Multiscale mapping of transcriptomic signatures for cardiotoxic drugs

Jens Hansen,[1,2] Yuguang Xiong,[1,2] Priyanka Dhanan,[1,2] Bin Hu,[1,2] Arjun S. Yadaw,[1,2] Gomathi Jayaraman,[1,2] Rosa Tolentino,[1,2] Yibang Chen,[1,2] Kristin G. Beaumont,[3] Robert Sebra,[3] Dusica Vidovic,[4] Stephan C. Schürer,[4] Joseph Goldfarb,[1, 2] James Gallo,[1,2,5] Marc R. Birtwistle,[1,2,6] Eric A. Sobie,[1,2] Evren U. Azeloglu,[1,2,7] Seth Berger,[8] Angel Chan,[1,2,9] Christoph Schaniel,[1,10] Nicole C. Dubois, [1,11] Ravi Iyengar[1,2]

1 Mount Sinai Institute for Systems Biomedicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
2 Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
3 Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
4 Institute for Data Science and Computing, University of Miami, Coral Gables, FL 33146, USA
5 School of Pharmacy and Pharmaceutical Sciences, University of Buffalo SUNY System, Buffalo NY 14260
6 Chemical and Biomolecular Engineering, Clemson University, Clemson, SC, 29634, USA
7 Department of Medicine, Division of Nephrology, Icahn School of Medicine at Mount Sinai, New, York, NY 10029, USA
8 Departments of Pediatrics and Genomics & Precision Medicine, George Washington University School of medicine and Health Sciences, Washington DC USA, 20012
9 Cardiology Division, Department of Medicine, Memorial Sloan Kettering Cancer Center New York NY 10065
10 Department of Medicine, Division of Hematology and Medical Oncology, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
11 Department of Cell, Developmental and Regenerative Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

Address correspondence to

Jens Hansen – jens.hansen@mssm.edu
or
Ravi Iyengar - ravi.iyengar@mssm.edu

Abstract

Drug-induced gene expression profiles can identify potential mechanisms of toxicity. We focused on obtaining signatures for cardiotoxicity of FDA-approved tyrosine kinase inhibitors (TKIs) in human induced pluripotent stem cell-derived cardiomyocytes. Using bulk transcriptomics profiles, we applied singular value decomposition to identify drug-selective patterns in cell lines obtained from multiple healthy human subjects. Cellular pathways affected by highly cardiotoxic TKIs include energy metabolism, contractile, and extracellular matrix dynamics. Projecting these pathways to single cell expression profiles indicates that TKI responses can be evoked in both cardiomyocytes and fibroblasts. Whole genome sequences of the cell lines, using outlier responses enabled us to correctly reidentify a genomic variant associated with anthracycline cardiotoxicity and predict genomic variants potentially associated with TKI cardiotoxicity. We conclude that mRNA expression profiles when integrated with publicly available genomic, pathway, and single cell transcriptomic datasets, provide multiscale predictive understanding of cardiotoxicity for drug development and patient stratification.

One sentence summary:

Genes, pathways, and cell types of the human heart associated with antineoplastic drug cardiotoxicity.

Introduction

Adverse side-effects of therapeutically useful drugs continue to be a substantial problem [1], post-approval pharmacovigilance studies often revealing adverse events that lead to warning labels mandated by the FDA [2]. Early indications of a potential for adverse events will be useful in drug development [3,4] and the use of whole genome sequence data from individuals can predict who might be susceptible to adverse events. These assertions are based on the premise that various functions at the molecular and cellular levels drive adverse events in different cell types [5]. Preclinical studies at the molecular level are useful, as has been demonstrated by HERG channel protein interacting drugs and the potential for arrhythmias [6,7]. Drug-related adverse events are often organ-selective. Many efficacious antineoplastic drugs, such as tyrosine kinase inhibitors (TKI) that are used for targeted therapy, are associated with cardiac insufficiencies and development of heart failure [8,9]. We have shown that drug-induced transcriptomic profiles in adult human heart cells can be associated with clinical adverse event propensity as assessed from FDA pharmacovigilance data [10]. However, a systematic understanding of the molecular pathways, cell origins and genomic determinants associated with drug therapy-related cardiotoxicity are still not well understood.

Human induced pluripotent stem cell (hiPSC)-derived cardiomyocytes [11] have been useful for understanding cardiotoxicity [12]. Hence, transcriptional profiles in human cardiac cells can form starting points for studies focused on mechanism-based drug signatures that could be used for prediction of cardiotoxicity potential. Here, we have used six hiPSC-derived cardiac cell lines from healthy individuals to study 54 FDA-approved drugs and identify drug-selective transcriptomic signatures across the different human subject lines by singular value decomposition (SVD)-based analysis of total transcriptomic responses. Single cell analyses of hiPSC-derived “ventricular cardiomyocytes” indicated that they are composed of multiple clusters. They include a cluster very similar to adult cardiomyocytes. Additional clusters show varying levels of similarity to different cell types in the human heart, including fibroblasts. Projections of drug affected pathways inferred from bulk transcriptomic data of hiPSC-derived cardiomyocytes on to single cell expression profiles of human hearts from healthy and heart failure patients indicate that TKIs can affect both cardiomyocytes as well as other cell types to produce adverse events, and that the drug-induced pathways are similar to those altered in hiPSC-derived cardiomyocytes from patients with failing hearts [13,14]. Using outlier analyses and Whole Genome Sequencing (WGS) from the cell lines used for the transcriptomic studies, we correctly re- identified the genomic variant associated with anthracycline cardiotoxicity that was originally discovered in genome-wide association studies (GWAS) [15]. We used this approach to identify and predict potential effects of genomic variants associated with TKI action as identified by the transcriptomic signature. These findings indicate that integration of experimental data with public data sets can be powerful drivers of multiscale understanding of organ level adverse events associated with drug therapy.

Results

To identify pathway activities and potential genomic variants associated with TKI-induced cardiotoxicity (TIC), we treated six hiPSC-derived cardiomyocyte cell lines from six healthy individuals [16] (Suppl. Figures 1A/B) (Suppl. Table 1) with one of 27 TKIs, 4 anthracyclines and 26 other cardiac and noncardiac acting drugs (Suppl. Table 2) for 48 hours (Suppl. Table 3 [16]). All drugs are FDA-approved. The group of TKIs contained 23 small molecule TKIs and 4 monoclonal antibodies against TKs. They could be separated into 10 cardiotoxic and 17 non-cardiotoxic TKIs based on the results of clinical studies (Suppl. Table 4) and FAERS data analysis (Suppl. Figure 2). Bulk transcriptomic analysis of control [16] and treated cell lines generated 266 lists of differentially expressed genes (DEGs) (Suppl. Table 5) (Suppl. Table 6 shows averaged DEGs). All lists of DEGs, each representing the response observed for one sample, i.e. a unique cell line/drug combination (Suppl. Figure 3A), were subjected to pairwise correlation and hierarchical clustering (Suppl. Figure 3B).

Singular value decomposition to unmask drug-selective expression responses

To identify drug-selective expression responses we developed a computational pipeline that uses Singular Value Decomposition (SVD) to search for shared gene expression components in multiple cell lines treated with the same drug (Figure 1A, Suppl. Figure 4). Clustering of the transcriptomic data only grouped a few samples treated with the same drugs into the same cluster (Figure 1B, Suppl. Figure 3B). On the contrary, the treated cell line or the amplitude of the drug response, i.e., the number of significant DEGs, mainly determined the clustering results (Suppl. Figure 5A). To quantify drug-selective clustering, we calculated one F1 score for each drug that documents how close all cell lines treated with that drug cluster together (Suppl. Figure 5B). With possible values from 0 to 1, the median F1-score of 0.116 documents very low drug-selective clustering efficiencies (Figure 1C).

SVD (Suppl. Figure 6A) of all 266 samples identified 266 eigenarrays (Suppl. Figure 6B) whose linear combination using sample-specific coefficients gives the complete gene expression profiles. The first eigenarray correlates (Suppl. Figure 6C) with the amplitude of the response (Suppl. Figure 6D), and its most prominent genes (Suppl. Table 7) enrich for muscle contraction (Suppl. Figure 6E, Suppl. Table 8). These results indicate that the first eigenarray describes a general response to perturbation that masks drug-selective effects. Removal of the first eigenarray (Suppl. Table 9) disrupts the clustering by the number of significant DEGs (Suppl. Figure 7). Since the clustering is still dominated by cell-line-selective effects, the removal did not markedly improve the drug-selective clustering efficiencies (Suppl. Figure 8).

We ranked the remaining eigenarrays by their ability to separate samples of each drug or cell line from all other samples (Suppl. Figure 9A). Our algorithm allowed a clear separation of cell-line and drug-selective effects (Suppl. Figure 9B). For each drug, we combined unique sets of drug-selective eigenarrays to drug-selective subspaces with the intention to optimize the drug's clustering efficiency (Suppl. Figure 10A) while preserving a maximum of the original information, as quantified by cosine similarity. Changing the relative contribution of F1 scores (Suppl. Figure 10B) and cosine similarities (Suppl. Figure 10C) we screened the potential

subspaces for outlier responses where one cell line showed a significantly different response to the drug of interest than all other cell lines (Suppl. Figures 10D, 11, 12). Such outlier responses will be linked to our genomic analysis at a later stage. For 24 drugs, we could identify subspaces with outlier responses (Suppl. Figure 10E), for the other 30 drugs, we selected subspaces based on high relative contribution of the F1 score (95%).

Drug-selective gene expression profiles are similar for drugs with similar mechanisms

Projection of gene expression profiles into the final drug-selective subspaces generated drug-selective gene expression profiles (Suppl. Table 10) (averaged DEGs in Suppl. Table 11) with high clustering efficiencies (Figure 1D, Suppl. Figure 10F). Drug-selective DEGs were combined into a new matrix. Clustering of this matrix identifies clusters that mostly contained samples that were treated with the same drug and documented preservation of outlier characteristics (Figure 1E, Suppl. Figure 13). Eleven of the 24 identified outlier samples are clearly separated from the cell lines treated with the same drug (closed circles in Figure 1E and Suppl. Figure 13). Another eight outlier samples are grouped together with the samples treated with the same drug in a larger cluster (open circles in Suppl. Figure 13), but are separated from those, once that cluster is subclustered. Only four identified outliers reside in the same cluster (open rectangles) as the other cell lines treated with the same drug. Comparison of Figures 1B and 1E in terms of the breadth of the colored bars allows for visualization of this drug selective clustering.

Our algorithm identified drug-selective gene expression profiles for each drug independently. Nevertheless, drugs with similar mechanisms and overlapping targets still cluster together (Figure 1E, Suppl. Figure 13). All eight TKIs targeting EGFR signaling, either by inhibiting the receptor (EGFR, ERBB2) or its intracellular ERK signaling cascade (MAP2K1, MAP2K2) [17] are part of the same cluster. One of the four non-TKI drugs in this cluster, phenylephrine, and isoprenaline which is part of a slightly bigger cluster, stimulate adrenergic signaling that can cross-activate ERK signaling in the heart [18], as directly documented for phenylephrine in the perfused rat heart [19]. Two additional non-TKI drugs in this cluster, verapamil and amiodarone, influence adrenergic signaling as antagonists [20-22] and verapamil has been shown to antagonize ERK signaling in rat cardiomyocytes [23]. Similarly, the JAK inhibitors ruxolitinib and tofacitinib and the proteasome inhibitors bortezomib and carflizomib are part of two independent clusters that contain no other drugs. Two of three antidiabetics, rosiglitazone and saxagliptin are part of a smaller cluster that first merges with a cluster containing decitabine, a drug with hyperglycemia as a major side effect [24] and then with a cluster containing cyclosporine and olmesartan. Cyclosporine increases insulin resistance in humans [25] and olmesartan -/+ amlodipine ameliorates it in rats [26] and patients [27], respectively. The anthracyclines epirubicin, daunorubicin and doxorubicin are part of larger cluster, while the anthracycline idarubicin clusters together with their identified outlier samples.

In summary, our SVD pipeline uncovers similar expression profiles induced by similar drugs. Since it does not actively search for such similarities, but only aims at maximizing the overlap

between DEGs induced by the same drug, this finding serves as internal validation for the extraction of drug-selective components by our computational pipeline.
The grouping of drugs with similar mechanisms into higher-level clusters explains eleven non-outlier samples that are separated from the other samples treated with the same drug (closed rhombuses), leaving only four samples that are not part of drug-selective clusters for no obvious biological reason (open rhombuses)(Suppl. Fig. 13).

SVD enables identification of subcellular processes relevant to TKI cardiotoxicity

Complete and decomposed DEGs were subjected to pathway enrichment analysis using the Molecular Biology of the Cell Ontology (MBCO) (20) (Suppl. Figure 14A). MBCO subcellular processes (SCPs) are organized in three to four levels where higher-level SCPs (i.e., levels with lower numbers) describe more general, and lower-level SCPs more detailed, cell biological processes. Predicted up- or downregulated SCPs of each drug/cell line combination and SCP level were ranked by significance (Suppl. Tables 12/13/14). We will refer to these ranks as enrichment ranks in the following.
Overall comparison of the enrichment results before and after decomposition documents a great increase in the number of overlapping SCPs in different cell lines treated with the same drug (Suppl. Figures 14B/C/D/E, 15A).
In addition, our decomposition pipeline allowed identification of SCPs that were masked in the complete dataset. For example, we could document downregulation of SCPs involved in single protein degradation as an anthracycline-specific effect (Suppl. Figure 15B), results that we did not obtain using the complete DEGs (Suppl. Figure 15C). For a discussion of these SCPs, their clinical relevance for supportive treatment and examples that demonstrate agreement with prior knowledge from small-scale experiments see Supplemental Information.

Subcellular processes and cell types indicative of cardiotoxic response to TKI treatment

To identify signatures associated with TKI-induced cardiotoxicity we searched for SCPs that are almost exclusively up- or downregulated at higher enrichment ranks by cardiotoxic TKIs as compared to noncardiotoxic TKIs, using F1 score and Area under the Curve (AUC) statistics (Figure 2A, Suppl. Figure 16). SCPs were ranked by decreasing AUC. The top 10, 10, 25 and 10 level-1, -2, -3 and -4 SCPs were grouped by functional similarities (Figure 2B, Suppl. Figure 18) and integrated into the MBCO hierarchy (Figure 2C, Suppl. Figure 19). In addition, we mapped identified SCPs back to the cardiotoxic TKIs that up- or downregulate them (Suppl. Figure 20).
Besides being elevated or decreased above or below a threshold that determines an SCP's association with a cardiotoxic response, an SCP's activity might already be beyond that threshold at baseline level, i.e., before TKI treatment. Candidates for this group should be those SCPs that are regulated by non-cardiotoxic TKIs. A non-cardiotoxic TKI might change the baseline SCP activity from sufficient to insufficient for a cardiotoxic response. Consequently, our second focus was the investigation of the SCPs that our algorithm identified, if applied to non-cardiotoxic TKIs in the same manner (Suppl. Figures 18, 19, 21).

Single cell RNA sequencing of four of our six cardiomyocyte cell lines identified at least two major subtypes – adult cardiomyocytes and an epicardium derived cell type that is similar to fibroblasts [16] (Suppl. Figure 22A-D, Suppl. Tables 15,16A/B). The functions of SCPs we identified by bulk transcriptomics often represent canonical functions of cell types other than cardiomyocytes. We determined how the top SCPs in our data map to our subtypes (Suppl. Figure 23, Suppl. Tables 16C-F), and to the major cell types of the adult human heart analyzed by others using single cell transcriptomics [28] (Figure 3A, Suppl. Figure 23, Suppl. Tables 17, 18A-D).

Additionally, we mapped our SCPs to SCPs in hiPSC-derived cardiomyocytes obtained from infants with Dilated Cardiomyopathy (DCM) (GSE184899) (Suppl. Table 19) or in adult heart cells of patients with DCM or Hypertrophic Cardiomyopathy (HCM) [13,14] (Figure 3A, Suppl. Figure 24, Suppl. Tables 20A/B/C). Many of the identified SCPs described well-known biology involved in inherited and acquired cardiomyopathies that develop independently of chemotherapy and map to multiple cell types of the adult human heart. Our results suggest that cardiotoxic TKIs might mimic these mechanisms by early (within 48 hrs) transcriptional regulation.

Four SCPs involved in muscle contraction and sarcomere renewal were simultaneously identified for the cardiotoxic and noncardiotoxic drugs (Figure 2B, Suppl. Figure 18). Another two SCPs were identified for each group independently. These SCPs that span all four SCP levels (Figure 2C, Suppl. Figure 19) are regulated by the two TKI toxicity groups in the opposite directions. The difference in directionality suggests insufficient sarcomere renewal as relevant for TKI-induced cardiotoxicity. All cardiotoxic TKIs, except vandetinib and trastuzumab downregulate (Suppl. Figure 20), while 13 of 17 non-cardiotoxic TKIs upregulate (Suppl. Figure 21) at least one of these SCPs, respectively. The SCPs map to cardiomyocyte clusters in cell lines and the adult human heart (Figure 3B, Suppl. Figure 23). In agreement with our data, proper heart functioning depends on a continuous turnover of cardiomyocyte sarcomere proteins [29]. About half of hypertrophic and dilated cardiomyopathies are associated with mutations in sarcomeric proteins [30]. Further supporting this line of reasoning, we found downregulation of muscle contraction-related SCPs in adult and hiPSC-derived DCM cardiomyocytes (Figure 3C, Suppl. figure 24). These SCPs might serve as a starting point for supportive therapy. Four of the six regulated genes of the SCP ‘Thin myofilament organization’ (Suppl. Table 12) are involved in blocking the binding of the myosin head to the thin myofilament during muscle contraction. This mechanism is targeted by the new drug mavacamten that was recently approved by the FDA to treat obstructive Hypertrophic Cardiomyopathy (HCM) [31].

Our data associates the upregulation of SCPs involved in the citric acid cycle and mitochondrial energy generation with a cardiotoxic response (Figures 2B/C, Suppl. Figures 18/19). They are

almost exclusively upregulated by pazopanib (Suppl. Figure 20), a TKI with a high rate of cardiotoxicity (>10%) (Suppl. Table 4). Many studies document an overall reduction in oxidative phosphorylation during heart failure [32] and identified SCPs are downregulated in adult DCM and HCM cardiomyocytes as well as in hiPSC-derived DCM cardiomyocytes (Figure 3B, Suppl. Figure 24). Nevertheless, in support of our findings, compensatory upregulation of oxidative phosphorylation was suggested for patients belonging to a large DCM subgroup that is caused by truncating titin variants [33-35].

Increasing evidence links ferroptosis, an iron-dependent accumulation of lipid peroxides that triggers cell death, to heart failure [36]. Some identified SCPs in our data might converge on ferroptosis as an endpoint. The SCP 'Desaturation of fatty acids' was associated with a cardiotoxic response (Figures 2B/C, Suppl. Figures 18/19) due to its upregulation by pazopanib (Suppl. Figure 20). Upregulated genes mapping to this SCP (Suppl. Table 12) can generate polyunsaturated fatty acids (PUVAs) [37,38] that are precursors of lipid peroxides [36]. A lower activity of the level-2 SCP 'Cellular antioxidant systems' and its level-1 parent 'Cellular redox homeostasis', pathways offering protection against ferroptosis [36], is associated with a cardiotoxic response (Suppl. Figures 18/19). They were downregulated by the cardiotoxic TKIs vandetinib and bevacizumab (Suppl. Figure 20). In addition, downregulation of 'Cellular iron storage' by the anthracyclines doxorubicin and daunorubicin identified by us (Suppl. Figures 14D, 15B) and others [39] suggests stimulation of ferroptosis by an increase of intracellular free iron [39]. In agreement, supportive therapy with iron chelators protects against anthracycline-induced cardiotoxicity [40].

The clinical relevance of our findings is further supported by identification of SCPs involved in cholesterol metabolism and natriuretic peptide signaling (Figure 2B, Suppl. Figure 18). This is supported by recommended treatment of cardiomyopathy and drug-induced cardiomyopathy with statins [41] and neprilysin [42,43], respectively. Prediction of extracellular collagen-crosslinking (Suppl. Figure 18) agrees with histopathological observations that correlate with disease progression [44,45]. Involvement of identified signaling pathways in TIC i.e., PDGF, HIF-1alpha, Oncostatin M and Hippo signaling (Figure 2B, Suppl. Figure 18) is supported by their involvement in drug-independent cardiomyopathy. These and other findings regarding SCPs are discussed in further detail in the supplemental information.
Taken together the data from our experiments, when integrated with data from the literature, indicate the TKI are likely to have their cardiotoxic effects by regulating SCPs in multiple cell types of the heart.

Candidates for genomic variants associated with drug-induced cardiotoxicity.
Using whole genome sequencing of the six cell lines [16], our next analysis step focused on the identification of potential genomic variants that might be associated with a higher or lower risk for drug- or more specifically TKI-induced cardiotoxicity (DIC, TIC, respectively) (Figure 4A, Suppl. Figure 25). Cardiotoxic responses to TKI treatment are observed in less than 10% to

20% of the treated patients (Suppl. Table 4). Consequently, we hypothesized that the population-wide frequency of a potential monogenic allele associated with DIC should be in a similar range. Relevance for cardiac tissue was assumed for those variants that map to gene coding regions or are part of cis-expression (e) or -splicing (s) QTLs in the adult human heart [46]. We hypothesized that potentially cardiotoxic or -protective variants could interfere with a drug's pharmacokinetics (PK) or -dynamics (PD) or with pathways that are targeted by that drug.

Variants of the first interference group should induce a different transcriptomic response, allowing us to link cell-line-specific variants to identified 24 outlier responses. Since, based on statistical likelihoods, either none or only one of our six healthy volunteers should suffer from cardiotoxicity induced by a drug of interest, we focused on variants that are overexpressed in an outlier cell line for a drug of interest, if compared to the other five lines. Mechanistic knowledge is added by considering only variants that map to genes with potential involvement in a drug's PK/PD.

Variants of the second group should map to pathways that either contain drug target proteins or are induced or repressed by the drug of interest on the transcriptional level. Using the results of our transcriptomic analysis, we focused on the second group with a particular interest in those SCPs that are associated with a cardiotoxic or non-cardiotoxic response.

Identification of genomic variants interfering with a drug's PK/PD

Our PK/PD algorithm allowed re-identification of the variant *rs2229774* (Figure 4B, Suppl. Table 21) within the coding region of the *RARG* gene [15], one out of three genomic variants with the strongest evidence for anthracycline-induced cardiotoxicity (AIC) [47]. The suppressive activity of the transcription factor RARG on the expression of the anthracycline target TOP2B is reduced by this variant, leading to increased AIC [48]. Pathway enrichment analysis (Suppl. Figure 14D) suggests missing upregulation of WNT receptor signaling as a potential mechanism of *rs2229774* -triggered AIC (Suppl. Figures 14D, 26A). In agreement, inhibition or activation of WNT signaling enhances or mitigates AIC, respectively [49-51].

In total, our outlier-based screening approach identified 213 and 201 potential variants that map to 133 and 129 genes involved in the PK/PD of three anthracyclines and three cardiotoxic TKIs, respectively (Figure 4C, Suppl. Figure 26B, Suppl. Table 21). Identified variant candidates could trigger or protect from AIC or TIC. Seven of the variants map to drug target proteins of ponatinib and trametinib (Figure 4D), TKIs with medium (1-10%) and high (>10%) rates of TIC (Suppl. Table 4).

Considering all 24 drugs with documented outlier responses for this analysis we identified 847 variants (Suppl. Figure 26B) mapping to 378 genes (Suppl. Figure 26C) (Suppl. Table 21).

SCP-based identification of genomic variants associated with TKI -induced cardiotoxicity.

Identified level-2, -3 and -4 SCPs associated with the response to cardiotoxic and non-cardiotoxic TKIs enrich for genes that are mapped to inherited DCM or HCM, either within the

HuGe Phenopedia database [52] or obtained from GWAS [53,54] (Figure 4E). We left out level-1 SCPs from this analysis, since they were predicted based on genes mapping to a sub function of the general cellular functions they describe. The significant overlap supports our observation that TKIs induce cardiotoxicity by mimicking mechanisms involved in inherited or acquired cardiomyopathy, as discussed above and in the supplemental information.

To identify a second set of potential variants interfering with TIC we mapped all variants underrepresented in the general population to identified SCPs. In total, we identified 2383, 1340 and 219 non-overlapping variants mapping to level-2, -3 and -4 SCPs that are up- or downregulated by cardiotoxic TKIs (Figure 4F). Mapping identified SCPs back to the drugs that induce them (Suppl. Figure 20) enables identification of variants that might interfere with cardiotoxicity of a drug of interest (Suppl. Table 22). For example, we predicted 31 variants mapping to seven genes of the SCP 'Thin myofilament organization' that might interfere with the cardiac response to sunitinib and ponatinib downregulating this SCP in six (enrichment ranks 2x3, 4, 2x7) and five (5x1) of six treated cell lines, respectively (Suppl. Figure 20), but also to lapatinib, dabrafenib and bevacizumab (Figure 4G). Four of those genes are involved in the same function that is targeted by mavacamten, as described above.
Most variants identified in GWAS that map to our SCPs are too frequent to meet our cutoff criteria that aim at identification of single variants triggering TIC in a monogenetic manner. Inherited HCM and DCM exist as both monogenic and polygenic diseases. Similarly, single or multiple variants might be related to TIC, requiring development of more sophisticated genomic algorithms. Nevertheless, identified SCPs are a good starting point to reduce the burden of multiple testing hypothesis by restricting the focus on genes with a functional implication in TIC.

Discussion

A key question that arises in identifying transcriptional signatures from short term treatment of cell lines with drugs is whether early transcriptional events in a 1-2-day time frame can be predictive of later physiological (pathophysiological) states of an organ that takes weeks or months to manifest when a drug is used clinically. While it is not possible to answer this question for multiple organs and drug classes, in the case of TKI drugs and heart the answer is yes. Two types of relationships support this conclusion. We previously showed that transcriptomic signatures of TKI with differing toxicity profiles in FAERS could be associated with these profiles [55]. In this study, the pathway analyses of drug-selective gene expression lists across cell lines show that many TKIs with adverse event propensity selectively regulate pathways that have also been shown to be involved in different types of cardiomyopathies [56]. Additionally, the identity of the pathways themselves such as contraction dynamics and extracellular matrix are directly relevant to development of cardiac insufficiency. It is noteworthy that mavacamten, a drug approved in 2022 by the FDA to treat hypertrophic cardiomyopathy [31] regulates the level of actin-myosin bridge formation and thus controls contractile dynamics. Taken together, it appears reasonable to conclude that the pathways regulated by cardiotoxic

drugs can serve as pathway signatures for cardiotoxicity. Such signatures could be used in drug development. Transcriptomic profiles of drug candidates could predict risk of cardiotoxicity from early cell-based study (Figure5A). Additionally, this approach can also be used to predict new drug targets that could mitigate TKI-induced cardiotoxicity (Figure 5B) by reversing relevant pathway activities [57].

Our method for identifying drug selective DEGs could have wider applications since both bulk and single cell expression profiles often represent the combination of multiple DEG sets induced by different actions on the same cell type or mixtures of cell types (i.e., tissue). Our computational pipeline decomposes these mixed gene expression profiles into individual effect-selective DEG sets and identifies the DEG sets involved in the response of interest. For example, grouping patients with the same disease into subgroups based on shared disease clinical manifestations can enable identification of gene expression profiles associated with each selected subgroup that can lead to tailored treatment options. A particular advantage of our pipeline is that patient groupings do not need to be mutually exclusive, i.e. patients can be assigned to each disease subgroup independently of the other assignments.

A serendipitous finding from our studies on the hiPSC derived cardiomyocytes is that single cell RNA-Seq allowed us to resolve these cells into multiple clusters [16]. As single cell transcriptomic data from normal [28] and diseased [13,14] adult human heart became available in 2022 we compared the groups in our data to the different cell types in the adult human heart. This comparison allowed us to identify the group that is most like ventricular cardiomyocytes and to identify a group related to fibroblasts. Several of the top ranked SCPs are most likely expressed in multiple cell types, since they are not specifically enriched in any of them. Thus, the integration of bulk transcriptomic and single cell transcriptomic data allows us to project the SCPs onto different cell types of the heart. These integrated analyses suggest that the cardiotoxic drugs are likely to have their pathophysiological effects by affecting multiple cell types. The effects on fibroblasts and extracellular matrix dynamics [58] can readily affect the elasticity of the cardiac muscle and it has long been known that extracellular collagen-crosslinking is directly related to heart failure [44,45]. It appears likely that the effects of TKIs on fibroblasts can contribute in part to cardiotoxic effects. In evaluating drug candidates for cardiotoxicity, it may be necessary to consider the drugs' effects not only on cardiomyocytes but also on cardiac fibroblasts and other cell types in the heart.
Using the common differential expression patterns for each drug across cell lines, it is possible to identify outlier responses in a particular cell line (i.e., human subject). Such responses could provide information about genomic characteristics that could be associated with adverse events since these events occur in typically 1-20% of the population. Genomic variants associated with the proteins involved in PK/PD are likely candidates. We tested the validity of this assumption and found that a genomic variant associated with RARG can be identified by integration of whole genome wide sequencing of our six cell lines and outlier analysis of the anthracycline drugs. GWAS [15] and follow-up studies [48] had previously identified RARG to be causally linked to anthracycline-induced cardiotoxicity, since RARG regulates the expression

of the anthracycline target topoisomerase 2B. Based on this validation we searched for genomic variants associated with cardiotoxic TKIs. We considered PK, PD, and key pathways regulated by these drugs and developed a list of variants whose occurrence resemble those of cardiotoxic events. Here data integration identifies genomic variants that could serve as hypotheses for targeted genomic studies to identify risk determinants in different populations (Figure 5C). Overall, our study shows that integration of transcriptomic and genomic data with various publicly available datasets can provide an integrated understanding and serve as a powerful hypothesis generator for different laboratory and clinic-based studies to detect and avoid adverse events associated with drug therapy.

Data Availability. Transcriptomic and genomic data are available at NCBI GEO and dbGAP under the accession numbers GSE174773, GSE217421 and phs002088.v1.p1, respectively.

Acknowledgement This research was supported by a grant from NIH Common Fund 5U54HG008098 Drug Combination Signatures for Prediction and Mitigation of Toxicity as part of the LINCS program. We thank Jill Gregory for help with the flowchart figures.

References

1 de Vries, E. N., Ramrattan, M. A., Smorenburg, S. M., Gouma, D. J. & Boermeester, M. A. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* **17**, 216-223 (2008). https://doi.org:10.1136/qshc.2007.023622

2 Drozda, K., Pacanowski, M. A., Grimstein, C. & Zineh, I. Pharmacogenetic Labeling of FDA-Approved Drugs: A Regulatory Retrospective. *JACC Basic Transl Sci* **3**, 545-549 (2018). https://doi.org:10.1016/j.jacbts.2018.06.001

3 Smith, A. F., Klotz, A. & Wormstone, I. M. Improving the drug development process by reducing the impact of adverse events: the case of cataracts considered. *Drug Discov Today* **21**, 510-516 (2016). https://doi.org:10.1016/j.drudis.2016.01.001

4 Dorato, M. A. & Buckley, L. A. Toxicology in the drug discovery and development process. *Curr Protoc Pharmacol* **Chapter 10**, Unit10 13 (2006). https://doi.org:10.1002/0471141755.ph1003s32

5 Ma'ayan, A., Jenkins, S. L., Goldfarb, J. & Iyengar, R. Network analysis of FDA approved drugs and their targets. *Mt Sinai J Med* **74**, 27-32 (2007). https://doi.org:10.1002/msj.20002

6 Vandenberg, J. I. *et al.* hERG K(+) channels: structure, function, and clinical significance. *Physiol Rev* **92**, 1393-1478 (2012). https://doi.org:10.1152/physrev.00036.2011

7 Raschi, E., Vasina, V., Poluzzi, E. & De Ponti, F. The hERG K+ channel: target and antitarget strategies in drug development. *Pharmacol Res* **57**, 181-195 (2008). https://doi.org:10.1016/j.phrs.2008.01.009

8 Jain, D. & Aronow, W. Cardiotoxicity of cancer chemotherapy in clinical practice. *Hosp Pract (1995)* **47**, 6-15 (2019). https://doi.org:10.1080/21548331.2018.1530831

9 Garcia-Alvarez, A., Garcia-Albeniz, X., Esteve, J., Rovira, M. & Bosch, X. Cardiotoxicity of tyrosine-kinase-targeting drugs. *Cardiovasc Hematol Agents Med Chem* **8**, 11-21 (2010). https://doi.org:10.2174/187152510790796192

10 Yoshida, Y. & Yamanaka, S. Induced Pluripotent Stem Cells 10 Years Later: For Cardiac Applications. *Circ Res* **120**, 1958-1968 (2017). https://doi.org:10.1161/CIRCRESAHA.117.311080

11 Sharma, A. *et al.* High-throughput screening of tyrosine kinase inhibitor cardiotoxicity with human induced pluripotent stem cells. *Sci Transl Med* **9** (2017). https://doi.org:10.1126/scitranslmed.aaf2584

12 Wang, H. *et al.* Adaptation of Human iPSC-Derived Cardiomyocytes to Tyrosine Kinase Inhibitors Reduces Acute Cardiotoxicity via Metabolic Reprogramming. *Cell Syst* **8**, 412-426 e417 (2019). https://doi.org:10.1016/j.cels.2019.03.009
13 Koenig, A. L. *et al.* Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. *Nat Cardiovasc Res* **1**, 263-280 (2022). https://doi.org:10.1038/s44161-022-00028-6
14 Chaffin, M. *et al.* Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy. *Nature* **608**, 174-180 (2022). https://doi.org:10.1038/s41586-022-04817-8
15 Aminkeng, F. *et al.* A coding variant in RARG confers susceptibility to anthracycline-induced cardiotoxicity in childhood cancer. *Nat Genet* **47**, 1079-1084 (2015). https://doi.org:10.1038/ng.3374
16 Schaniel, C. *et al.* A library of induced pluripotent stem cells from clinically well-characterized, diverse healthy human individuals. *Stem Cell Reports* **16**, 3036-3049 (2021). https://doi.org:10.1016/j.stemcr.2021.10.005
17 Huang, L. & Fu, L. Mechanisms of resistance to EGFR tyrosine kinase inhibitors. *Acta Pharm Sin B* **5**, 390-401 (2015). https://doi.org:10.1016/j.apsb.2015.07.001
18 Patel, P. A., Tilley, D. G. & Rockman, H. A. Beta-arrestin-mediated signaling in the heart. *Circ J* **72**, 1725-1729 (2008). https://doi.org:10.1253/circj.cj-08-0734
19 Lazou, A., Sugden, P. H. & Clerk, A. Activation of mitogen-activated protein kinases (p38-MAPKs, SAPKs/JNKs and ERKs) by the G-protein-coupled receptor agonist phenylephrine in the perfused rat heart. *Biochem J* **332 (Pt 2)**, 459-465 (1998). https://doi.org:10.1042/bj3320459
20 Karliner, J. S., Motulsky, H. J., Dunlap, J., Brown, J. H. & Insel, P. A. Verapamil competitively inhibits alpha 1-adrenergic and muscarinic but not beta-adrenergic receptors in rat myocardium. *J Cardiovasc Pharmacol* **4**, 515-520 (1982). https://doi.org:10.1097/00005344-198205000-00025
21 Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* **46**, D1074-D1082 (2018). https://doi.org:10.1093/nar/gkx1037
22 Sloskey, G. E. Amiodarone: a unique antiarrhythmic agent. *Clin Pharm* **2**, 330-340 (1983).
23 Cheng, W., Zhu, Y. & Wang, H. The MAPK pathway is involved in the regulation of rapid pacing-induced ionic channel remodeling in rat atrial myocytes. *Mol Med Rep* **13**, 2677-2682 (2016). https://doi.org:10.3892/mmr.2016.4862
24 Ma, Y. Y. *et al.* Use of decitabine for patients with refractory or relapsed acute myeloid leukemia: a systematic review and meta-analysis. *Hematology* **24**, 507-515 (2019). https://doi.org:10.1080/16078454.2019.1632407
25 Sarno, G. *et al.* New-onset diabetes mellitus: predictive factors and impact on the outcome of patients undergoing liver transplantation. *Curr Diabetes Rev* **9**, 78-85 (2013).
26 Yanagihara, H., Ushijima, K., Arakawa, Y., Aizawa, K. & Fujimura, A. Effects of telmisartan and olmesartan on insulin sensitivity and renal function in spontaneously hypertensive rats fed a high fat diet. *J Pharmacol Sci* **131**, 190-197 (2016). https://doi.org:10.1016/j.jphs.2016.06.003
27 Derosa, G. *et al.* Olmesartan/amlodipine combination versus olmesartan or amlodipine monotherapies on blood pressure and insulin resistance in a sample of hypertensive patients. *Clin Exp Hypertens* **35**, 301-307 (2013). https://doi.org:10.3109/10641963.2012.721841
28 Litvinukova, M. *et al.* Cells of the adult human heart. *Nature* **588**, 466-472 (2020). https://doi.org:10.1038/s41586-020-2797-4
29 Hnia, K., Clausen, T. & Moog-Lutz, C. Shaping Striated Muscles with Ubiquitin Proteasome System in Health and Disease. *Trends Mol Med* **25**, 760-774 (2019). https://doi.org:10.1016/j.molmed.2019.05.008
30 Vikhorev, P. G. & Vikhoreva, N. N. Cardiomyopathies and Related Changes in Contractility of Human Heart Muscle. *Int J Mol Sci* **19** (2018). https://doi.org:10.3390/ijms19082234
31 Dalo, J. D., Weisman, N. D. & White, C. M. Mavacamten, a First-in-Class Cardiac Myosin Inhibitor for Obstructive Hypertrophic Cardiomyopathy. *Ann Pharmacother*, 10600280221117812 (2022). https://doi.org:10.1177/10600280221117812

32 Lopaschuk, G. D., Karwi, Q. G., Tian, R., Wende, A. R. & Abel, E. D. Cardiac Energy Metabolism in Heart Failure. *Circ Res* **128**, 1487-1513 (2021). https://doi.org:10.1161/CIRCRESAHA.121.318241
33 Verdonschot, J. A. J. *et al.* Titin cardiomyopathy leads to altered mitochondrial energetics, increased fibrosis and long-term life-threatening arrhythmias. *Eur Heart J* **39**, 864-873 (2018). https://doi.org:10.1093/eurheartj/ehx808
34 Schafer, S. *et al.* Titin-truncating variants affect heart function in disease cohorts and the general population. *Nat Genet* **49**, 46-53 (2017). https://doi.org:10.1038/ng.3719
35 Herman, D. S. *et al.* Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* **366**, 619-628 (2012). https://doi.org:10.1056/NEJMoa1110186
36 Yang, X., Kawasaki, N. K., Min, J., Matsui, T. & Wang, F. Ferroptosis in heart failure. *J Mol Cell Cardiol* **173**, 141-153 (2022). https://doi.org:10.1016/j.yjmcc.2022.10.004
37 Lee, J. Y. *et al.* Polyunsaturated fatty acid biosynthesis pathway determines ferroptosis sensitivity in gastric cancer. *Proc Natl Acad Sci U S A* **117**, 32433-32442 (2020). https://doi.org:10.1073/pnas.2006828117
38 Blahova, Z., Harvey, T. N., Psenicka, M. & Mraz, J. Assessment of Fatty Acid Desaturase (Fads2) Structure-Function Properties in Fish in the Context of Environmental Adaptations and as a Target for Genetic Engineering. *Biomolecules* **10** (2020). https://doi.org:10.3390/biom10020206
39 Huang, J. *et al.* Understanding Anthracycline Cardiotoxicity From Mitochondrial Aspect. *Front Pharmacol* **13**, 811406 (2022). https://doi.org:10.3389/fphar.2022.811406
40 Voest, E. E., van Acker, S. A., van der Vijgh, W. J., van Asbeck, B. S. & Bast, A. Comparison of different iron chelators as protective agents against acute doxorubicin-induced cardiotoxicity. *J Mol Cell Cardiol* **26**, 1179-1185 (1994). https://doi.org:10.1006/jmcc.1994.1136
41 Kim, J., Nishimura, Y., Kewcharoen, J. & Yess, J. Statin Use Can Attenuate the Decline in Left Ventricular Ejection Fraction and the Incidence of Cardiomyopathy in Cardiotoxic Chemotherapy Recipients: A Systematic Review and Meta-Analysis. *J Clin Med* **10** (2021). https://doi.org:10.3390/jcm10163731
42 McMurray, J. J. *et al.* Angiotensin-neprilysin inhibition versus enalapril in heart failure. *N Engl J Med* **371**, 993-1004 (2014). https://doi.org:10.1056/NEJMoa1409077
43 Velazquez, E. J. *et al.* Angiotensin-Neprilysin Inhibition in Acute Decompensated Heart Failure. *N Engl J Med* **380**, 539-548 (2019). https://doi.org:10.1056/NEJMoa1812851
44 Lopez, B., Querejeta, R., Gonzalez, A., Larman, M. & Diez, J. Collagen cross-linking but not collagen amount associates with elevated filling pressures in hypertensive patients with stage C heart failure: potential role of lysyl oxidase. *Hypertension* **60**, 677-683 (2012). https://doi.org:10.1161/HYPERTENSIONAHA.112.196113
45 Lopez, B. *et al.* Myocardial Collagen Cross-Linking Is Associated With Heart Failure Hospitalization in Patients With Hypertensive Heart Failure. *J Am Coll Cardiol* **67**, 251-260 (2016). https://doi.org:10.1016/j.jacc.2015.10.063
46 Consortium, G. T. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318-1330 (2020). https://doi.org:10.1126/science.aaz1776
47 Aminkeng, F. *et al.* Recommendations for genetic testing to reduce the incidence of anthracycline-induced cardiotoxicity. *Br J Clin Pharmacol* **82**, 683-695 (2016). https://doi.org:10.1111/bcp.13008
48 Magdy, T. *et al.* RARG variant predictive of doxorubicin-induced cardiotoxicity identifies a cardioprotective therapy. *Cell Stem Cell* **28**, 2076-2089 e2077 (2021). https://doi.org:10.1016/j.stem.2021.08.006
49 Liang, L. *et al.* Dkk1 exacerbates doxorubicin-induced cardiotoxicity by inhibiting the Wnt/beta-catenin signaling pathway. *J Cell Sci* **132** (2019). https://doi.org:10.1242/jcs.228478
50 El-Ela, S. R. A., Zaghloul, R. A. & Eissa, L. A. Promising cardioprotective effect of baicalin in doxorubicin-induced cardiotoxicity through targeting toll-like receptor 4/nuclear factor-kappaB and Wnt/beta-catenin pathways. *Nutrition* **102**, 111732 (2022). https://doi.org:10.1016/j.nut.2022.111732

51 Feng, D. *et al.* DDX3X alleviates doxorubicin-induced cardiotoxicity by regulating Wnt/beta-catenin signaling pathway in an in vitro model. *J Biochem Mol Toxicol* **36**, e23077 (2022). https://doi.org:10.1002/jbt.23077
52 Yu, W., Clyne, M., Khoury, M. J. & Gwinn, M. Phenopedia and Genopedia: disease-centered and gene-centered views of the evolving knowledge of human genetic associations. *Bioinformatics* **26**, 145-146 (2010). https://doi.org:10.1093/bioinformatics/btp618
53 Pirruccello, J. P. *et al.* Analysis of cardiac magnetic resonance imaging in 36,000 individuals yields genetic insights into dilated cardiomyopathy. *Nat Commun* **11**, 2254 (2020). https://doi.org:10.1038/s41467-020-15823-7
54 Harper, A. R. *et al.* Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. *Nat Genet* **53**, 135-142 (2021). https://doi.org:10.1038/s41588-020-00764-0
55 van Hasselt, J. G. C. *et al.* Transcriptomic profiling of human cardiac cells predicts protein kinase inhibitor-associated cardiotoxicity. *Nat Commun* **11**, 4809 (2020). https://doi.org:10.1038/s41467-020-18396-7
56 Burke, M. A., Cook, S. A., Seidman, J. G. & Seidman, C. E. Clinical and Mechanistic Insights Into the Genetics of Cardiomyopathy. *J Am Coll Cardiol* **68**, 2871-2886 (2016). https://doi.org:10.1016/j.jacc.2016.08.079
57 Hansen, J. *et al.* Systems pharmacology-based integration of human and mouse data for drug repurposing to treat thoracic aneurysms. *JCI Insight* **4** (2019). https://doi.org:10.1172/jci.insight.127652
58 Porter, K. E. & Turner, N. A. Cardiac fibroblasts: at the heart of myocardial remodeling. *Pharmacol Ther* **123**, 255-278 (2009). https://doi.org:10.1016/j.pharmthera.2009.05.002
59 Xiong, Y. *et al.* A Comparison of mRNA Sequencing with Random Primed and 3'-Directed Libraries. *Sci Rep* **7**, 14626 (2017). https://doi.org:10.1038/s41598-017-14892-x
60 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013). https://doi.org:10.1093/bioinformatics/bts635
61 Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* **47**, e47 (2019). https://doi.org:10.1093/nar/gkz114
62 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010). https://doi.org:10.1093/bioinformatics/btp616
63 Alter, O., Brown, P. O. & Botstein, D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* **97**, 10101-10106 (2000). https://doi.org:10.1073/pnas.97.18.10101
64 Hansen, J., Meretzky, D., Woldesenbet, S., Stolovitzky, G. & Iyengar, R. A flexible ontology for inference of emergent whole cell function from relationships between subcellular processes. *Sci Rep* **7**, 17689 (2017). https://doi.org:10.1038/s41598-017-16627-4
65 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* (2021). https://doi.org:10.1016/j.cell.2021.04.048
66 Asp, M. *et al.* A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. *Cell* **179**, 1647-1660 e1619 (2019). https://doi.org:10.1016/j.cell.2019.11.025
67 Lachmann, A. *et al.* ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* **26**, 2438-2444 (2010). https://doi.org:10.1093/bioinformatics/btq466
68 Lachmann, A. & Ma'ayan, A. KEA: kinase enrichment analysis. *Bioinformatics* **25**, 684-686 (2009). https://doi.org:10.1093/bioinformatics/btp026
69 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97 (2016). https://doi.org:10.1093/nar/gkw377

70 Herrmann, J. Adverse cardiac effects of cancer therapies: cardiotoxicity and arrhythmia. *Nat Rev Cardiol* **17**, 474-502 (2020). https://doi.org:10.1038/s41569-020-0348-1
71 Gammella, E., Maccarinelli, F., Buratti, P., Recalcati, S. & Cairo, G. The role of iron in anthracycline cardiotoxicity. *Front Pharmacol* **5**, 25 (2014). https://doi.org:10.3389/fphar.2014.00025
72 Cole, D. C. & Frishman, W. H. Cardiovascular Complications of Proteasome Inhibitors Used in Multiple Myeloma. *Cardiol Rev* **26**, 122-129 (2018). https://doi.org:10.1097/CRD.0000000000000183
73 Fabarius, A. *et al.* Centrosome aberrations after nilotinib and imatinib treatment in vitro are associated with mitotic spindle defects and genetic instability. *Br J Haematol* **138**, 369-373 (2007). https://doi.org:10.1111/j.1365-2141.2007.06678.x
74 Giehl, M. *et al.* Detection of centrosome aberrations in disease-unrelated cells from patients with tumor treated with tyrosine kinase inhibitors. *Eur J Haematol* **85**, 139-148 (2010). https://doi.org:10.1111/j.1600-0609.2010.01459.x
75 Sethunath, V. *et al.* Targeting the Mevalonate Pathway to Overcome Acquired Anti-HER2 Treatment Resistance in Breast Cancer. *Mol Cancer Res* **17**, 2318-2330 (2019). https://doi.org:10.1158/1541-7786.MCR-19-0756
76 Bienengraeber, M. *et al.* ABCC9 mutations identified in human dilated cardiomyopathy disrupt catalytic KATP channel gating. *Nat Genet* **36**, 382-387 (2004). https://doi.org:10.1038/ng1329
77 Dupont, E. *et al.* Altered connexin expression in human congestive heart failure. *J Mol Cell Cardiol* **33**, 359-371 (2001). https://doi.org:10.1006/jmcc.2000.1308
78 Perea-Gil, I. *et al.* Serine biosynthesis as a novel therapeutic target for dilated cardiomyopathy. *Eur Heart J* **43**, 3477-3489 (2022). https://doi.org:10.1093/eurheartj/ehac305
79 Gorabi, A. M. *et al.* Statins Attenuate Fibrotic Manifestations of Cardiac Tissue Damage. *Curr Mol Pharmacol* **14**, 782-797 (2021). https://doi.org:10.2174/1874467214666210210123206
80 Okuyama, H. *et al.* Statins stimulate atherosclerosis and heart failure: pharmacological mechanisms. *Expert Rev Clin Pharmacol* **8**, 189-199 (2015). https://doi.org:10.1586/17512433.2015.1011125
81 Rotariu, D. *et al.* Oxidative stress - Complex pathological issues concerning the hallmark of cardiovascular and metabolic disorders. *Biomed Pharmacother* **152**, 113238 (2022). https://doi.org:10.1016/j.biopha.2022.113238
82 Groenendyk, J., Sreenivasaiah, P. K., Kim, D. H., Agellon, L. B. & Michalak, M. Biology of endoplasmic reticulum stress in the heart. *Circ Res* **107**, 1185-1197 (2010). https://doi.org:10.1161/CIRCRESAHA.110.227033
83 Medamana, J., Clark, R. A. & Butler, J. Platelet-Derived Growth Factor in Heart Failure. *Handb Exp Pharmacol* **243**, 355-369 (2017). https://doi.org:10.1007/164_2016_80
84 Cheng, M., Park, H., Engelmayr, G. C., Moretti, M. & Freed, L. E. Effects of regulatory factors on engineered cardiac tissue in vitro. *Tissue Eng* **13**, 2709-2719 (2007). https://doi.org:10.1089/ten.2006.0414
85 Vantler, M. *et al.* PDGF-BB protects cardiomyocytes from apoptosis and improves contractile function of engineered heart tissue. *J Mol Cell Cardiol* **48**, 1316-1323 (2010). https://doi.org:10.1016/j.yjmcc.2010.03.008
86 McGrath, M. F., de Bold, M. L. & de Bold, A. J. The endocrine function of the heart. *Trends Endocrinol Metab* **16**, 469-477 (2005). https://doi.org:10.1016/j.tem.2005.10.007
87 Ding, K., Gui, Y., Hou, X., Ye, L. & Wang, L. Transient Receptor Potential Channels, Natriuretic Peptides, and Angiotensin Receptor-Neprilysin Inhibitors in Patients With Heart Failure. *Front Cardiovasc Med* **9**, 904881 (2022). https://doi.org:10.3389/fcvm.2022.904881
88 Ong, S. G. & Hausenloy, D. J. Hypoxia-inducible factor as a therapeutic target for cardioprotection. *Pharmacol Ther* **136**, 69-81 (2012). https://doi.org:10.1016/j.pharmthera.2012.07.005
89 Kubin, T. *et al.* The Role of Oncostatin M and Its Receptor Complexes in Cardiomyocyte Protection, Regeneration, and Failure. *Int J Mol Sci* **23** (2022). https://doi.org:10.3390/ijms23031811
90 Wang, P. *et al.* The alteration of Hippo/YAP signaling in the development of hypertrophic cardiomyopathy. *Basic Res Cardiol* **109**, 435 (2014). https://doi.org:10.1007/s00395-014-0435-8

91 Horn, M. A. & Trafford, A. W. Aging and the cardiac collagen matrix: Novel mediators of fibrotic remodelling. *J Mol Cell Cardiol* **93**, 175-185 (2016). https://doi.org:10.1016/j.yjmcc.2015.11.005
92 Baiocchini, A. *et al.* Extracellular Matrix Molecular Remodeling in Human Liver Fibrosis Evolution. *PLoS One* **11**, e0151736 (2016). https://doi.org:10.1371/journal.pone.0151736
93 Shangzu, Z. *et al.* Aquaporins: Important players in the cardiovascular pathophysiology. *Pharmacol Res* **183**, 106363 (2022). https://doi.org:10.1016/j.phrs.2022.106363

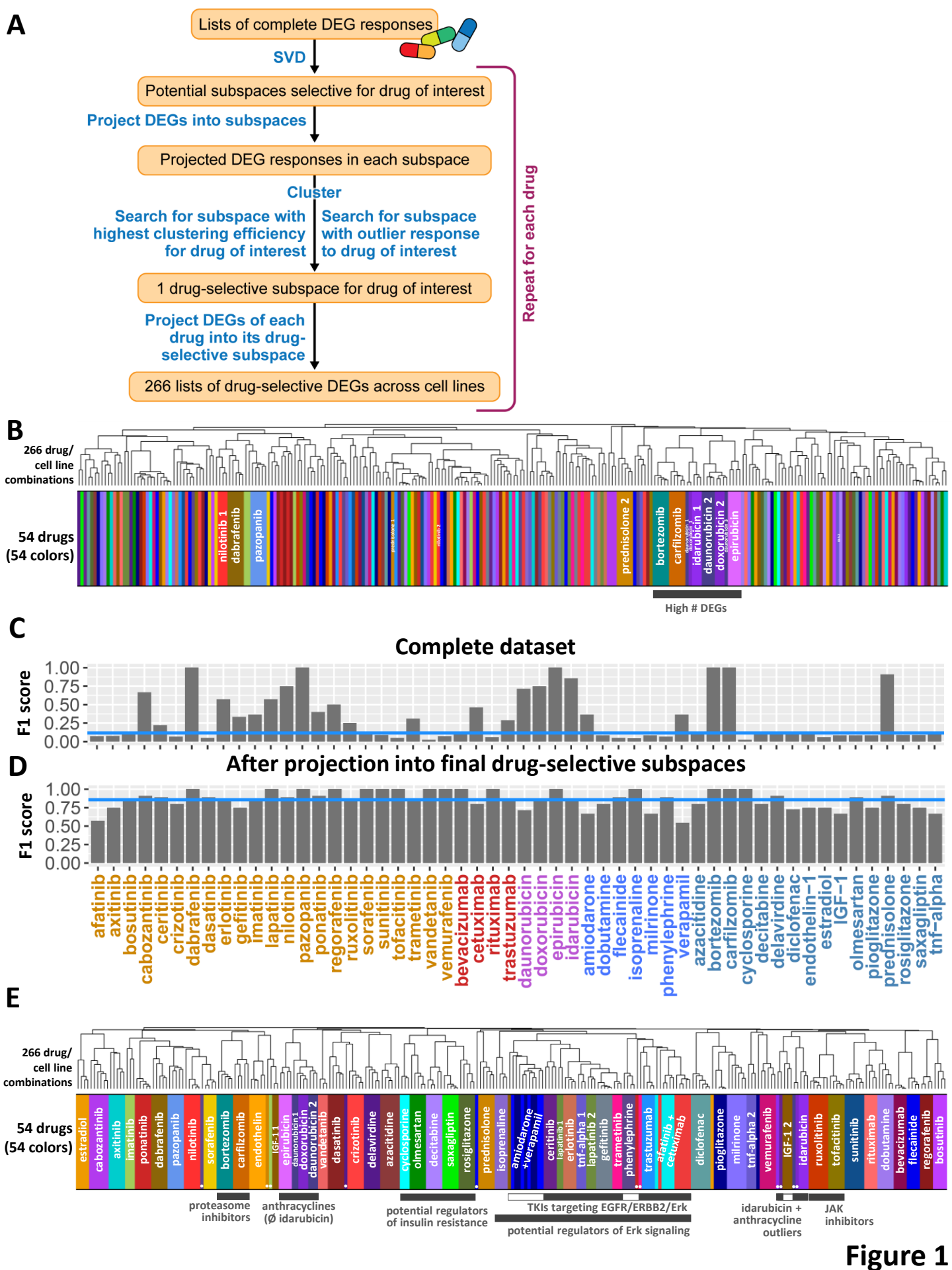

A
Lists of complete DEG responses
SVD
Potential subspaces selective for drug of interest
Project DEGs into subspaces
Projected DEG responses in each subspace
Cluster
Search for subspace with highest clustering efficiency for drug of interest
Search for subspace with outlier response to drug of interest
1 drug-selective subspace for drug of interest
Project DEGs of each drug into its drug-selective subspace
266 lists of drug-selective DEGs across cell lines
Repeat for each drug
B
266 drug/ cell line combinations
54 drugs (54 colors)
nilotinib 1
dabrafenib
pazopanib
prednisolone 2
bortezomib
carfilzomib
idarubicin 1
daunorubicin 2
doxorubicin 2
epirubicin
High # DEGs
C
Complete dataset
F1 score
1.00
0.75
0.50
0.25
0.00
D
After projection into final drug-selective subspaces
F1 score
1.00
0.75
0.50
0.25
0.00
afatinib
axitinib
bosutinib
cabozantinib
ceritinib
crizotinib
dabrafenib
dasatinib
erlotinib
gefitinib
imatinib
lapatinib
nilotinib
pazopanib
ponatinib
regorafenib
ruxolitinib
sorafenib
sunitinib
tofacitinib
trametinib
vandetanib
vemurafenib
bevacizumab
cetuximab
rituximab
trastuzumab
daunorubicin
doxorubicin
epirubicin
idarubicin
amiodarone
dobutamine
flecainide
isoprenaline
milrinone
phenylephrine
verapamil
azacitidine
bortezomib
carfilzomib
cyclosporine
decitabine
delavirdine
diclofenac
endothelin−1
estradiol
IGF−1
olmesartan
pioglitazone
prednisolone
rosiglitazone
saxagliptin
tnf−alpha
E
266 drug/ cell line combinations
54 drugs (54colors)
estradiol
cabozantinib
axitinib
imatinib
ponatinib
dabrafenib
pazopanib
nilotinib
sorafenib
bortezomib
carfilzomib
endothelin
IGF-1 1
epirubicin
daunorubicin 1
doxorubicin
daunorubicin 2
vandetanib
dasatinib
crizotinib
delavirdine
azacitidine
cyclosporine
olmesartan
decitabine
saxagliptin
rosiglitazone
prednisolone
isoprenaline
amiodarone +verapamil
ceritinib
lapatinib 1
erlotinib
tnf-alpha 1
lapatinib 2
gefitinib
trametinib
phenylephrine
trastuzumab
afatinib +
cetuximab
diclofenac
pioglitazone
milrinone
tnf-alpha 2
vemurafenib
IGF-1 2
idarubicin
ruxolitinib
tofacitinib
sunitinib
rituximab
dobutamine
bevacizumab
flecainide
regorafenib
bosutinib
proteasome inhibitors
anthracyclines (Ø idarubicin)
potential regulators of insulin resistance
TKIs targeting EGFR/ERBB2/Erk
potential regulators of Erk signaling
idarubicin + anthracycline outliers
JAK inhibitors

Figure 1

Figure 1. Singular value decomposition identifies drug-selective gene expression responses. 266 samples, each representing a unique combination of one out of three to six hiPSC-derived cardiomyocyte cell lines treated with one out of 54 drugs, were subjected to bulk RNAseq analysis, followed by calculation of 266 lists of drug-induced, differentially expressed genes (DEGs). **(A)** Our computational pipeline uses singular value decomposition (SVD) to identify drug-selective gene expression responses that are components of the complete responses. See methods section and Suppl. Figure 4 for details. **(B)** Pairwise correlation analysis followed by hierarchical clustering reveals that most complete DEGs are dominated by cell-line-selective effects hiding drug-selective effects. See Suppl. Figure 5 for larger dendrogram. **(C)** We used F1 score statistics to document the clustering efficiency, i.e. how close samples treated with the same drug cluster together. Low clustering efficiencies quantitatively describe the finding that only a few complete DEG responses are dominated by drug-selective effects. orange: Small molecule kinase inhibitors (KI), red: monoclonal antibodies against KIs, purple: anthracyclines, blue: cardiac acting drugs, gray-blue: noncardiac acting drugs. **(D)** Projection of the complete DEG responses into each of the identified 54 drug-selective subspaces greatly increases the clustering efficiencies for all 54 drugs. **(E)** Pairwise correlation of 266 merged drug-selective responses, followed by hierarchical clustering, documents that SVD allows identification of components induced in all cell lines treated with the same drug. Clusters that contain drugs with similar mechanisms are labeled with gray bars. White insets indicate drugs in those clusters that are not part of the outlined mechanisms written below the bars. White circles indicate outlier samples that were identified by our pipeline and cluster as outliers in the merged dataset as well. See Suppl. Figure 13 for larger dendrogram.

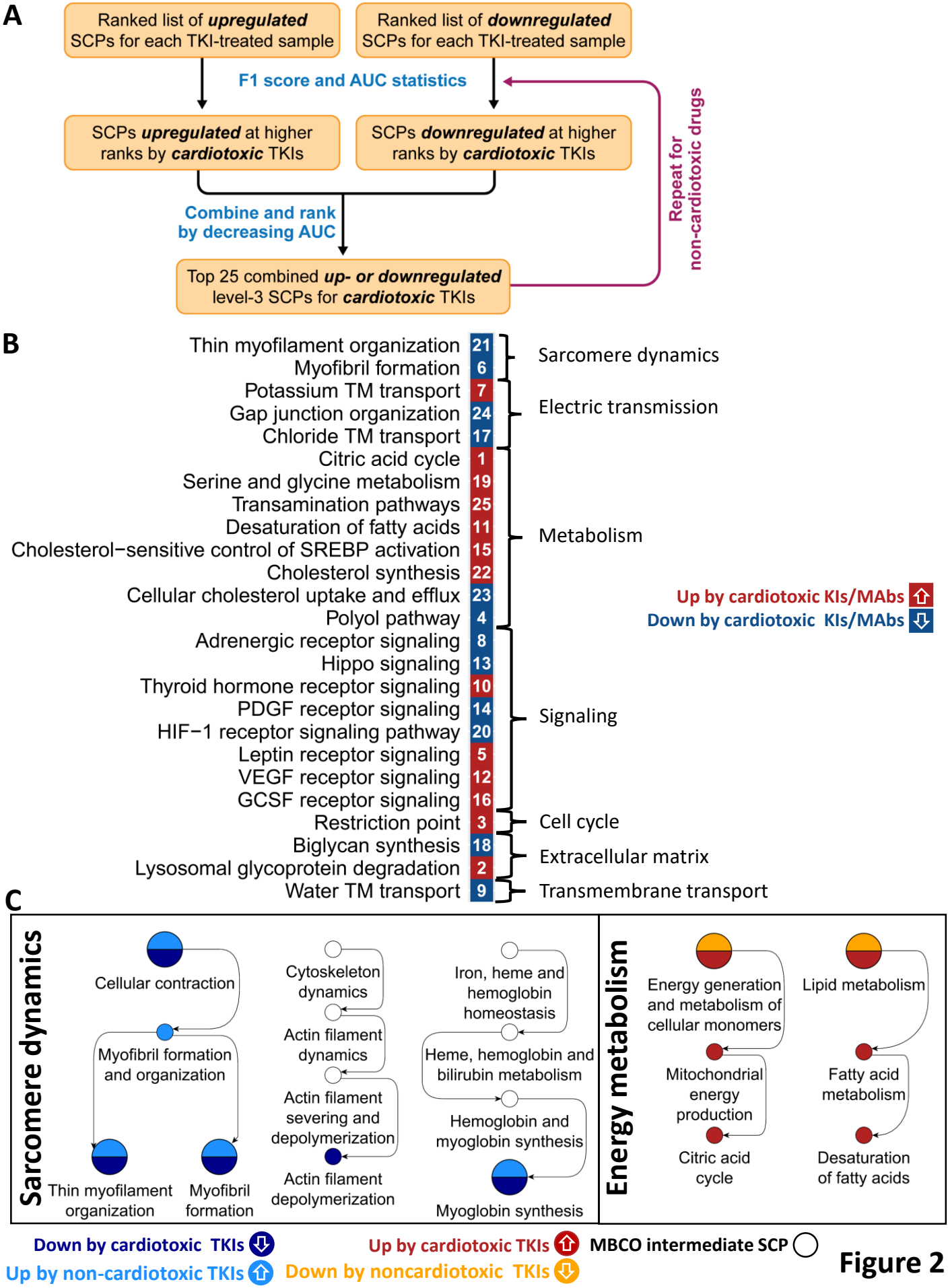
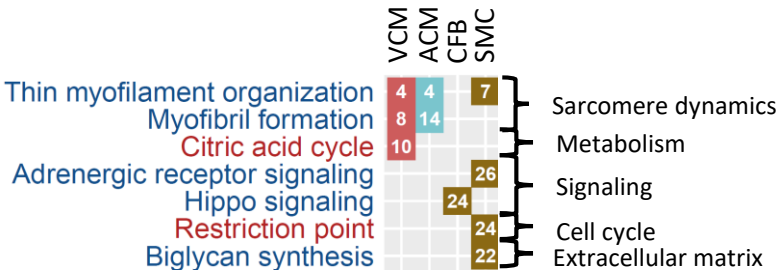

A
Ranked list of upregulated SCPs for each TKI-treated sample
Ranked list of downregulated SCPs for each TKI-treated sample
F1 score and AUC statistics
SCPs upregulated at higher ranks by cardiotoxic TKIs
SCPs downregulated at higher ranks by cardiotoxic TKIs
Combine and rank by decreasing AUC
Top 25 combined up- or downregulated level-3 SCPs for cardiotoxic TKIs
Repeat for non-cardiotoxic drugs
B
Thin myofilament organization 21
Myofibril formation 6
Sarcomere dynamics
Potassium TM transport 7
Gap junction organization 24
Chloride TM transport 17
Electric transmission
Citric acid cycle 1
Serine and glycine metabolism 19
Transamination pathways 25
Desaturation of fatty acids 11
Cholesterol−sensitive control of SREBP activation 15
Cholesterol synthesis 22
Cellular cholesterol uptake and efflux 23
Polyol pathway 4
Metabolism
Up by cardiotoxic KIs/MAbs
Down by cardiotoxic KIs/MAbs
Adrenergic receptor signaling 8
Hippo signaling 13
Thyroid hormone receptor signaling 10
PDGF receptor signaling 14
HIF−1 receptor signaling pathway 20
Leptin receptor signaling 5
VEGF receptor signaling 12
GCSF receptor signaling 16
Signaling
Restriction point 3
Cell cycle
Biglycan synthesis 18
Lysosomal glycoprotein degradation 2
Extracellular matrix
Water TM transport 9
Transmembrane transport
C
Sarcomere dynamics
Cellular contraction
Myofibril formation and organization
Thin myofilament organization
Myofibril formation
Cytoskeleton dynamics
Actin filament dynamics
Actin filament severing and depolymerization
Actin filament depolymerization
Iron, heme and hemoglobin homeostasis
Heme, hemoglobin and bilirubin metabolism
Hemoglobin and myoglobin synthesis
Myoglobin synthesis
Energy metabolism
Energy generation and metabolism of cellular monomers
Mitochondrial energy production
Citric acid cycle
Lipid metabolism
Fatty acid metabolism
Desaturation of fatty acids
Down by cardiotoxic TKIs
Up by cardiotoxic TKIs
MBCO intermediate SCP
Up by non-cardiotoxic TKIs
Down by noncardiotoxic TKIs

Figure 2

Figure 2. Potential subcellular processes indicative of drug induced cardiotoxicity. Up- and downregulated genes among the top 600 drug-selective gene expression profiles were subjected to pathway enrichment analysis using the Molecular Biology of the Cell Ontology (MBCO). Significantly up- or downregulated subcellular processes (SCPs) (p-value ≤ 0.05) were ranked separately by significance for each sample and SCP level. **(A)** We developed a computational pipeline that screens for SCPs that are up- or downregulated at higher ranks by cardiotoxic (or non-cardiotoxic) TKIs. Up- and downregulated SCPs are ranked based on F1 and Area under the Curve statistics. See methods for details. **(B)** Top 25 level-3 SCPs predicted for the cardiotoxic TKIs were grouped based on the higher-level functions. **(C)** SCPs identified for cardiotoxic and non-cardiotoxic TKIs for all levels were integrated into the MBCO hierarchy. Selected branches are shown.

A


VCM
ACM
CFB
SMC
Thin myofilament organization
Myofibril formation
Citric acid cycle
Adrenergic receptor signaling
Hippo signaling
Restriction point
Biglycan synthesis
4
4
7
8
14
10
26
24
24
22
Sarcomere dynamics
Metabolism
Signaling
Cell cycle
Extracellular matrix

B

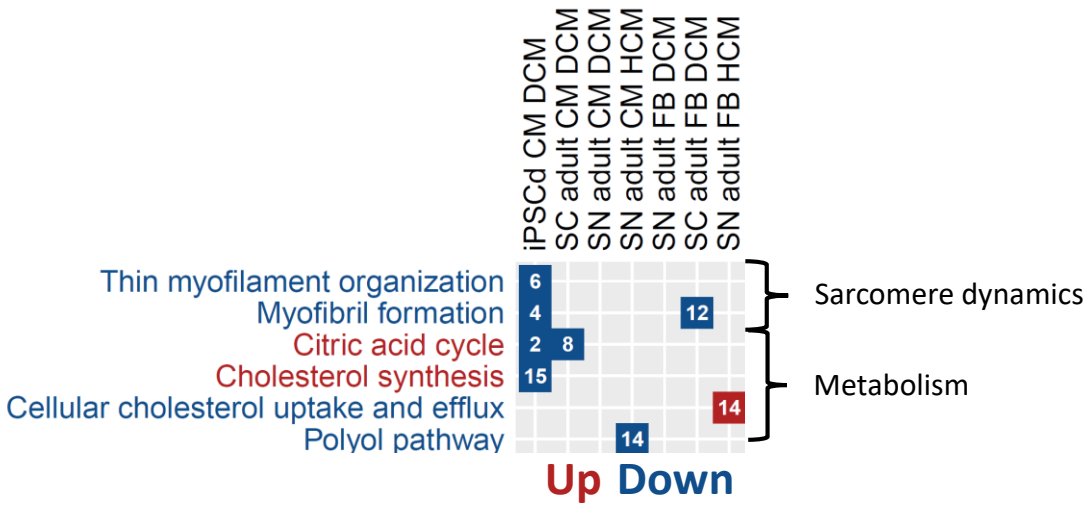

iPSCd CM DCM
SC adult CM DCM
SN adult CM DCM
SN adult CM HCM
SN adult FB DCM
SC adult FB DCM
SN adult FB HCM
Thin myofilament organization
Myofibril formation
Citric acid cycle
Cholesterol synthesis
Cellular cholesterol uptake and efflux
Polyol pathway
6
4
12
2
8
15
14
14
Sarcomere dynamics
Metabolism
Up
Down

Figure 3

Figure 3. SCPs can be mapped to cellular subtypes and known cardiomyopathy disease mechanisms. (A) We subjected marker genes for ventricular and atrial cardiomyocytes (VCM and ACM, respectively), cardiac fibroblasts (CFB) and smooth muscle cells (SMC) obtained from single nucleus RNAseq of the adult human heart [28] to pathway enrichment analysis using MBCO. Significant SCPs of each cell type (p-value ≤ 0.05) were ranked by significance (numbers in the diagram). Names of SCPs whose higher and lower activities favor a cardiotoxic response are colored red and blue, respectively. **(B)** Differentially expressed genes (DEGs) in heart cells obtained by single cell (SC) [13] or nucleus (SN) [14] RNAseq from patients with Dilated or Hypertrophic Cardiomyopathy (DCM or HCM, respectively) as well as in hiPSC-derived cardiomyocytes obtained from an infant patient with DCM (GSE184899) were subjected to pathway enrichment analysis. Significantly up- (red fields) or downregulated (blue fields) SCPs of each cell type (p-value ≤ 0.05) were ranked by significance (numbers in the diagram).

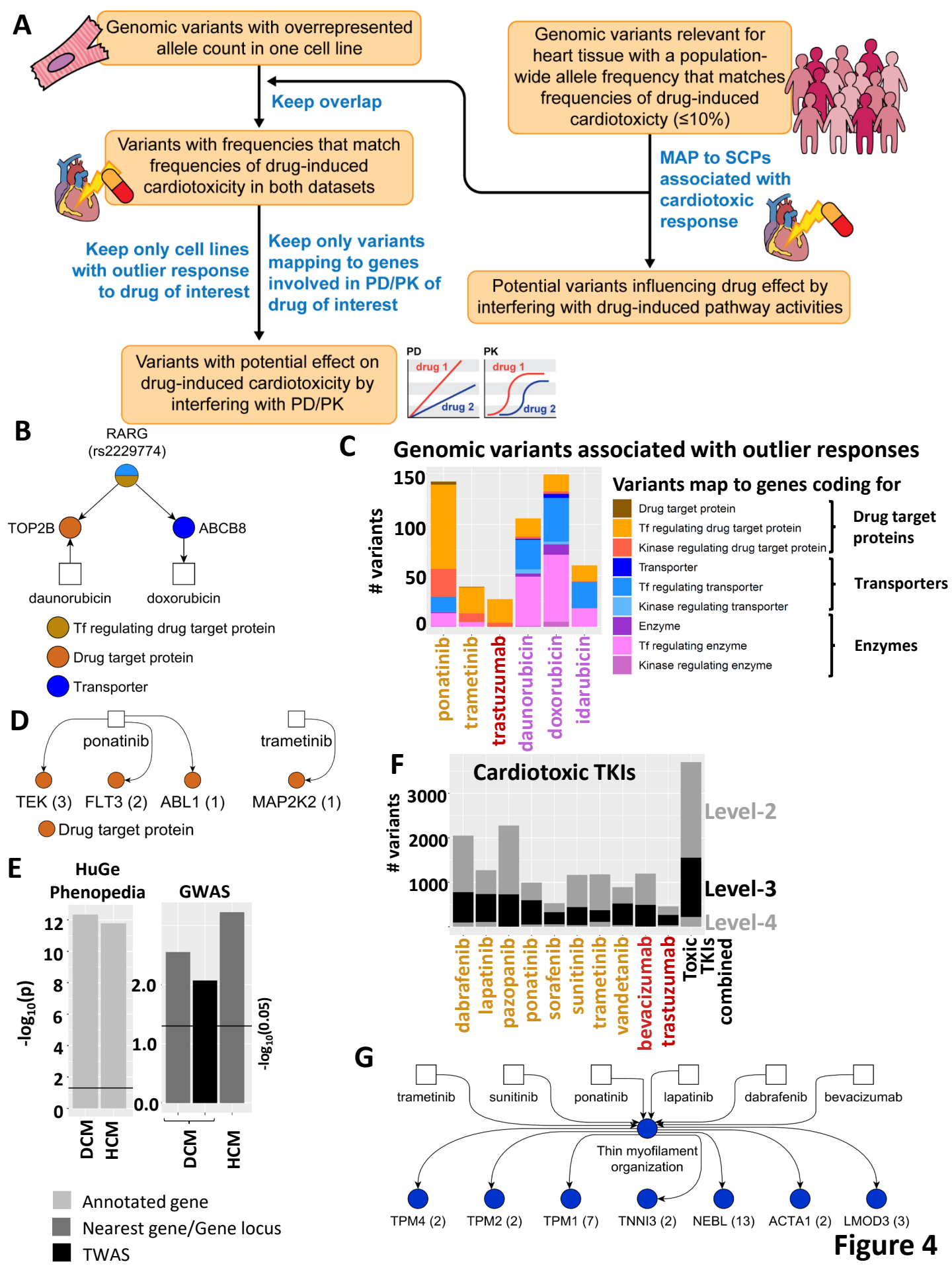

A
Genomic variants with overrepresented allele count in one cell line
Keep overlap
Variants with frequencies that match frequencies of drug-induced cardiotoxicity in both datasets
Keep only cell lines with outlier response to drug of interest
Keep only variants mapping to genes involved in PD/PK of drug of interest
Variants with potential effect on drug-induced cardiotoxicity by interfering with PD/PK
PD
drug 1
drug 2
PK
drug 1
drug 2
Genomic variants relevant for heart tissue with a population-wide allele frequency that matches frequencies of drug-induced cardiotoxicity (≤10%)
MAP to SCPs associated with cardiotoxic response
Potential variants influencing drug effect by interfering with drug-induced pathway activities
B
RARG (rs2229774)
TOP2B
ABCB8
daunorubicin
doxorubicin
Tf regulating drug target protein
Drug target protein
Transporter
C
Genomic variants associated with outlier responses
variants
150
100
50
0
ponatinib
trametinib
trastuzumab
daunorubicin
doxorubicin
idarubicin
Variants map to genes coding for
Drug target protein
Tf regulating drug target protein
Kinase regulating drug target protein
Transporter
Tf regulating transporter
Kinase regulating transporter
Enzyme
Tf regulating enzyme
Kinase regulating enzyme
Drug target proteins
Transporters
Enzymes
D
ponatinib
trametinib
TEK (3)
FLT3 (2)
ABL1 (1)
MAP2K2 (1)
Drug target protein
E
HuGe Phenopedia
GWAS
-log10(p)
12
10
8
6
4
2
0
2.0
1.0
0.0
-log10(0.05)
DCM
HCM
DCM
HCM
Annotated gene
Nearest gene/Gene locus
TWAS
F
Cardiotoxic TKIs
variants
3000
2000
1000
Level-2
Level-3
Level-4
dabrafenib
lapatinib
pazopanib
ponatinib
sorafenib
sunitinib
trametinib
vandetanib
bevacizumab
trastuzumab
Toxic TKIs combined
G
trametinib
sunitinib
ponatinib
lapatinib
dabrafenib
bevacizumab
Thin myofilament organization
TPM4 (2)
TPM2 (2)
TPM1 (7)
TNNI3 (2)
NEBL (13)
ACTA1 (2)
LMOD3 (3)

Figure 4

Figure 4. Identification of genomic variants that are potentially associated with a cardiotoxic response. Whole genome sequencing of our six cell lines [16] was used to identify alleles in our cell lines at known variant positions. **(A)** Pipeline for identification of potential genomic variants involved in PK/PD or induced SCP activities for a drug of interest. **(B)** The identified variant rs2229774 in cell line MSN09 maps to the coding region of the transcription factor RARG regulating the expression of TOP2B and ABCB8, both involved in PK/PD of daunorubicin and doxorubicin that induce an outlier response in this cell line. **(C)** In total, we identified 213 and 201 potential variants associated with TKI- or anthracycline-induced cardiotoxicity by interference with PK/PD mechanisms, respectively. Variants mapping to multiple gene classes are split equally among them to prevent double counting. **(D)** Numbers in brackets indicate how many variants map to individual drug target proteins of cardiotoxic TKIs. **(E)** We compared the overlap of identified SCP genes associated with a cardiotoxic or non-cardiotoxic response to genes associated with inherited DCM or HCM, either within the HuGe Phenopedia database or identified in GWAS. In the DCM, GWAS identified variants were associated with either the nearest genes or genes with the connected strongest cis-regulated expression as predicted by transcription-wide-association studies (TWAS). **(F)** Variants with a low frequency in the whole population that are part of e/sQTLs or lie within coding regions were mapped to up- or downregulated level -2, -3 and -4 SCPs that we predicted as indicative for TKI-induced cardiotoxicity. Variants that map to identified SCPs of multiple levels are only counted for the lowest level SCPs (higher level numbers) to prevent double counting. **(G)** Up- and downregulated SCPs associated with a cardiotoxic response were mapped back to the cardiotoxic TKIs that induce them. Numbers in brackets show identified variants for each SCP gene.

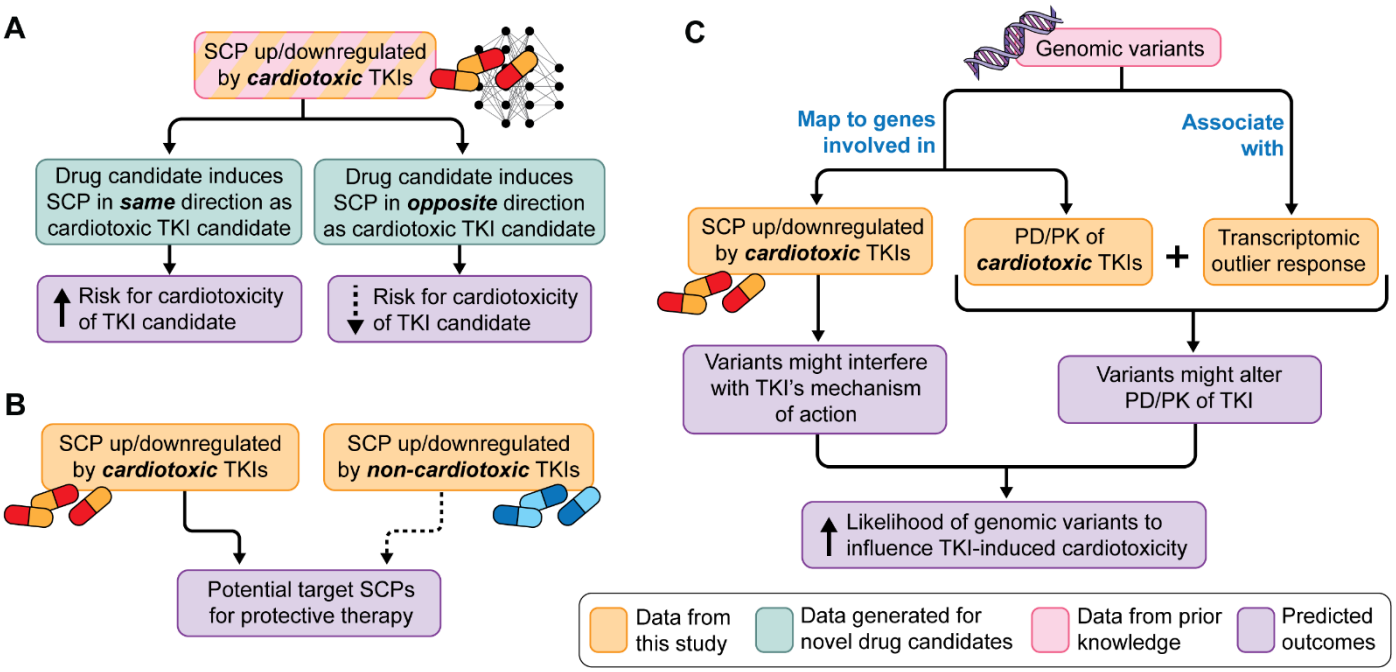

A
SCP up/downregulated by cardiotoxic TKIs
Drug candidate induces SCP in same direction as cardiotoxic TKI candidate
Drug candidate induces SCP in opposite direction as cardiotoxic TKI candidate
Risk for cardiotoxicity of TKI candidate
Risk for cardiotoxicity of TKI candidate
B
SCP up/downregulated by cardiotoxic TKIs
SCP up/downregulated by non-cardiotoxic TKIs
Potential target SCPs for protective therapy
C
Genomic variants
Map to genes involved in
Associate with
SCP up/downregulated by cardiotoxic TKIs
PD/PK of cardiotoxic TKIs
+
Transcriptomic outlier response
Variants might interfere with TKI's mechanism of action
Variants might alter PD/PK of TKI
Likelihood of genomic variants to influence TKI-induced cardiotoxicity
Data from this study
Data generated for novel drug candidates
Data from prior knowledge
Predicted outcomes

Figure 5

Figure 5. Potential use of cell based transcriptomic data for drug therapy induced adverse events. The flowcharts summarize the how integration of experimentally gathered transcriptomic data with publicly available pathway and genomic data bases can **(A**) help predict toxicity of drug candidates, (**B**) identify potential new drug targets to mitigate cardiotoxicity and (**C)** enable design of clinical studies to associate genomic variants with cardiotoxicity propensity.