



Tri-Con

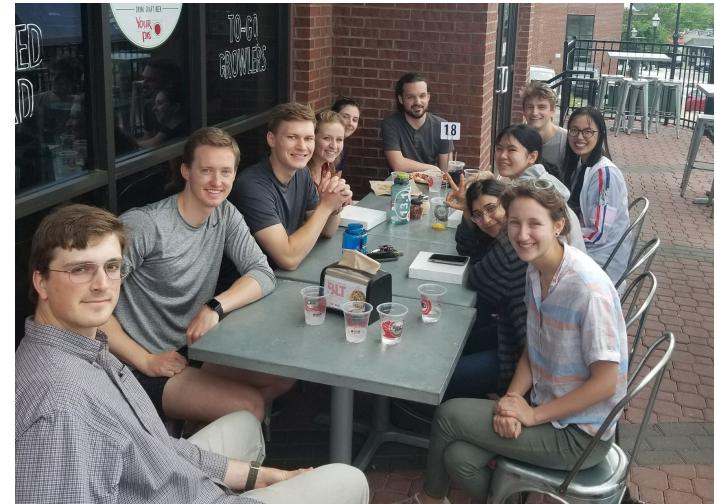
Preclinical Drug Discovery Hackathon

Reed Bender, Benafsh Husain Sapra, Ben Shealy,
Jeannette Koschmann



Who are we?

- The Systems Genetics lab at Clemson
- Reed Bender
 - M.S. Student in Biomedical Data Science and Informatics
- Benafsh Husain Sapra
 - PhD Candidate in Biomedical Data Science and Informatics
- Ben Shealy
 - PhD Candidate in Computer Engineering



Who's the Patient?



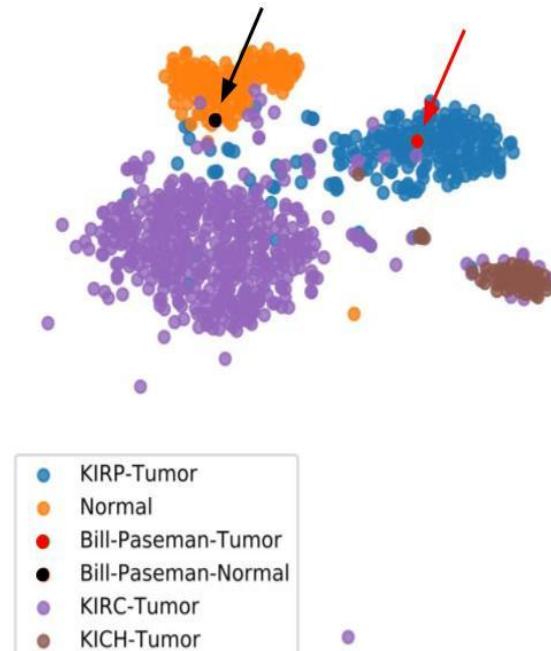
Tony Stark: “I shouldn’t do anything. They could kill you, they’re gonna kill me, either way, and even if they don’t, I’ll probably be dead in a week.”

Yinsen: “Then this is a very important week for you, isn’t it?”

What do we have?

t-SNE Clustering of Unified RNA-sequencing data

- Unified RNA-sequencing data
 - TCGA kidney cancer data (KIRP, KICH, KIRC)
 - GTEx normal tissue data
 - Bill Paseman's tumor and normal RNAseq
- Aligned BAM files with Bill Paseman's tumor DNA sequences
- Candidate gene lists for potential kidney-cancer specific aberrations



What are our goals?

Short-term Goals

- Process the unified gene expression matrix to identify novel kidney cancer markers
- Analyze candidate gene lists and relationships to search for potential targets
- Identify potential therapeutics for those targets

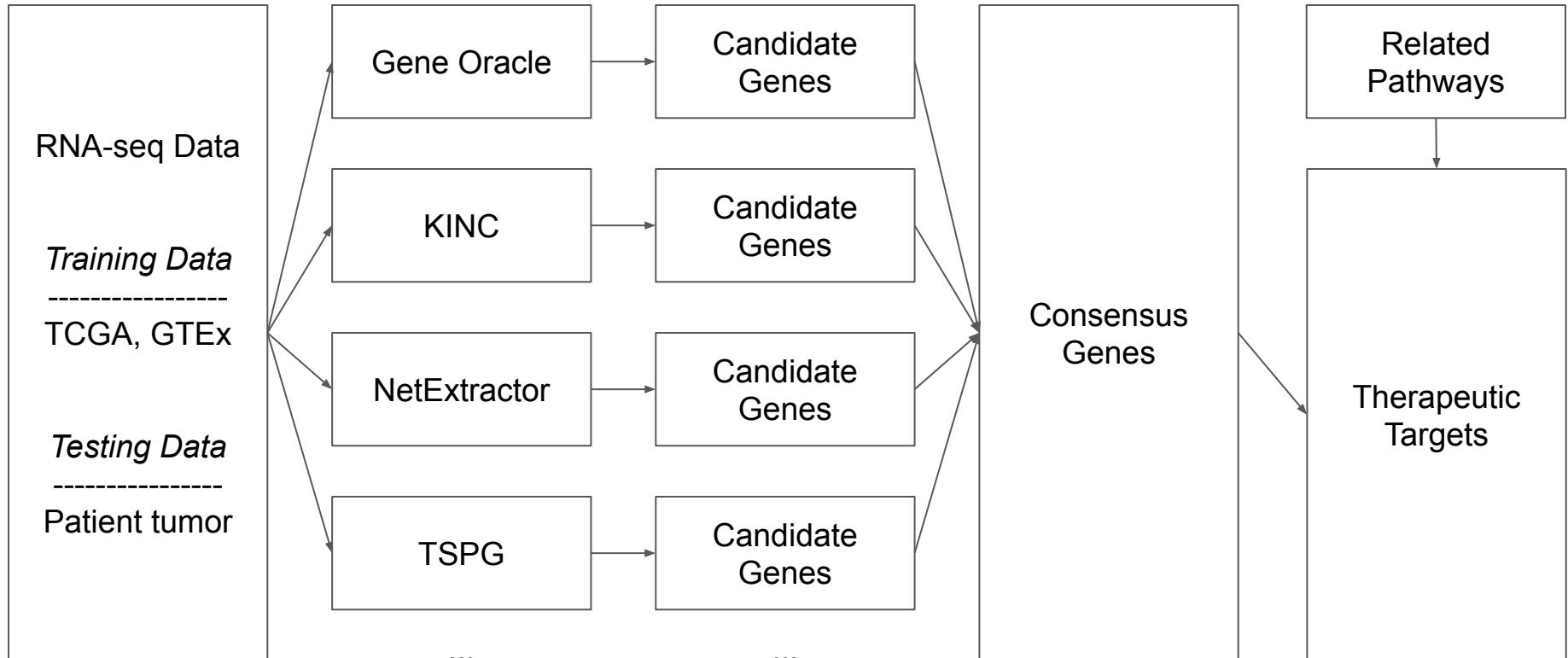
Long-term Goals

- Discover methods of comparing RNA-seq data with DNA
- Test efficacy of potential therapeutics
- Create precision medicine centric pipeline to apply to a variety of cancers

Come talk to us!

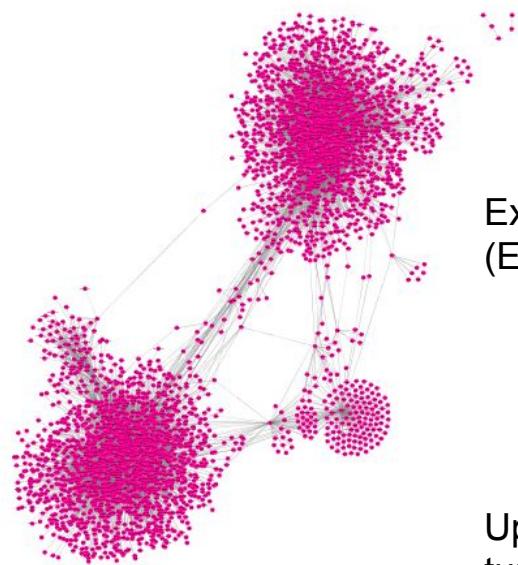
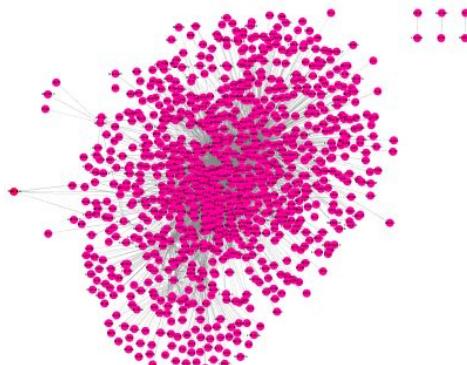
- Use our GEM and/or Bill's RNA seq data
- Find candidate genes to compare with ours
- Talk to Bill Paseman

Workflow Diagram



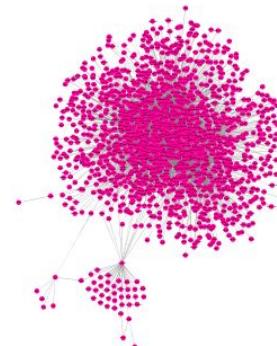
Gene Relationship Networks

Bill RNAseq Network (EdgeCrafting)

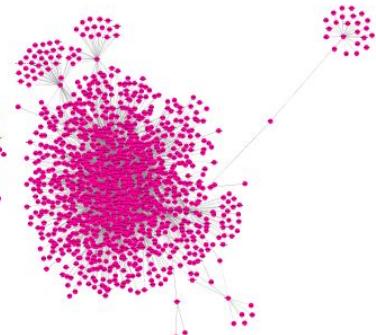


Example Kidney RNAseq Network
(EdgeCrafting)

Up-regulated tumor samples



Down-regulated tumor samples

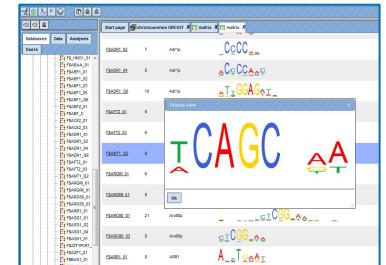
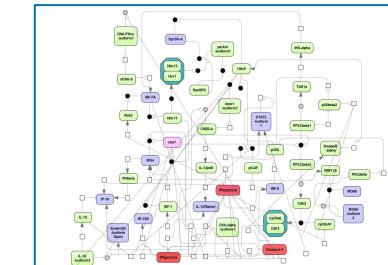
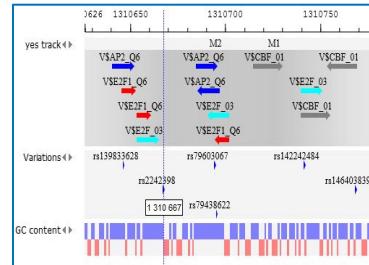
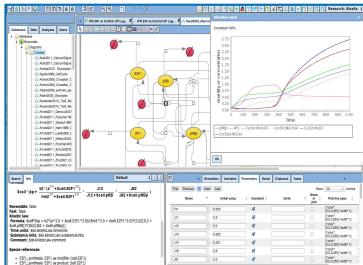
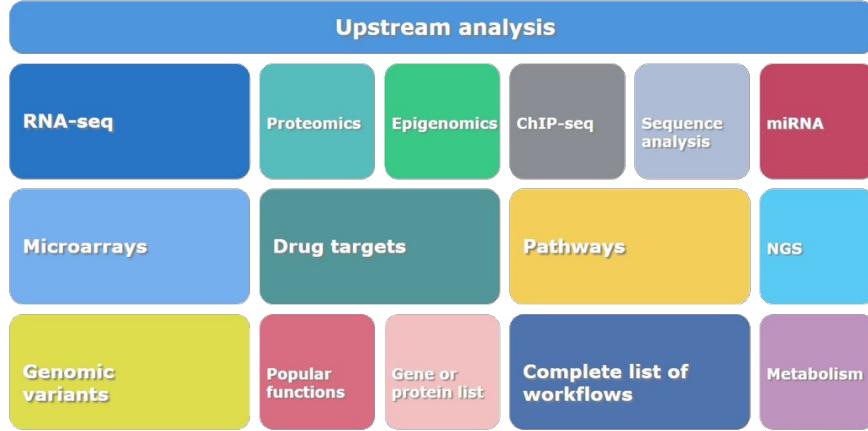




genExplain platform

gene  plain

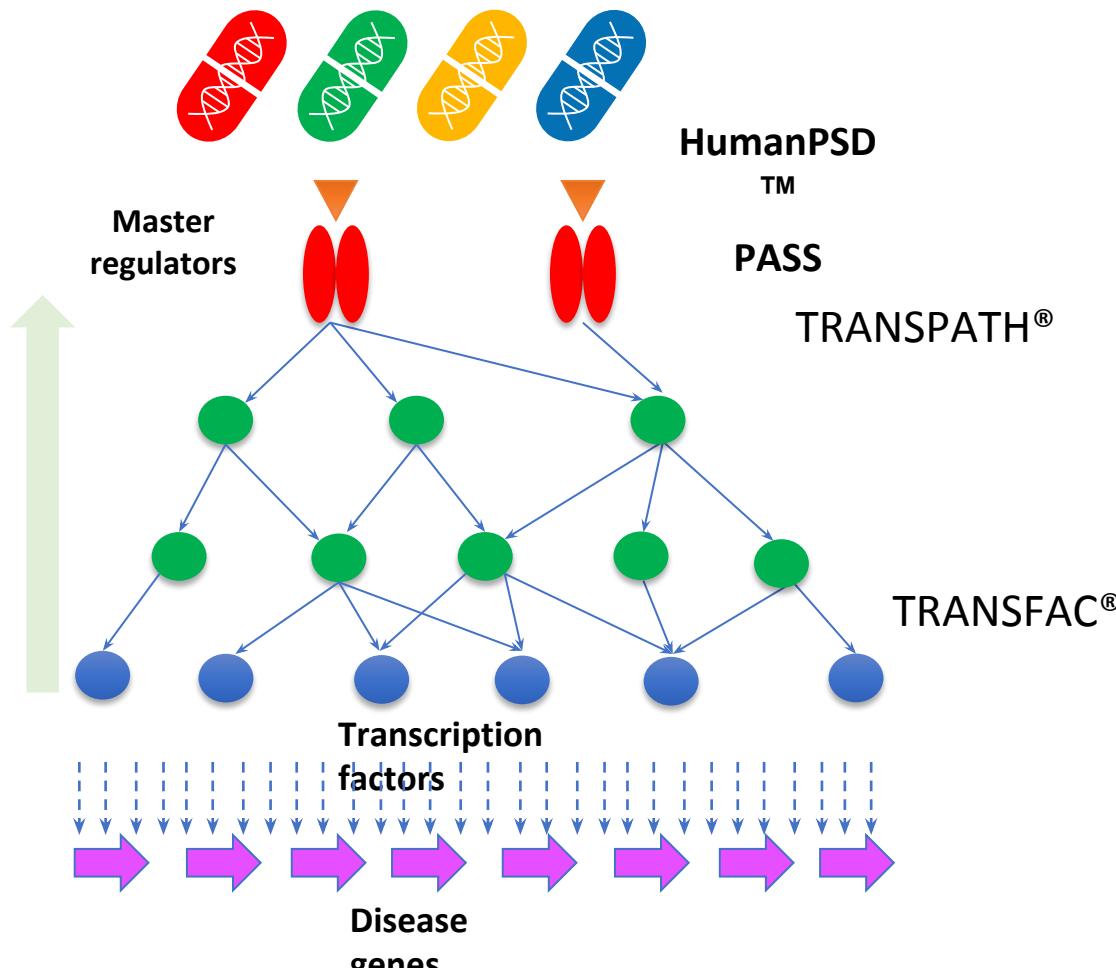
Tools



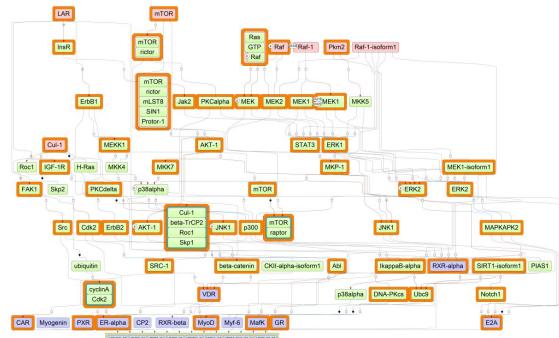
Key features

- ✓ Integrated databases and analysis tools
- ✓ Ready-made workflows for an easy start
- ✓ Upstream analysis (integration of genome regulatory regions analysis with pathway analysis)
- ✓ Knowledge-based data analysis
- ✓ Group project work including chat function
- ✓ JavaScript and R scripts for advanced users
- ✓ Simulation engine inside

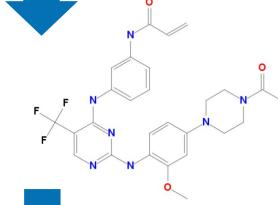
Search for master-regulators in signaling pathways



Upstream Analysis

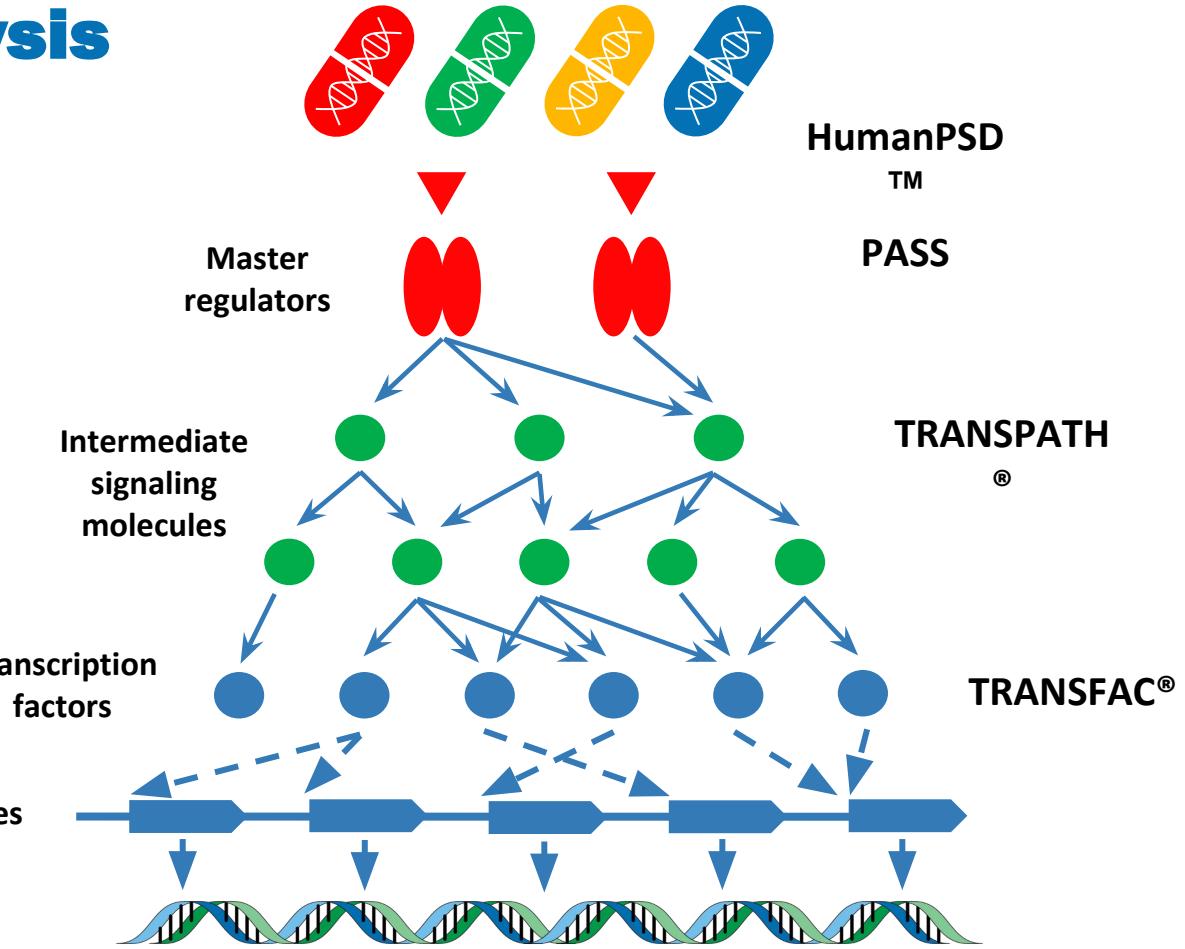


- ✓ Cetuximab
- ✓ Gefitinib



Genes

Genome



Sequence and Pathway analysis

HCK and DUSP9 are promising druggable targets for treating acute kidney injury, chronic kidney disease-mineral and bone disorder, fused kidney, kidney calculi, kidney diseases, cystic kidney diseases, chronic kidney failure, kidney neoplasms, kidney transplantation, acute kidney tubular necrosis, medullary sponge kidney, multicystic dysplastic kidney, polycystic kidney diseases, autosomal dominant polycystic kidney and autosomal recessive polycystic kidney that control activity of RELA, HMGA1 and THRA transcription factors on promoters of differentially expressed genes

Jeannette Koschmann

j.koschmann79@gmail.com

Data received on 03/03/2020 ; Run on 03/03/2020 ; Report generated on 04/03/2020

Genome Enhancer release 1.9 (TRANSFAC®, TRANSPATH® and HumanPSD™ release 2020.1)





Genome Enhancer

Genome Enhancer is a ready-to-use pipeline for multi-omics data analysis, which allows completely automated identification of potential drug targets by running the Upstream analysis combined with search of corresponding perspective drug compounds.

Sequence and Pathway analysis

Identification of master-regulators in gene regulatory and signal transduction pathways.

Your name
Your organization

Report generated on 01/06/2017. Run on 01/06/2017. Report generated on 01/06/2017.

Summary

In this report we present the results of causal multi-omics data analysis, which was performed by the automated pipeline system "From genome to target". The goal of the pipeline is to identify master regulators in gene regulatory networks as potential drug targets for the studied pathological process. On the first step of analysis, the most differentially expressed genes in pathological genes in particular are identified. The second step of analysis performs the search for so-called master-regulators, which control transcription factors that were found on the first step. The identified master-regulators are potential targets for the studied disease. After the druggability checkup, the most promising master-regulators are chosen as potential drug targets.

Here we applied the pipeline "From genome to target" for analysis of multi-omics data set that contains transcriptomics data in **carcinoma**. The results of this analysis help us to better understand the molecular mechanisms of the studied pathological state of the disease. Such approach promises to be very effective for rapid and accurate identification of disease state drug targets with potential.

Introduction

Multiple "omics" data are generated worldwide measuring gene and protein expression, identifying genetic and epigenetic changes and discovering diseases causing mutations and variations for various pathological states of multiple organisms. Still the challenge remains to reveal deep molecular mechanisms of diseases and to find new therapeutic targets. The "From genome to target" pipeline in comparison to the norm. The causal molecular mechanisms of diseases on the level of cellular regulatory networks can be described by specific pathological epigenetic changes in genomes. The molecular regulation of cells can be analyzed by combining transcriptomic and proteomic data sets related to pathology progression. Reconstruction of the disease-specific regulatory networks and identification of potential master regulators of such networks can give us a clue on what potential ways of blocking the pathogenic regulatory cascades are. Selection of certain molecular targets can even allow us to stop the pathological process and cure the disease. The analysis of multi-omics data analysis cannot reconstruct the cell regulatory networks due to the inability of detection of complex signal hierarchy. Thus such approaches provide only a very limited due to the causes of the observed phenomena and actually do not lead to the understanding of the pathology molecular mechanism.

Unlike common approaches, the "upstream analysis" method [1-5], integrated in the pipeline system "From genome to target", uses direct interpretation of raw data without any pre-processing steps. This approach comprises two major steps: (1) analysis of enhancers of identified differentially expressed genes (DEGs) to reveal transcription factors (TFs) involved in the process under study; (2) reconstruction of signaling pathways activated there and identification of master-regulators as the top components of these pathways. The first step is based on the idea that the TFs binding sites distribution and signal activation algorithms - Match [6] and CMA [7]. The second step is based on the help of intracellular signal transduction databases and special graph search algorithms implemented in the pipeline system "From genome to target". The pipeline system is interactive and the user can make his own improvements such as dynamical simulation of the constructed signal transduction network and druggability check of the revealed targets (with the use of QSPR approaches). Therefore the applied in the pipeline system "From genome to target" opens new perspectives to process valuable omics data by complete automatization of such complex tasks as disease molecular mechanism identification and drug target selection.

Table 11. The resulting list of drugs for the studies pathology						
ID	Name	Structure	Target names	Target activity score	Toxicity score	Disease activity score
PC:17905389	AMG 900		MAPK10, SRC, DUSP6	0.798	0.16	0.427
PC:17904561	Osimertinib		MAPK10, SRC, EGF	0.793	0.171	0.474

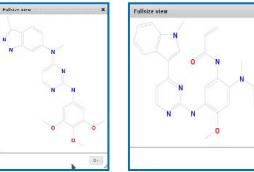


Table 8. Identified master regulators, responsible for the regulation of upregulated genes in wild type; g1 treatment vs. wild type; control treatment		
ID	Master molecule name	Weight
M00000019071	TAB1	0.49158
M000002220	p38beta-isoform1(h)	0.47206
M0000131941	ACLP(h)	0.44456
M0000209228	DLSP14(h)	0.43801
M0000022168	MLK2(h)	0.43595
M0000057036	Ref-1-isoform1(h)	0.42105
M0000118076	EGF-ErbB1(pY1)-ErbB2(pY);Src	0.42064
M000009415	MKK6-isoform1(h)	0.42022
M0000036550	MKP-7(h)	0.41758
M0000041399	MKK3(h);pS1891(pT193)	0.41435

gene plain

Tools

- ✓ A clear result, issued in the form of a finished publication
- ✓ Suitable for direct use by physicians and biologists
- ✓ Processes all types of omics data
- ✓ Does not require special skills
- ✓ Generates a ready report on the predicted therapeutic targets and drug compounds which inhibit them
- ✓ Easy data upload and annotation
- ✓ One-click run

Key benefits



- ✓ Identifies activated targets in the examined patient data and selects the best fit therapy for every studied case
 - ✓ Issues a clear result in the form of a scientific paper
 - ✓ Suites for use by medical doctors and biologists
 - ✓ Does not require special skills
 - ✓ Processes all types of omics data
 - ✓ Easy data upload and annotation
 - ✓ One-click run
 - ✓ Generates a comprehensive report on identified drug targets and prospective therapies

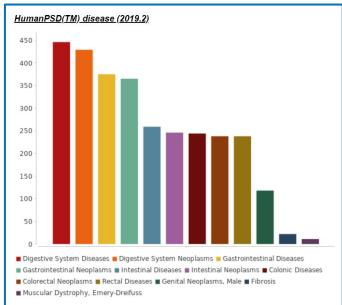
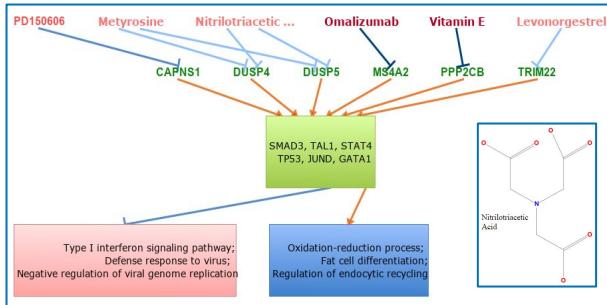
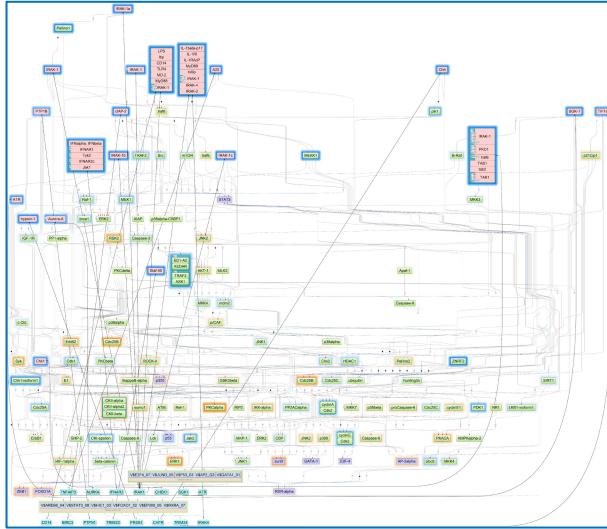


Table 11. The chemical compounds and known drugs (from Human PSD) potentially acting on corresponding master regulators.									
<u>See full table →</u>									
ID	Name	Target names	Target activity score	NA	Phase 1	Phase 2	Phase 3	Phase 4	
D802983	Fomoterol	ADRB2	0.47	Asthma, Bronchitis, Diabetes Mellitus, Type I, Lung Diseases, Obstructive, Patient Satisfaction, Pulmonary Disease, Chronic Obstructive	Asthma, Lung Diseases, Obstructive, Lung Neoplasm, Precursors Conditions, Pulmonary Disease, Chronic Obstructive	Allergy, Obsession, Alzheimer Disease, Asthma, Obsessive, Exercise-Induced, Exercise-Induced, Bronchial Spasms, Bronchitis, Chronic, Endocrinoma...	Smoking, Asthma, Alzheimer Disease, Exercise-Induced, Bronchitis, Bronchitis, Chronic, Emphysema, Lung Disease...	3 8	
D803084	Fica	CASP7	0.6	Acidosis, Acidosis, Renal Tubular, Affect, Arterial Occlusion, Brain Abscess, Diabetic Mellitus, Heart Disease...	Bites and Stings, Dystonia, Immunologic Syndromes, Psychotic Syndromes, Migraine Disorders, Muscle Spasticity, Paroxysm, Smallpox...	Bites and Stings, Burkitt Lymphoma, Cerv. Dysplasia, HIV Infections, Hypertension, Immunologic Deficiency Syndromes...	Diabetes, Heart Failure, Lymphoma, T-Cell Lymphoma, T-Cell, T-Cell, Myositis, Infarction, Pneumonia...	Acute Coronary Syndrome, Arteriosclerosis, Arteritis, Arteritis, HIV Infection, Hyperlipidemias, Hyperglycideridemia, Infarction...	0 11



The workflow

Genome Enhancer uses *Upstream Analysis*, an integrated promoter and pathway analysis, to identify potential drug targets for the studied pathology.

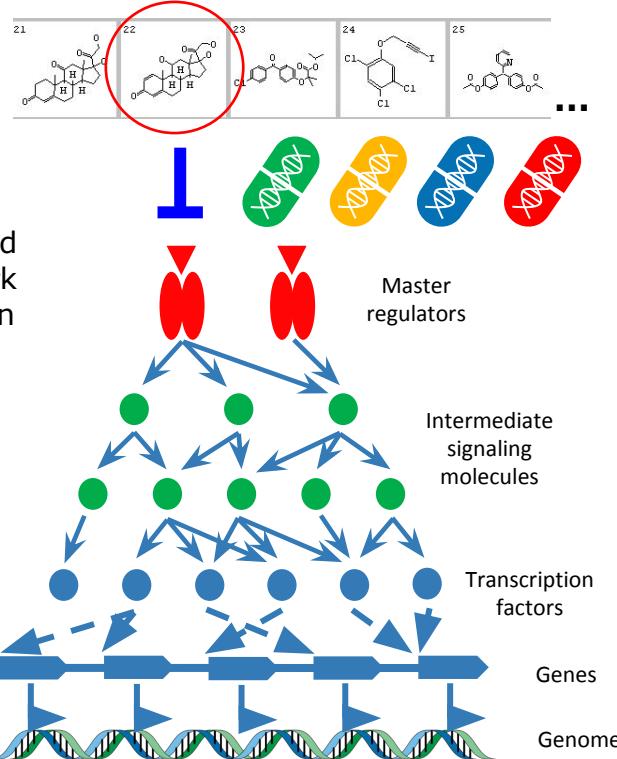
In the first step of this analysis the transcription factors that regulate differentially expressed or mutated genes are identified with the use of the [TRANSFAC®](#) database of transcription factors binding sites.

The second step searches for common master-regulators of the identified transcription factors by building a personalized signal transduction network of the studied pathology using the [TRANSPATH®](#) database of mammalian signal transduction and metabolic pathways.

The identified master regulators are prospective drug targets candidates and they are used for further selection of chemical compounds that can bring therapeutic benefit for the studied clinical case. In this step the [HumanPSD™](#) database is employed to identify drugs that have been tested in clinical trials and [PASS](#) tool is used to perform cheminformatic analysis and predict small molecules that can affect the identified targets.

As the result of analysis a comprehensive report is generated.

Upstream analysis schema



Identification of potential drug targets

In the last step of the analysis the software identify known drugs as well as new potentially active chemical compounds that are potentially suitable for inhibition (or activation) of the identified molecular targets in the context of specified human disease.

Table 8. Known drug targets for known drugs revealed in this study. The column **Druggability score** contains the number of drugs that are potentially suitable for inhibition (or activation) of the target. **Total rank** is the sum of the ranks of the master molecules sorted by keynode score, CMA score, transcriptomics data.

[See full table](#)

ID	Gene symbol	Gene description	Druggability score	Total rank
ENSG00000101336	HCK	HCK proto-oncogene, Src family tyrosine kinase	4	43
ENSG00000166851	PLK1	polo like kinase 1	5	43
ENSG00000100385	IL2RB	interleukin 2 receptor subunit beta	4	47
ENSG00000132334	PTPRE	protein tyrosine phosphatase, receptor type E	1	
ENSG00000138798	EGF	epidermal growth factor	3	
ENSG00000182866	LCK	LCK proto-oncogene, Src family tyrosine kinase	14	
ENSG00000117650	NEK2	NIMA related kinase 2	1	
ENSG00000113263	ITK	IL2 inducible T-cell kinase	2	
ENSG00000113721	PDGFRB	platelet derived growth factor receptor beta	9	
ENSG00000112062	MAPK14	mitogen-activated protein kinase 14	53	

Table 9. The list of drugs (from Human-PDG) approved or used in clinical trials for the application in acute kidney injury, chronic kidney disease-mineral and bone disorder, failed kidney, kidney calculus, kidney diseases, cystic kidney diseases, chronic kidney diseases, kidney transplantation, acute kidney tubular necrosis, medullary sponge kidney, multicystic dysplastic kidney, polycystic kidney diseases, autosomal dominant polycystic kidney and renal tubular acidosis, and on rare cases of other diseases where the drug is known to be applied for the disease. The column Target activity screen contains the names of screened diseases, where the drug is known to be applied. We use sum of clinical trials phases as the weight of the disease. Drug rank column contains total rank of giving drugs among all found. See Methods section for details.

[See full table](#)

ID	Name	Target names	Target activity score	NA	Phase 1	Phase 2	Phase 3	Phase 4	Disease activity score
D801254	Dasatinib	LOX, ABL1, PDGFRA	0.72	Brain, Nervous, Carcinoma, Small Cell, Carcinoma, Transitional Cell, Gastrointestinal Stromal Tumors, Glialoma, Leukemia, Lymphoma, Leukemia, Lymphoid...	Kidney Neoplasia, Adenocarcinoma, Adenocarcinoma, Clear Cell, Adenocarcinoma, Mucinous, Brain Diseases, Brain Neoplasms...	Kidney Diseases, Kidney Neoplasia, Adenocarcinoma, Adenocarcinoma, Clear Cell, Blodd Cells, Brain Diseases...	Leukemia, Leukemia, Lymphoid, Leukemia, Myelogenous, Chronic, BCR-ABL Positive, Leukemia, Myeloid, Accelerated Phase, Leukemia, Hyperacute, Leukemia, Myeloid, Chronic-Phase...	Leukemia, Leukemia, Lymphoid, Leukemia, Myelogenous, Chronic, BCR-ABL Positive, Leukemia, Myeloid, Precursor Cell Lymphoblastic Leukemia, Thrombosis...	5
D800398	Sorafenib	PDGFRB, FGFR1, BRAF	0.67	Kidney, Adenocarcinoma, Adenoma, Brain, Adrenocortical, Breast, Ductal Neoplasms, Endocrine, Carcinoma, Hepatocellular...	Kidney Diseases, Kidney Neoplasms, Adenocarcinoma, Adenoma, Adenoma, Adenoma, Liver Cell, Astrocytoma, Bile Duct Neoplasms...	Kidney Diseases, Kidney Neoplasms, Adenocarcinoma, Adenoma, Adenoma, Liver Cell, Astrocytoma, Bile Duct Neoplasms...	Kidney Neoplasia, Adenocarcinoma, Breast, Carcinoma, Carcinoma, Carcinoma, Hepatocellular, Hepatocellular, Liver Neoplasm, Non-Small-Cell Lung, Carcinoma, Renal Cell, Thrombosis...	Carcinoma, Hepatocellular, Carcinoma, Renal Cell, Liver Neoplasm, Neoplasia, Non-Hodgkin Lymphoma, Thrombosis...	50

Table 7. Master regulators that may govern the regulation of the list of genes provided as input in Experiment. **Total rank** is the sum of the ranks of the master molecules sorted by keynote score, CMA score, transcriptomics data.

[See full table →](#)

ID	Master molecule name	Gene symbol	Gene description	Total rank
MO000032712	MKP-4(h)	DUSP9	dual specificity phosphatase 9	7
MO000019256	ZAP-70(h)	ZAP70	zeta chain of T-cell receptor associated protein kinase 70	10
MO000019253	Lck(h)	LCK	LCK proto-oncogene, Src family tyrosine kinase	14
MO000022403	plk1(h)	PLK1	polo like kinase 1	32
MO000022400	Cdc25C(h)	CDC25C	cell division cycle 25C	33
MO000020274	(FasL:Fas)6: (Daxx{pS667}:ASK1)2	DAXX, FAS, FASLG, MAP3K5	Fas cell surface death receptor, Fas ligand, death domain associated protein, mitogen-activated prot...	34
MO000038590	Rac1:GTP:MEKK4	CYBA, CYBB, MAP3K4, NCF1, NCF2, NCF4, RAC1, SYTL1	cytochrome b-245 alpha chain, cytochrome b-245 beta chain, mitogen-activated protein kinase kinase k...	34
MO000043863	prlr(h):tec(h):Vav(h)	PRLR, TEC, VAV1	prolactin receptor, tec protein tyrosine kinase, vav guanine nucleotide exchange factor 1	38
MO000096187	plk1(h)	PLK1	polo like kinase 1	38
MO000032652	MKP-2(h)	DUSP4	dual specificity phosphatase 4	41

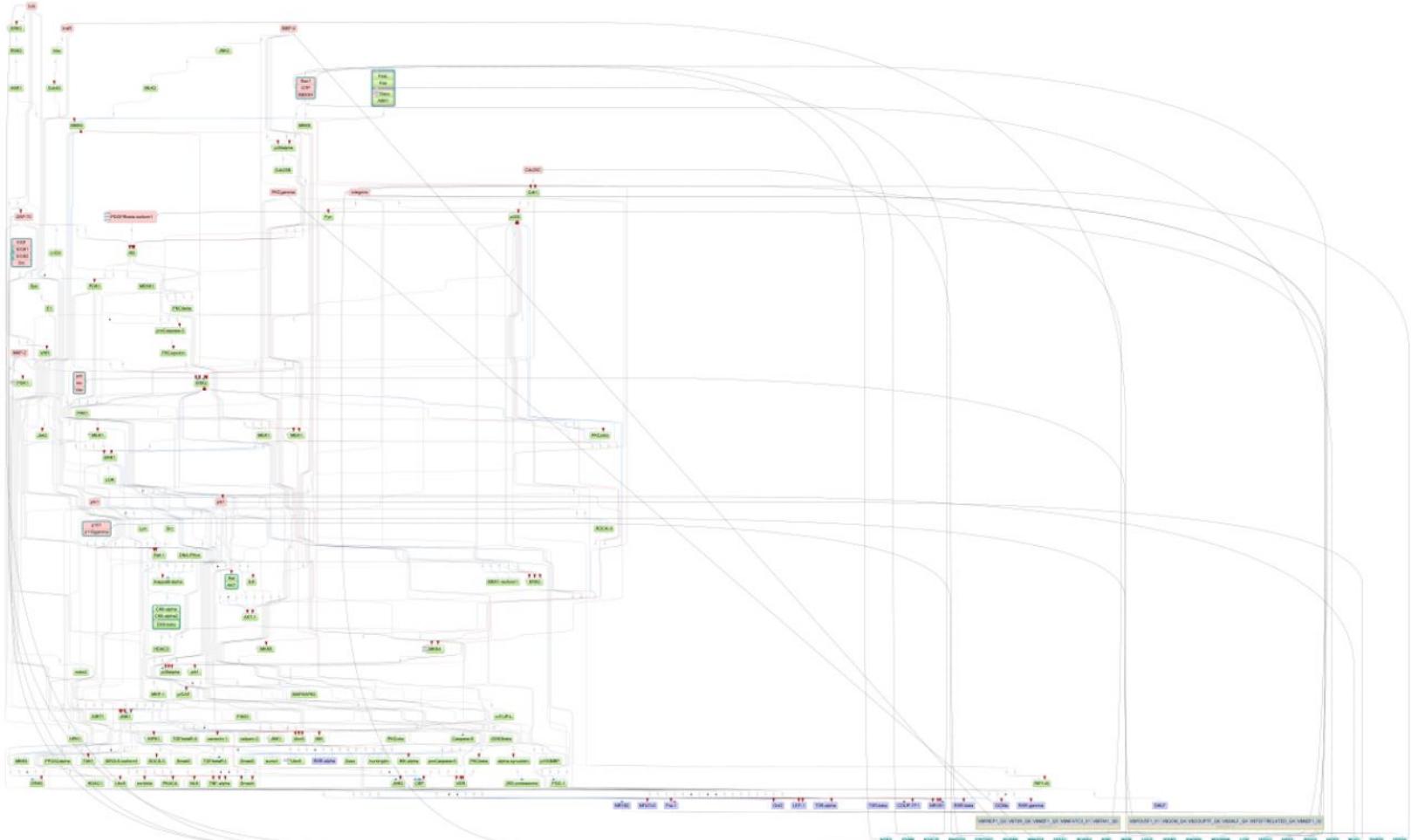


Figure 9. Diagram of intracellular regulatory signal transduction pathways of the list of genes provided as input in Experiment. Master regulators are indicated by red rectangles, transcription factors are blue rectangles, and green rectangles are intermediate molecules, which have been added to the network during the search for master regulators from selected TFs. Orange and blue frames highlight molecules that are encoded by up- and downregulated genes, resp.

4. Identification of potential drugs

In the last step of the analysis we strived to identify known drugs as well as new potentially active chemical compounds that are potentially suitable for inhibition (or activation) of the identified molecular targets in the context of specified human disease. First, we identify known drugs using information from HumanPSD™ database [5] about their targets and about clinical trials where the drugs have been tested for the treatment of various human diseases. Table 8 shows the resulting list of druggable master regulators that represent the predicted drug targets of the studied pathology. Table 9 lists chemical compounds and known drugs (from the HumanPSD™ database) potentially acting on corresponding master regulators.

Table 8. Known drug targets for known drugs revealed in this study. The column **Druggability score** contains the number of drugs that are potentially suitable for inhibition (or activation) of the target. **Total rank** is the sum of the ranks of the master molecules sorted by keynode score, CMA score, transcriptomics data.

[See full table →](#)

ID	Gene symbol	Gene description	Druggability score	Total rank
ENSG00000101336	HCK	HCK proto-oncogene, Src family tyrosine kinase	4	43
ENSG00000166851	PLK1	polo like kinase 1	5	43
ENSG00000100385	IL2RB	interleukin 2 receptor subunit beta	4	47
ENSG00000132334	PTPRE	protein tyrosine phosphatase, receptor type E	1	49
ENSG00000138798	EGF	epidermal growth factor	3	55
ENSG00000182866	LCK	LCK proto-oncogene, Src family tyrosine kinase	14	61
ENSG00000117650	NEK2	NIMA related kinase 2	1	65
ENSG00000113263	ITK	IL2 inducible T-cell kinase	2	69
ENSG00000113721	PDGFRB	platelet derived growth factor receptor beta	9	71
ENSG00000112062	MAPK14	mitogen-activated protein kinase 14	53	76

Table 9. The list of drugs (from Human PSD) approved or used in clinical trials for the application in acute kidney injury, chronic kidney disease-mineral and bone disorder, fused kidney, kidney calculi, kidney diseases, cystic kidney diseases, chronic kidney failure, kidney neoplasms, kidney transplantation, acute kidney tubular necrosis, medullary sponge kidney, multicystic dysplastic kidney, polycystic kidney diseases, autosomal dominant polycystic kidney and autosomal recessive polycystic kidney and acting on master regulators revealed in our study. The column **Target activity score** contains the value of numeric function that depends on ranks of all targets that were found for the drug. The column **Disease activity score** contains the weighted sum of user selected diseases where the drug is known to be applied. We use sum of clinical trials phases as the weight of the disease. **Drug rank** column contains total rank of given drug among all found. See [Methods](#) section for details.

[See full table →](#)

ID	Name	Target names	Target activity score	NA	Phase 1	Phase 2	Phase 3	Phase 4	Disease activity score
DB01254	Dasatinib	LCK, ABL1, PDGFRB	0.72	Brain Neoplasms, Carcinoma, Squamous Cell, Carcinoma, Transitional Cell, Gastrointestinal Stromal Tumors, Glioblastoma, Leukemia, Leukemia, Lymphoid...	Kidney Neoplasms, Adenocarcinoma, Adenocarcinoma, Clear Cell, Adenocarcinoma, Mucinous, Brain Abscess, Brain Diseases, Breast Neoplasms...	Kidney Diseases, Kidney Neoplasms, Adenocarcinoma, Adenocarcinoma, Clear Cell, Blast Crisis, Brain Abscess, Brain Diseases...	Leukemia, Lymphoid, Leukemia, Myelogenous, Chronic, BCR-ABL Positive, Leukemia, Myeloid, Accelerated Phase, Leukemia, Myeloid, Acute, Leukemia, Myeloid, Chronic-Phase...	Leukemia, Leukemia, Lymphoid, Leukemia, Myelogenous, Chronic, BCR-ABL Positive, Leukemia, Myeloid, Precursor Cell Lymphoblastic Leukemia-Lymphoma	5
DB00398	Sorafenib	PDGFRB, FGFR1, BRAF	0.67	Kidney Neoplasms, Adenocarcinoma, Ascites, Brain Abscess, Brain Neoplasms, Breast Neoplasms, Carcinoma, Hepatocellular...	Kidney Diseases, Kidney Neoplasms, Adenocarcinoma, Adenoma, Adenoma, Liver Cell, Astrocytoma, Bile Duct Neoplasms...	Kidney Diseases, Kidney Neoplasms, Adenocarcinoma, Adenoma, Adenoma, Liver Cell, Adrenocortical Carcinoma, Bile Duct Neoplasms...	Kidney Neoplasms, Adenocarcinoma, Breast Neoplasms, Carcinoma, Adenoma, Liver Cell, Carcinoma, Hepatocellular, Carcinoma, Non-Small-Cell Lung, Carcinoma, Renal Cell...	Carcinoma, Hepatocellular, Carcinoma, Renal Cell, Liver Neoplasms, Neoplasms, Noma, Thrombosis	10

Table 10. The list of drugs (from HumanPSD) known to be acting on master regulators revealed in our study that can be proposed as a drug repurposing initiative for the treatment of acute kidney injury, chronic kidney disease-mineral and bone disorder, fused kidney, kidney calculi, kidney diseases, cystic kidney diseases, chronic kidney failure, kidney neoplasms, kidney transplantation, acute kidney tubular necrosis, medullary sponge kidney, multicystic dysplastic kidney, polycystic kidney diseases, autosomal dominant polycystic kidney and autosomal recessive polycystic kidney. **Target activity score** column contains value of numeric function that depends on ranks of all targets that were found for the drug. **Drug rank** column contains total rank of given drug among all found. See [Methods](#) section for details.

ID	Name	Target names	Target activity score	NA	Phase 1	Phase 2	Phase 3	Phase 4	Drug rank	
DB08896	Regorafenib	ABL1, PDGFRB, FGFR1, BRAF	0.72	Adenocarcinoma, Carcinoma, Hepatocellular, Cholangiocarcinoma, Colorectal Neoplasms, Gastrointestinal Stromal Tumors, Glioblastoma, Liver Neoplasms...	Carcinoma, Hepatocellular, Carcinoma, Small Cell, Colorectal Neoplasms, Esophageal Neoplasms, Gastrointestinal Neoplasms, Gastrointestinal Stromal Tumors, Intestinal Neoplasms...	Adenocarcinoma, Bile Duct Neoplasms, Brain Abscess, Breast Neoplasms, Carcinoid Tumor, Carcinoma, Adenoid Cystic, Carcinoma, Islet Cell...	Carcinoma, Hepatocellular, Colonic Neoplasms, Colorectal Neoplasms, Esophageal Neoplasms, Gastrointestinal Stromal Tumors, Neoplasms, Noma...	Carcinoma, Hepatocellular, Colonic Neoplasms, Colorectal Neoplasms, Esophageal Neoplasms, Gastrointestinal Stromal Tumors, Neoplasms,	Colorectal Neoplasms, Gastrointestinal Stromal Tumors, Neoplasms, Rectal Neoplasms	36
DB09079	Nintedanib	LCK, FGFR3, FGFR1	0.6	Carcinoma, Carcinoma, Non- Small-Cell Lung, Colorectal Neoplasms, Endometrial Neoplasms, Fallopian Tube Neoplasms, Idiopathic Pulmonary Fibrosis, Lung Diseases...	Adenocarcinoma, Breast Neoplasms, Carcinoma, Hepatocellular, Carcinoma, Non- Small-Cell Lung, Carcinoma, Renal Cell, Carcinoma, Small Cell, Colonic Neoplasms...	Adenocarcinoma, Adenocarcinoma, Clear Cell, Adenocarcinoma, Mucinous, Angiomyoma, Appendiceal Neoplasms, Breast Neoplasms, Carcinoid Tumor...	Carcinoma, Non- Small-Cell Lung, Colorectal Neoplasms, Idiopathic Pulmonary Fibrosis, Lung Diseases, Lung Diseases, Interstitial, Mesothelioma, Neoplasms...	Idiopathic Pulmonary Fibrosis, Pulmonary Fibrosis	45	

Next, new potential small molecular ligands were predicted for the revealed targets and a general druggability check was run using a pre-computed database of spectra of biological activities of chemical compounds from a library of 13040 most pharmaceutically active known compounds. The spectra of biological activities has been computed using the program PASS [11-13] on the basis of a (Q)SAR approach. Table 11 shows the resulting list of druggable master regulators, which represent the predicted drug targets of the studied pathology. Table 12 lists chemical compounds and known drugs potentially acting on the corresponding master regulators.

*Table 11. Extended list of drug targets revealed in this study (targets that are predicted by PASS program potentially targeted by an extended list of known drugs and pharmaceutically active chemical compounds). The column **Druggability score** contains a numeric value which indicates how suitable this target is to be inhibited (or activated) by a drug. See [Methods](#) section for details.*

[See full table →](#)

ID	Name	Gene symbol	Gene description	Druggability score	Total rank
ENSG00000130829	DUSP9	DUSP9	dual specificity phosphatase 9	13.33	7
ENSG00000159352	PSMD4	PSMD4	proteasome 26S subunit, non-ATPase 4	0.85	10
ENSG00000161057	PSMC2	PSMC2	proteasome 26S subunit, ATPase 2	0.85	10
ENSG00000158402	CDC25C	CDC25C	cell division cycle 25C	19.77	33
ENSG00000117560	FASLG	FASLG	Fas ligand	2.66	34
ENSG00000120875	DUSP4	DUSP4	dual specificity phosphatase 4	13.33	41
ENSG00000101336	HCK	HCK	HCK proto-oncogene, Src family tyrosine kinase	0.67	43
ENSG00000166851	PLK1	PLK1	polo like kinase 1	0.75	43
ENSG00000100385	IL2RB	IL2RB	interleukin 2 receptor subunit beta	9.32	47
ENSG00000112576	CCND3	CCND3	cyclin D3	8.3	49

