基于网页的药物再利用工具

中国制造2025中与生信密切相关的是生物医药及高性能医疗器械，『发展针对重大疾病的化学药、中药、生物技术药物新产品，重点包括新机制和新靶点化学药、抗体药物、抗体偶联药物、全新结构蛋白及多肽药物、新型疫苗、临床优势突出的创新中药及个性化治疗药物。』。正好今天看到一篇药物再利用相关的文章，给大家分享一下。

这篇文章主要介绍了药物再利用的网页工具。大家可以利用这些工具，看自己研究的基因或蛋白是否与某些药物有关，或者是否可以作为某些药物的作用靶点等~~~

> 文章：
> Web-based drug repurposing tools: a survey. 『Briefings in Bioinformatics』

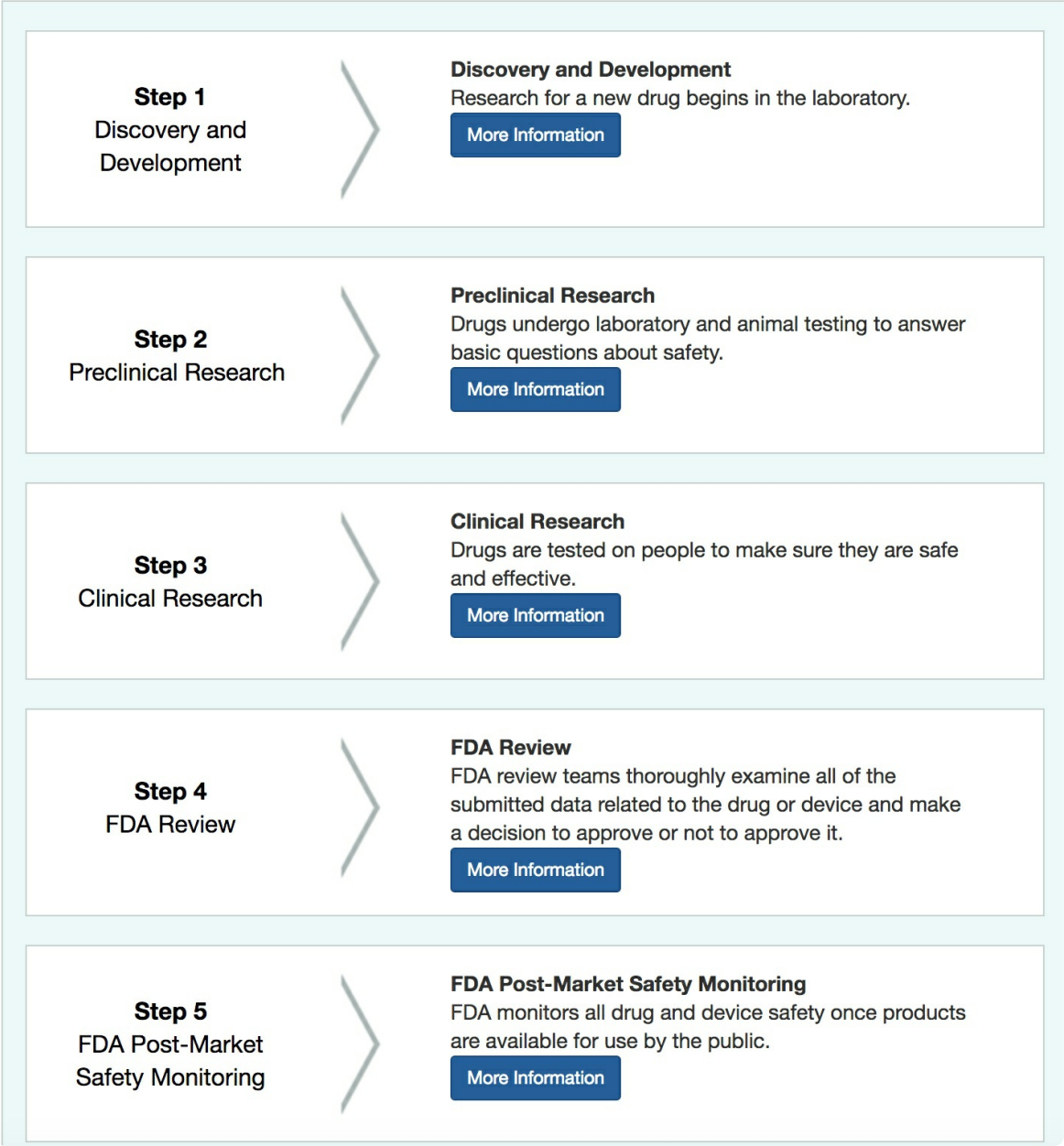
```
drug repurposing/drug repositioning: 药物再利用
wikipedia解释: Drug repositioning (also known as drug repurposing, re-profiling,
re-tasking or therapeutic switching) is the application of known drugs and compou
nds to treat new indications (i.e., new diseases). [https://en.wikipedia.org/wiki
/Drug_repositioning]
_____________________________________________________
ncbi解释: Repurposing generally refers to studying drugs that are already approve
d to treat one disease or condition to see if they are safe and effective for tre
ating other diseases. [https://ncats.nih.gov/preclinical/repurpose#learn-more]
简而言之，研究已有药物的性质，看是否能安全有效的应用于其他疾病。
```

```
Eroom’s law: 倒摩尔定律
wikipedia解释: Eroom's law is the observation that drug discovery is becoming slo
wer and more expensive over time, despite improvements in technology (such as hig
h throughput screening, biotechnology, combinatorial chemistry, and computational
 drug design), a trend first observed in the 1980s. The cost of developing a new
drug roughly doubles every nine years (inflation-adjusted)
倒摩尔定律：  尽管技术不断进步，但药物研发的速度越来越慢、越来越快
摩尔定律：  当价格不变时，集成电路上可容纳的元器件的数目，约每隔18-24个月便会增加一倍，性能也将提升一倍
```

> 药物研发流程
> https://www.fda.gov/ForPatients/Approvals/Drugs/default.htm

The Drug Development Process

SHARE | TWEET | LINKEDIN | PIN IT | EMAIL | PRINT



随着药物研发的越来越快、越来越贵，而且临床试验阶段就中止的药物比例越来越高，药物再利用渐渐成为对抗药物失败风险的有效途径。一些计算分析和挖掘方法常被用于探索大量的生物学、生物医学数据。但很多计算工具常常难以使用，限制了非计算背景的科学家使用。所以用户友好的网页工具变的非常必要。

> The ideal candidates for repurposing are leads which have made it past Phase III, in terms of the American Food and Drug Administration (FDA) system, as this implies they are proven to be efficacious in larger populations and verified to be safe

> The recent times have seen many successes in repositioning old drugs (see **RepurposeDB** for a list of repurposed drugs),

有以下特点的网页工具将被排除在外：

1. Web services that only provide a collection of repurposed drugs: RepurposeDB , The Drug Repurposing Hub and repoDB;

2. Tools whose outputs do not provide a direct way of inferring the repurposing prediction (e.g. they may generate pharmacophores, weighted or non-weighted inter action networks): PharmMapper , PharmaGist, ProSMoS and VisANT;

3. Studies that focus on any one single family, or some functional subset of pro teins, or a single disease: iCDI-PseFpt, AlzPlatform and ACTP;

4. Resources that only aggregate databases and provide associ- ations by connect ing them, without using any predictive algorithm-based analysis: Pharos, SIDER, D Tome, ChEMBL and PubChem;

5. Web-UI tools that predict interactions of molecules with non-protein targets: ChemiRs;

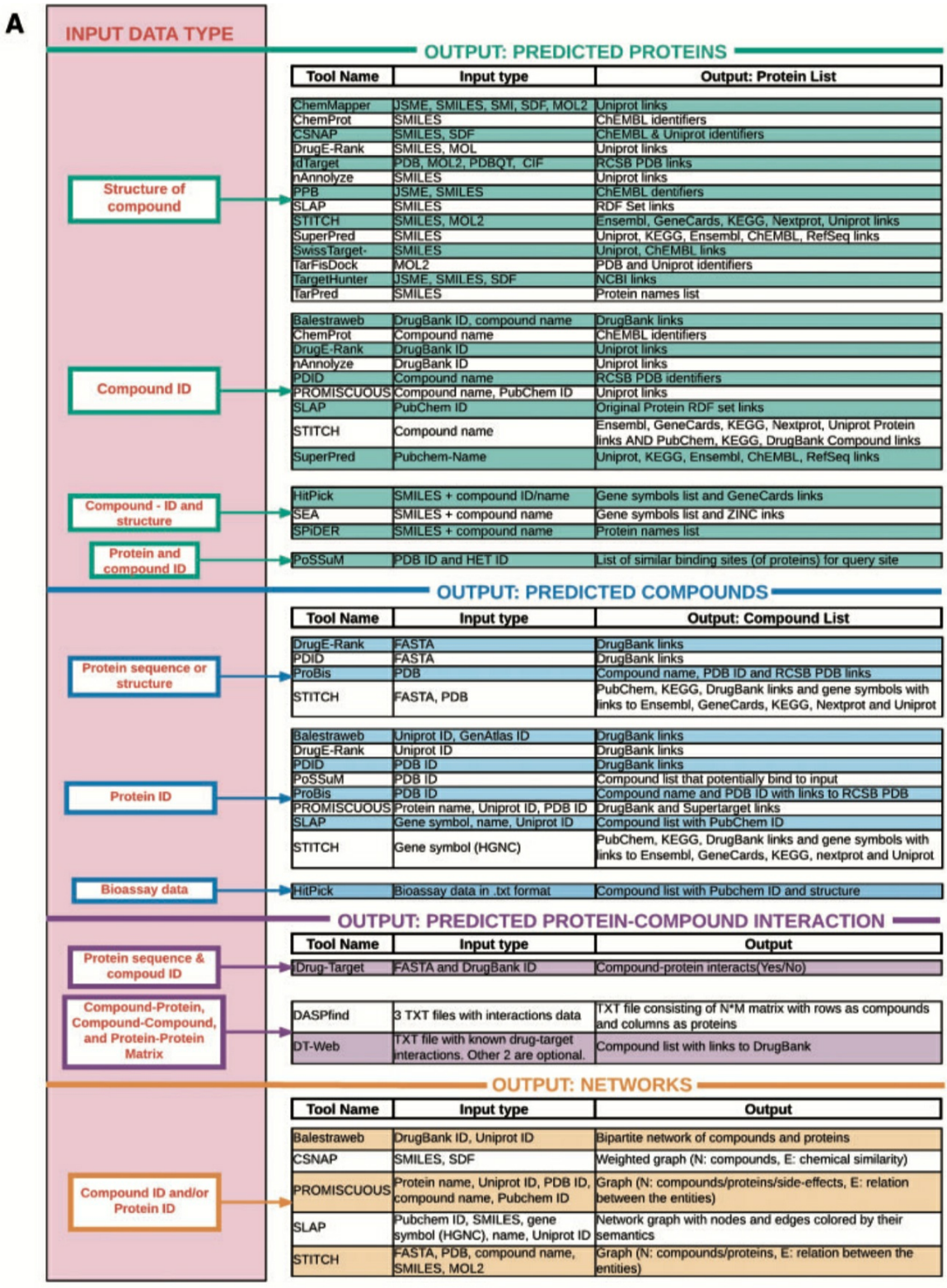
6. Web portals not accessible during the time of this review (authors have been contacted around the month of May 2017, if the Web site was down)

7. Studies that are not published in peer-reviewed journals;

作者将网页工具分为三大类：

1. Predicting drug–target interactions
    1. Ligand similarity using fingerprint encoding
    2. 3D structures of drug and targets
    3. Biological networks
    4. Binding site parametrization
    5. Others
2. Linking drugs to disease
3. Using drug–induced gene expression to predict new connections

Web servers predicting drug–target interactions.



Tool Name	Input type	Output: Protein List
ChemMapper	JSME, SMILES, SMI, SDF, MOL2	Uniprot links
ChemProt	SMILES	ChEMBL identifiers
CSNAP	SMILES, SDF	ChEMBL & Uniprot identifiers
DrugE-Rank	SMILES, MOL	Uniprot links
idTarget	PDB, MOL2, PDBQT, CIF	RCSB PDB links
nAnnolyze	SMILES	Uniprot links
PPB	JSME, SMILES	ChEMBL dentifiers
SLAP	SMILES	RDF Set links
STITCH	SMILES, MOL2	Ensembl, GeneCards, KEGG, Nextprot, Uniprot links
SuperPred	SMILES	Uniprot, KEGG, Ensembl, ChEMBL, RefSeq links
SwissTarget-	SMILES	Uniprot, ChEMBL links
TarFisDock	MOL2	PDB and Uniprot identifiers
TargetHunter	JSME, SMILES, SDF	NCBI links
TarPred	SMILES	Protein names list
Balestraweb	DrugBank ID, compound name	DrugBank links
ChemProt	Compound name	ChEMBL identifiers
DrugE-Rank	DrugBank ID	Uniprot links
nAnnolyze	DrugBank ID	Uniprot links
PDID	Compound name	RCSB PDB identifiers
PROMISCUOUS	Compound name, PubChem ID	Uniprot links
SLAP	PubChem ID	Original Protein RDF set links
STITCH	Compound name	Ensembl, GeneCards, KEGG, Nextprot, Uniprot Protein links AND PubChem, KEGG, DrugBank Compound links
SuperPred	Pubchem-Name	Uniprot, KEGG, Ensembl, ChEMBL, RefSeq links
HitPick	SMILES + compound ID/name	Gene symbols list and GeneCards links
SEA	SMILES + compound name	Gene symbols list and ZINC inks
SPiDER	SMILES + compound name	Protein names list
PoSSuM	PDB ID and HET ID	List of similar binding sites (of proteins) for query site

Tool Name	Input type	Output: Compound List
DrugE-Rank	FASTA	DrugBank links
PDID	FASTA	DrugBank links
ProBis	PDB	Compound name, PDB ID and RCSB PDB links
STITCH	FASTA, PDB	PubChem, KEGG, DrugBank links and gene symbols with links to Ensembl, GeneCards, KEGG, Nextprot and Uniprot
Balestraweb	Uniprot ID, GenAtlas ID	DrugBank links
DrugE-Rank	Uniprot ID	DrugBank links
PDID	PDB ID	DrugBank links
PoSSuM	PDB ID	Compound list that potentially bind to input
ProBis	PDB ID	Compound name and PDB ID with links to RCSB PDB
PROMISCUOUS	Protein name, Uniprot ID, PDB ID	DrugBank and Supertarget links
SLAP	Gene symbol, name, Uniprot ID	Compound list with PubChem ID
STITCH	Gene symbol (HGNC)	PubChem, KEGG, DrugBank links and gene symbols with links to Ensembl, GeneCards, KEGG, nextprot and Uniprot
HitPick	Bioassay data in .txt format	Compound list with Pubchem ID and structure

Tool Name	Input type	Output
iDrug-Target	FASTA and DrugBank ID	Compound-protein interacts(Yes/No)
DASPfind	3 TXT files with interactions data	TXT file consisting of N*M matrix with rows as compounds and columns as proteins
DT-Web	TXT file with known drug-target interactions. Other 2 are optional.	Compound list with links to DrugBank

Tool Name	Input type	Output
Balestraweb	DrugBank ID, Uniprot ID	Bipartite network of compounds and proteins
CSNAP	SMILES, SDF	Weighted graph (N: compounds, E: chemical similarity)
PROMISCUOUS	Protein name, Uniprot ID, PDB ID, compound name, Pubchem ID	Graph (N: compounds/proteins/side-effects, E: relation between the entities)
SLAP	Pubchem ID, SMILES, gene symbol (HGNC), name, Uniprot ID	Network graph with nodes and edges colored by their semantics
STITCH	FASTA, PDB, compound name, SMILES, MOL2	Graph (N: compounds/proteins, E: relation between the entities)

Web servers linking drugs to disease

B

INPUT DATA TYPE

Compound name

Disease name

Protein ID

OUTPUT: PREDICTED DISEASES

Tool Name	Input type	Output
MeSHDD	Compound name	Disease list
RE:fine drugs	Compound name	Disease list
MeSHDD	Disease name	Compound list
RE:fine drugs	Disease name	Compound list

OUTPUT: PREDICTED COMPOUNDS AND DISEASES

Tool Name	Input type	Output
RE:fine drugs	Gene symbol (HGNC)	List of disease name + compound name

Web servers using drug-induced gene expressions to predict new connections

C

INPUT DATA TYPE

Gene expression

OUTPUT: PREDICTED COMPOUNDS

Tool Name	Input type	Output: Compund List
CMap	List of up/down regulated genes	Names
DeSigN	List of up/down regulated genes	With protein names
GoPredict	List of gene ID (Ensembl) and status	Names
L1000CDS2	List of up/down regulated genes	With Pubchem links
NFFinder	List of up/down regulated genes	With disease name
PDOD	List of gene ID (Entrez) and status	Names
MANTRA 2.0	CEL, Gene symbols(HGNC)	With network (Nodes: compounds, E: distance)

所有网页工具的简单信息

Table 1. List of web servers

Predicting drug–target interactions							
Name	Method	Databases	URL	Date of last update	No. of citations	Batch queries	API
Ligand-similarity using fingerprint encoding							
ChemMapper[a]	3D similarity (chemotype features and molecular shape) calculation	ChEMBL [24], DrugBank [27], BindingDB [28], KEGG [29] and PDB [30]	http://lilab.ecust.edu.cn/chemmapper/	Dec 2016	52	No	No
ChemProt 3.0	2D similarity-based approach	ChEMBL [24], DrugBank [27], BindingDB [28], STITCH [31], PharmGKB [32], IUPHAR [33], Ki Database [34], CTD [35], WOMBAT [36]	http://potentia.cbs.dtu.dk/ChemProt/	Jan 2015	105[b]	No	No
HitPick	1NN similarity search and Laplacian-modified Naïve Bayesian target models	STITCH 3.1 [37]	http://mips.helmholtz-muenchen.de/proj/hitpick	May 2013[c]	28	Yes	No
iDrug-Target	Fingerprint-based approach with machine learning	KEGG [38]	http://www.jci-bioinfo.cn/iDrug-Target/	Jan 2015[c]	72	Yes	No
PPB	Multi fingerprint-based approach. Ten different fingerprints	ChEMBL [24]	http://gdbtools.unibe.ch:8080/PPB/	Nov 2016	0	No	No
SEA	2D similarity-based approach	MDDR [39], ChEMBL [24], WOMBAT [36]	http://sea.bkslab.org/	Feb 2007	826	No	No
SuperPred	2D similarity-based approach	SuperTarget [40], ChEMBL [24], BindingDB [28]	http://prediction.charite.de	Apr 2014	107[b]	No	No
SwissTarget-Prediction	Combination of 2D and 3D similarity approach	ChEMBL [24]	http://www.swisstargetprediction.ch	Jul 2014[c]	52	No	No
TarPred	KNN-based data fusion with 2D fingerprint-based similarity	DrugBank [27], BindingDB [28], TTD [41]	http://www.dddc.ac.cn/tarpred	Jun 2015[c]	8	No	No
TargetHunter	2D similarity-based approach	ChEMBL [24]	http://www.cbligand.org/TargetHunter/	Oct 2016	46	No	No
3D structures of drug and targets							
idTarget	Divide-and-conquer-based docking approach	PDB [42]	http://idtarget.rcas.sinica.edu.tw/	Aug 2015	59	No	No
PDID	Predictions generated using ILbind, SMAP and eFindSite	PDB [42]	http://biomine.ece.ualberta.ca/PDID/	May 2015	3	No	No
TarFisDock	Reverse ligand-protein docking approach	PDB [42]	http://www.dddc.ac.cn/tarfisdock/	Aug 2014	234	No	No
Biological networks							
Balestraweb	PMF method	DrugBank [27]	http://balestra.csb.pitt.edu/	Jun 2015	8	No	No
CSNAP	CSN-based approach	ChEMBL [24]	https://services.mbi.ucla.edu/CSNAP/index.html	Aug 2015	12	No	No
DASPfind	Network-based approach	BRENDA [43], SuperTarget [40], DrugBank [27] and KEGG [38]	http://www.cbrc.kaust.edu.sa/daspfind/	Mar 2016[c]	6	Yes	No
DT-Web	Recommendation-based approach	DrugBank [27]	http://alpha.dmi.unict.it/dtweb/	Jun 2015[c]	11	Yes	No
PROMISCUOUS	Network-based approach evolving around target–target and drug–target interactions along with side effects	SuperDrug [44], UniProt [45], PDB [46] and SIDER [21]	http://bioinformatics.charite.de/promiscuous/	May 2011	134	Yes	No
SLAP	Semantic Link Association-based Prediction	Chem2Bio2RDF [47], Chem2Bio2OWL ontology [48], DrugBank [27]	http://chem2bio2rdf.org/slap	Dec 2010	66	Yes	Yes
STITCH	Search Tool for Interacting Chemicals	DrugBank [27], GLIDA [49], MATADOR [50], TTD [41], CTD [35], KEGG [38], PID [51],	http://stitch.embl.de/	Dec 2016	890 [b]	Yes	Yes

Name	Method	Databases	URL	Date of last update	No. of citations	Batch queries	API
		Reactome [52], BioCyc [53], ChEMBL [24], PDSP Ki Database [34], PDB [45]					
Binding site parameterization							
ProBis	Fast maximum clique algorithm	ProBiS-Database [54]	http://probis.cmm.ki.si/	Jul 2015	60	No	No
PoSSuM	All-pairs similarity	PDB [42], ChEMBL [24], UniProt [45], GO [55], PDBSprotEC [56], CATH [57], SCOP [58], SCOPe [59]	http://possum.cbrc.jp/PoSSuM/	Oct 2014	31 [b]	No	No
Other approaches							
DrugE-Rank	Feature-based and similarity-based approach	DrugBank [27]	http://datamining-iip.fudan.edu.cn/service/DrugE-Rank	Jun 2016	6	No	No
DR.PRODIS[a]	FINDSITEcomb algorithm	DrugBank [27], ChEMBL [24], PDB [46], SIDER [21], COSMIC [60], OMIM [61]	http://cssb.biology.gatech.edu/dr.prodis/	Jun 2015[c]	20	No	No
SPiDER	Self-organizing map-based prediction	COBRA [62]	http://modlabcadd.ethz.ch/software/spider/	Jun 2017	56	No	No
LINKING DRUGS TO DISEASE							
MeSHDD	Repositioning based on drug–drug similarity	MEDLINE, DrugBank [27]	http://apps.chiragjpgroup.org/MeSHDD/	Dec 2016[c]	3	No	No
RE:fine drugs[a]	Repositioning based on drug–protein and Gene–Disease interaction	DrugBank [27], GWAS [63], PheWAS [64]	http://drug-repurposing.nationwidechildrens.org	Dec 2016[c]	8	No	No
Using drug-induced gene expression profiles to predict new connections							
CMap	Discovering patterns of association between drug sensitivity and gene expression signatures	GEO [65]	http://www.broad.mit.edu/cmap	Jun 2017	2801[b]	Yes	Yes
DeSigN	Global baseline DEGs to drug response	GDSC [66]	http://design.cancerresearch.my/	May 2016	0	Yes	No
GoPredict	Integration of genomic, transcriptomic and signaling pathway data to allow drug repurposing	TCGA [67], Tumorscape [68], COSMIC [60], DrugBank [27], KEGGDrug, GO [55]	http://csblcanges.fimm.fi/GOPredict/	Aug 2014	2	Yes	No
L1000CDS[2]	Predict compounds that either reverse or mimic an input gene expression signature	GEO [65], CCLE [69], CMap [70]	http://amp.pharm.mssm.edu/L1000CDS2/	Mar 2017	9	Yes	Yes
MANTRA 2.0	Network-based and nonparametric statistics on gene expression data	CMap [70]	http://mantra.tigem.it/	Feb 2014 [c]	431[b]	Yes	No
NFFinder	Matching similar gene expression signatures and discovering new connections between drugs, genes and diseases related to the same biological process	GEO [65], CMap [70], DrugMatrix [71]	http://nffinder.cnb.csic.es/	Jul 2015[c]	12	Yes	No
PDOD	Predict drugs having opposite effects on altered states of disease genes	KEGG [38], DrugBank [27], CTD [35], GEO [65]	http://gto.kaist.ac.kr/pdod/index.php/main	Jan 2016	1	Yes	No

[a]Need to contact authors for commercial usage.
[b]Sum total of citations of references of all versions of the tool.
[c]Date of publication is provided because the authors were unreachable to confirm the date of last update.

具体到每个工具的信息，还需要有兴趣的朋友自己看相关论文。

如果各位对计算方法、工具感兴趣的话，可以看

『On the Integration of In Silico Drug Design Methods for Drug Repurposing』