HEALTH AND MEDICINE

Rare genetic variability in human drug target genes modulates drug response and can guide precision medicine

Yitian Zhou[1], Gabriel Herras Arribas[1], Ainoleena Turku[1,2], Tuuli Jürgenson[3], Souren Mkrtchian[1], Kristi Krebs[3,4], Yi Wang[5], Barbora Svobodova[6], Lili Milani[3], Gunnar Schulte[1], Jan Korabecny[6], Stefano Gastaldello[1], Volker M. Lauschke[1,7]*

Interindividual variability in drug response constitutes a major concern in pharmacotherapy. While polymorphisms in genes involved in drug disposition have been extensively studied, drug target variability remains underappreciated. By mapping the genomic variability of all human drug target genes onto high-resolution crystal structures of drug target complexes, we identified 1094 variants localized within 6 Å of drug-binding pockets and directly affecting their geometry, topology, or physicochemical properties. We experimentally show that binding site variants affect pharmacodynamics with marked drug- and variant-specific differences. In addition, we demonstrate that a common *BCHE* variant confers resistance to tacrine and rivastigmine, which can be overcome by the use of derivatives based on squaric acid scaffolds or tryptophan conjugation. These findings underscore the importance of genetic drug target variability and demonstrate that integration of genomic data and structural information can inform personalized drug selection and genetically guided drug development to overcome resistance.

INTRODUCTION

The response to medications differs drastically between individuals with many patients not experiencing the intended treatment outcomes due to the manifestation of adverse drug reactions (ADRs) or insufficient treatment response. Lack of response constitutes the most common reason for project terminations in drug development, with more than half of all candidate drugs that enter clinical trials failing because of efficacy issues (*1*, *2*). In addition, ethnogeographic or interindividual variability in drug response is also frequently observed for marketed drugs (*3*, *4*). Overall, it is estimated that 10 to 45% of patients do not respond appropriately to pharmacological treatment (*5*), and the fraction of nonresponders can be even higher for certain diseases, such as major depressive disorder where more than half of all patients do not respond to first-line antidepressant therapy (*6*).

Genetic polymorphisms in genes involved in drug disposition, such as drug-metabolizing enzymes and drug transporters, have long been described as biomarkers to predict therapeutic efficacy or ADR risks (*7*, *8*). Well-established examples include *DPYD* genotyping to guide fluorouracil treatment (*9*), *CYP2C19* genotype–guided antiplatelet therapy (*10*), and *CYP2D6* profiling to derive therapeutic recommendations for tamoxifen (*11*). In contrast, how polymorphisms in genes encoding drug targets affect drug response is less studied, at least in part, because of their extensive molecular diversity (*12*, *13*). While the frequencies of loss-of-function variation in human drug targets have been recently mapped (*14*), the extent to which drug target variability affects drug efficacy has not been evaluated.

For a given drug to be effective, the parent molecule or one or more of its metabolites must bind to the target protein and modulate its activity. Genetic variations in drug target genes can affect expression levels or functionality of the gene product, thus altering the stoichiometry of drug molecules to functional target proteins at a given dose. Furthermore, variants that affect amino acids in the drug-binding pocket can directly influence drug-target associations by altering pocket geometry, topology, or physiochemical interactions. Recent seminal work showed that naturally occurring missense variants in drug-binding sites within genes encoding members of the G protein–coupled receptor (GPCR) family of drug targets constitute important determinants of drug response whose consideration could increase prescription precision and reduce socioeconomic costs by up to 500 million British pounds annually (*15*). However, how drug target variability in genes other than GPCRs might modulate drug response has not been investigated.

Here, we systematically profiled the genetic variability of all human protein drug targets based on whole-exome and whole-genome sequencing data from 138,632 unrelated individuals across seven ethnogeographic groups. In total, we identified 479,860 naturally occurring exonic variants, of which 82,884 were predicted to affect gene function using a set of 17 partly orthogonal computational algorithms. By leveraging available high-resolution crystal structures, we identified 1094 variations that affected amino acids within 6 Å of the binding drug molecule. To evaluate the relevance of these findings, we experimentally validated the effects of variants on the response to drugs across diverse classes, including angiotensin-converting enzyme (ACE) inhibitors (ACEis), cholinesterase inhibitors, and microtubule polymerization blockers. Binding site variability significantly affected the activity of ACEis with >10-fold differences in response between tested ACE variants. Similarly, tacrine strongly inhibited reference sequence cholinesterase, whereas the common D98G binding site variant (minor allele frequency = 1.2%) was resistant

[1]Department of Physiology and Pharmacology, Karolinska Institutet, 171 77 Stockholm, Sweden. [2]Orion Pharma R&D, P.O. Box 65 (Orionintie 1), FI-02101 Espoo, Finland. [3]Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia. [4]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. [5]Pharmaceutical Informatics Institute, College of Pharmaceutical Sciences, Zhejiang University, 310058 Hangzhou, China. [6]Biomedical Research Centre, University Hospital Hradec Králové, Sokolská 581, 500 05 Hradec Králové, Czech Republic. [7]Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology, Stuttgart, Germany.
*Corresponding author. Email: volker.lauschke@ki.se

to tacrine. By using a library of tacrine-squaramide derivatives, we identified compounds with strong inhibitory effects on both reference and mutant enzyme, thus overcoming drug resistance. Combined, these data suggest that drug target variability provides an important determinant of interindividual variability in drug response and that integration of genomic and structural information can provide promising means to guide personalized drug selection to optimize drug response.

RESULTS

Overview of the genetic variability profile in human drug targets

To characterize human drug target variability, we first obtained all drug targets ($n = 893$) for 1578 U.S. Food and Drug Administration (FDA)–approved drugs (*16*). We excluded all nonhuman (e.g., directly targeting pathogens; $n = 129$) and nonprotein targets (e.g.,

targeting DNA or chelating metal ions; $n = 10$) as well as all target variants that only arise because of somatic mutations ($n = 60$), resulting in a total of 606 target proteins targeted by 1155 drugs that give rise to a total of unique 3346 drug-target pairs (Fig. 1A and table S1). The majority of these drugs target the nervous system ($n = 242$; 21%), alimentary tract and metabolic system ($n = 178$; 15%), and cardiovascular system ($n = 171$; 15%), whereas only few anti-infectives ($n = 4$; 0.3%) and antiparasitic drugs ($n = 3$; 0.3%) had human host cell protein targets (Fig. 1B). Most human drug targets were enzymes ($n = 204$; 34%), ion channels ($n = 131$; 22%), and membrane receptors ($n = 113$; 19%). Notably, 45.7% of all drugs are associated with different target-encoding genes (Fig. 1C). To corroborate the accuracy of these annotations, we matched the target annotations by Santos *et al.* (*16*) against the Therapeutic Target Database (*17*). Notably, we find good overlap between both resources with 86.2% of drugs being annotated with the same target in both databases, albeit sometimes different isoforms (table S2). However,
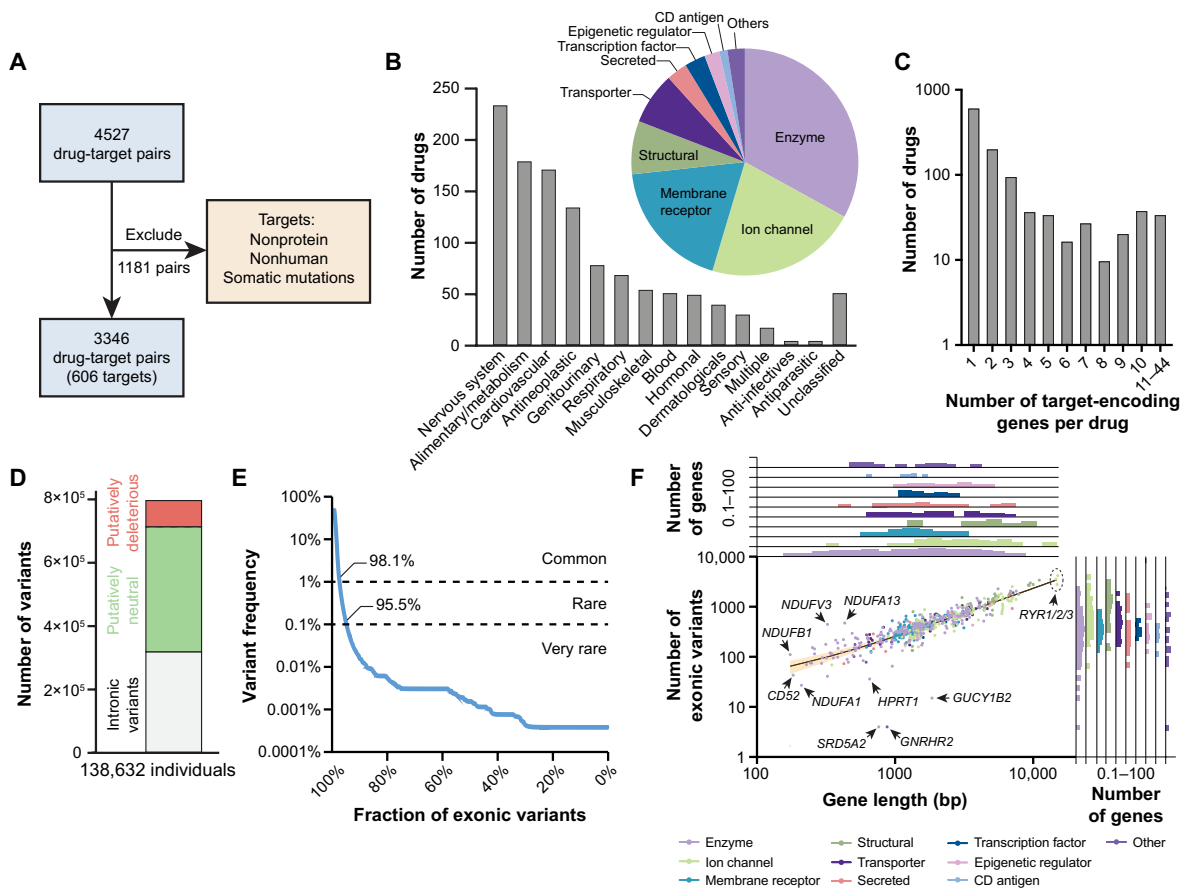
Fig. 1. The genetic landscape of human drug targets. (**A**) After exclusion of drugs with nonprotein or nonhuman targets or drugs targeting only specific somatic mutations, we obtained 606 genes encoding the target proteins of 1155 FDA-approved drugs, resulting in a total of 3346 unique drug-target pairs. (**B**) The analyzed drugs were distributed across anatomical therapeutic chemical (ATC) classifications. The most common targets were enzymes and ion channels followed by membrane receptors and structural components. (**C**) Column plot showing the number of target encoding genes per drug. (**D**) Across 138,632 individuals, we identified a total of 798,842 variants of which 479,860 were exonic. Using stringent computational assessments of pathogenicity (see Materials and Methods), 82,884 variants were identified as putatively deleterious. (**E**) The majority of exonic variants in drug target genes are rare with minor allele frequency (MAF) <1%. (**F**) Variant numbers differed >100-fold between drug targets, primarily because of differences in gene length ($R^2 = 0.87$). Confidence bands (95%) are shown in yellow. The gene density distributions for gene length and variant number are shown across the different protein classes as histograms on top and on the right of the scatter plot, respectively.

for some drugs, particularly those targeting the central nervous system, the exact mechanism of action is not firmly established.

Next, we analyzed the genetic variability in all 606 human drug target genes. Using sequencing data from 138,632 individuals, we identified a total of 798,842 variants, of which 479,860 were exonic (Fig. 1D). Computational predictions further indicated that 82,884 (24.3%) exonic variants were likely to have impact on gene functions. The vast majority of these variants were rare (98.1%) or very rare (95.9%) with minor allele frequencies <1 and 0.1%, respectively (Fig. 1E). Furthermore, more than half of all exonic variants (n = 246,768 variants; 51.4%) were singletons. Target genes that encode structural proteins harbored the highest number of exonic variants (on average, 820 missense variants and 70 nonsense variants per target), whereas genes that encode CD antigens were the least variable drug targets (on average, 159 missense variants and 11 nonsense variants per target; fig. S1). Few proteins were devoid of missense variants, including *SRD5A2* (targeted by dutasteride and finasteride), *GUCY1B2* (various guanylate cyclase modulators), and *GNRHR2* (danazol); no nonsense variants were found in *NDUFA1* (metformin), *PRPS1* (nelarabine), *KCNA4* (dalfampridine and guanidine), *CACNG8* (bepridil), and *CD52* (alemtuzumab).

The majority of drug target genes harbored more than 100 variants with the highest number of variants found in the ryanodine receptor family (*RYR1*, *RYR2*, and *RYR3*; n > 2700 variants each; Fig. 1F). The calcium channels RyR1 and RyR3 (encoded by *RYR1* and *RYR3*) are targeted by the muscle relaxant dantrolene, whereas RyR2 (*RYR2*) is targeted by the vasodilator hydralazine. In contrast, only five variants were found in *SRD5A2* and *GNRHR2*, which encode the targets of the antiandrogens dutasteride and finasteride and the antiestrogen danazol, respectively.

While the number of variants correlated overall well with gene length [coefficient of determination (R^2) = 0.87], 27 targets differed more than twofold from the average variability (Fig. 1F). Even within closely related targets, the extent of genetic variability could differ extensively. For instance, 444 and 473 variants were identified in the NADH (reduced form of nicotinamide adenine dinucleotide) dehydrogenase complex components *NDUFV3* [1.4 variants per kilo–base pair (kbp)] and *NDUFA13* (1.1 variants/kbp), whereas *NDUFA1* only harbored 27 variants (0.1 variants/kbp).

Systematic analysis of drug-binding site variability

We hypothesized that variations localized directly within the drug-binding pockets might be most likely to affect drug action by affecting pocket geometry or topology or by altering the physiochemical interactions between drug and target, which might affect the thermodynamic stability of the protein-ligand complex. To estimate the impact of genetically encoded drug-binding site variability, we mapped all variants onto the corresponding three-dimensional protein structures of the respective drug target. To this end, we first extracted all drug targets for which high-resolution crystal structures of the drug bound to its human target were available or could be modeled with high confidence (see Materials and Methods for details), thus assuring exact drug- and target-specific binding site information (table S3).

In total, such high-quality structural information of targets complexed with approved drugs was available for 638 drug-target pairs (486 drugs and 110 targets; Fig. 2A). Across these targets, 1094 naturally occurring genetic variations resulted in amino acid exchanges in close proximity (≤6 Å) of the bound drug (table S4). Alanine to

threonine (A > T, 2.5%) or valine (A > V, 2.5%) as well as arginine to cysteine (R > C, 2.7%) or histidine (R > H, 2.7%) were the most common substitution types (fig. S2). In contrast, the inverse amino acid exchanges were less common with the most notable differences being observed for cysteine to arginine, which was >10-fold less abundant (C > R, 0.2%). Overall, most variants were identified in enzyme-binding sites with an aggregated allele frequency of 3.6% (n = 582), whereas the variability in the binding pockets of CD antigens (n = 15) and epigenetic regulators (n = 3) was very low (aggregated allele frequency < 0.1%; Fig. 2B). While ion channels were the second most common drug target class, the total number (n = 34) and the aggregated frequency (<0.1%) of variants in the binding sites of drugs targeting in ion channels were very low. Most binding site variants affected targets of antineoplastic and immunomodulating agents (aggregated allele frequency = 5.6%), whereas only few variants were found in the binding pockets of proteins targeted by antiparasitic drugs or hormonal modulators (aggregated allele frequency < 0.1%; Fig. 2C). Across all genes variability in binding sites was substantially lower than outside of binding pockets (P < 0.0001, paired t test), suggesting that drug-binding sites are enriched in functionally important domains (Fig. 2D). Furthermore, variability was slightly but significantly lower in the binding pocket of small molecules compared to the epitopes recognized by therapeutic antibodies (fig. S3). In total, the carrier frequency of all 1094 binding site variants was 17.5%, suggesting that approximately one in six individuals carries at least one missense variant in the binding sites, potentially affecting the binding of at least one FDA-approved medication (Fig. 2E).

The most frequent binding site variants were R241W in *ITGAL* (rs1064524), D98G in *BCHE* (rs1799807), and T164I in *ADRB2* (rs1800888) within the binding pockets of efalizumab, tacrine, and albuterol, respectively (Fig. 2F). To evaluate whether these variants might manifest in phenotypes, we conducted a phenome-wide association study (PheWAS) analysis using genetic data linked with longitudinal electronic medical records from 51,138 individuals of the Estonian Biobank [Fig. 2G and (*18*)]. Notably, we found that rs1064524 in *ITGAL* [minor allele frequency (MAF)$_{Biobank}$ = 4.3%] associated with phenome-wide significance (Bonferroni correction) with pleural plaques (β = 1.9; P = 7.5 × 10^{-7}) and diseases of the thymus (β = 1.5; P = 4.5 × 10^{-5}). These results suggest that rs1064524 localizes to a functionally important domain and modulates protein function. By contrast, the other variants were too rare in the Estonian population to allow for reliable PheWAS. Hence, this association provides proof of concept that the integration of structural and genetic data can flag functionally relevant variations, likely due to substantial overlap of drug-binding pockets with active sites or structurally important protein domains.

Ethnogeographic differences in drug-binding site variability

Next, we interrogated the genetic variability in drug targets stratified by ethnogeographic groups, as interethnic differences can contribute considerably to differences in pharmacological response. By analyzing drug-binding site variability across seven major human populations, we found that drug-binding site variation differed approximately threefold across groups (Fig. 3A). Variability was highest in individuals of African ancestry where 1 in 4 individuals harbored at least one binding pocket variant and lowest in East Asian population with one variant per 14 individuals.

Among the most abundant variants, large ethnogeographic differences were observed. While rs1064524 in *ITGAL* and rs1799807 in
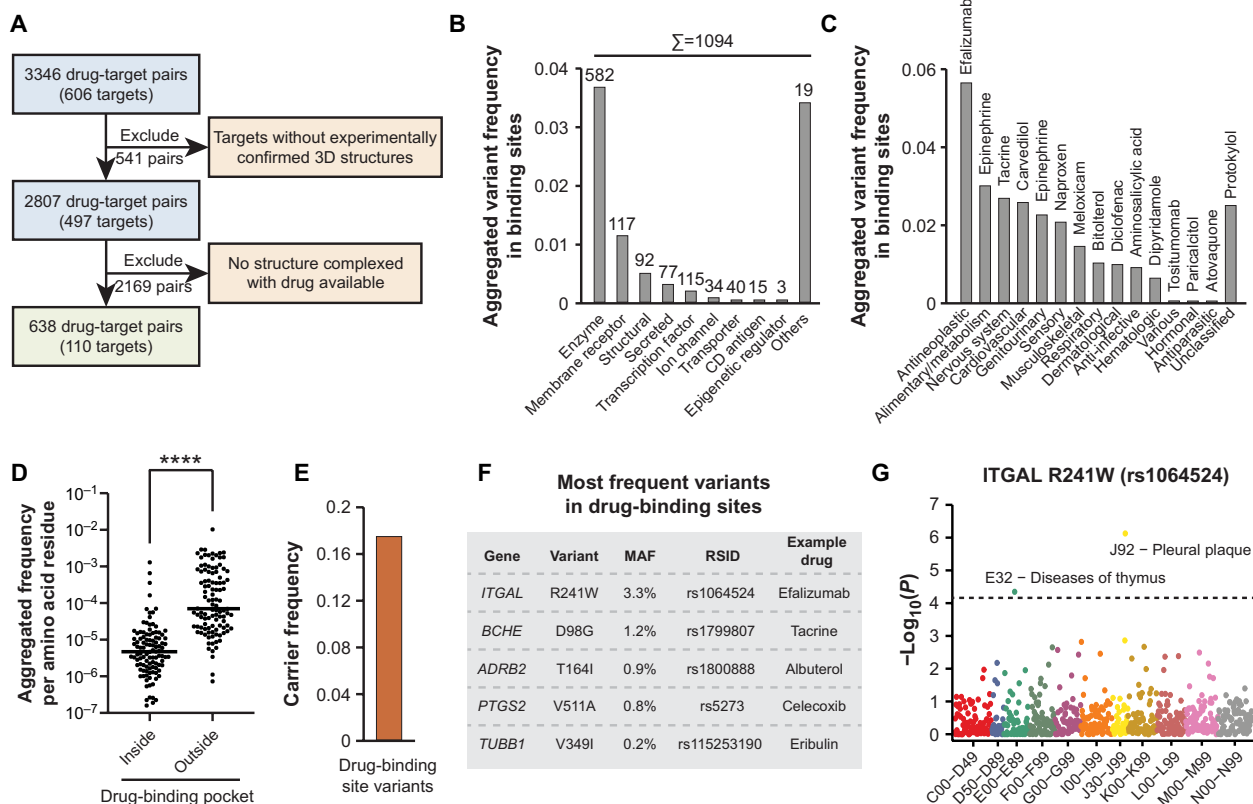
Fig. 2. Characterization of the genetic variability in drug-binding sites. Only variants that affected an amino acid within 6 Å of the bound drug as determined by crystallographic data are considered. (**A**) High-quality structural information of the target protein complexed with the respective drug was available for 110 of 606 targets, corresponding to 638 drug-target pairs. (**B**) The aggregated variant frequency in binding sites is shown for the different target categories. The number of binding site variants is indicated on top of the respective columns. Note that drug-binding sites in enzymes were most variable ($n = 582$), whereas only few variants ($n = 34$) were identified in the binding sites of drugs targeting ion channels despite being similarly common drug targets (compare Fig. 1B). (**C**) Drug-binding site variants were most common in targets of antineoplastic drugs and drugs targeting the alimentary system, whereas variants affecting hormonal and antiparasitic medications targeting the host were very rare. Drugs targeting the most variable binding site in each category are indicated. (**D**) Comparison of aggregated variant frequency within and outside of drug-binding sites. (**E**) Overall, approximately one in six individuals (aggregated variant frequency = 17.5%) carries at least one drug-binding site variant. (**F**) The most frequent binding site variants and their corresponding drugs are shown. MAF, minor allele frequency; RSID, rs number. (**G**) For the most common variant rs1064524 in *ITGAL*, a PheWAS in the Estonian population identified associations with two biologically plausible phenotypes with phenome-wide significance. ATC codes are shown on the abscissa and significance as –log(*P*) on the ordinate. Dashed line indicates phenome-wide significance threshold after Bonferroni correction. Other variants were too rare for PheWAS analyses.

BCHE were relatively common in all populations studied except East Asians, rs115253190 in *TUBB1* and rs138139116 in *TUBB6* were almost specific to African populations and Ashkenazim, respectively (Fig. 3B). Overall, the highest number of homozygote-binding site variant carriers was found in the Finnish population (Fig. 3C). Here, 39 per 10,000 individuals were homozygous for binding pocket variants in *ITGAL* and *BCHE*. In contrast, the rates of homozygotes in Ashkenazi Jews, Latinos, and East Asians were much lower with 8, 6, and 2 per 10,000 individuals, respectively (Fig. 3C). These results demonstrate that binding site variability signatures are distinctly different across populations and could contribute to ethno-geographic differences in drug response.

Binding site variants affect drug pharmacodynamics

Heterologous expression of reference and variant genes coupled to functional assays using appropriate end points is considered as the gold standard to evaluate the functional impact of pharmacogenetic variants (*19*, *20*). To elucidate the functional effect of binding site variants, we experimentally evaluated the effect of binding site variants

in targets from diverse therapeutic areas (cardiology, oncology, and neurology). Specifically, we tested five *ACE* variants in the binding pocket for ACEis, which most guidelines recommend as the first-line treatment for the management of hypertension, seven tubulin variants (*TUBB1*) in the binding site of the polymerization inhibitor eribulin, and seven variants in the cholinesterase (*BCHE*)–binding site for the anti–Alzheimer's disease (AD) medications tacrine and rivastigmine.

Using a recently developed live cell assay based on aggregation-induced emission in which ACE catalyzes the enzymatic cleavage of the nonfluorescent peptide probe tetraphenylethene (TPE)–Ser-Asp-Lys-Pro (SDKP) into fluorescent TPE (*21*), we first evaluated the baseline activity of ACE and its variants and found no significant differences (Fig. 4A). Subsequently, we evaluated the inhibitory effects of the commonly used ACEis—captopril, enalapril, lisinopril, quinapril, and fosinopril. Notably, we observed drastic drug- and variant-specific differences (Fig. 4, B to F). The largest variability in drug response was found in fosinopril, which inhibited only 19 ± 4% SEM of Q288R, 69 ± 1% of the reference enzyme, and 97 ± 3% of the
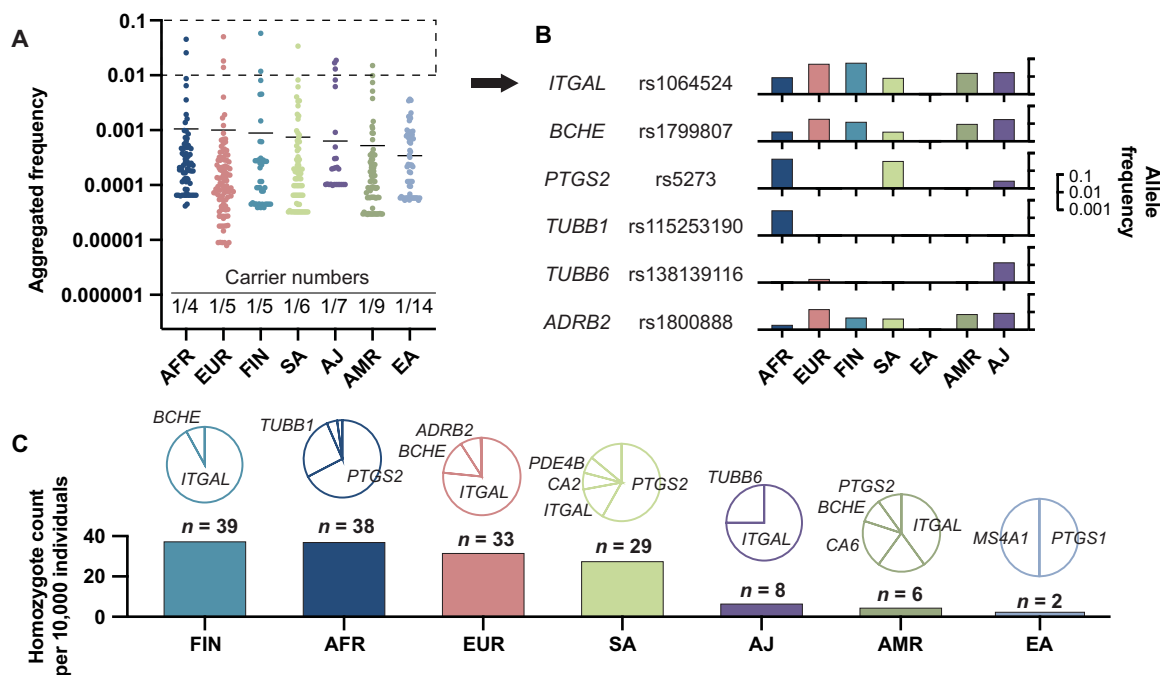
Fig. 3. Ethnogeographic variability in human drug-binding sites. (**A**) The aggregated frequencies of binding site variants is shown for every drug target across seven human populations (AFR, African; EUR, European; FIN, Finnish; SA, South Asian; AJ, Ashkenazi Jews; AMR, Latino; and EA, East Asian). The number of binding site variant carriers is shown for each population at the bottom of the plot. (**B**) Interethnic differences in population-specific frequencies are shown for all variants that are common (allele frequency > 1%) in at least one population. Note the considerable differences in population frequencies for most variants. (**C**) The number of homozygous-binding site variant carriers per 10,000 individuals is shown across ethnogeographic groups. Pie charts indicate the genes most commonly affected in the respective population.

H520N variant. Similarly, large differences were observed for lisinopril (inhibition between 78 ± 1% of Y527C and 11 ± 9% of H520R) and quinapril (inhibition between 81 ± 6% of Y527C and 17 ± 2% of Q288R).

Y527C (average inhibition, 71 ± 4%) and the reference enzyme (average inhibition, 61 ± 9%) were overall most sensitive to ACEi inhibition, whereas Q288R was most resistant (average inhibition, 32 ± 8%). Captopril was the most potent inhibitor of the reference enzyme and of H520R, Y527C, and Y530C variants, inhibiting 72 ± 2% (reference enzyme), 82 ± 3%, 73 ± 2%, and 66 ± 1% of ACE activity, respectively [median inhibitory concentration (IC_{50}) = 4.2 to 8.6 µM, HillSlope: −0.5 to −0.4], whereas fosinopril (97 ± 3% inhibition) and quinapril (81 ± 6% inhibition) were most effective for H520N (IC_{50} = 2.4 µM, HillSlope = −1.1) and Y527C (IC_{50} = 6.5 µM, HillSlope = −0.7), respectively (Fig. 4G). Enalapril constituted the least effective ACEi, inhibiting only 26 to 60% of ACE activity across reference and variant ACEs (IC_{50} > 50 µM; Fig. 4G). Docking analyses showed that for captopril as the most potent ACEi, Y527C and Y530C, which did not significantly affect captopril inhibition ($P_{HillSlope}$ and P_{IC50} > 0.2), had only minor impacts on ligand interaction stability and shape similarity (Fig. 4H). In contrast, docking indicated larger effects on binding stability for Q288R and H520N that significantly reduced sensitivity to captopril ($P_{HillSlope}$ and P_{IC50} < 0.05).

In addition, we experimentally tested the sensitivity of seven variants in *TUBB1*, encoding tubulin β1, a common target for chemotherapeutic drugs that act by disrupting mitotic spindle dynamics (*22*) on the cytotoxicity of the microtubule-destabilizing halichondrin eribulin. While eribulin at 1 nM concentration already reduced

viability in cells expressing reference tubulin or C12Y by 92 ± 0.6 and 89 ± 0.5%, respectively, drug effects were significantly reduced in cells expressing the tubulin β1 variants Y208C (14 ± 1%), S176P (21 ± 3%), R241W (40 ± 5%), L311F (48 ± 2%), G223R (55 ± 3%), and C211Y (58 ± 4%), suggesting partial resistance to eribulin action (Fig. 4I).

Last, we tested the inhibitory effect of tacrine on seven variants in its BCHE-binding pocket. Notably, six of these variants abrogated enzymatic activity and were thus not tested for their effect on tacrine pharmacodynamics (Fig. 4J). The remaining variant D98G reduced BCHE baseline activity by 51 ± 41% and conferred strong tacrine resistance (tacrine IC_{50} for D98G >15 µM compared to <1 µM for the reference enzyme; Fig. 4K). Combined, these data reveal drastic differences in drug sensitivity between naturally occurring variants across therapeutic areas and target classes and demonstrate that consideration of genetic drug target variability in drug selection can optimize personalized pharmacodynamics.

Genetically guided selection of pharmacological alternatives can overcome drug resistance in vitro

Next, we investigated whether drug resistance could be overcome using structurally related derivates to pave the way for genetically guided drug development for otherwise resistant drug targets. To this end, we focused on $BCHE_{D98G}$ because of its high frequency (MAF = 1.2%). First, we tested the inhibitory effect of the two approved nonselective cholinesterase inhibitors, tacrine and rivastigmine, and found that while both drugs inhibited reference BCHE by 87 and 68%, respectively, their inhibitory effect was drastically reduced for $BCHE_{D98G}$ (52% for tacrine and 12% for rivastigmine;
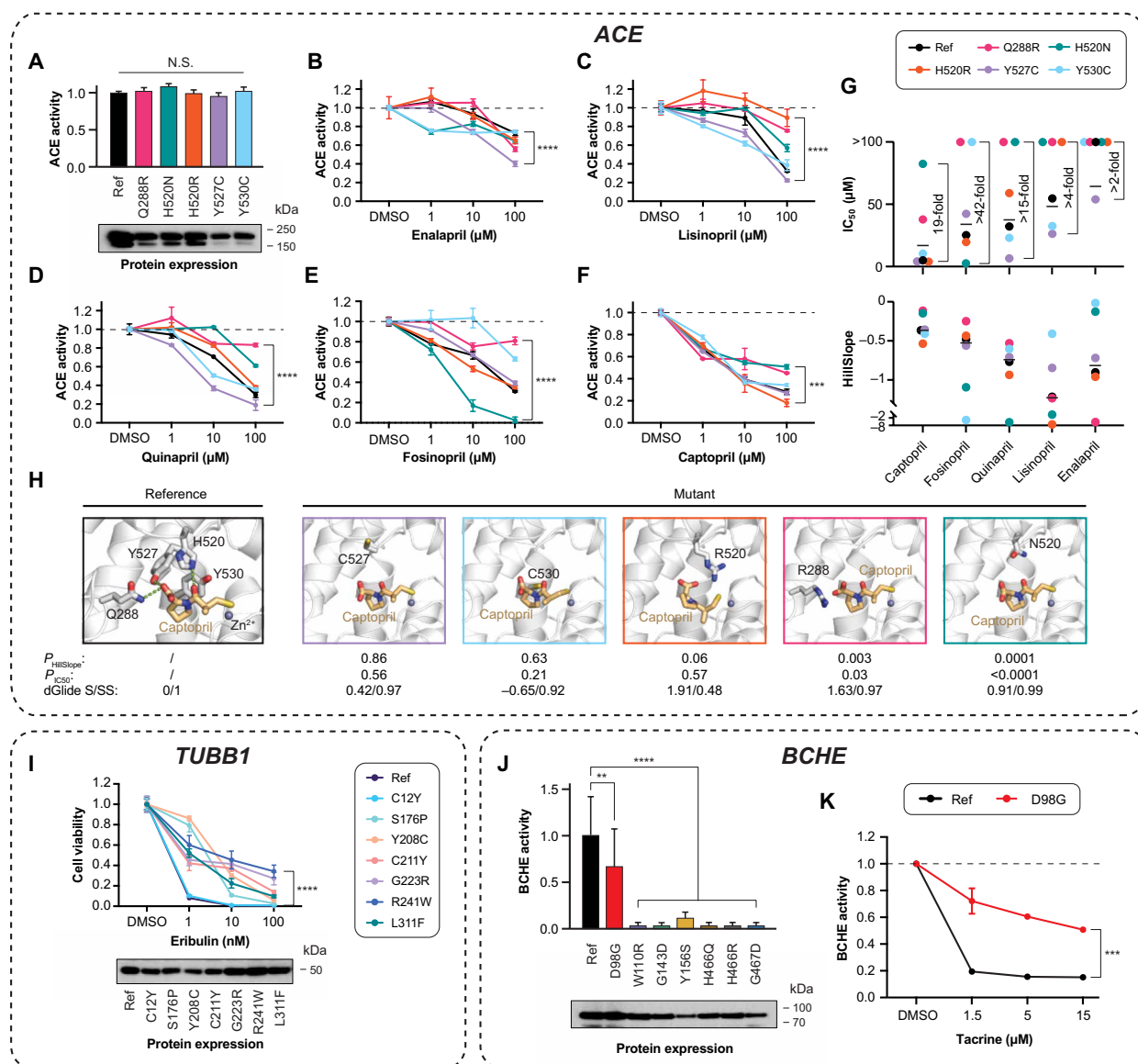
Fig. 4. Naturally occurring drug-binding site variability affects drug response in vitro. Effects of binding site variants in *ACE* (**A** to **H**), *TUBB1* (**I**), and *BCHE* (**J** and **K**) were evaluated using functional assays (see Materials and Methods). (A) None of the *ACE* drug-binding site variants had major impacts on ACE expression or baseline ACE activity. Data are shown as means ± SEM; $n = 3$. Dose-response curves of reference ACE and its variants are shown to the clinically approved ACEis enalapril (B), lisinopril (C), quinapril (D), fosinopril (E), and captopril (F). ***$P < 0.001$; ****$P < 0$ (F test). (G) The IC_{50} values and HillSlope coefficients are shown for each ACE variant–ACEi pair, and the largest fold change across variants is indicated for each drug. (H) Cocrystallized structure of captopril binding to ACE [Protein Data Bank (PDB) ID: 1UZF] and docking poses of captopril with the five ACE variant structures are shown. Atom color code in sticks: oxygen (red), nitrogen (blue), and sulfur (yellow). The zinc ion and hydrogen bonds are shown as violet sphere and green dashed lines, respectively. Parameters indicating the differences between wild-type and variants are shown below the structures. (I) Eribulin effect in cells transfected with reference *TUBB1* or naturally occurring *TUBB1* variants. Note that viability in variant carriers is strongly increased indicating eribulin resistance. (J) With the exception of D98G, drug-binding site variants in *BCHE*-abrogated enzymatic function (**$P < 0.01$; ****$P < 0.0001$; heteroscedastic two-tailed t test; $n = 3$). (K) Inhibitory effects of the cholinesterase inhibitor tacrine were strongly reduced in D98G compared to reference enzyme (***$P < 0.001$; F test). N.S., not significant; DMSO, dimethyl sulfoxide.

Fig. 5A). We thus tested a panel of 20 tacrine, 6-chlorotacrine, or 7-methoxytacrine derivates using squaramide as a structural scaffold [table S5 and (*23*)]. Notably, three compounds (K1514, K1524, and K1526) that share a similar homodimer structure with methylene linker lengths between five and eight carbons inhibited both the reference and D98G variant enzyme with equal potency (Fig. 5A). Similarly, K1035, a chlorotacrine-tryptophan heterodimer previously shown

to inhibit BCHE at low nanomolar concentrations (*24*), strongly inhibited both reference and D98G variant. Notably, the promising tacrine derivatives are predicted to be central nervous system available by parallel artificial membrane permeation assays (*25*) and have shown strong potential to be centrally active (*24*).

To analyze the molecular details of these observations, we conducted structural analyses using K1035 and tacrine for which
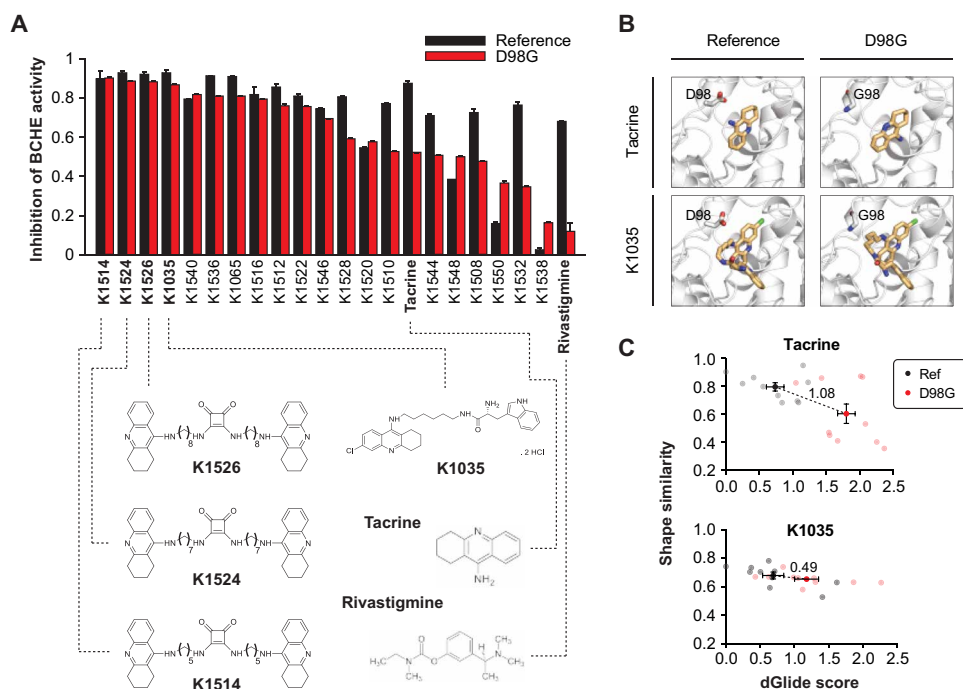
Fig. 5. Resistance caused to drug-binding site variability can be overcome using drug derivatives. (A) Inhibitory effect on reference and D98G BCHE were evaluated for the cholinesterase inhibitors tacrine and rivastigmine as well as 20 tacrine derivatives. Data are shown as means ± SEM; $n = 4$. Chemical structures of the most potent tacrine derivatives (K1504, K1524, K1526, and K1035) are shown below. Structures of tacrine and rivastigmine are shown for reference. **(B)** Matrix representation of reference and D98G BCHE structures interacting with tacrine (PDB ID: 4BDS) and K1035 (PDB ID: 6I0C). Note that while the positioning of tacrine is strongly altered in D98G, K1035 poses remain invariant. **(C)** Shape similarity and docking score (indicated by dGlide score) of the docked ligand poses in D98G compared to the cocrystallized ligands in (B). Note that the divergence of reference and variant structures (indicated by dashed lines) are considerably higher for tacrine than for K1035. Error bars represent SEM of the 10 best ligand-protein poses for both reference and D98G BCHE.

experimentally determined crystal structures with the reference enzyme were available. Docking studies of tacrine to reference and D98G BCHE revealed substantial steric flexibility, resulting in a 180° reorientation of the drug molecule in the pocket of the D98G variant (Fig. 5B). In contrast, the configuration of K1035 was not majorly altered between reference and variant, and its shape similarity to the cocrystallized ligand pose remained largely unaffected by the D98G variant (Fig. 5C). These results demonstrate that structural derivation using squaramide motifs or other scaffolds can overcome variant-mediated drug resistance, thus opening previously unidentified avenues for genetically guided drug development specifically targeted at resistant variants.

DISCUSSION

While genetic variability in genes involved in drug disposition is included in the labels of >130 drugs, drug target variations are only included in the labels of warfarin (targeting VKORC1), cystic fibrosis transmembrane conductance regulator–targeting medications, such as ivacaftor and lumacaftor, and drugs targeting specific somatic mutations in oncology. These differences might be explained, at least in part, because of the larger heterogeneity of drug targets compared to genes involved in drug transport and metabolism, which complicate the experimental interrogation of pharmacodynamic variants. Few studies have systematically evaluated the genetic variability in drug targets and found that the genetic variability in drug-related genes is extensive with a multitude of variants predicted to

result in functional consequences (26–29). However, systematic investigations using structural mapping approaches of population-scale genomic data across approved medications and their targets have not been performed. We here present a comprehensive map of naturally occurring drug-binding site variability of all human drug targets with available experimentally derived high-resolution crystal structures based on genomic information from 138,632 individuals, covering 42% of all approved drugs.

We find that approximately one in six individuals carries at least one variant within a drug-binding pocket. However, as not all targets had structural binding information, the true total frequency of binding site variant carriers can be expected to be even higher. To evaluate the functional consequences of genetic variations, a plethora of computational algorithms that base their predictions on a multitude of features, including sequence conservation, structural information, and functional genomic data have been developed (30, 31). While these tools are considered reasonably accurate in estimating variant pathogenicity, i.e., variant effects on endogenous functions, they are not equipped to predict variant effects on drug action. Here, we showed that although the binding sites of drugs and endogenous substrates commonly overlap, computational algorithms showed only minor overrepresentations in the fraction of putatively deleterious binding site variants compared to variants outside of binding pockets for most genes (fig. S4). Furthermore, while variant effect predictors provided overall reliable predictions for the assessment of variant pathogenicity with the best algorithm showing 98% sensitivity and 92% specificity, predictive accuracy drops <76% for

variants that were experimentally shown to alter drug pharmacodynamics (fig. S5 and tables S6 and S7). In contrast, PheWAS analysis of a common binding site variant in *ITGAL* (rs1064524), encoding integrin α L, revealed strong associations with pleural plaques, dense hyalinized collagen fibers formed in parietal pleura, and diseases of the thymus. *ITGAL* encodes the lymphocyte antigen CD11A, an integrin that has been shown to be necessary for thymus development (*32*) and transmigration of leukocytes across the mesothelium lining the pleura (*33*), thus providing a mechanistic fundament for the genetic association. While sufficiently powered PheWASs could not be performed for other binding site variants because of their lower frequencies, these findings provide proof of concept that the identification of binding site variability can also pinpoint variations important for endogenous functions.

Our experimental results showed that with one exception (C12Y in *TUBB1*), the tested binding site variants either abrogated endogenous functions or affected drug pharmacodynamics (18 of 19; 95%), thus demonstrating the overall relevance of drug-binding site variability. For ACE-binding site variants, we observe >40-fold differences in IC_{50} between marketed ACEis. For instance, ACE_{H520N} was inhibited by fosinopril with an IC_{50} of 2.4 μM, whereas this variant was highly resistant to all other ACEis ($IC_{50} > 82.4$ μM; Fig. 4G). Similarly, efficacy of quinapril was sixfold higher in cells expressing ACE_{Y527C} compared to cells expressing the reference enzyme. These results strongly suggest that drug target variability might explain part of the interindividual variability in the management of hypertension and incentivize the adoption of genotype-informed drug selection. We want to emphasize that we tested drug effects in heterologous overexpression systems, resulting in in vitro IC_{50} values (average IC_{50} around 40 μM) that are approximately 10-fold higher than the therapeutic exposure levels in vivo (clinical c_{max} values of the different ACEis = 0.2 to 6 μM). While these models allow for the comparative analysis of pharmacogenomic effects of different variants in vitro, results are not intended for direct quantitative translation into the in vivo setting.

The vast majority of binding site variants in drug targets are extremely rare. An exception is $BCHE_{D98G}$, which is relatively common, particularly in Europeans (MAF = 1.8%) and Ashkenazi Jews (MAF = 1.7%). This variant was first described as the major determinant of prolonged apnea in response to muscle relaxants (*34*). Here, we experimentally show that this variant also confers resistance to the AD medications tacrine and rivastigmine. Thus, developing compounds that can improve the pharmacological effects for D98G variant carriers might be an important advancement to improve therapeutic outcomes. Here, we demonstrate that four tacrine derivatives featured improved inhibitory effect on $BCHE_{D98G}$. Furthermore, these compounds have previously been demonstrated to also strongly inhibit acetylcholinesterase (ACHE) (*23*). While ACHE constitutes the primary cholinesterase in the cerebral cortex, BCHE activity increases in the hippocampus and temporal cortex with progressing AD and BCHE inhibition correlates with improved cognition (*35*). Given the distinct expression patterns of the two genes in different brain regions of patients with AD, strong inhibition of both cholinesterases constitutes an important goal for AD management (*36*, *37*). These findings suggest that K1035, K1514, K1524, and K1526 might constitute promising tacrine alternatives for further preclinical development, particularly for D98G carriers (>3% of the European population), and demonstrate that drug derivation can provide appealing strategies to optimize drug efficacy for individuals with genetic variable binding sites.

The presented data demonstrate that genetic variability affecting amino acids in drug-binding pockets can have pronounced effects on drug efficacy, but the clinical validation of these findings remains challenging because of the low frequency of individual binding site variants. The variants rs56040400 and rs373359894 in *C5* that both affect arginine-885 are predicted by our analysis to affect the binding of the complement inhibitor eculizumab, which aligns with trial data from 345 Japanese patients with paroxysmal nocturnal hemoglobinuria, in which all 11 nonresponders carried a variant in p.Arg[885] (*38*). Similarly, our data suggested that carriers of the *ADRB2* variant rs1800888 exhibit altered response to β2-adrenergic receptor agonists, which is corroborated by the finding that carriers of this variant exhibit a fivefold lower response to isoproterenol (*39*). Furthermore, previous studies showed a substantial enrichment of rare genetic variants in drug targets in nonresponders to antiseizure medications (*40*, *41*). Furthermore, germline rare variant burden in multidrug resistance transporters was negatively correlated with outcomes in patients with breast cancer undergoing chemotherapy (*42*). Jointly, these and our data incentivize further mechanistic studies, as well as the design of prospective trials that evaluate the added value and cost-effectiveness of genotype-informed drug selection, ideally across indications, drugs, and health care systems. Before these data are available, we suggest that information about binding pocket variability could already flag individuals at increased risk of nonresponse for increased surveillance, particularly in therapeutic areas where timely response is critical, such as oncology.

In summary, this study comprehensively profiled human drug target variability on an unprecedented scale and highlights the importance of drug target variability for personalized medicine. Our results demonstrate that integration of population-scale genomic data and structural target information can flag variants with important effects on drug pharmacodynamics and inform personalized prescribing and targeted drug development for carriers of variants resistant to conventional treatment.

MATERIALS AND METHODS
Data sources
Genetic variants of all drug target proteins and their frequencies were derived from whole-genome and whole-exome sequencing data of 138,632 individuals provided by the Genome Aggregation Database from seven ethnogeographic groups (European, African, Latinos, East Asian, South Asian, Ashkenazi Jews, and Finnish) (*43*). Experimentally determined high-resolution crystal structures of target proteins crystallized with the respective drug were obtained from the Protein Data Bank (PDB) (*44*). The putative functional consequences of genetic variants in drug targets were assessed using 17 partly orthogonal computational algorithms (SIFT, PolyPhen-2, LRT, MutationTaster, MutationAssessor, FATHMM, PROVEAN, MetaSVM, MetaLR, M-CAP, REVEL, CADD, FATHMM-MKL, VEST3, DANN, Eigen, and GERP++). Variants were classified as deleterious if >80% of algorithms predicted functional impacts. In addition, all frameshifts, start-lost, and stop-gain variants as well as all variations affecting canonical splice sites were considered as deleterious.

Drug target–binding site identification
Drug targets were identified on the basis of recent comprehensive maps of the human druggable genome (*16*). In total, 606 human drug target proteins were identified for which the respective drug

did not target a specific somatic mutation. The evaluation of variants within drug-binding sites relies on the accurate identification of binding pockets. Thus, we only considered those targets for our evaluations of which (i) the structure of the human protein complexed with the respective drug was available, (ii) the structure of a homologous animal protein complexed with the respective drug was available, (iii) the structure of a paralogous human protein complexed with the respective drug was available, (iv) the structure of the human protein complexed with a drug with the same backbone chemical structure was available, or (v) the structure of a paralogous human protein complexed with a drug with the same backbone chemical structure was available. For those targets for which only homologous or paralogous protein structures were available, structural agreement was evaluated using Phyre2 (45). Model accuracy was considered sufficient if confidence was >90% and sequence identity was ≥30%. Binding pockets were defined as all amino acids within 6 Å of the respective drug molecule using PyMOL (version 2.1.1, Schrödinger LLC).

PheWAS using the Estonian Biobank
A PheWAS was conducted for the most common drug-binding site variant, rs1064524 in *ITGAL*. Less abundant variants had to be excluded because of the low number of carriers and, accordingly, low statistical power in the Estonian Biobank. International Classification of Diseases (ICD)-10 codes were obtained from the electronic health records of the participants with the number of unrelated individuals ranging from 18,601 to 51,138, depending on the phenotype. PheWAS codes were first grouped into three-character identifiers, and only PheWAS codes C00-N99 (excluding D78, F70-F89, H, and J00-J22) were selected for further analyses. PheWAS codes with less than five cases in total were excluded, resulting in a total of 725 PheWAS codes. PheWAS analysis was performed using the R package PheWAS (46) using logistic regression models with gender, birth year, and the first five principle components as covariates. Bonferroni correction was used to account for multiple testing with a P value significance threshold of $0.05/725 \approx 6.9 \times 10^{-5}$ (where 725 is the number of PheWAS codes analyzed).

Molecular docking studies
Cocrystallized structure of captopril with ACE [PDB identification (ID): 1UZF; (47)], tacrine with BCHE [4BDS; (48)], and its derivate K1035 with BCHE [6I0C; (24)] were obtained from the Research Collaboratory for Structural Bioinformatics PDB database. To prepare these structures for the molecular docking, we used the Schrödinger Maestro molecular modeling platform (release, January 2019, Schrödinger, LLC, New York, NY) and PyMOL software. The protein-ligand complexes were first refined in Maestro using Protein Preparation Wizard, and the ligand-binding sites were optimized with Prime using a 7-Å threshold. After these refinement steps, the ligands were removed, and selected point mutations were introduced to each protein structure with the Mutagenesis Wizard in PyMOL.

The cocrystallized ligands were docked back to the corresponding wild-type and mutated protein structures to a box located and sized on the basis of the location and size of the ligand in the original (optimized) complex using Schrödinger's Glide software. Standard precision (SP) parameters were modified by enhancing the planarity of the conjugated P groups. The postdocking minimization was conducted for 20 best ligand poses, and the final output was set to 10 best poses per each ligand-protein pair. As binding of captopril

to ACE (PDB ID: 1UZF) involves a metal coordination, an additional location restraint for the coordinating thiol group was added to the docking grid to enhance the enrichment of the biologically meaningful docking poses. The g score of each pose was compared to the best-scoring wild-type pose of the corresponding drug-receptor complex, and the dGlide score was calculated as Glide g score$_{\text{best-scoring wild-type pose}}$ − Glide g score$_{\text{studied pose}}$. For shape and location analysis, we used ShaEP software (49) using "onlyshape" and "noOptimization" parameters; the docking poses of each ligand were compared to the shape and binding location of the cocrystallized ligand at the corresponding crystal structure. All docking poses are provided as data file S1 in mol2 format.

Experimental evaluation of variants in *ACE, BCHE,* and *TUBB1*
Site-directed mutagenesis
Reference ACE, BCHE, and TUBB open reading frames were inserted into eukaryotic c-Myc–tagged plasmids, and drug target variants were generated using the QuikChange II Site-Directed Mutagenesis Kit (200524, Agilent Technologies AH Diagnostic). The sequences of the primers used to generate the variant proteins are listed in table S8. Following mutagenesis, variant constructs were sequenced to confirm the correct nucleotide mutations.
Functional assays
Human embryonic kidney 293 cells were grown in a humidified atmosphere at 37°C and 5% CO_2 in in Dulbecco's modified Eagle's medium (DMEM) without phenol red supplemented with 10% heat-inactivated fetal bovine serum and penicillin and streptomycin (100 IU/ml). Cells were transiently transfected with the respective plasmids using Viromer RED (Lipocalyx, Germany). Expression levels of protein variants were analyzed by Western blots using antibodies against c-Myc (9E10, Santa Cruz Biotechnology).

ACE activity was measured as previously reported (21). Cells were seeded in a 96-well plate and transfected with the corresponding ACE reference and variant plasmids. After 24 hours, the medium was replaced with fresh medium containing ACEi (captopril, enalapril, lisinopril, quinapril, and fosinopril) at final concentrations of 1, 10, and 100 μM as indicated. Notably, because enalapril, quinapril, and fosinopril are prodrugs, we used their respective active metabolites (enalaprilat, quinaprilat, and fosinoprilat) instead of the parent compounds for exposure. After 4 hours of exposure, the medium was replaced with the peptide probe TPE-SDKP (50 μM) that emits fluorescence upon ACE-dependent cleavage in tris buffer and incubated for 2 hours at 37°C. Subsequently, $ZnCl_2$ was added to a final concentration of 3 mM and incubated for another 60 min after which the results were measured.

BCHE functionality was quantified using Ellman's assay. Transfected cells were seeded in 96-well format, and after 24 hours, the medium was replaced with fresh DMEM without phenol red. After 48 hours, 50 μl of medium containing the secreted cholinesterase proteins was transferred into a new 96-well plate and was incubated with 50 μl of cholinesterase inhibitors (tacrine, rivastigmine, or tacrine-squaramide derivatives) at a final concentration of 1.5 μM for 30 min. Ellman solution (300 μl) was added to each well, and absorbance was measured at 412 nm using a SpectraMax M2 microplate reader from 1 to 5 min.

To evaluate the effect of tubulin variants, cells were seeded for transfection with the corresponding TUBB1 reference and variant plasmids. After 24 hours, the medium was replaced with fresh DMEM without phenol red containing eribulin at a final concentration of 1,

10, or 100 nM. After 72 hours of exposure, adenosine 5′-triphosphate levels were measured using the CellTiter-Glo kit (Promega) to quantify cell viability.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abi6856

REFERENCES AND NOTES

1. R. K. Harrison, Phase II and phase III failures: 2013–2015. *Nat. Rev. Drug Discov.* **15**, 817–818 (2016).
2. K. Smietana, M. Siatkowski, M. Møller, Trends in clinical success rates. *Nat. Rev. Drug Discov.* **15**, 379–380 (2016).
3. J. A. Johnson, Ethnic differences in cardiovascular drug response: Potential contribution of pharmacogenetics. *Circulation* **118**, 1383–1393 (2008).
4. S. Porcelli, A. Drago, C. Fabbri, S. Gibiino, R. Calati, A. Serretti, Pharmacogenetics of antidepressant response. *J. Psychiatry Neurosci.* **36**, 87–113 (2011).
5. P. S. Lacombe, L. Vicente, J. C. Pagès, P. L. Morselli, Causes and problems of nonresponse or poor response to drugs. *Drugs* **51**, 552–570 (1996).
6. M. H. Trivedi, A. J. Rush, S. R. Wisniewski, A. A. Nierenberg, D. Warden, L. Ritz, G. Norquist, R. H. Howland, B. Lebowitz, P. J. McGrath, K. Shores-Wilson, M. M. Biggs, G. K. Balasubramani, M. Fava; STAR*D Study Team, Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. *Am. J. Psychiatry* **163**, 28–40 (2006).
7. U. M. Zanger, M. Schwab, Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* **138**, 103–141 (2013).
8. V. M. Lauschke, Y. Zhou, M. Ingelman-Sundberg, Novel genetic and epigenetic factors of importance for inter-individual differences in drug disposition, response and toxicity. *Pharmacol. Ther.* **197**, 122–152 (2019).
9. L. M. Henricks, C. A. T. C. Lunenburg, F. M. de Man, D. Meulendijks, G. W. J. Frederix, E. Kienhuis, G.-J. Creemers, A. Baars, V. O. Dezentjé, A. L. T. Imholz, F. J. F. Jeurissen, J. E. A. Portielje, R. L. H. Jansen, P. Hamberg, A. J. ten Tije, H. J. Droogendijk, M. Koopman, P. Nieboer, M. H. W. van de Poel, C. M. P. W. Mandigers, H. Rosing, J. H. Beijnen, E. Werkhoven, A. B. P. van Kuilenburg, R. H. N. van Schaik, R. H. J. Mathijssen, J. J. Swen, H. Gelderblom, A. Cats, H. J. Guchelaar, J. H. M. Schellens, DPYD genotype-guided dose individualisation of fluoropyrimidine therapy in patients with cancer: A prospective safety analysis. *Lancet Oncol.* **19**, 1459–1467 (2018).
10. M. D. Klein, A. K. Williams, C. R. Lee, G. A. Stouffer, Clinical utility of *CYP2C19* genotyping to guide antiplatelet therapy in patients with an acute coronary syndrome or undergoing percutaneous coronary intervention. *Arterioscler. Thromb. Vasc. Biol.* **39**, 647–652 (2019).
11. W. He, F. Grassmann, M. Eriksson, S. Eliasson, L. Margolin, L. Thorén, P. Hall, K. Czene, CYP2D6 genotype predicts tamoxifen discontinuation and prognosis in patients with breast cancer. *J. Clin. Oncol.* **38**, 548–557 (2020).
12. J. P. Overington, B. Al-Lazikani, A. L. Hopkins, How many drug targets are there? *Nat. Rev. Drug Discov.* **5**, 993–996 (2006).
13. P. Imming, C. Sinning, A. Meyer, Drugs, their targets and the nature and number of drug targets. *Nat. Rev. Drug Discov.* **5**, 821–834 (2007).
14. E. V. Minikel, K. J. Karczewski, H. C. Martin, B. B. Cummings, N. Whiffin, D. Rhodes, J. Alföldi, R. C. Trembath, D. A. van Heel, M. J. Daly; Genome Aggregation Database Production Team; Genome Aggregation Database Consortium, S. L. Schreiber, D. G. MacArthur, Evaluating drug targets through human loss-of-function genetic variation. *Nature* **581**, 459–464 (2020).
15. A. S. Hauser, S. Chavali, I. Masuho, L. J. Jahn, K. A. Martemyanov, D. E. Gloriam, M. M. Babu, Pharmacogenomics of GPCR drug targets. *Cell* **172**, 41–54.e19 (2018).
16. R. Santos, O. Ursu, A. Gaulton, A. P. Bento, R. S. Donadi, C. G. Bologa, A. Karlsson, B. al-Lazikani, A. Hersey, T. I. Oprea, J. P. Overington, A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).
17. Y. Wang, S. Zhang, F. Li, Y. Zhou, Y. Zhang, Z. Wang, R. Zhang, J. Zhu, Y. Ren, Y. Tan, C. Qin, Y. Li, X. Li, Y. Chen, F. Zhu, Therapeutic target database 2020: Enriched resource for facilitating research and early development of targeted therapeutics. *Nucleic Acids Res.* **48**, D1031–D1041 (2020).
18. S. Reisberg, K. Krebs, M. Lepamets, M. Kals, R. Mägi, K. Metsalu, V. M. Lauschke, J. Vilo, L. Milani, Translating genotype data of 44,000 biobank participants into clinical pharmacogenetic recommendations: Challenges and solutions. *Genet. Med.* **21**, 1345–1354 (2019).
19. Y. Zhou, K. Fujikura, S. Mkrtchian, V. M. Lauschke, Computational methods for the pharmacogenetic interpretation of next generation sequencing data. *Front. Pharmacol.* **9**, 1437 (2018).
20. V. M. Lauschke, M. Ingelman-Sundberg, Emerging strategies to bridge the gap between pharmacogenomic research and its clinical implementation. *NPJ Genom. Med.* **5**, 9 (2020).
21. H. Wang, Y. Huang, X. Zhao, W. Gong, Y. Wang, Y. Cheng, A novel aggregation-induced emission based fluorescent probe for an angiotensin converting enzyme (ACE) assay and inhibitor screening. *Chem. Commun.* **50**, 15075–15078 (2014).
22. E. Mukhtar, V. M. Adhami, H. Mukhtar, Targeting microtubules by natural agents for cancer therapy. *Mol. Cancer Ther.* **13**, 275–284 (2014).
23. B. Svobodova, E. Mezeiova, V. Hepnarova, M. Hrabinova, L. Muckova, T. Kobrlova, D. Jun, O. Soukup, M. L. Jimeno, J. Marco-Contelles, J. Korabecny, Exploring structure-activity relationship in tacrine-squaramide derivatives as potent cholinesterase inhibitors. *Biomolecules* **9**, 379 (2019).
24. K. Chalupova, J. Korabecny, M. Bartolini, B. Monti, D. Lamba, R. Caliandro, A. Pesaresi, X. Brazzolotto, A. J. Gastellier, F. Nachon, J. Pejchal, M. Jarosova, V. Hepnarova, D. Jun, M. Hrabinova, R. Dolezal, J. Zdarova Karasova, M. Mzik, Z. Kristofikova, J. Misik, L. Muckova, P. Jost, O. Soukup, M. Benkova, V. Setnicka, L. Habartova, M. Chvojkova, L. Kleteckova, K. Vales, E. Mezeiova, E. Uliassi, M. Valis, E. Nepovimova, M. L. Bolognesi, K. Kuca, Novel tacrine-tryptophan hybrids: Multi-target directed ligands as potential treatment for Alzheimer's disease. *Eur. J. Med. Chem.* **168**, 491–514 (2019).
25. L. Di, E. H. Kerns, K. Fan, O. J. McConnell, G. T. Carter, High throughput artificial membrane permeability assay for blood-brain barrier. *Eur. J. Med. Chem.* **38**, 223–232 (2003).
26. M. R. Nelson, D. Wegmann, M. G. Ehm, D. Kessner, P. St. Jean, C. Verzilli, J. Shen, Z. Tang, S. A. Bacanu, D. Fraser, L. Warren, J. Aponte, M. Zawistowski, X. Liu, H. Zhang, Y. Zhang, J. Li, Y. Li, L. Li, P. Woollard, S. Topp, M. D. Hall, K. Nangle, J. Wang, G. Abecasis, L. R. Cardon, S. Zöllner, J. C. Whittaker, S. L. Chissoe, J. Novembre, V. Mooser, An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* **337**, 100–104 (2012).
27. F. E. Dewey, M. F. Murray, J. D. Overton, L. Habegger, J. B. Leader, S. N. Fetterolf, C. O'Dushlaine, C. V. Van Hout, J. Staples, C. Gonzaga-Jauregui, R. Metpally, S. A. Pendergrass, M. A. Giovanni, H. L. Kirchner, S. Balasubramanian, N. S. Abul-Husn, D. N. Hartzel, D. R. Lavage, K. A. Kost, J. S. Packer, A. E. Lopez, J. Penn, S. Mukherjee, N. Gosalia, M. Kanagaraj, A. H. Li, L. J. Mitnaul, L. J. Adams, T. N. Person, K. Praveen, A. Marcketta, M. S. Lebo, C. A. Austin-Tse, H. M. Mason-Suares, S. Bruse, S. Mellis, R. Phillips, N. Stahl, A. Murphy, A. Economides, K. A. Skelding, C. D. Still, J. R. Elmore, I. B. Borecki, G. D. Yancopoulos, F. D. Davis, W. A. Faucett, O. Gottesman, M. D. Ritchie, A. R. Shuldiner, J. G. Reid, D. H. Ledbetter, A. Baras, D. J. Carey, Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science* **354**, aaf6814 (2016).
28. C. P. I. Schärfe, R. Tremmel, M. Schwab, O. Kohlbacher, D. S. Marks, Genetic variation in human drug-related genes. *Genome Med.* **9**, 117–115 (2017).
29. C. Yan, N. Pattabiraman, J. Goecks, P. Lam, A. Nayak, Y. Pan, J. Torcivia-Rodriguez, A. Voskanian, Q. Wan, R. Mazumder, Impact of germline and somatic missense variations on drug binding sites. *Pharmacogenomics J.* **17**, 128–136 (2017).
30. T. A. Peterson, E. Doughty, M. G. Kann, Towards precision medicine: Advances in computational approaches for the analysis of human variants. *J. Mol. Biol.* **425**, 4047–4063 (2013).
31. Y. Zhou, V. M. Lauschke, Computational tools to assess the functional consequences of rare and noncoding pharmacogenetic variability. *Clin. Pharmacol. Ther.* 10.1002/cpt.2289 , (2021).
32. T. O. Bose, S. L. Colpitts, Q.-M. Pham, L. Puddington, L. Lefrancois, CD11a is essential for normal development of hematopoietic intermediates. *J. Immunol.* **193**, 2863–2872 (2014).
33. S. E. Mutsaers, C. M. Prele, A. R. Brody, S. Idell, Pathogenesis of pleural fibrosis. *Respirology* **9**, 428–440 (2004).
34. M. C. McGuire, C. P. Nogueira, C. F. Bartels, H. Lightstone, A. Hajra, A. F. van der Spek, O. Lockridge, B. N. la du, Identification of the structural mutation responsible for the dibucaine-tesistant (atypical) variant form of human serum cholinesterase. *Proc. Natl. Acad. Sci.* **86**, 953–957 (1989).
35. N. Kandiah, M. C. Pai, V. Senanarong, I. Looi, E. Ampil, K. W. Park, A. K. Karanam, S. Christopher, Rivastigmine: The advantages of dual inhibition of acetylcholinesterase and butyrylcholinesterase and its role in subcortical vascular dementia and Parkinson's disease dementia. *Clin. Interv. Aging* **12**, 697–707 (2017).
36. E. K. Perry, R. H. Perry, G. Blessed, B. E. Tomlinson, Changes in brain cholinesterases in senile dementia of Alzheimer type. *Neuropathol. Appl. Neurobiol.* **4**, 273–277 (1978).
37. A. Ciro, J. Park, G. Burkhard, N. Yan, C. Geula, Biochemical differentiation of cholinesterases from normal and Alzheimer's disease cortex. *Curr. Alzheimer Res.* **9**, 138–143 (2012).
38. J.-i. Nishimura, M. Yamamoto, S. Hayashi, K. Ohyashiki, K. Ando, A. L. Brodsky, H. Noji, K. Kitamura, T. Eto, T. Takahashi, M. Masuko, T. Matsumoto, Y. Wano, T. Shichishima, H. Shibayama, M. Hase, L. Li, K. Johnson, A. Lazarowski, P. Tamburini, J. Inazawa, T. Kinoshita, Y. Kanakura, Genetic variants in C5 and poor response to eculizumab. *N. Eng. J. Med.* **370**, 632–639 (2014).

39. V. Dishy, R. Landau, G. G. Sofowora, H. G. Xie, R. M. Smiley, R. B. Kim, D. W. Byrne, A. J. J. Wood, C. M. Stein, Beta2-adrenoceptor Thr164Ile polymorphism is associated with markedly decreased vasodilator and increased vasoconstrictor sensitivity in vivo. *Pharmacogenetics* **14**, 517–522 (2004).

40. S. Wolking, C. Moreau, A. T. Nies, E. Schaeffeler, M. McCormack, P. Auce, A. Avbersek, F. Becker, M. Krenn, R. S. Møller, M. Nikanorova, Y. G. Weber, S. Weckhuysen, G. L. Cavalleri, N. Delanty, C. Depondt, M. R. Johnson, B. P. C. Koeleman, W. S. Kunz, A. G. Marson, J. W. Sander, G. J. Sills, P. Striano, F. Zara, F. Zimprich, M. Schwab, R. Krause, S. M. Sisodiya, P. Cossette, S. L. Girard, H. Lerche; EpiPGX Consortium, Testing association of rare genetic variants with resistance to three common antiseizure medications. *Epilepsia* **61**, 657–666 (2020).

41. K. A. Myers, M. F. Bennett, B. E. Grinton, G. Dabscheck, E. K. Chan, L. E. Bello-Espinosa, L. G. Sadleir, S. D'Alfonso, A. L. Schneider, J. A. Damiano, M. S. Hildebrand, M. Bahlo, S. F. Berkovic, J. Buchhalter, I. E. Scheffer, Contribution of rare genetic variants to drug response in absence epilepsy. *Epilepsy Res.* **170**, 106537 (2021).

42. Q. Xiao, Y. Zhou, S. Winter, F. Büttner, E. Schaeffeler, M. Schwab, V. M. Lauschke, Germline variant burden in multidrug resistance transporters is a therapy-specific predictor of survival in breast cancer patients. *Int. J. Cancer* **146**, 2475–2487 (2020).

43. M. Lek, K. J. Karczewski, E. V. Minikel, K. E. Samocha, E. Banks, T. Fennell, A. H. O'Donnell-Luria, J. S. Ware, A. J. Hill, B. B. Cummings, T. Tukiainen, D. P. Birnbaum, J. A. Kosmicki, L. E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D. N. Cooper, N. Deflaux, M. De Pristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M. I. Kurki, A. L. Moonshine, P. Natarajan, L. Orozco, G. M. Peloso, R. Poplin, M. A. Rivas, V. Ruano-Rubio, S. A. Rose, D. M. Ruderfer, K. Shakir, P. D. Stenson, C. Stevens, B. P. Thomas, G. Tiao, M. T. Tusie-Luna, B. Weisburd, H.-H. Won, D. Yu, D. M. Altshuler, D. Ardissino, M. Boehnke, J. Danesh, S. Donnelly, R. Elosua, J. C. Florez, S. B. Gabriel, G. Getz, S. J. Glatt, C. M. Hultman, S. Kathiresan, M. Laakso, S. M. Carroll, M. I. McCarthy, D. M. Govern, R. M. Pherson, B. M. Neale, A. Palotie, S. M. Purcell, D. Saleheen, J. M. Scharf, P. Sklar, P. F. Sullivan, J. Tuomilehto, M. T. Tsuang, H. C. Watkins, J. G. Wilson, M. J. Daly, D. G. MacArthur; Exome Aggregation Consortium, Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).

44. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

45. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).

46. J. C. Denny, M. D. Ritchie, M. A. Basford, J. M. Pulley, L. Bastarache, K. Brown-Gentry, D. Wang, D. R. Masys, D. M. Roden, D. C. Crawford, PheWAS: Demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics* **26**, 1205–1210 (2010).

47. R. Natesh, S. L. U. Schwager, H. R. Evans, E. D. Sturrock, K. R. Acharya, Structural details on the binding of antihypertensive drugs captopril and enalaprilat to human testicular angiotensin I-converting enzyme. *Biochemistry* **43**, 8718–8724 (2004).

48. F. Nachon, E. Carletti, C. Ronco, M. Trovaslet, Y. Nicolet, L. Jean, P.-Y. Renard, Crystal structures of human cholinesterases in complex with huprine W and tacrine: Elements of specificity for anti-Alzheimer's drugs targeting acetyl- and butyryl-cholinesterase. *Biochem. J.* **453**, 393–399 (2013).

49. M. J. Vainio, J. S. Puranen, M. S. Johnson, ShaEP: Molecular overlay based on shape and electrostatic potential. *J. Chem. Inf. Model.* **49**, 492–502 (2009).

Citation: Y. Zhou, G. H. Arribas, A. Turku, T. Jürgenson, S. Mkrtchian, K. Krebs, Y. Wang, B. Svobodova, L. Milani, G. Schulte, J. Korabecny, S. Gastaldello, V. M. Lauschke, Rare genetic variability in human drug target genes modulates drug response and can guide precision medicine. *Sci. Adv.* **7**, eabi6856 (2021).