer Activates Master Transcription Factor NR3C1 Expression and Promotes 5-FU Resistance in Gastric Cancer. *Adv Sci*. 2025;12:e2409050.
75. Chen L, Wen A. Unveiling the role of O(6)-methylguanine-DNA methyltransferase in cancer therapy: insights into alkylators, pharmacogenomics, and others. *Front Oncol*. 2024;14:1424797.
76. Bignucolo A, Scarabel L, Toffoli G, Cecchin E, De Mattia E. Predicting drug response and toxicity in metastatic colorectal cancer: the role of germline markers. *Expert Rev Clin Pharm*. 2022;15:689–713.
77. Park JH, Kim NS, Park JY, Chae YS, Kim JG, Sohn SK, et al. MGMT -535G>T polymorphism is associated with prognosis for patients with metastatic colorectal cancer treated with oxaliplatin-based chemotherapy. *J Cancer Res Clin Oncol*. 2010;136:1135–42.
78. Dong Y, Wang Z, Xie GF, Li C, Zuo WW, Meng G, et al. Pregnane X receptor is associated with unfavorable survival and induces chemotherapeutic resistance by transcriptional activating multidrug resistance-related protein 3 in colorectal cancer. *Mol Cancer*. 2017;16:71.
79. Lancaster CS, Sprowl JA, Walker AL, Hu S, Gibson AA, Sparreboom A. Modulation of OATP1B-type transporter function alters cellular uptake and disposition of platinum chemotherapeutics. *Mol Cancer Ther*. 2013;12:1537–44.
80. Theile D, Grebhardt S, Haefeli WE, Weiss J. Involvement of drug transporters in the synergistic action of FOLFOX combination chemotherapy. *Biochem Pharm*. 2009;78:1366–73.
81. De Mattia E, Roncato R, Dalle Fratte C, Ecça F, Toffoli G, Cecchin E. The use of pharmacogenetics to increase the safety of colorectal cancer patients treated with fluoropyrimidines. *Cancer Drug Resist*. 2019;2:116–30.
82. Maslarinou A, Manolopoulos VG, Ragia G. Pharmacogenomic-guided dosing of fluoropyrimidines beyond DPYD: time for a polygenic algorithm?. *Front Pharm*. 2023;14:1184523.
83. De Mattia E, Cecchin E, Montico M, Labriet A, Guillemette C, Dreussi E, et al. Association of STAT-3 rs1053004 and VDR rs11574077 With FOLFIRI-Related Gastrointestinal Toxicity in Metastatic Colorectal Cancer Patients. *Front Pharm*. 2018;9:367.
84. De Mattia E, Polesel J, Roncato R, Labriet A, Bignucolo A, Dreussi E, et al. Germline Polymorphisms in the Nuclear Receptors PXR and VDR as Novel Prognostic Markers in Metastatic Colorectal Cancer Patients Treated With FOLFIRI. *Front Oncol*. 2019;9:1312.
85. De Mattia E, Toffoli G, Polesel J, D'Andrea M, Corona G, Zagonel V, et al. Pharmacogenetics of ABC and SLC transporters in metastatic colorectal cancer patients receiving first-line FOLFIRI treatment. *Pharmacogenet Genomics*. 2013;23:549–57.
86. Labriet A, De Mattia E, Cecchin E, Levesque E, Jonker D, Couture F, et al. Improved Progression-Free Survival in Irinotecan-Treated Metastatic Colorectal Cancer Patients Carrying the HNF1A Coding Variant p.I27L. *Front Pharm*. 2017;8:712.
87. Russell LE, Zhou Y, Almousa AA, Sodhi JK, Nwabuo CK, Lauschke VM. Pharmacogenomics in the era of next generation sequencing - from byte to bedside. *Drug Metab Rev*. 2021;53:253–78.
88. Gong L, Klein CJ, Caudle KE, Moyer AM, Scott SA, Whirl-Carrillo M, et al. Integrating Pharmacogenomics into the Broader Construct of Genomic Medicine: Efforts by the ClinGen Pharmacogenomics Working Group (PGxWG). *Clin Chem*. 2025;71:36–44.
89. Zhou Y, Lauschke VM. Population pharmacogenomics: an update on ethnogeographic differences and opportunities for precision public health. *Hum Genet*. 2022;141:1113–36.
90. Zhou Y, Park Y, Camara MD, Lauschke VM. Opportunities and Challenges of Population Pharmacogenomics. *Ann Hum Genet*. 2025;89:384–97.

AUTHOR CONTRIBUTIONS

EDM, EC and VML conceptualized and designed the study. MS and FP provided and collected clinical samples and recorded clinical data. EDM directed and supervised the execution of sequencing experiments. EP, RR, and LS carried out sample processing and sequencing experiments. JP created the dataset by selecting matched cases and controls. YP and YZ carried out bio-informatic data processing and computational analysis. EDM, YP, VML and EC carried out interpretation of data. EDM and YP wrote the manuscript. EC, VML, GT, MS, HJG and JJS edited and revised the manuscript. EC and VML supervised the overall study. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Italian Ministry of Health (Ricerca Corrente) and European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 668353. We also acknowledge funding from the Swedish Research Council [grant numbers: 2021-02801 and 2023-03015], by the SciLifeLab and Wallenberg National Program for Data-Driven Life Science [WASPDDL22:006], Cancerfonden [23-0763PT and 24-3735PJ], the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2180—390900677, the Robert Bosch Foundation, Stuttgart, Germany, and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2023-00209528).

COMPETING INTERESTS

VML is co-founder, CEO and shareholder of HepaPredict AB, as well as co-founder and shareholder of Shanghai Hepo Biotechnology Ltd. All other authors declare that they have no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study population was selected from the prospective PREPARE trial (NCT03093818, ClinicalTrials.gov). The PREPARE study adhered to the principles outlined in the 1975 Declaration of Helsinki (with the 1983 revision) and was granted ethical approval by the local ethic committee. All patients provided written informed consent before entering the study to donate their anonymized blood samples for further studies including NGS analysis. All experiments were carried out in accordance with the relevant guidelines and regulations of Centro di Riferimento Oncologico di Aviano.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41388-025-03587-7>.

Correspondence and requests for materials should be addressed to Volker M. Lauschke or Erika Cecchin.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.