Expert Opinion on Therapeutic Targets

ISSN: 1472-8222 (Print) 1744-7631 (Online) Journal homepage: https://www.tandfonline.com/loi/iett20

The Therapeutic Target Database: an Internet resource for the primary targets of approved, clinical trial and experimental drugs

Xin Liu, Feng Zhu, Xiaohua Ma, Lin Tao, Jingxian Zhang, Shengyong Yang, Yuquan Wei & Yu Zong Chen

To cite this article: Xin Liu, Feng Zhu, Xiaohua Ma, Lin Tao, Jingxian Zhang, Shengyong Yang, Yuquan Wei & Yu Zong Chen (2011) The Therapeutic Target Database: an Internet resource for the primary targets of approved, clinical trial and experimental drugs, Expert Opinion on Therapeutic Targets, 15:8, 903-912, DOI: 10.1517/14728222.2011.586635

To link to this article: https://doi.org/10.1517/14728222.2011.586635

Expert Opinion

The Therapeutic Target Database: an Internet resource for the primary targets of approved, clinical trial and experimental drugs

Xin Liu, Feng Zhu, Xiaohua Ma, Lin Tao, Jingxian Zhang, Shengyong Yang, Yuquan Wei & Yu Zong Chen†

†*Bioinformatics and Drug Design Group, Department of Pharmacy, Centre for Computational Science and Engineering, National University of Singapore, Singapore*

Increasing numbers of proteins, nucleic acids and other molecular entities have been explored as therapeutic targets. A challenge in drug discovery is to decide which targets to pursue from an increasing pool of potential targets, given the fact that few innovative targets have made it to the approval list each year. Knowledge of existing drug targets (both approved and within clinical trials) is highly useful for facilitating target discovery, selection, exploration and tool development. The Therapeutic Target Database (TTD) has been developed and updated to provide information on 358 successful targets, 251 clinical trial targets and 1254 research targets in addition to 1511 approved drugs, 1118 clinical trials drugs and 2331 experimental drugs linked to their primary targets (3257 drugs with available structure data). This review briefly describes the TTD database and illustrates how its data can be explored for facilitating target and drug searches, the study of the mechanism of multi-target drugs and the development of *in silico* target discovery tools.

Keywords: *in silico* target discovery, multi-target drugs, primary target, TTD

Expert Opin. Ther. Targets (2011) 15(8):903-912

1. Introduction

Target discovery efforts have led to the discovery of hundreds of successful targets (targeted by at least one approved drug) and > 1000 research targets (targeted by experimental drugs only) [1,2]. A challenge in drug development is to define a clear strategy to decide which targets to pursue from an ever-growing selection of potential targets [3]. The selection of an appropriate target or multiple targets is not only important for therapeutic efficacy but also affects the drug development odds, given that only 12 innovative targets have made it to the approval list during 1994 ~ 2005 [4].

Resources that provide comprehensive information about the targets of approved, clinical trial and experimental drugs are highly useful for facilitating the discovery, assessment and selection of targets, the development of target discovery strategy and technology, and decision making [2,5]. To facilitate the access of information about drug targets, several publicly accessible databases such as the Therapeutic Target Database (TTD) [6], DrugBank [7] and Potential Drug Target Database (PDTD) [8] have been developed. These databases complement each other to provide target and drug profiles. TTD provides information about the primary targets of approved and experimental drugs [6]. DrugBank is an excellent source for comprehensive drug data with information about both drug actions and drug

informa healthcare

10.1517/14728222.2011.586635 © 2011 Informa UK, Ltd. ISSN 1472-8222

targets [7]. PDTD contains active-sites as well as functional information for potential targets with available 3D structures [8].

In particular, TTD [9] provides the information of the primary targets of approved, clinical trial and investigational drugs and the drugs active against these targets. While drugs typically modulate the activities of multiple proteins [10] and up to 14,000 drug-targeted-proteins have been reported [11], the reported number of primary targets directly related to the therapeutic efficacies of approved drugs is limited to 324 [1]. Information about the primary targets of more comprehensive sets of approved, clinical trial and experimental drugs is highly useful for facilitating focused investigations and discovery efforts against the most relevant and proven targets [5,12]. TTD is intended to provide the relevant information and data. The usefulness of the data in TTD is illustrated by four case studies. These are similarity target search, similarity drug search, the study of the mechanism of multi-target drugs and the development of *in silico* target discovery tools.

2. TTD data collection and access

The literature reported data about the approved, clinical trial and experimental drugs and their primary targets were collected and verified from multiple sources including the FDA Drugs@FDA webpage [13] with information about FDA approved drugs, NIH ClinicalTrial.gov website, CenterWatch Drugs in Clinical Trials Database, latest company reports (accessible from company websites) and review papers of 384 pharmaceutical companies including companies that jointly develop drugs with another company (e.g., AstraZeneca, Bayer, Boehringer Ingelheim, Genentech, GSK, Idenix, Incyte, ISIS, Merck, Novartis, Pfizer, Roche, Sanofi-Aventis, Schering-Plough, Spectrum, Takeda, Teva), 230 recent articles in reputable journals (Expert Opin Investig Drugs, Nature Rev Drug Discov, TiPS, Drug Discov Today, Curr Opin Pharmacol, Curr Drug Targets, Curr Topics Mechem, Mini Rev Mechem, Anticancer Agent Medchem and Science) and 2008 Reports of Medicines in Development (biotechnology, HIV/AIDS, cancer, children, diabetes, neurological disorders, rare diseases and women). The full list of 384 companies and the 230 recent articles are in supplementary Tables S1 and S2, and these can be accessed from the TTD database download page.

The primary targets of 211 drugs and drug binding modes of 79 drugs are not specified in our searched documents. Additional literature search has been conducted to first identify the target reported to be directly modulated by the drug and through which the drug reportedly mediates its claimed therapeutic activities, and then find additional evidence to confirm a target if cell line or *in vivo* studies have been reported, and the drug modulates the target with potencies < 500 nM, or lower potencies but modulates cell lines at < 1 μM. All identified targets have been evaluated for evidences of *in vitro*, *in vivo* and knockout studies linking a target to a clinical trial drug.

TTD data can be accessed by keyword and customized search. Customized search fields include target name, drug name, disease indication (436 indications), target biochemical class (61 classes), drug therapeutic class (156 classes) and drug mode of action. Further information about each target can be accessed via crosslink to SwissProt/UniProt, PDB, KEGG, OMID and PubMed. Related target entries can be recursively searched by clicking a disease or drug name. Similarity targets of an input protein sequence in FASTA format can be searched by using the BLAST sequence alignment tool [14]. Similar drugs of an input drug structure can be searched by using molecular descriptor based Tanimoto similarity searching method [15]. The whole TTD data, target sequences and drug structures can be downloaded via the download link. Target and drug entries are assigned standardized TTD IDs for easy identification, analysis and linkage to other related databases.

3. Application of TTD data in searching similarity drugs and targets

The usefulness of TTD data can be illustrated by two case studies. One is the search of similarity drugs of an investigative agent, which may be used for drug-likeness evaluation based on the level of similarity to existing drugs [16] and for target assessment based on the targets of similarity drugs [17]. The drug similarity search tool in TTD is based on the Tanimoto similarity searching method [15] that ranks similarity drugs based on the level of similarity of their molecular descriptors with respect to those of an input agent. Drug similarity search by inputting the structure of a preclinical PI3K and mTOR dual kinase inhibitor GSK2126458 [18] identified several clinical trial drugs with similar molecular descriptors. Three of the top five drugs are of relevance to PI3K and mTOR inhibition, BEZ235 is a PI3K and mTOR dual inhibitor in a Phase II clinical trial [19], GSK1059615 is a PI3K inhibitor entered but terminated in Phase I trial [20] and AZD8055 is an mTOR inhibitor in Phase I – II trial [21].

The second case is the search of the similarity targets of a research target, which may be used for such applications as the design of active compounds based on the known active compounds against similarity targets [22]. The target similarity search tool in TTD is based on the BLAST program from NCBI [23] that ranks similarity targets based on the level of sequence similarity to the sequence of an input target with an E-value cutoff at 1 [24]. Target similarity search by inputting the sequence of yellow fever virus nonstructural protein NS3 identified the HCV nonstructural protein NS3 as a similarity target with an E-value of 2×10^{-4}. Both proteins harbor a serine protease domain responsible for most of the processing events of the nonstructural region of the HCV polyprotein [25] and yellow fever virus polyprotein [26,] respectively. Three

HCV NS3 inhibitors BILN-2061, ITMN-191 and VX-950 [27] were found in TTD, which may be considered as potential structural templates for designing yellow fever virus nonstructural protein NS3 inhibitors.

4. Application of TTD data in searching for the mechanism of multi-target drugs

The usefulness of the data in TTD can be further illustrated by studying to what extent problems such as the study of the mechanism of multi-target drugs can be facilitated based on the available data in TTD [28]. One example is the mechanism of imatinib mesilate, a revolutionary anticancer drug for the treatment of chronic myeloid leukemia primarily via selective inhibition of BCR/ABL and platelet-derived growth factor receptor (PDGFR) [29]. Both targets are listed in TTD. From expanded literature searches starting from the references provided in TTD, one finds that the oncogenic BCR/ABL kinase drives the pathogenesis of chronic myeloid leukemia by activating a variety of cellular signaling and transformation processes [30]. In addition to its role in proliferation and survival, PDGFR is crucial in regulating interstitial fluid pressure (IFP), which is a common feature of solid tumors and is thought to impede transcapillary transport of chemotherapy [31]. Simultaneous targeting of BCR/ABL and PDGFR by imatinib produces enhanced anti-leukemic effects via three collective actions. One is the main therapeutic action of BCR/ABL inhibition. The other two are facilitating actions in: i) improved tumor microenvironment via reduction of a negative factor and ii) enhanced drug transport.

Cancer is significantly influenced in tumor microenvironment, which typically involves such physiological conditions as hypoxia, low extracellular pH and high IFP [32]. Elevated IFP have been reported to contribute to tumor progression [33] and survival [34] possibly via such as mechanisms as blocking the infiltration of immune cells [32] and antibodies [35] into the tumor. Therefore, inhibition of PDGFR-β helps lowering the IFP and thus reducing a factor that favors tumor progression and survival. Many anticancer drugs, particularly high-molecular-mass compounds such as many kinase inhibitors, are transported from the circulatory system to tumor cells through the interstitial space [35]. Increased IFP decreases transcapillary transport in tumors, thereby reduces the uptake of drugs or therapeutic antibodies into the tumor [35]. Cancer cells are, therefore, exposed to a lower effective concentration of therapeutic agents than normal cells, lowering the therapeutic efficiency. Inhibition of PDGFR may enhance drug transport by reducing the IFP-mediated barrier.

5. Application of TTD data in the development of *in silico* target discovery tools

Another example of the application of TTD data is the use of the relevant data for developing sequence-based [36], structure-based [37], machine learning-based (using physicochemical and sequence features) [38], systems-based [2,4,39] *in silico* methods and the combination of these methods [5] for predicting the druggability of potential targets. The developed *in silico* methods can be tested by their evaluation results of the clinical trial targets [5].

Sequence similarity to the drug-binding domain of a successful target has been frequently used for searching potential targets on the basis that high sequence similarity to a successful target may indicate structural and functional properties suitable for drug modulation [36]. Drug-binding site structural folds tend to be more conserved than sequences, and thus structural and binding energetic analysis in comparison with those of a successful target provides useful clues to target druggability [37]. Drug-binding and modulation is strongly influenced by distinguished target-site physicochemical properties, which can be recognized by machine learning methods for classifying potential targets [38]. Moreover, several studies have shown that targets tend to show distinguished systems profiles that can be explored for druggability assessment [2,4,39]. Four *in silico* target prediction methods have been developed based on each of these four profiles generated from the relevant data of up to 316 successful targets in TTD database. Method A measures drug-binding domain sequence similarity against those of 168 successful targets with identifiable drug-binding domain. Method B studies drug-binding domain structural similarity against those of 129 successful targets with available structure. Method C predicts druggable proteins from a machine learning model trained by 316 successful targets [38]. Method D evaluates whether the systems-level druggability rules [2,4] are satisfied. More detailed descriptions about these methods are given in the supplementary material.

Each of these *in silico* methods has its unique advantages and limitations. Sequence-based methods link druggability to similarity to the drug-binding domain of a target [40], which may not fully capture druggable features un-reflected by homology [40] and tend to indiscriminately select homologous proteins. Structure-based methods evaluate druggability by structural comparison with target binding-site features [37,40], which are less effective for targets of unknown structure and for accounting systems profiles. Machine learning-based methods classify druggable targets based on the structural and physicochemical properties that separate target and non-target proteins [2,4,38], which cannot fully capture systems profiles and may disproportionately interpret certain physicochemical properties due to biases in protein descriptors or training data sets. Simple systems-level druggability rules have shown usefulness in target prediction [2,4,39], which are not intended for describing structural, physicochemical and functional aspects of druggability. These limitations may be reduced if these methods are combined [5].

The collective predictive performance of the four methods have been tested against clinical trial targets in

Table 1. List of Phase III targets in the 2008 release of the CenterWatch Drugs in Clinical Trials Database identified as a promising target by combinations of at least three of the methods A, B, C and D.

Target	Predicted as promising by combination	Number of target affiliated pathways	Number of human similarity proteins outside target family	Number of tissues target is primarily distributed	Targeted disease conditions	Target exploration status (tested drug)	Past and current development or approval status (year of report)
CCK-A receptor*	Combination of A, B, C, D	2	1	1	Irritable bowel syndrome	Phase III (dexloxiglumide)	Positive results in Phase III (2007) and a large European Phase III (2010), in talks with the FDA and EMEA for possible approval (2010)
Coagulation factor IIa*	Combination of A, B, C, D	3	0	5	Venous thromboembolism	Phase II–III (SR-123781A)	Positive results in a large European Phase III trial (2008)
NTRK1*	Combination of A, B, C, D	3	6	2	Acute myeloid leukemia	Phase II–III (lestaurtinib)	Lestaurtinib approved by the FDA as an orphan drug (2006)
5-HT_3 receptor	Combination of A, C, D	1	0	2	Irritable bowel syndrome	Phase III (cilansetron)	Positive Phase III results (2004), filed but withdrawn for FDA approval (2005), still in talks with MHRA and EU (2010)
BK-2 receptor*	Combination of A, C, D	4	0	P	Hereditary angioedema, traumatic brain injuries	Phase III (icatibant) Phase II (anatibant)	Positive Phase III results (2006), icatibant approved in the EU (2008)
Heparanase*	Combination of A, C, D	2	0	2	Hepatocellular cancer	Phase III (PI-88)	PI-88 fast tracked by the FDA (2008)
MDR 3	Combination of A, C, D	1	0	3	Acute myeloid leukemia	Phase III (LY335979)	
Orexin-OX1/OX2 receptor*	Combination of A, C, D	1	0	2	Sleep disorders	Phase III (almorexant)	Positive Phase III result (2010)
Somatostatin receptor 1	Combination of A, C, D	1	0	5	Cushing's disease, renal cell carcinoma	Phase III (pasireotide), Phase II (CAP-232)	
NK-2 receptor*	Combination of A, C, D	2	0	3	Depression	Phase III (saredutant)	Positive Phase III result (2007), trial discontinued (2009)
Thrombin receptor*	Combination of A, B, C	4	0	5	Cardiovascular disorders	Phase III (SCH-530348)	
CXCR4	Combination of A, B, D	3	2	P	Non-Hodgkin's lymphoma, late-stage solid tumors	Phase III (plerixafor), Phase I–II (AMD-070), Phase I (MSX-122)	Positive results for the treatment of multiple myeloma (2007), plerixafor approved by the FDA (2008)
C1 esterase*	Combination of A, B, D	1	3	P	Hereditary angioedema	Phase III (cinryze)	Positive results for treating hereditary angioedema, significantly decreases the number of attacks in patients (2007), cinryze approved by the FDA (2008)
NPYR5	Combination of A, B, D	1	0	2	Obesity	Phase III (CGP71683A)	
Plasma kallikrein*	Combination of A, B, D	1	0	5	Hereditary angioedema	Phase III (ecallantide)	Positive Phase III results (2007), ecallantide approved by the FDA (2009)
Sphingosine 1-phosphate receptor 1*	Combination of A, B, D	1	0	5	MS	Phase III (gilenya)	Positive Phase III results (2008), the FDA granted priority review (2010), gilenya approved by the FDA (2010)

Tissue distribution 'P' represents cases where target is distributed in > 5 tissues but the disease relevant targets are located within blood vessels or cells lining the arteries where they have higher priority to bind drugs.
*Are innovative targets without a protein subtype as a successful target.
EMEA: European Medicines Agency; MHRA: Medicine and Healthcare products Regulatory Agency.

Table 2. List of Phase I and II targets in the 2008 release of the CenterWatch Drugs in Clinical Trials Database identified as a promising target by combinations of at least three of the methods A, B, C and D.

Research target	Identified by combination	Number of target affiliated pathways	Number of human similarity proteins outside target family	Number of tissues target is primarily distributed	Targeted disease conditions	Target exploration stage (testing drug)
C-C chemokine receptor 2*	Combination of A, B, C, D	1	0	1	Rheumatoid arthritis, MS	Phase II (INCB3284), Phase I (CCX915)
ErbB-4	Combination of A, B, C, D	3	4	2	Breast cancer	Phase II (CI-1033)
FGFR-3	Combination of A, B, C, D	3	0	4	Solid tumors, multiple myeloma	Phase II (XL999), Phase I (CHIR-258)
Guanylate cyclase B*	Combination of A, B, C, D	3	0	1	Heart disease	Phase Ia (CD-NP), Preclinical (guanilib)
HDAC4	Combination of A, B, C, D	1	1	P	Basal cell carcinoma, melanoma, cancer	Phase II (avugane, romidepsin, MS-275, PXD101)
Neuropeptide Y receptor 2	Combination of A, B, C, D	1	0	4	Obesity	Phase II (obinepitide)
Neuropeptide Y receptor 4	Combination of A, B, C, D	1	0	3	Schizophrenia, schizoaffective disorders	Phase I – II (TM30339)
Toll-like receptor 3	Combination of A, B, C, D	1	0	2	Human papillomavirus infections	Phase II (HspE7)
FGFR-1	Combination of A, B, C	5	0	> 10	Coronary heart disease, solid tumors	Phase II (XL999), Phase II (FGF-1)
PKC-γ	Combination of A, B, C	16	0	4	Acute myocardial infarction	Phase II (midostaurin), Phase I – II (KAI-9803)
Tyrosine-protein kinase receptor HTK*	Combination of A, B, C	1	4	> 10	Lung cancer, solid tumors	Phase II (XL647)
Histamine H_3 receptor	Combination of A, C, D	1	0	4	Attention-deficit hyperactivity disorder, Alzheimer's disease, schizophrenia	Phase II (cipralisant), Phase I (ABT-239)
Leukotriene B4 receptor 1*	Combination of A, C, D	1	0	4	Cancer, renal cell carcinoma	Phase II (LY293111), Phase I (Biomed 101)
Motilin receptor*	Combination of A, C, D	1	0	1	Irritable bowel syndrome, gastrointestinal disorders	Phase IIb (mitemcinal), Phase I (KOS-2187)
NK-3 receptor*	Combination of A, C, D	2	0	1	Schizophrenia, schizoaffective disorders	Phase IIb (osanetant), Phase II (talnetant)
Somatostatin receptor type 4	Combination of A, C, D	1	2	3	Solid tumors	Phase II (CAP-232)
Tissue kallikrein-2*	Combination of A, C, D	1	0	2	Atopic dermatitis	Phase II (dermolastin)
Toll-like receptor 8	Combination of A, C, D	1	0	5	Genital warts, systemic lupus erythematosus	Phase II (resiquimod), Phase I (CPG 52364)
CDK7	Combination of A, B, D	1	1	P	B-cell malignancies	Phase I (SNS-032)
Coagulation factor IX*	Combination of A, B, D	1	5	1	Thrombosis, venous thromboembolism	Phase IIa (REG1), Phase I completed (TTP889)
Melanocortin receptor*	Combination of A, B, D	1	0	3	Sexual (female) and erectile dysfunction	Phase IIb (bremelanotide)
Metabotropic glutamate receptor 2/3*	Combination of A, B, D	1	0	1	Psychosis	Phase II (LY2140023, LY354740)
PPAR-δ	Combination of A, B, D	3	0	P	Obesity	Phase II (MBX-8025), Phase I (KD3010)
Serine/threonine-protein kinase Chk2	Combination of A, B, D	2	0	4	Solid tumors	Phase I (XL844), Phase I (UCN-01)
Serine/threonine-protein kinase PLK*	Combination of A, B, D	1	1	P	Pancreatic, prostate and a number of other cancers	Phase I (HMN-214)

Tissue distribution 'P' represents cases where target is distributed in > 5 tissues but the disease relevant targets are located within blood vessels or cells lining the arteries where they have higher priority to bind drugs.
*Are innovative targets without a protein subtype as a successful target.

Table 3. Statistics of identified promising targets from the 155 clinical trial and 1019 non-clinical trial targets and clinical trial target enrichment factors with respect to the 1019 research targets by combinations of method A, B, C and D used in this study.

Method or combination		No and percentage of the 30 Phase III targets predicted by method	No and percentage of the 84 Phase II targets predicted by method	No and percentage of the 41 Phase I targets predicted by method	No and percentage of the 864 non-clinical trial targets predicted as target by method	Target prediction enrichment factor for Phase II and III targets	Target prediction enrichment factor for all clinical trial targets
Combination of A, B, C, D		3 (10%)	7 (8.3%)	1 (2.4%)	4 (0.5%)	6	4.8
Any 3-combination of A, B, C, D		15 (50%)	20 (23.8%)	6 (14.6%)	30 (3.5%)	4.4	3.8
	Combination of A, B, C	4 (13.3%)	8 (9.5%)	1 (2.4%)	7 (0.8%)	5.4	4.3
	Combination of A, B, D	7 (23.3%)	12 (14.3%)	4 (9.8%)	19 (2.2%)	4	3.6
	Combination of A, C, D	10 (33.3%)	14 (16.7%)	3 (7.3%)	10 (1.2%)	5.8	4.8
	Combination of B, C, D	3 (10%)	7 (8.3%)	1 (2.4%)	5 (0.6%)	5.6	4.5
Any of A, B, C, D		26 (86.7%)	54 (64.3%)	24 (58.5%)	257 (29.7%)	2	1.9
	A	18 (60%)	39 (46.4%)	17 (41.5%)	126 (14.6%)	2.5	2.4
	B	10 (33.3%)	26 (31%)	8 (19.5%)	94 (10.9%)	2.3	2.1
	C	13 (43.3%)	21 (25%)	5 (12.2%)	42 (4.9%)	3.8	3.2
	D	21 (70%)	32 (38.1%)	13 (31.7%)	140 (16.2%)	2.3	2.1

Targets that have drugs tested in multiple phases are only included in the highest phase category.

Table 4. List of Phase III targets in the 2008 release of the CenterWatch Drugs in Clinical Trials Database dropped by combinations of at least three of the methods A, B, C and D.

Research target	Identified as promising by method or combination	Number of target affiliated pathways	Number of human similarity proteins outside target family	Number of tissues target is primarily distributed	Targeted disease conditions	Target exploration status (tested drug)
AKT	Combination of A, B	25	1	P	Non-Hodgkin's lymphoma, multiple myeloma, renal cell carcinoma	Phase III (enzastaurin) no superior efficacy vs lomustine, Phase II (perifosine), Phase II (XL880), Phase I completed (RX-0201), Phase I (XL418)
CDK2	Combination of A, B	4	0	P	Lymphocytic leukemia, lung cancer (NSCLC), non-Hodgkin's lymphoma	Phase III (flavopiridol), Phase II completed (seliciclib), Phase I–II (AT7519), Phase I (SNS-032), Preclinical (capridine-β)
Squalene synthetase	Combination of C, D	2	0	4	Hyperlipidemia	Phase III (TAK-475) discontinued
Arachidonate 5-lipoxygenase-activating protein	Only D	1	0	1	Coronary artery disease, heart attack, cardiovascular disorders	Phase III (DG031) veliflapon discontinued Phase I (AM803, AM103)
Heme oxygenase*	Only D	1	0	1	Neonatal hyperbilirubinemia, jaundice	Phase III completed (stannsoporfin)
Farnesyl protein transferase	Only D	2	0	P	Myeloid leukemia	Phase III (R115777) tipifarnib filed but deemed not approvable by the FDA
Lipoprotein-associated phospholipase A2	Only D	1	0	P	Atherosclerosis, cardiovascular disorders	Phase II – III (darapladib), phase I (659032)
MMP-12	Only D	1	0	4	Lung cancer (NSCLC)	Phase III (AE-941) neovastat does not improve chemotherapy
Myophosphorylase	Only D	2	0	1	Lymphocytic leukemia, diabetes mellitus	Phase III (flavopiridol), Phase IIa (PSN357)
Neutral endopeptidase	Only D	3	0	P	Hypertension, congestive heart failure	Phase II – III (ilepatril) discontinued, Phase II (SLV 306)
Sphingosine kinase	Only D	3	0	4	Ovarian cancer	Phase III (phenoxodiol) failed Phase III trial 2010
Heat-shock protein HSP90	Only C	1	0	> 10	Multiple myeloma, metastatic breast cancer, prostate cancer	Phase III (tanespimycin) discontinued, Phase II (alvespimycin hydrochloride, IPI-504), Phase I (CNF1010, SNX-5422, STA-9090), IND filed (AT13387)
Cathepsin K	None	No info	0	4	Osteoporosis, bone metastases	Phase III (odanacatib), Phase II (relacatib), Phase I – II (MIV-701),
MMP-2/-9	None	3	0	6	Lung cancer (NSCLC), osteoarthritis	Phase III (neovastat) neovastat does not improve chemotherapy Phase II (PG-530742)

*Has a positive Phase III result reported in 2004, but since then there has been no report about the further progress of the Phase III drug. Tissue distribution 'P' represents cases where target is distributed in > 5 tissues, but the disease relevant targets are located within blood vessels or cells lining the arteries where they have higher priority to bind drugs.

the 2008 release of the CenterWatch Drugs in Clinical Trials Database and non-clinical trial research targets [5]. It is noted that two targets thrombopoietin receptor and IL-1β approved in 2008 [41] are not in the Phase III target list, and one target α-glucosidase approved in 1995 is misplaced into the Phase III target list. These three targets are thus not analyzed here. Clinical trial targets that have drugs in multiple phases are only included in the highest phase category.

The best overall performance is produced by the combination of at least three methods, which maximize the collective predictive capability of the methods and minimize the impact of limited structural availability [5]. This combination identifies 53% of the 30 Phase III (**Table 1**), 24 and 15% of the 84 Phase II and 41 Phase I (**Table 2**), and 4% of the 864 non-clinical trial research targets as promising. There is no published report about target success rates in different drug development stages. It is noted that the reported probabilities of successes in developing systemic broad spectrum antibacterials are 67, 50, 25 and 3% in Phase III, Phase II, Phase I and preclinical stages [42]. The percentages of the identified promising clinical trial targets are lower than but roughly follow a similar descending trend as the reported drug development rates. The overall performance of different combinations is given in **Table 3**. These combinations enriched Phase II and III target identification rate by 4 ~ 6-fold over random selection, with the combination of all four methods producing the highest enrichment.

Of the 16 predicted 'promising' Phase III targets in the 2008 release of the CenterWatch Drugs in Clinical Trials Database in **Table 1**, 6 (37%) targets have been approved and another 5 (31%) have shown positive Phase III results, respectively. Of the 15 predicted 'non-promising' Phase III targets in the 2008 release of the CenterWatch Drugs in Clinical Trials Database in **Table 4**, 13 targets has no positive results reported so far. One target (farnesyl protein transferase) with a drug (tipifarnib) filed for approval but was deemed not approvable by FDA. On the other hand, three targets (heat-shock protein 90, squalene synthetase and arachidonate 5-lipoxygenase-activating protein) are with their Phase III drugs (TAK-475, veliflapon and tanespimycin) discontinued, and five targets (AkT, MMP-12, MMP-2, MMP-9, and sphingosine kinase) with their Phase III drugs (enzastaurin, neovastat and phenoxodiol) produced negative results. Therefore, target druggability and level of difficulty in its exploration appear to be strongly associated with its genetic, structural, physicochemical and systems profiles [5].

6. Expert opinion

TTD database is intended to be a useful resource in complementing other related databases by providing comprehensive information about the primary targets and other drug data for the approved, clinical trial and experimental drugs [6]. The primary target data in TTD are useful for facilitating such studies as the mechanism of drug actions [43] and for the development of multiple *in silico* target discovery methods [2,4,36-39] that are in combination capable of identifying high percentages of Phase III targets including most of the targets of positive Phase III results and eliminating difficult and un-promising targets [5]. Comparative analysis of multiple profiles of the targets in TTD also provides useful clues to the identification of promising targets [2,4,39]. TTD data and their usefulness can be further enhanced along with expanded knowledge of the genomic, proteomic, structural, functional and systems profiles of therapeutic targets [2,5,44-46] and with further development of target discovery and validation methods [5,47]. As in the case of PDTD [8], some of the virtual screening methods [48] and data sets may be included in drug discovery databases for facilitating target oriented drug lead discovery.

Given the low number of innovative targets that have made to the approved list [4], there is a need to carefully evaluate and select a target from the pool of > 1000 research targets [1,2] before the start of a drug discovery process so as to increase the odds of successful development. The druggability of a target is determined by multiple factors such as the structural and physicochemical properties that accommodate the binding and modulation by drug-like molecules [36-38] and the systems profiles that influence the toxicity and pharmacokinetic features of the binding drugs [2,4,39]. Evaluation of these profiles with respect to successful targets is important for facilitating the selection of druggable targets suitable for target discovery [5].

Rapid progress in genomics [45], structures [40] and proteomics is revolutionizing target discovery, which is further facilitated by advances in high-throughput technologies [49], and cellular [50] and physiological studies [51]. These progresses combined with increased molecular understanding of diseases and their targets [2] can be incorporated into drug discovery databases to facilitate the access of the relevant information and the development of efficient tools for identifying innovative targets of new therapies. The knowledge and analysis of the multiple targets of multi-target drugs [46,52] and the individual variations in response to the targeting of different combinations of these multiple targets [53] can also facilitate the development of personalized medicine for enhanced therapeutic efficacies.

Declaration of interest

The authors state no conflict of interest and have received no payment in preparation of this manuscript.

Bibliography

1. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov 2006;5(12):993-6
2. Zheng CJ, Han LY, Yap CW, et al. Therapeutic targets: progress of their exploration and investigation of their characteristics. Pharmacol Rev 2006;58(2):259-79
3. Hewitt DJ, Hargreaves RJ, Curtis SP, et al. Challenges in analgesic drug development. Clin Pharmacol Ther 2009;86(4):447-50
4. Zheng C, Han L, Yap CW, et al. Progress and problems in the exploration of therapeutic targets. Drug Discov Today 2006;11(9-10):412-20
5. Zhu F, Han L, Zheng C, et al. What are next generation innovative therapeutic targets? Clues from genetic, structural, physicochemical, and systems profiles of successful targets. J Pharmacol Exp Ther 2009;330(1):304-15
6. Zhu F, Han B, Kumar P, et al. Update of TTD: Therapeutic Target Database. Nucleic Acids Res 2010;38(Database issue):D787-91
7. Wishart DS, Knox C, Guo AC, et al. DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res 2008;36(Database issue):D901-6
8. Gao Z, Li H, Zhang H, et al. PDTD: a web-accessible protein database for drug target identification. BMC Bioinformatics 2008;9:104
9. Therapeutic target database. Available from: http://bidd.nus.edu.sg/group/cjttd/TTD.asp
10. Yildirim MA, Goh KI, Cusick ME, et al. Drug-target network. Nat Biotechnol 2007;25(10):1119-26
11. Wishart DS, Knox C, Guo AC, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res 2006;34(Database issue):D668-72
12. Lindsay MA. Target discovery. Nat Rev Drug Discov 2003;2(10):831-8
13. Drugs@FDA. Available from: http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm
14. Altschul SF, Madden TL, Schaffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25(17):3389-402
15. Willett P, Barnard JM, Downs GM. Chemical Similarity Searching. J Chem Inf Comput Sci 1998;38(6):983-96
16. Hull RD, Fluder EM, Singh SB, et al. Chemical similarity searches using latent semantic structural indexing (LaSSI) and comparison to TOPOSIM. J Med Chem 2001;44(8):1185-91
17. Hert J, Keiser MJ, Irwin JJ, et al. Quantifying the relationships among drug classes. J Chem Inf Model 2008;48(4):755-65
18. Hoang B, Frost P, Shi Y, et al. Targeting TORC2 in multiple myeloma with a new mTOR kinase inhibitor. Blood 2010;116(22):4560-8
19. Santiskulvong C, Konecny GE, Fekete M, et al. Dual targeting of phosphoinositide 3-kinase and mammalian target of rapamycin using NVP-BEZ235 as a novel therapeutic approach in human ovarian carcinoma. Clin Cancer Res 2011;17(8):2373-84
20. Brachmann S, Fritsch C, Maira SM, et al. PI3K and mTOR inhibitors: a new generation of targeted anticancer agents. Curr Opin Cell Biol 2009;21(2):194-8
21. Chresta CM, Davies BR, Hickson I, et al. AZD8055 is a potent, selective, and orally bioavailable ATP-competitive mammalian target of rapamycin kinase inhibitor with in vitro and in vivo antitumor activity. Cancer Res 2010;70(1):288-98
22. Anand K, Ziebuhr J, Wadhwani P, et al. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. Science 2003;300(5626):1763-7
23. BLAST: Basic Local Alignment Search Tool. Available from: http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD = Web&PAGE_TYPE = BlastHome
24. Gerstein M. Measurement of the effectiveness of transitive sequence comparison, through a third 'intermediate' sequence. Bioinformatics 1998;14(8):707-14
25. Steinkuhler C, Biasiol G, Brunetti M, et al. Product inhibition of the hepatitis C virus NS3 protease. Biochemistry 1998;37(25):8899-905
26. Chambers TJ, Weir RC, Grakoui A, et al. Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. Proc Natl Acad Sci USA 1990;87(22):8898-902
27. Geitmann M, Dahl G, Danielson UH. Mechanistic and kinetic characterization of hepatitis C virus NS3 protein interactions with NS4A and protease inhibitors. J Mol Recognit 2011;24(1):60-70
28. Zhu F, Zheng CJ, Han LY, et al. Trends in the exploration of anticancer targets and strategies in enhancing the efficacy of drug targeting. Curr Mol Pharmacol 2008;1(3):213-32
29. Gill AL, Verdonk M, Boyle RG, et al. A comparison of physicochemical property profiles of marketed oral drugs and orally bioavailable anti-cancer protein kinase inhibitors in clinical development. Curr Top Med Chem 2007;7(14):1408-22
30. Quintas-Cardama A, Kantarjian H, Cortes J. Flying under the radar: the new wave of BCR-ABL inhibitors. Nat Rev Drug Discov 2007;6(10):834-48
31. Bauman JE, Eaton KD, Martins RG. Antagonism of platelet-derived growth factor receptor in non small cell lung cancer: rationale and investigations. Clin Cancer Res 2007;13(15 Pt 2):s4632-6
32. Liao YP, Schaue D, McBride WH. Modification of the tumor microenvironment to enhance immunity. Front Biosci 2007;12:3576-600
33. Milosevic M, Fyles A, Hedley D, et al. The human tumor microenvironment: invasive (needle) measurement of oxygen and interstitial fluid pressure. Semin Radiat Oncol 2004;14(3):249-58
34. Milosevic M, Fyles A, Hedley D, et al. Interstitial fluid pressure predicts survival in patients with cervix cancer independent of clinical prognostic factors and tumor oxygen measurements. Cancer Res 2001;61(17):6400-5
35. Heldin CH, Rubin K, Pietras K, et al. High interstitial fluid pressure - an

obstacle in cancer therapy. Nat Rev Cancer 2004;4(10):806-13

36. Orth AP, Batalov S, Perrone M, et al. The promise of genomics to identify novel therapeutic targets. Expert Opin Ther Targets 2004;8(6):587-96
37. Cheng AC, Coleman RG, Smyth KT, et al. Structure-based maximal affinity model predicts small-molecule druggability. Nat Biotechnol 2007;25(1):71-5
38. Han LY, Zheng CJ, Xie B, et al. Support vector machines approach for predicting druggable proteins: recent progress in its exploration and investigation of its usefulness. Drug Discov Today 2007;12(7-8):304-13
39. Yao L, Rzhetsky A. Quantitative systems-level determinants of human genes targeted by successful drugs. Genome Res 2008;18(2):206-13
40. Hajduk PJ, Huth JR, Tse C. Predicting protein druggability. Drug Discov Today 2005;10(23-24):1675-82
41. Hughes B. 2008 FDA drug approvals. Nat Rev Drug Discov 2009;8(2):93-6
42. Payne DJ, Gwynn MN, Holmes DJ, et al. Drugs for bad bugs: confronting the challenges of antibacterial discovery. Nat Rev Drug Discov 2007;6(1):29-40
43. Canfield ML, Walker WR, Brown LG. Contingency interaction analysis in psychotherapy. J Consult Clin Psychol 1991;59(1):58-66
44. Edwards A. Large-scale structural biology of the human proteome. Annu Rev Biochem 2009;78:541-68
45. Kramer R, Cohen D. Functional genomics to new drug targets. Nat Rev Drug Discov 2004;3(11):965-72
46. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. Nat Chem Biol 2008;4(11):682-90
47. Lee GM, Craik CS. Trapping moving targets with small molecules. Science 2009;324(5924):213-15
48. Li H, Yap CW, Ung CY, et al. Machine learning approaches for predicting compounds that interact with therapeutic and ADMET related proteins. J Pharm Sci 2007;96(11):2838-60
49. Ilag LL, Ng JH, Beste G, et al. Emerging high-throughput drug target validation technologies. Drug Discov Today 2002;7(18 Suppl):S136-42
50. Jackson PD, Harrington JJ. High-throughput target discovery using cell-based genetics. Drug Discov Today 2005;10(1):53-60
51. Lindsay MA. Finding new drug targets in the 21st century. Drug Discov Today 2005;10(23-24):1683-7
52. Ma XH, Shi Z, Tan C, et al. In-silico approaches to multi-target drug discovery : computer aided multi-target drug design, multi-target virtual screening. Pharm Res 2010;27(5):739-49
53. Stegmeier F, Warmuth M, Sellers WR, et al. Targeted cancer therapies in the twenty-first century: lessons from imatinib. Clin Pharmacol Ther 2010;87(5):543-52

Affiliation

Xin Liu[1,2] PhD, Feng Zhu[1] PhD, Xiaohua Ma[1] PhD, Lin Tao[1] PhD, Jingxian Zhang[1] PhD, Shengyong Yang[2] PhD, Yuquan Wei[2] PhD & Yu Zong Chen[†1,2] PhD
[†]Author for correspondence
[1]Bioinformatics and Drug Design Group, Department of Pharmacy, Centre for Computational Science and Engineering, National University of Singapore, Blk S16, Level 8, 3 Science Drive 2, Singapore 117543, Tel: +65 6516 6877; Fax: +65 6774 6756; E-mail: phacyz@nus.edu.sg
[2]State Key Laboratory of Biotherapy, Sichuan University, Chengdu 610064, P.R.China

Supplementary material available online

Tables S1 and S2
Materials and Methods