



From molecular signatures to predictive biomarkers: modeling disease pathophysiology and drug mechanism of action

Andreas Heinzl¹, Paul Perco¹, Gert Mayer², Rainer Oberbauer³, Arno Lukas¹ and Bernd Mayer^{1*}

¹ emergentec biodevelopment GmbH, Vienna, Austria

² Department of Internal Medicine IV, Medical University of Innsbruck, Innsbruck, Austria

³ Department of Internal Medicine III, KH Elisabethinen Linz and Medical University of Vienna, Vienna, Austria

Edited by:

Christine Nardini, Partner Institute for Computational Biology, China

Reviewed by:

Zahraa Naji Sabra, American University of Beirut, Lebanon
Satyaprakash Nayak, Pfizer Inc., USA

*Correspondence:

Bernd Mayer, emergentec biodevelopment GmbH, Gersthoferstrasse 29-31, Vienna, 1180, Austria
e-mail: bernd.mayer@emergentec.com

Omics profiling significantly expanded the molecular landscape describing clinical phenotypes. Association analysis resulted in first diagnostic and prognostic biomarker signatures entering clinical utility. However, utilizing Omics for deepening our understanding of disease pathophysiology, and further including specific interference with drug mechanism of action on a molecular process level still sees limited added value in the clinical setting. We exemplify a computational workflow for expanding from statistics-based association analysis toward deriving molecular pathway and process models for characterizing phenotypes and drug mechanism of action. Interference analysis on the molecular model level allows identification of predictive biomarker candidates for testing drug response. We discuss this strategy on diabetic nephropathy (DN), a complex clinical phenotype triggered by diabetes and presenting with renal as well as cardiovascular endpoints. A molecular pathway map indicates involvement of multiple molecular mechanisms, and selected biomarker candidates reported as associated with disease progression are identified for specific molecular processes. Selective interference of drug mechanism of action and disease-associated processes is identified for drug classes in clinical use, in turn providing precision medicine hypotheses utilizing predictive biomarkers.

Keywords: omics, integration, molecular model, biomarker, target, systems biology, systems pharmacology, precision medicine

INTRODUCTION

Despite a continuously rising number of clinical trials the rate of bringing novel medication to the clinic is stalling (Pammolli et al., 2011). Here, Omics profiling and high throughput drug screening technologies at the interface of large scale clinical data have triggered novel conceptual strategies aimed at improved patient stratification for enabling precision medicine (Trusheim et al., 2011; Hollebecque et al., 2014). For implementing such approaches a number of issues need to be addressed including: (i) mirroring the clinical categorization of a phenotype on a molecular level description, (ii) spotting molecular factors mechanistically driving disease progression, (iii) drug-based intervention specifically addressing such progression mechanisms, and (iv) predictive biomarkers allowing fit-for-purpose analysis regarding a match of relevant pathophysiology and drug mechanism of action on the individual patient level (Heinzl et al., 2012).

A clinically well-established example is HER2 positive breast cancer characterized by overexpression of a member of the epidermal growth factor receptor family (ERBB2) playing a mechanistic role in progressive disease. In case the factor is proving positive for a patient the specific presentation is amenable for treatment tackling growth signaling (Hicks and Kulkarni, 2008). Still,

the clinical presentation of breast cancer shows heterogeneous pathophysiologies apart HER2 positive subtypes. In consequence, when aiming at a comprehensive assessment of progressive breast cancer phenotypes multimarker panels are needed, e.g., implemented by a multiplexed assay holding 70 individual molecular features (Buyse et al., 2006). Such multimarker panels have generally become a promising strategy for characterizing complex clinical presentations, e.g., utilizing a serum marker panel for predicting coronary artery disease in symptomatic patients, or a urinary proteomics profile for early diagnosis of diabetic kidney disease (LaFramboise et al., 2012; Züribig et al., 2012).

Failure for identifying a single causative factor as proxy for determining progression of a complex clinical phenotype becomes apparent when comparing the performance of marker panels with single markers, with the latter e.g., reviewed by Hellemons et al. for onset and progression of diabetic kidney disease (Hellemons et al., 2012). In clinical practice a different type of biomarker may be utilized, providing a phenotypic readout primarily reflecting the functional status of an organ in contrast to the pathophysiological characteristics. In kidney disease such functional markers are used in patient management as well as clinical trial design, including the estimated glomerular filtration rate (eGFR) and proteinuria (reflecting glomerular filtration and

permeation of macromolecules across the glomerular capillary wall, respectively).

Association of these parameters with worsening of diabetic kidney disease, together with increasing incidence of endpoints as cardiovascular events is undisputed (Adler et al., 2003). However, these markers do not provide information on the specific molecular characteristics of the disease. Functional markers render stratification for tailored therapy in the concept of precision medicine essentially impossible.

The molecular pathway of primary interest in the present clinical setting of diabetic kidney disease is the renin-angiotensin system (RAS), in its activity at foremost controlling blood pressure and fluid balance. Blockade of the RAS has been able to reduce the incidence of renal events in patients with and without diabetes mellitus (Ruggenti et al., 1998; Brenner et al., 2001). In a study by Lewis et al. angiotensin receptor blockade by Irbesartan reduced the risk of a primary composite endpoint (doubling of baseline serum creatinine concentration, development of end-stage renal disease or death from any cause) during a follow up period of 2.6 years by 20% when compared to the placebo (Lewis et al., 2001). Nevertheless, 50% of patients in the Irbesartan group reached the primary endpoint after 54 months. In an effort to increase the efficacy of RAS antagonistic therapy an angiotensin receptor blocker was combined with placebo or the angiotensin converting enzyme (ACE) inhibitor Lisinopril (Fried et al., 2013). The combination therapy did not reduce the incidence of a combined renal endpoint. On the contrary an increased risk of hyperkalemia and acute kidney injury was observed confirming other reports questioning the safety of this approach (Mann et al., 2008; Parving et al., 2008).

Next to addressing RAS, organ-specific molecular processes involving inflammation and oxidative stress have been implicated in progressive tubulointerstitial fibrosis, the best histological, hence molecular mechanistic predictor of an adverse renal disease prognosis (Rodríguez-Iturbe and García García, 2010). Bardoxolone, a nuclear factor-erythroid-2-related factor 2 activator with anti-oxidative capacity increased eGFR in patients with advanced diabetic renal disease (Pergola et al., 2011). However, a large prospective controlled randomized trial with hard endpoints had to be stopped because of severe side effects (De Zeeuw et al., 2013).

As given with these examples for chronic kidney disease (but in its conceptual fundament holding true for a multitude of highly prevalent chronic diseases), many of the recent interventional studies failed to achieve their goals. Here biomarkers promise to take a key role in selecting patients for studies and/or to predict the long term effects of a drug on hard endpoints. Upfront stratification in randomized controlled trials by separating patients by drug response as measured by biomarkers serving as endpoint surrogate and then randomizing the groups separately is an approach which is, at least from a statistical point of view, preferable to *post-hoc* analysis (De Leon, 2012). Such an enrichment strategy is currently e.g., tested in the SONAR study (clinicaltrials.gov reference NCT01858532) addressing diabetic nephropathy (DN).

However, with respect to fit of specific drugs biomarkers need to carry predictive value, i.e., a biomarker shall on a

patient-specific level identify responders benefitting from drug effect. In this setting various levels need to be considered involving genetic and environmental components defining disease presentation and progression. The drug target may see genetic polymorphism impacting drug binding, but polymorphism may further involve drug transport and drug metabolism (Johnson, 2001). A significant number of genetic polymorphisms have in the meantime become drug label-relevant regarding drug efficacy, but also toxicity and side effects (U.S. Food and Drug Administration, 2014). Pharmacogenomics has clearly demonstrated that the genetic background of an individual introduces heterogeneity in drug response.

Still, this setting assumes a homogeneous patient population with respect to the molecular mechanistic factors determining disease progression, only exhibiting differences in genetic peculiarities of one and the same molecular mechanistic context. In such setting functional biomarkers appear sufficient for identifying progressive disease, and drug variance is fully explained by the genetic background in regard to the mechanism of action of a specific drug.

A complementary perspective may be that the molecular mechanistic background and progression-relevant molecular factors are *per se* diverse and patient-specific, naturally determining drug response (Mayer et al., 2012). In such scenario a biomarker needs to serve as proxy of key mechanistic factors characterizing and driving a disease on a patient-specific level, combined with educating on the specific interference of disease mechanism with drug mechanism of action. For capturing these constraints a detailed molecular map of a clinical phenotype and its interference with a drug mechanism of action is needed, and here integration of Omics profiling adds to identifying such mechanisms (Fechete et al., 2011; Mühlberger et al., 2012).

An a priori stratification of patients based on an appropriately chosen biomarker panel reflecting the pathophysiology of a given patient (group) allowing to determine a match with a specific drug's mechanism of action appears as promising approach. As recently discussed by Himmelfarb et al. fresh approaches are critical in finding therapies to kidney disease benefiting patients, outlining the importance of improving the translational aspect in clinical research (Himmelfarb and Tuttle, 2013). Here, omics technologies have added significantly to the data landscape characterizing chronic kidney disease, however, in a first instance mainly expanding the candidate set of apparently relevant processes and pathways, going in hand with a large number of biomarker candidates, which individually hamper clinically relevant assessment on disease progression (Fechete et al., 2011; Hellemons et al., 2012).

Integrative approaches in the realm of Systems Biology have been proposed for reaching a consensus description of chronic kidney disease pathophysiology, including molecular models of DN as well as of the reno-cardial axis (He et al., 2012; Komorowsky et al., 2012; Mayer et al., 2012; Heinzel et al., 2013). Still, a translation process needs to be followed, joining disease pathophysiology, stratification markers allowing enrichment strategies, combined with on a molecular mechanistic level matching drugs for allowing precision medicine (Mirnezami et al., 2012). In this work we exemplify such procedure on DN

being the major clinical presentation leading to end stage renal disease.

MATERIALS AND METHODS

GENERAL DATA SOURCES

Protein coding genes identified as associated with DN were collected from public domain transcriptomics data sources, complemented with molecular features reporting such association in scientific literature. Molecular signatures educating on ACE inhibitor mechanism of action were extracted from public domain transcriptomics sources. Proteins discussed as biomarkers or drug target candidates in the context of DN were extracted from scientific literature, with the set of targets further extended with known drug targets of drugs currently utilized in clinical trials including renal endpoints. Protein-protein interaction information and molecular pathway maps were retrieved from public domain databases.

Clinical phenotype molecular data

A literature search in NCBI Pubmed utilizing the query string *diabetic nephropathies[majr] AND (microarray analysis[mh] OR gene expression profiling[mh]) AND humans[mh] NOT review* resulted in 37 transcriptomics studies. Explicitly restricting to explorative, array-based mRNA expression studies on human kidney tissue yielded four studies as suitable for inclusion in further analysis. For Berthier et al. and Cohen et al. expression signatures could be retrieved directly from the publications (Cohen et al., 2008; Berthier et al., 2009). For Woroniecka et al. and Baelde et al. the raw expression profiles were retrieved from Gene Expression Omnibus (GSE30122, GSE1009) (Baelde et al., 2004; Woroniecka et al., 2011). Robust Multi-array Average (RMA) normalization

for the data set of Woroniecka et al. and MAS5 normalization for the data set of Baelde et al., followed by Significance Analysis of Microarrays (SAM) was employed for identifying features showing differential regulation comparing diabetic kidney disease and healthy control samples. In case of microdissected sample material separate analysis was done for the glomerular and tubulointerstitial compartment.

To further complement the set of DN-associated features a literature mining approach based on Pubmed Medical Subject Headings (MeSH) annotation and publication to gene links provided in gene2pubmed (<ftp://ftp.ncbi.nlm.nih.gov/gene2pubmed/>) was executed. A Pubmed search using *diabetic nephropathies[majr] AND human[mh]* as query string was performed for identifying publications of relevance in the context of DN, resulting in 10,766 publications. Protein coding genes explicitly discussed in these publications were extracted from gene2pubmed by filtering based on Pubmed ID and Taxonomy ID (9606 for human).

Finally, the sets of differentially regulated features identified in the individual transcriptomics studies as well as the set of genes from literature extraction were consolidated on the Ensembl gene namespace (Table 1).

Biomarker and target annotation from scientific literature

A NCBI Pubmed search for publications holding *Diabetic Nephropathies* further qualified by one of the following qualifiers *pathology, physiopathology, enzymology, metabolism, complications, blood, diagnosis, urine, and epidemiology* as major MeSH concept, further demanding one of the MeSH concepts *Biological Markers* or *Tumor Markers, Biological* was performed for identifying publications discussing biomarker candidates. For

Table 1 | Diabetic nephropathy molecular data space.

Data type	Study setup	# Protein coding genes	References
Transcriptomics, tissue biopsies	Comparison of healthy references (GFR > 60) and established DN (GFR 30-59);		Berthier et al., 2009
	Glomerular compartment:	5	
	Tubulointerstitial compartment:	7	
Transcriptomics, tissue biopsies	Comparison of healthy references (GFR > 60) and established DN (GFR 30-59);		Woroniecka et al., 2011
	Glomerular compartment:	164	
	Tubulointerstitial compartment:	183	
Transcriptomics, tissue biopsies	Comparison of healthy references (GFR > 60) and patients with type 2 diabetes > 5 years;		Baelde et al., 2004
	Glomerular compartment:	167	
Transcriptomics, tissue biopsies	Comparison of healthy references and established DN (no further details provided)		Cohen et al., 2008
	Tubulointerstitial compartment:	69	
Literature extraction	PubMed MeSH query as defined in main text	415	–
Total number of unique protein coding genes		881	

Provided is the data type, study setup details, number of protein coding genes identified as DN-associated, and literature reference for a study.

retrieving drug target candidates the term *Diabetic Nephropathies* with the qualifiers *drug therapy* and *therapy* was used, respectively. The search revealed 615 publications for biomarkers and 2,692 for drug targets. Their respective Pubmed IDs were subsequently used for extracting human genes from the gene2pubmed file, resulting in 54 biomarker candidates and 19 drug target candidates.

Target annotation via drugs under investigation

Clinical trial data for completed and currently ongoing clinical trials were retrieved from ClinicalTrials.gov (<http://clinicaltrials.gov/>). The advanced search as provided on the ClinicalTrials.gov webpage was used for identifying studies that fulfilled the following two criteria: *Study Type* equals *Interventional Studies* and *Condition* contains *Diabetic Nephropathy*, revealing 206 clinical studies. Title and trial description were manually reviewed for focus on renal disease, resulting in 124 studies further considered. Respective drug interventions were mapped to their DrugBank entries (Law et al., 2014), extracting human drug targets as listed, being further mapped on the Ensembl gene namespace. In total 86 drug targets were identified using this approach, of which one was also part of the 19 target candidates retrieved from mining of scientific literature essentially covering basic and translational research activities.

Drug mechanism of action molecular data

A set of ACE inhibitors was retrieved from the Anatomical Therapeutic Chemical (ATC) classification system maintained by the World Health Organization (WHO). 16 compounds classified under *ACE inhibitors, plain* (ATC code: C09AA) were identified and used for subsequent data extraction from DrugMatrix (<https://ntp.niehs.nih.gov/drugmatrix/index.html>). For six out of the 16 drugs sets of genes being affected by drug presence in rat kidney tissue after drug administration were available within DrugMatrix. Obtained rat gene sets were subsequently mapped from Unigene IDs (Sayers et al., 2009) to Ensembl rat IDs and from there further to human ortholog genes according to Ensembl (Table 2).

Table 2 | Drug mechanism of action data space.

Drug name	# Protein coding genes	Database references
Benazepril	442	ICX5600735
Captopril	535	ICX5602791
Enalapril	526	ICX5601254
Lisinopril	558	ICX5601689
Quinapril	572	ICX5602295
Ramipril	519	ICX5602317
Total number of unique protein coding genes	2058	

Given is the drug name, number of associated human protein coding genes identified as significantly affected by drug presence in transcriptomics profiling, and DrugMatrix reference identifier.

MOLECULAR PATHWAY AND PROTEIN INTERACTION DATA

KEGG and Panther pathway membership information for protein coding genes was obtained via KEGG's REST service and from the plain-text database file available on the Panther web site, respectively (Thomas et al., 2003; Kanehisa et al., 2014). Human protein-protein interaction data from BioGRID, INTACT and Reactome were extracted from the respective plain-text files provided by the individual data sources (Stark et al., 2006; Kerrien et al., 2012; Croft et al., 2014). Gene and protein identifiers provided in the original sources were mapped to their respective Ensembl gene IDs. Protein-protein interaction data were further merged into a protein-protein interaction network using Ensembl gene IDs as common denominator of the individual networks.

MOLECULAR PATHWAY AND PROCESS IDENTIFICATION

Molecular pathways and processes were analyzed on the one hand on the basis of a literature review of KEGG and Panther pathways already discussed as relevant in the context of DN. In a second approach de-novo identification of DN molecular processes was performed utilizing the DN pathophysiology feature set. A segmentation algorithm for the identification of processes in the DN protein-protein interaction network was pursued for assembling a molecular process model for DN. Utilizing an analogous procedure a molecular mechanism of action model for ACE inhibitors was constructed utilizing expression signatures obtained from DrugMatrix.

DN pathways from literature

A NCBI Pubmed search for publications utilizing the query string "*diabetic nephropathy*"[ti] OR "*diabetic nephropathies*"[ti] AND (*pathway*[ti] OR *pathways*[ti]) was performed resulting in 53 publications holding the keywords in the title. Subsequently, named entity recognition was performed to annotate occurrence of pathway names according to KEGG and Panther entries in the title and abstract of these publications. Finally, abstracts holding a pathway name were manually reviewed to ensure an association of the identified pathway in the context of DN, leading to 27 individual pathways discussed in literature as being afflicted with DN. Relations between pathways were inferred based on shared genes and the number of protein-protein interactions spanning across pathway boundaries.

Molecular process models

Computing molecular process models followed the procedure described in Mayer et al. (2012); Heinzel et al. (2014). In essence, three main steps are performed: (i) mapping of a feature signature being either the DN pathophysiology association (Table 1) or the ACE mechanism of action set (Table 2) on the consolidated protein interaction network, followed by induced subgraph extraction. Nodes with a degree of zero are removed from the subgraph. (ii) molecular process identification via utilizing a segmentation algorithm (MCODE with default settings, Bader and Hogue, 2003), and (iii) determining inter-process relations defined by the number of protein-protein interactions observed between any actual two molecular processes contrasted against the number of interactions between two random sets of nodes with matching node set size.

Enrichment analysis

For identifying significance of enrichment of molecular feature sets in molecular processes and pathways a Fisher's exact test with a significance level set to 0.05 was used. Benjamini Hochberg correction was employed to adjust for multiple testing.

RESULTS

DN MOLECULAR PATHWAYS

Screening scientific literature resulted in 27 molecular pathways being observed in the context of DN according to KEGG and Panther pathway annotation (**Figure 1**). The pathway map is dominated by linked signaling components, with major elements being MAPK-VEGF, and Jak-STAT-cytokine-cytokine receptor interaction further interacting with TGF-beta signaling, covering among others mechanisms of hypoxia response and fibrosis, respectively (Rudnicki et al., 2009; Loeffler and Wolf, 2014). Additional mechanistic aspects include stress response and involvement of extracellular matrix (McLennan et al., 2013; Tan and de Haan, 2014). Further, a number of specific pathways in the context of metabolism are included, as well as the RAS, with the latter however showing no direct links to other pathways on the molecular feature overlap or direct protein interaction level.

Screening for biomarker candidates in scientific literature resulted in 54 protein coding genes, extraction of drug target candidates from literature as well as clinical trials brought forward 104 such genes. Of the 54 biomarker candidates 23 are assigned to the DN pathway map, for the 104 target candidates 52 are involved (**Table 3**).

Significant coverage regarding biomarker as well as target candidates is again seen for central signaling components

including chemokine signaling, cytokine-cytokine receptor interaction, complemented by MAPK and PI3K-Akt signaling. Also mechanisms are addressed including key features as VEGFA and TGFBI. No specific targeting is seen for counteracting structural changes in ECM, and minor efforts appear to be assigned to adapting stress response. For seven out of 20 pathways discussed no biomarker or target annotation is identified, and complementary a large number of such features are assigned also outside the pathway landscape presented in **Figure 1**. Prominent examples for void biomarker assignment include connective tissue growth factor (CTGF) as factor in fibrosis not being assigned in KEGG, the same being true for uromodulin (UMOD) shown to be associated with progressive disease including genetic polymorphisms (Deshmukh et al., 2013; James et al., 2013). CTGF is also discussed in the therapeutic context via utilizing a monoclonal antibody-based approach (Adler et al., 2010).

Testing the DN pathophysiology feature set retrieved from consolidation of transcriptomics profiles regarding enrichment in the given DN pathway landscape identified seven such pathways as significant, however, missing central mechanisms as hypoxia response or TGFβ signaling. In contrast other pathways beyond the map given in **Figure 1** appeared significantly enriched, including focal adhesion, cell adhesion molecules and adherence junctions, linking to the signaling aspects involved in the disease.

DN MOLECULAR MODEL

Complementary to analysis on molecular pathways as defined in KEGG and Panther we performed a network segmentation

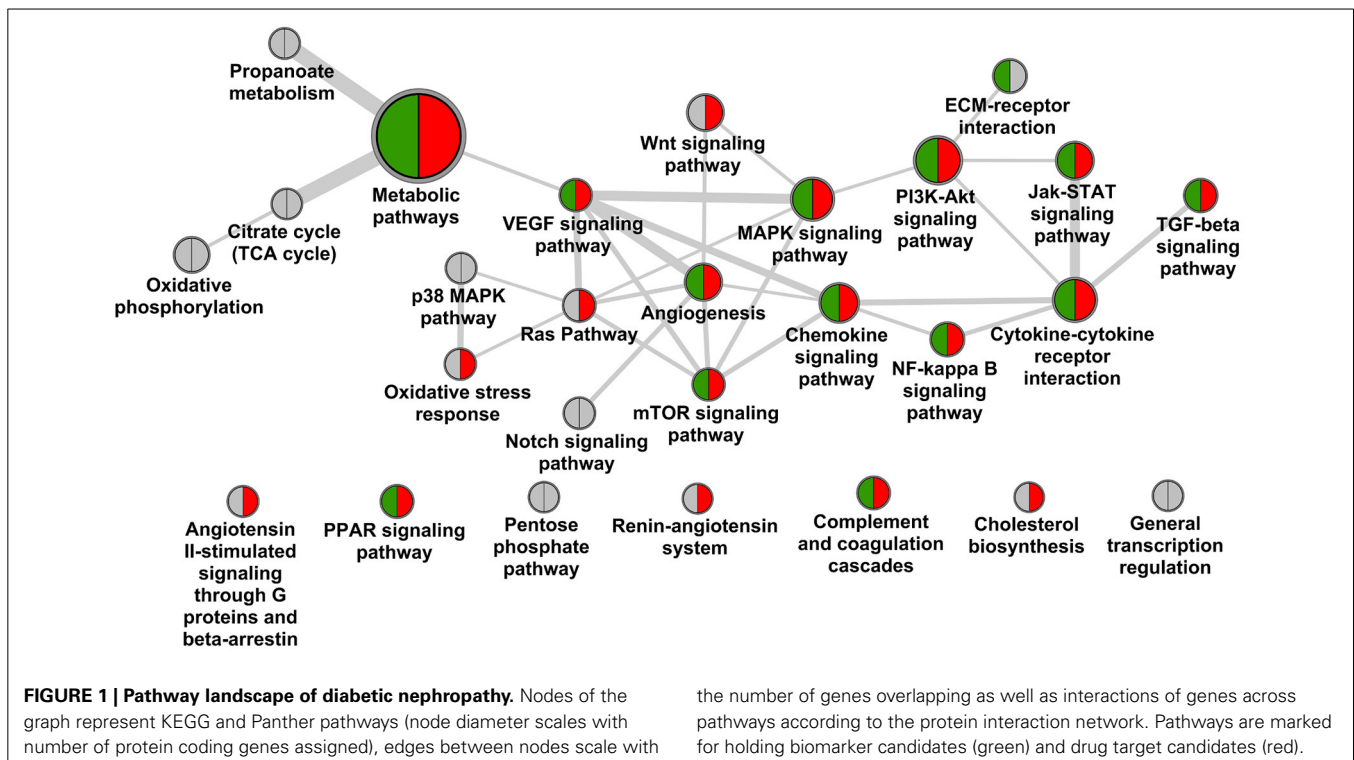


Table 3 | Molecular pathway annotation, diabetic nephropathy.

Pathway name	# Genes	Biomarker	Drug target	Enrichment
Angiogenesis	148	HSPB2, VEGFA, HSPB2-C11orf52	JUN, VEGFA	No
Angiotensin II-stimulated signaling through G proteins and beta-arrestin	35	–	AGTR1	No
Chemokine signaling	190	CCL2, NFKB1, CCL5	CCL2	Yes
Cholesterol biosynthesis	11	–	HMGCR	No
Complement and coagulation cascades	69	F2, FGB, MBL2	SERPIND1, SERPINC1	Yes
Cytokine-cytokine receptor interaction	272	CCL2, LEP, VEGFA, TNFRSF11B, CCL5, PRL, TGFB1	CCL2, TGFB1, VEGFA, TNFSF12, IL18, IL1B, FLT1	Yes
ECM-receptor interaction	87	SPP1, FN1	–	Yes
Jak-STAT signaling	158	LEP, PRL	SOCS1	No
MAPK signaling	256	TGFB1, FGF23, NFKB1	CACNA1H, CACNA1I, CACNB4, CACNA1S, CASP3, CACNA2D3, TGFB1, CACNB3, CACNA1A, CACNA1B, CACNA1C, CACNA1D, CACNA1F, CACNA1G, JUN, CACNB2, CACNG1, IL1B, CACNA2D1, CACNB1	No
Metabolic pathways	1165	XYLT2, PTGDS, KL, PON1, PON2	PTGS2, PDXK, QPRT, ALOX5, NT5E, IMPDH1, ACSL4, XDH, CES1, NNMT, ANPEP, HMGCR, IMPDH2, CYP11B2	No
mTOR signaling	61	VEGFA	PDPK1, VEGFA, INS	No
NF-kappa B signaling	90	NFKB1	PTGS2, IL1B	No
Oxidative stress response	44	–	JUN	No
PI3K-Akt signaling	345	SPP1, VEGFA, FN1, NFKB1, PRL, FGF23	PDPK1, FLT1, VEGFA, INS	Yes
PPAR signaling	71	ADIPOQ	PPARG, ACSL4, FABP1, PDPK1, PPARA, ADIPOQ	No
Ras Pathway	69	–	PDPK1, JUN	No
Renin-angiotensin system	17	–	ACE2, AGTR1, REN, ANPEP, ACE	Yes
TGF-beta signaling	80	TGFB1, SMAD1	TGFB1	No
VEGF signaling	62	VEGFA	PTGS2, VEGFA	No
Wnt signaling	139	–	JUN	Yes
–	–	SPON2, WTAP, UMOD, LCN2, HP, VNN1, AGER, TGFB1, RBP4, NPHS1, HBA1, HBA2, DEFA1B, LPA, CST3, CTGF, ACTA1, PGC, S100A9, DPP4, ALB, CCKAR, GSTP1, DEFA3, S100A8, DEFA1, MMP9, CDH1, S100A4, NPPB, HAVCR1	SOAT1, SLC6A4, ADORA1, MC2R, SIRT1, CYCS, RETN, EDNRA, CRH, EDNRB, KCNA1, ADORA2A, CALM2, CALM3, CALM1, PTX3, PDE3A, KCNMA1, P2RY12, SLC12A1, SLC12A3, GLP1R, DPP4, PDE5A, NR3C2, KCNJ11, ITGB2, KIF6, MMP9, CA12, TUBB1, NAMPT, HCAR3, HCAR2, AR, HBA1, HBA2, CA9, KCNH2, CA2, CA1, CASP1, TUBB, CA4, AHR, CTGF, ABCA1, PDE4A, PDE4B, SCN5A, MMP2, NPC1L1	
Citrate cycle (TCA cycle)	31	–	–	No
General transcription regulation	30	–	–	No
Notch signaling	48	–	–	No
Oxidative phosphorylation	122	–	–	No
p38 MAPK	34	–	–	No
Pentose phosphate	27	–	–	No
Propanoate metabolism	32	–	–	No

Provided is the KEGG pathway name, number of genes assigned to the pathway according to the pathway source, biomarker, and drug target candidates included in the pathway (gene symbols), and indication of significance of enrichment of such pathway on the basis of the consolidated DN kidney tissue transcriptomics data.

procedure aimed at identifying DN molecular process segments defined by topological characteristics of the DN-specific subgraph. From the in total 881 protein coding genes included in the DN molecular pathophysiology gene set (Table 1) 880 were also part of the consolidated interaction network, and 634 were identified as member of the induced subgraph (Figure 2A). From the total set of 880 features 246 protein coding genes had no interaction to any other feature of the DN consensus set, hence being disregarded in molecular model computation. Apparent is the relatively minor overlap of features extracted from literature when compared to signatures from transcriptomics. From the in total 516 unique features consolidated from four transcriptomics profiling experiments and 414 features derived from scientific literature 49 are shared.

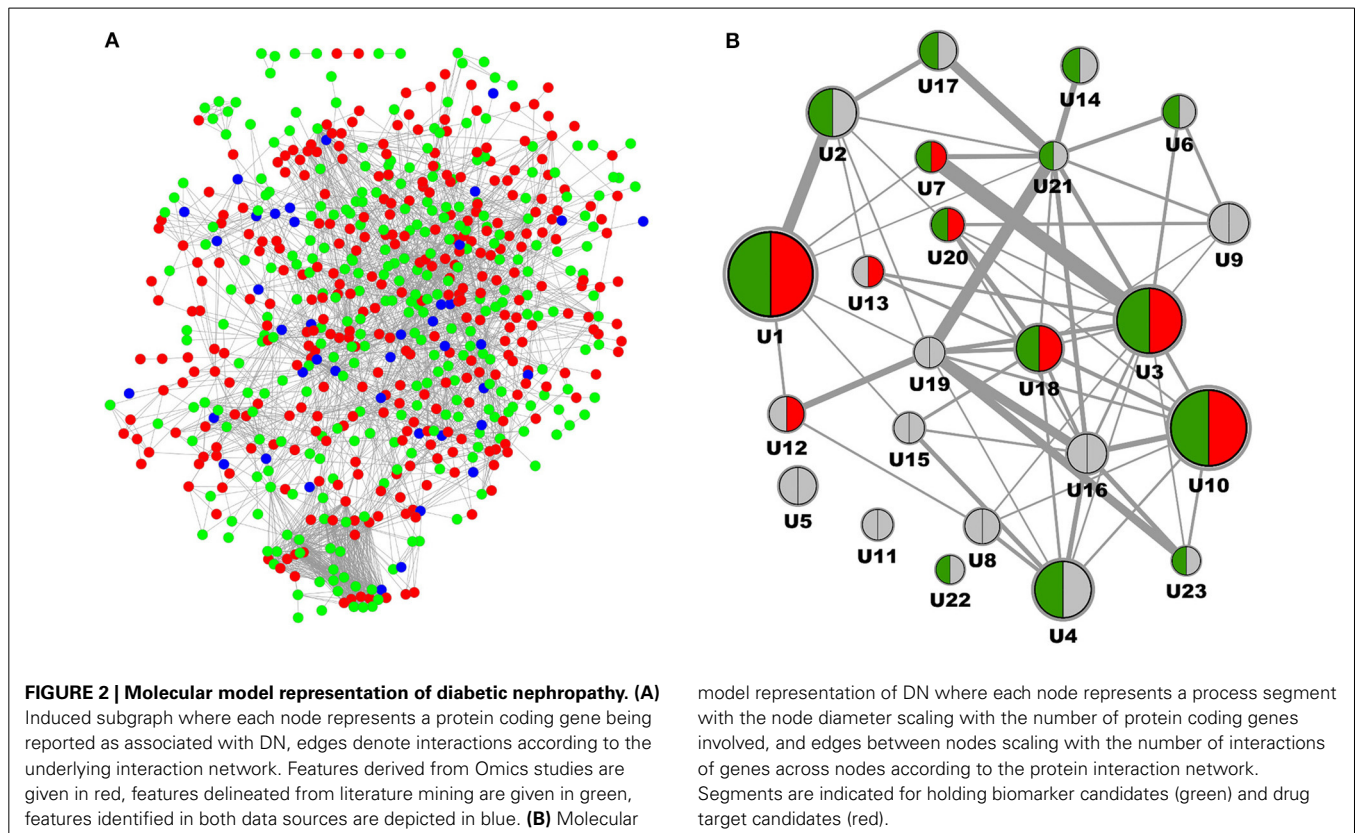
After MCODE segmentation 200 molecular features remained in process segments, forming a molecular model holding 23 process segments (Figure 2B). Median number of protein coding genes per process segment is 6, with the largest segment encoding 29 features, the smallest 3. Equivalently to the pathway graph in Figure 1 a process graph serves as approximation of individual molecular process characteristics together with their dependencies. Six process segments of the process model hold both, biomarker as well as target candidate annotation, with others encoding just one of the two or none. Of the 54 biomarker candidates 22 are included in the molecular model, the respective number for the 104 targets candidates is 16.

DN MOLECULAR MODEL AND DRUG MECHANISM OF ACTION MODEL INTERFERENCE

Consolidating transcriptomics signatures reflecting the impact of ACE inhibitors on the kidney interactome in a rat model utilizing six representative drugs resulted in 2058 molecular features (Table 2), with 661 features being identified in a least two of the six drug signatures. Mapping this consensus ACE feature subset on the consolidated interaction network allowed representation of 656 features. The induced subgraph included 332 features, after segmentation resulting in 12 process segments holding in total 92 molecular features (Figure 3, left). Median process feature set size was 8, with a maximum of 19 and a minimum of 3.

Interfering the ACE mechanism of action molecular model with the DN molecular model on the level of feature overlap (Figure 3) identified specific process segments of the DN molecular model also holding biomarker candidates (Table 4).

All four process segments of DN showing interference with the ACE drug mechanism of action model hold biomarker candidates. Two segments provide significant enrichment also on the level of molecular pathways, showing an integration of chemokine and cytokine signaling, RAS and complement and coagulation cascades for one process segment, the second process segment reflects components of PI3K-Akt signaling in the context of TGF β signaling and ECM receptor interaction.



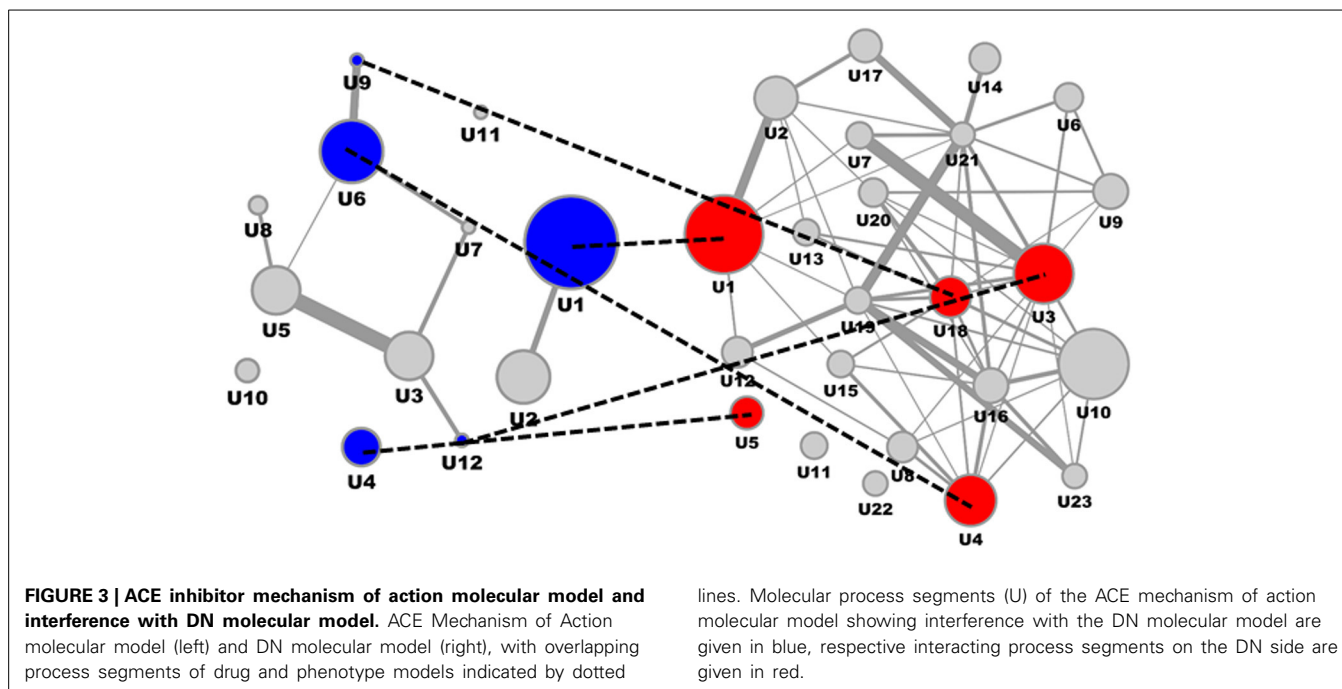


Table 4 | Diabetic nephropathy process segment interference.

Segment	# Genes in segment	Interference overlap	Biomarker candidates	Enriched pathways
1	29	7	CCL5	Chemokine signaling; Cytokine-cytokine receptor interaction; Renin-angiotensin system; Complement and coagulation cascades
18	11	2	HBA1, NFKB1, HP, HBA2	–
3	20	3	TGFB1	ECM-receptor interaction; TGF-beta signaling; PI3K-Akt signaling
4	16	2	ACTA1	–

Provided is the process segment number of the DN molecular model, number of genes assigned to the segment, number of features identified as affected according to the drug mechanism of action model, biomarkers involved in the segment (gene symbols), and relevant pathways from the DN pathway map being enriched in such segment.

Biomarker candidates serving as proxy for the interference of ACE and DN molecular models involve the chemokine (C-C motif) ligand 5 involved in immunoregulators and inflammatory processes, hemoglobin alpha 1 and 2 together with haptoglobin, the cytokine transforming growth factor, beta 1, along with the transcription factor NFKB1, finally including actin, alpha 1 involved in cell motility, structure and integrity.

DISCUSSION

For a large spectrum of clinical presentations an impressive number of drug targets have been proposed out of translational and preclinical research, with a significant number further proceeding into clinical trials. Just in the first half of 2014 close to 10,000 new clinical studies were recorded on the platform clinicaltrials.gov. Taking a specific look at diabetic nephropathy as clinical phenotype, 124 interventional trials in any status are identified at clinicaltrials.gov specifically involving the disease term,

covering 45 individual drug entities addressing 86 known targets. Via mining scientific literature additional 18 drug targets are identified.

Next to a number of trials utilizing drugs and drug combinations addressing known factors impacting DN progression as the RAS, drug targets are disparately distributed across molecular pathways, hence mechanisms assigned to the disease.

From literature mining 27 different pathways according to KEGG and Panther pathway annotation are discussed as associated with DN, of which 19 hold drug targets. These include well known mechanisms of relevance in DN including hypoxia response or fibrosis, combined with a large set of signaling components. On top, 52 drug targets are embedded in molecular context outside this literature-derived DN pathway landscape.

For biomarker candidates an equivalent situation is found. 54 unique proteins extracted from scientific literature are discussed in any biomarker context, covering 14 of the 27 pathways, with 31

biomarker candidates not assigned to any of the members of the extracted DN pathway map.

Interestingly, predictive performance regarding disease progression of any of the individual biomarker candidates proved limited value. For example, in a review by Hellemons et al. 13 relevant markers were found in the context of nephropathy in diabetes, of which five were found as significantly associated with onset as well as progression of DN again covering various mechanisms including inflammation (e.g., C-reactive protein), cell surface interaction and homeostasis (e.g., E-selectin, ICAM1) and metabolism (triglyceride levels) (Hellemons et al., 2012).

Apparently, individual biomarkers reflecting the status of an individual molecular process, pathway or mechanism cannot capture disease prognosis for the comprehensive DN population. In alternative approaches multimarker panels were included in classifiers on disease diagnosis and prognosis demonstrating improved performance also in blinded validation. In Roscioni et al. a signature of 273 peptides determined in urine were included in a support vector machine-based classifier (Roscioni et al., 2013). The signature held fragments of collagen eventually mirroring alterations in the extracellular matrix turnover and fibrosis together with markers of inflammation as e.g., the pro-inflammatory protein S100-A9, as well as uromodulin shown to be associated with interstitial fibrosis and tubular atrophy (Nkuipou-Kenfack et al., 2014).

One contributing factor for needing multimarker panels may be individual variance of baseline biomarker levels, where inclusion of multiple markers specifically in non-linear classification methods adds to robustness. However, a second factor may be generic heterogeneity of the patient population. Specific disease presentation may significantly vary not only across stages of disease progression eventually seeing a transition from protective to damaging mechanisms, but even within a specific chronic kidney disease category as defined by present clinical classification provided by KDIGO guidelines (KDIGO Board Members, 2013).

Improved prognostic performance of multimarker panels on top of strict functional classification of stage transitions in DN utilizing albuminuria but also eGFR as clinically used progression parameters clearly support the case of pathophysiological heterogeneity of a, in present clinical terms homogeneous, patient population. However, specifically for albuminuria the role of functional marker vs. factor in disease is discussed (Roscioni et al., 2014).

Deriving robust diagnostic or prognostic classifiers from e.g., proteomics or metabolomics profiling may add to clinical patient management regarding onset as well as intensity of therapeutic measures (Roscioni et al., 2013; Pena et al., 2014). Also in clinical trial design such enrichment strategies may be utilized by e.g., identifying individuals prone to fast disease progression, and randomizing in this high risk cohort into medication and placebo arm (e.g., Priority trial, clinicaltrials.gov reference NCT02040441).

Prognostic biomarkers in contrast to diagnostic parameters with known assignment to molecular processes and pathways further allow an approximation of what specific mechanisms are associated with disease progression. The DN pathway landscape discussed in this work is solely a cross-sectional representation

of the disease, in a first place not allowing deciphering which of the 27 individual pathways drive disease progression, and which other pathways are just bystanders or downstream consequences of mechanistic factors of disease. Hence, evaluating biomarker candidates for their association with progressive disease in turn allows determining mechanisms associated with progressive disease. Such knowledge is vital e.g., for determining novel drug targets, demanding to be embedded in disease mechanisms being factors for progressive disease. Remaining question however is if such mechanisms are relevant to the same extent or at all for a specific patient assigned to a clinical phenotype.

A prognostic biomarker set covering all potentially relevant processes enables specific molecular phenotyping of individual patients, being however not sufficient in terms of predicting drug response as a drug mechanism of action is not factored in. Here Systems Pharmacology aims at identifying drug response also on the level of molecular processes and pathways. Rationale is to not only focus on the specific drug target and its assignment to specific mechanisms, but to include the systemic molecular changes triggered by the drug including off-target effects as well as downstream molecular changes. Having a drug mechanism of action as well as a clinical phenotype represented on a molecular process or pathway level allows intersecting both molecular states. If from prognostic biomarker profiling of a patient specific progression-associated molecular disease mechanisms are identified, and a drug exhibits functional interference in such specific mechanisms such patient may be more prone for showing response to the drug. With such setting including knowledge on molecular phenotype composition, molecular process relevance in progressive disease and knowledge on interference of drug mechanism of action biomarker candidates initially serving a prognostic purpose can be rendered into predictive biomarkers on drug response.

Omics profiling has a major contribution to characterizing both, clinical phenotypes as well as drug mechanism of action. Integrating profiling results from clinical samples frequently sees minor overlap of individual studies, being in part driven by insufficient sample size combined with diverging inclusion criteria and sample material used (Fechete et al., 2011). In the example presented here 1010 features in total are identified as differentially regulated in transcriptomics or are being assigned to DN according to literature mining, with 880 unique features. An equivalent imbalance in feature coherence across studies is also found for the ACE inhibitor transcriptomics data. All these drugs address the same functional context, but from the in total 3152 features identified for six drugs included the total number of unique features are still 2058, with 661 being identified in at least two drug signatures.

Next divergence becoming apparent is the limited overlap of enrichment analysis based on signatures from profiling and feature-based literature mining compared to explicit literature mining for molecular pathways. Of the 27 pathways extracted from scientific references only seven are confirmed, however, seeing other pathways enriched not found via literature mining. On top, a major shortcoming is restricted representation of protein coding genes in such pathway maps, e.g., for KEGG covering 6491 and for Panther 2163 protein coding genes, respectively. This

limitation not only affects pathway enrichment but also assignment of biomarker and target candidates. Of the in total 104 drug target and 54 biomarker candidates 29 are neither assigned in any KEGG or Panther pathway.

Here a different approach may be followed, namely segmentation of protein interaction networks exhibiting improved coverage of the protein coding gene set. Consolidation of INTACT, Reactome, and BioGRID allows representation of in total 13,907 protein coding genes, clearly expanding beyond public domain pathway databases. In alternative approaches hybrid interaction networks are utilized for further expanding coverage of protein coding genes, but also for improving false negative rates regarding protein-protein interactions and relations (Fechete et al., 2013).

Computing a DN-specific as well as ACE inhibitor-specific induced subgraph followed by topology-based segmentation allows an alternative representation of a molecular process landscape for the clinical presentation as well as the drug mechanism of action. Interference analysis on the level of overlapping protein coding genes resulted in four process segments holding central aspects of DN pathophysiology. Seven biomarker candidates were identified in these interfering molecular processes. CCL5 (RANTES), involved in recruiting monocytes and macrophages to the renal cortex was shown to be suppressed by ACE inhibition, indicating that RANTES expression is mediated via Angiotensin II type 2 receptor (Kashiwagi et al., 2002). Equivalently, in animal models TGF β 1 expression was shown to be reduced by ACE inhibitors. Activation of NF κ B1 by angiotensin II was shown in vascular smooth muscle and mesangial cells (Hernández-Presa et al., 1997). In a study by Dong et al. analyzing cost effectiveness of ACE inhibitor treatment for patients with type 1 diabetes mellitus the level of glycosylated HbA1c showed clear impact on cost effectiveness of drug use per quality-adjusted life year (QALY) (Dong et al., 2004). The authors concluded that next to patient age also other factors need to be included in therapy considerations.

Apparently, drug mechanism of action affects numerous molecular processes, as exemplified for ACE inhibitors, many of these also afflicted with DN progression. Analyzing the molecular process interface of disease progression-relevant pathophysiology and drug mechanism of action allows proposing predictive markers. Testing such predictive biomarker candidates may educate on relevance of individual processes on a patient level, directly linking to likelihood of drug response.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme under the grant agreement no. 241544 (SysKid).

REFERENCES

- Adler, A. I., Stevens, R. J., Manley, S. E., Bilous, R. W., Cull, C. A., and Holman, R. R. (2003). Development and progression of nephropathy in type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS 64). *Kidney Int.* 63, 225–232. doi: 10.1046/j.1523-1755.2003.00712.x
- Adler, S. G., Schwartz, S., Williams, M. E., Arauz-Pacheco, C., Bolton, W. K., Lee, T., et al. (2010). Phase 1 study of anti-CTGF monoclonal antibody in patients with diabetes and microalbuminuria. *Clin. J. Am. Soc. Nephrol.* 5, 1420–1428. doi: 10.2215/CJN.09321209
- Bader, G. D., and Hogue, C. W. V. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4:2. doi: 10.1186/1471-2105-4-2
- Baelde, H. J., Eikmans, M., Doran, P. P., Lappin, D. W. P., de Heer, E., and Bruijn, J. A. (2004). Gene expression profiling in glomeruli from human kidneys with diabetic nephropathy. *Am. J. Kidney Dis.* 43, 636–650. doi: 10.1053/j.ajkd.2003.12.028
- Berthier, C. C., Zhang, H., Schin, M., Henger, A., Nelson, R. G., Yee, B., et al. (2009). Enhanced expression of Janus kinase-signal transducer and activator of transcription pathway members in human diabetic nephropathy. *Diabetes* 58, 469–477. doi: 10.2337/db08-1328
- Brenner, B. M., Cooper, M. E., de Zeeuw, D., Keane, W. F., Mitch, W. E., Parving, H. H., et al. (2001). Effects of losartan on renal and cardiovascular outcomes in patients with type 2 diabetes and nephropathy. *N. Engl. J. Med.* 345, 861–869. doi: 10.1056/NEJMoa011161
- Buyse, M., Loi, S., van't Veer, L., Viale, G., Delorenzi, M., Glas, A. M., et al. (2006). Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J. Natl. Cancer Inst.* 98, 1183–1192. doi: 10.1093/jnci/djj329
- Cohen, C. D., Lindenmeyer, M. T., Eichinger, F., Hahn, A., Seifert, M., Moll, A. G., et al. (2008). Improved elucidation of biological processes linked to diabetic nephropathy by single probe-based microarray data analysis. *PLoS ONE* 3:e2937. doi: 10.1371/journal.pone.0002937
- Croft, D., Mundo, A. F., Haw, R., Milacic, M., Weiser, J., Wu, G., et al. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Res.* 42, D472–D477. doi: 10.1093/nar/gkt1102
- De Leon, J. (2012). Evidence-based medicine versus personalized medicine: are they enemies? *J. Clin. Psychopharmacol.* 32, 153–164. doi: 10.1097/JCP.0b013e3182491383
- Deshmukh, H. A., Palmer, C. N. A., Morris, A. D., and Colhoun, H. M. (2013). Investigation of known estimated glomerular filtration rate loci in patients with type 2 diabetes. *Diabet. Med.* 30, 1230–1235. doi: 10.1111/dme.12211
- De Zeeuw, D., Akizawa, T., Audhya, P., Bakris, G. L., Chin, M., Christ-Schmidt, H., et al. (2013). Bardoxolone methyl in type 2 diabetes and stage 4 chronic kidney disease. *N. Engl. J. Med.* 369, 2492–2503. doi: 10.1056/NEJMoa1306033
- Dong, F. B., Sorensen, S. W., Manninen, D. L., Thompson, T. J., Narayan, V., Orians, C. E., et al. (2004). Cost effectiveness of ACE inhibitor treatment for patients with type 1 diabetes mellitus. *Pharmacoeconomics* 22, 1015–1027. doi: 10.2165/00019053-200422150-00005
- Fechete, R., Heinzel, A., Perco, P., Mönks, K., Söllner, J., Stelzer, G., et al. (2011). Mapping of molecular pathways, biomarkers and drug targets for diabetic nephropathy. *Proteomics Clin. Appl.* 5, 354–366. doi: 10.1002/prca.201000136
- Fechete, R., Heinzel, A., Soellner, J., Perco, P., Lukas, A., and Mayer, B. (2013). Using information content for expanding human protein coding gene interaction networks. *J. Comput. Sci. Syst. Biol.* 6, 73–82. doi: 10.4172/jcsb.1000102
- Fried, L. F., Emanuele, N., Zhang, J. H., Brophy, M., Conner, T. A., Duckworth, W., et al. (2013). Combined angiotensin inhibition for the treatment of diabetic nephropathy. *N. Engl. J. Med.* 369, 1892–1903. doi: 10.1056/NEJMoa1303154
- He, J. C., Chuang, P. Y., Ma'ayan, A., and Iyengar, R. (2012). Systems biology of kidney diseases. *Kidney Int.* 81, 22–39. doi: 10.1038/ki.2011.314
- Heinzel, A., Fechete, R., Mühlberger, I., Perco, P., Mayer, B., and Lukas, A. (2013). Molecular models of the cardiorenal syndrome. *Electrophoresis* 34, 1649–1656. doi: 10.1002/elps.201200642
- Heinzel, A., Fechete, R., Söllner, J., Perco, P., Heinze, G., Oberbauer, R., et al. (2012). Data graphs for linking clinical phenotype and molecular feature space. *Int. J. Syst. Biol. Biomed. Technol.* 1, 11–25. doi: 10.4018/ijssbtt.2012010102
- Heinzel, A., Mühlberger, I., Fechete, R., Mayer, B., and Perco, P. (2014). Functional molecular units for guiding biomarker panel design. *Methods Mol. Biol.* 1159, 109–133. doi: 10.1007/978-1-4939-0709-0_7
- Hellemons, M. E., Kerschbaum, J., Bakker, S. J. L., Neuwirt, H., Mayer, B., Mayer, G., et al. (2012). Validity of biomarkers predicting onset or progression of nephropathy in patients with Type 2 diabetes: a systematic review. *Diabet. Med.* 29, 567–577. doi: 10.1111/j.1464-5491.2011.03437.x
- Hernández-Presa, M., Bustos, C., Ortego, M., Tuñón, J., Renedo, G., Ruiz-Ortega, M., et al. (1997). Angiotensin-converting enzyme inhibition prevents arterial nuclear factor-kappa B activation, monocyte chemoattractant protein-1 expression, and macrophage infiltration in a rabbit model of early accelerated atherosclerosis. *Circulation* 95, 1532–1541.

- Hicks, D. G., and Kulkarni, S. (2008). HER2+ breast cancer: review of biologic relevance and optimal use of diagnostic tools. *Am. J. Clin. Pathol.* 129, 263–273. doi: 10.1309/99AE032R9FM8WND1
- Himmelfarb, J., and Tuttle, K. R. (2013). New therapies for diabetic kidney disease. *N. Engl. J. Med.* 369, 2549–2550. doi: 10.1056/NEJMe1313104
- Hollebecque, A., Massard, C., and Soria, J.-C. (2014). Implementing precision medicine initiatives in the clinic: a new paradigm in drug development. *Curr. Opin. Oncol.* 26, 340–346. doi: 10.1097/CCO.000000000000077
- James, L. R., Le, C., Doherty, H., Kim, H.-S., and Maeda, N. (2013). Connective tissue growth factor (CTGF) expression modulates response to high glucose. *PLoS ONE* 8:e70441. doi: 10.1371/journal.pone.0070441
- Johnson, J. A. (2001). Drug target pharmacogenomics: an overview. *Am. J. Pharmacogenomics* 1, 271–281. doi: 10.2165/00129785-200101040-00004
- Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* 42, D199–D205. doi: 10.1093/nar/gkt1076
- Kashiwagi, M., Masutani, K., Shinozaki, M., and Hirakata, H. (2002). MCP-1 and RANTES are expressed in renal cortex of rats chronically treated with nitric oxide synthase inhibitor. Involvement in macrophage and monocyte recruitment. *Nephron* 92, 165–173. doi: 10.1159/000064454
- KDIGO Board Members. (2013). KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int. Suppl.* 3, 1–150. doi: 10.1038/kisup.2012.73
- Kerrien, S., Aranda, B., Breuzau, L., Bridge, A., Broackes-Carter, F., Chen, C., et al. (2012). The IntAct molecular interaction database in 2012. *Nucleic Acids Res.* 40, D841–D846. doi: 10.1093/nar/gkr1088
- Komorowsky, C. V., Brosius, F. C., Pennathur, S., and Kretzler, M. (2012). Perspectives on systems biology applications in diabetic kidney disease. *J. Cardiovasc. Transl. Res.* 5, 491–508. doi: 10.1007/s12265-012-9382-7
- LaFramboise, W. A., Dhir, R., Kelly, L. A., Petrosko, P., Krill-Burger, J. M., Sciuilli, C. M., et al. (2012). Serum protein profiles predict coronary artery disease in symptomatic patients referred for coronary angiography. *BMC Med.* 10:157. doi: 10.1186/1741-7015-10-157
- Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A. C., Liu, Y., et al. (2014). DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.* 42, D1091–D1097. doi: 10.1093/nar/gkt1068
- Lewis, E. J., Hunsicker, L. G., Clarke, W. R., Berl, T., Pohl, M. A., Lewis, J. B., et al. (2001). Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. *N. Engl. J. Med.* 345, 851–860. doi: 10.1056/NEJMoa011303
- Loeffler, I., and Wolf, G. (2014). Transforming growth factor- β and the progression of renal disease. *Nephrol. Dial. Transplant.* 29(Suppl. 1), i37–i45. doi: 10.1093/ndt/gft267
- Mann, J. F. E., Schmieder, R. E., McQueen, M., Dyal, L., Schumacher, H., Pogue, J., et al. (2008). Renal outcomes with telmisartan, ramipril, or both, in people at high vascular risk (the ONTARGET study): a multicentre, randomised, double-blind, controlled trial. *Lancet* 372, 547–553. doi: 10.1016/S0140-6736(08)61236-2
- Mayer, P., Mayer, B., and Mayer, G. (2012). Systems biology: building a useful model from multiple markers and profiles. *Nephrol. Dial. Transplant.* 27, 3995–4002. doi: 10.1093/ndt/gfs489
- McLennan, S. V., Abdollahi, M., and Twiggs, S. M. (2013). Connective tissue growth factor, matrix regulation, and diabetic kidney disease. *Curr. Opin. Nephrol. Hypertens.* 22, 85–92. doi: 10.1097/MNH.0b013e32835b4889
- Mirnezami, R., Nicholson, J., and Darzi, A. (2012). Preparing for precision medicine. *N. Engl. J. Med.* 366, 489–491. doi: 10.1056/NEJMp1114866
- Mühlberger, I., Mönks, G., Fechele, R., Mayer, G., Oberbauer, R., Mayer, B., et al. (2012). Molecular pathways and crosstalk characterizing the cardiorenal syndrome. *OMICS* 16, 105–112. doi: 10.1089/omi.2011.0121
- Nkuiipou-Kenfack, E., Duranton, F., Gayraud, N., Argilés, A., Lundin, U., Weinberger, K. M., et al. (2014). Assessment of metabolomic and proteomic biomarkers in detection and prognosis of progression of renal function in chronic kidney disease. *PLoS ONE* 9:e96955. doi: 10.1371/journal.pone.0096955
- Pammolli, F., Magazzini, L., and Riccaboni, M. (2011). The productivity crisis in pharmaceutical R&D. *Nat. Rev. Drug Discov.* 10, 428–438. doi: 10.1038/nrd3405
- Parving, H.-H., Persson, E., Lewis, J. B., Lewis, E. J., and Hollenberg, N. K. (2008). Aliskiren combined with losartan in type 2 diabetes and nephropathy. *N. Engl. J. Med.* 358, 2433–2446. doi: 10.1056/NEJMoa0708379
- Pena, M. J., Lambers Heerspink, H. J., Hellemons, M. E., Friedrich, T., Dallmann, G., Lajer, M., et al. (2014). Urine and plasma metabolites predict the development of diabetic nephropathy in individuals with Type 2 diabetes mellitus. *Diabet. Med.* 31, 1138–1147. doi: 10.1111/dme.12447
- Pergola, P. E., Raskin, P., Toto, R. D., Meyer, C. J., Huff, J. W., Grossman, E. B., et al. (2011). Bardoxolone methyl and kidney function in CKD with type 2 diabetes. *N. Engl. J. Med.* 365, 327–336. doi: 10.1056/NEJMoa1105351
- Rodríguez-Iturbe, B., and García García, G. (2010). The role of tubulointerstitial inflammation in the progression of chronic renal failure. *Nephron Clin. Pract.* 116, c81–c88. doi: 10.1159/000314656
- Roscioni, S. S., de Zeeuw, D., Hellemons, M. E., Mischak, H., Zürbig, P., Bakker, S. J. L., et al. (2013). A urinary peptide biomarker set predicts worsening of albuminuria in type 2 diabetes mellitus. *Diabetologia* 56, 259–267. doi: 10.1007/s00125-012-2755-2
- Roscioni, S. S., Lambers Heerspink, H. J., and de Zeeuw, D. (2014). Microalbuminuria: target for renoprotective therapy PRO. *Kidney Int.* 86, 40–49. doi: 10.1038/ki.2013.490
- Rudnicki, M., Perco, P., Enrich, J., Eder, S., Heining, D., Bernthaler, A., et al. (2009). Hypoxia response and VEGF-A expression in human proximal tubular epithelial cells in stable and progressive renal disease. *Lab. Invest.* 89, 337–346. doi: 10.1038/labinvest.2008.158
- Ruggenti, P., Perna, A., Gherardi, G., Gaspari, F., Benini, R., and Remuzzi, G. (1998). Renal function and requirement for dialysis in chronic nephropathy patients on long-term ramipril: REIN follow-up trial. Gruppo Italiano di Studi Epidemiologici in Nefrologia (GISEN). Ramipril Efficacy in Nephropathy. *Lancet* 352, 1252–1256.
- Sayers, E. W., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetverin, V., et al. (2009). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 37, D5–D15. doi: 10.1093/nar/gkn741
- Stark, C., Breitkreutz, B.-J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 34, D535–D539. doi: 10.1093/nar/gkj109
- Tan, S. M., and de Haan, J. B. (2014). Combating oxidative stress in diabetic complications with Nrf2 activators: how much is too much? *Redox Rep.* 19, 107–117. doi: 10.1179/1351000214Y.00000000087
- Thomas, P. D., Campbell, M. J., Kejarawal, A., Mi, H., Karlak, B., Daverman, R., et al. (2003). PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* 13, 2129–2141. doi: 10.1101/gr.772403
- Trusheim, M. R., Burgess, B., Hu, S. X., Long, T., Averbuch, S. D., Flynn, A. A., et al. (2011). Quantifying factors for the success of stratified medicine. *Nat. Rev. Drug Discov.* 10, 817–833. doi: 10.1038/nrd3557
- U.S. Food and Drug Administration. (2014). *Table of Pharmacogenomic Biomarkers in Drug Labeling* [Internet].
- Woroniecka, K. I., Park, A. S. D., Mohtat, D., Thomas, D. B., Pullman, J. M., and Susztak, K. (2011). Transcriptome analysis of human diabetic kidney disease. *Diabetes* 60, 2354–2369. doi: 10.2337/db10-1181
- Zürbig, P., Jerums, G., Hovind, P., Macisaac, R. J., Mischak, H., Nielsen, S. E., et al. (2012). Urinary proteomics for early diagnosis in diabetic nephropathy. *Diabetes* 61, 3304–3313. doi: 10.2337/db12-0348

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 June 2014; accepted: 29 July 2014; published online: 22 August 2014.
 Citation: Heinzel A, Perco P, Mayer G, Oberbauer R, Lukas A and Mayer B (2014) From molecular signatures to predictive biomarkers: modeling disease pathophysiology and drug mechanism of action. *Front. Cell Dev. Biol.* 2:37. doi: 10.3389/fcell.2014.00037
 This article was submitted to *Systems Biology*, a section of the journal *Frontiers in Cell and Developmental Biology*.
 Copyright © 2014 Heinzel, Perco, Mayer, Oberbauer, Lukas and Mayer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.