Biomarker candidates for cardiovascular disease and bone metabolism disorders in chronic kidney disease: A systems biology perspective

Paul Perco[1,2,3], Julia Wilflingseder[2], Andreas Bernthaler[3], Martin Wiesinger[3], Michael Rudnicki[4], Barbara Wimmer[2], Bernd Mayer[3,5], and Rainer Oberbauer[1,2 *]

[1] Medical University of Vienna, Währinger Gürtel 18-20, A-1090 Vienna, Austria
[2] Krankenhaus der Elisabethinen, Fadingerstrasse 1, A-4020 Linz, Austria
[3] Institute for Theoretical Chemistry, University of Vienna, Währinger Strasse 17, A-1090 Vienna, Austria
[4] Medical University of Innsbruck, Anichstraße 35, A-6020 Innsbruck, Austria
[5] emergentec biodevelopment GmbH, Rathausstrasse 5/3, A-1010 Vienna, Austria

Sources of support: Austrian Science Fund (P-18325-B13 to R.O. and P-18130-B13 to B.M.)

Running title: Biomarkers of chronic kidney, cardiovascular and bone diseases

Correspondence:

Rainer Oberbauer, MD, MSc
KH Elisabethinen Linz and
Medical University of Vienna
Währinger Gürtel 18-20, 1090 Vienna, Austria
Phone: +43-732 7676 4300
Fax: +43-732 7676 4306
Mail: rainer.oberbauer@meduniwien.ac.at

Abstract

Patients with chronic kidney disease (CKD) show a panel of partially deregulated serum markers indicative for bone metabolism disorders and cardiovascular diseases. This review provides an overview of currently reported biomarker candidates at the interface of kidney disease, bone metabolism disorders and cardiovascular diseases, and gives details on their functional interplay on the level of protein-protein interaction data.

We retrieved 13 publications from 1999 to 2006 reporting 31 genes associated with cardiovascular diseases, and 46 genes associated with bone metabolism disorders in patients with CKD. We identified these genes to be functionally involved in signal transduction processes, cell communication, immunity and defense, as well as skeletal development. On the basis of the given set of 77 genes further 276 interacting proteins were identified using reference data on known protein interactions. Their functional interplay was estimated by linking properties reflected by gene expression data characterizing chronic kidney disease, gene ontology terms as provided by the gene ontology consortium, and transcription factor binding site profiles. Highly connected sub-networks of proteins associated with chronic kidney disease, cardiovascular diseases, or bone metabolism disorders were detected involving proteins like collagens (COL1A1, COL1A2), fibronectin (FN1), transforming growth factor beta 1 (TGFB1), or components of fibrinogen (FGA, FGB, FGG).

A systems biology approach provides a methodological framework for linking singular biomarker candidates towards deriving functional dependencies between clinically interlinked diseases.

Keywords

proteomics, genomics, systems biology, chronic kidney disease, cardiovascular disease, bone metabolism disorders

Main topics

Introduction

The use of serum biomarkers has been successfully demonstrated in the clinical context of cardiovascular diseases and bone metabolism disorders, and their predictive value as well as discriminatory power has been well established [1]. Both, - cardiovascular diseases and bone metabolism disorders - might be causally linked in patients with CKD, since the disturbances of the calcium-, phosphate-, vitamin D-, and parathyroid hormone (PTH) metabolism, as well as the incidence of cardiovascular events as myocardial infarction rises early in the course of kidney disease [2, 3].

Next to established diagnostic and prognostic parameters new biomarker candidates are currently arising with astonishing speed, in particular facilitated by genomic and proteomic techniques in principle allowing scans of whole transcriptomes and proteomes of clinical samples. Experimental procedures for deriving such initial marker profiles have traversed towards a routine procedure. The tough part however is the choice of those candidates with clinical relevance for further validation studies [4, 5]. Data integration, bioinformatics analyses, and functional testing of novel hypotheses drawn have been identified as a valuable strategy, commonly denoted in the context of systems biology [6].

Mondry and colleagues emphasized the potential of systems biology and quantitative models in their review on the molecular mechanisms of renal osteodystrophy [7]. Drake *et al.* focused on proteomic approaches and the use of protein-protein interaction (PPI) data for biomarker discovery in their review on systems biology of cardiovascular diseases [8].

This review provides an overview on the suspected link between cardiovascular diseases and bone metabolism disorders in patients with impaired renal function, and will furthermore characterize and analyze reported biomarkers associated with these particular diseases. Subsequently, the interdependency of reported biomarkers will be analyzed on a systems biology level taking into account data on gene expression in chronic kidney disease, functional gene annotation, protein-protein interactions, as well as gene regulatory elements reflected by joint transcription factor binding sites.

Kidney disease and cardiovascular risk

Chronic kidney disease is associated with increased risk for cardiovascular complications and all cause mortality. The risk of death and the prevalence of cardiovascular disease (CVD) start to rise significantly already in patients with early stage renal insufficiency, i.e. with a glomerular filtration rate (GFR) of less than 60 ml/min [9]. In dialysis patients the prevalence of CVD and the mortality due to CVD is even 10 to 30 times higher than in the general population [10]. Cardiovascular events in CKD patients are caused by traditional and non traditional risk factors and their interactions: Atherosclerosis, arteriosclerosis, and altered cardiac morphological characteristics are the main findings [11]. These complex characteristics impose a new challenge in identifying and treating patients with CVD in early stages of CKD towards improving outcome. So far there are no validated biomarkers for identifying the risk of CVD in CKD patients available. As for all biomarkers, CVD markers should be easily measurable and significantly deregulated in disease states. In statistical terms this constraint refers to adequate

discrimination (ROC (receiver operating characteristic)-AUC (area under curve)), as well as transportability, i.e. validity of a marker in different patient populations. The cardiovascular biomarkers which are discussed in the paper by Roberts *et al.* are involved in several pathophysiological processes such as endothelial dysfunction, vascular calcification, monocyte recruitment to the endothelium, inflammation, oxidative stress, sympathetic nervous system activation, glycosylation of proteins, bone marrow function, platelet activation, left ventricular structure and function, myocardial necrosis and other processes [11]. According to the authors an improvement in cardiovascular risk stratification might be achieved by measuring a combination of cardiovascular biomarkers, each representing a different aspect of CVD pathophysiology. Next to their function for assessing the level of risk of vascular disease, biomarkers could depict potential targets for the prevention of such disease [12]. However, the link between the given CVD biomarker candidates and CKD remains elusive.

Kidney disease and bone metabolism disorders

The kidney is involved in calcium/phosphate homeostasis which is tightly regulated by the phosphate-excretion regulating hormones (phosphatonins) fibroblast growth factor 23 (FGF23), parathyroid hormone (PTH) and by the action of the active form of vitamin D (1α,25-dihydroxy cholecalciferol or calcitriol) which is exclusively synthesized in the kidney. In the healthy subject PTH is secreted by the parathyroidea upon hypocalcemia and/or hyperphosphatemia. PTH stimulates the release of calcium and phosphate from bone tissue, the synthesis of calcitriol in the kidney and the reabsorption of calcium by the enteric mucosa and the distule tubule cells in the kidney. On the other hand, PTH

increases the excretion of phosphate in by the renal proximal tubule cells. The renal 1-α hydroxylation of vitamin D which transforms vitamin D into its active form calcitriol is PTH-dependent. Calcitriol increases the enteral and renal calcium and phosphate reabsorption and bone mineralization. Finally, FGF23 is a phosphatonin increased by high phosphate ingestion. FGF23 enhances fractional renal phosphate excretion and inhibits the 1-α hydroxylation of vitamin D thus directly interfering with calcitriol synthesis.

In stage I and stage II of kidney disease, i.e. when GFR is normal or only slightly reduced, the levels of calcium, phosphate and PTH in plasma are usually not different from healthy individuals [13]. During progression to stage III of chronic kidney disease fractional renal excretion of phosphate rises mainly due to phosphate retention and subsequent increased levels of the phosphatonins FGF23 and PTH, which keep the serum phosphate levels in the normal range [14]. Thus progression of kidney disease causes changes in phosphate homeostasis finally leading to a rise in serum levels of PTH which is called secondary hyperparathyroidism (sHPT). In addition patients with stage III kidney disease frequently suffer from a deficiency in 25-OH-vitamin D_3, which leads to diminished synthesis of active vitamin D_3 [15]. Furthermore, the action of the 1-α-hydroxylase in the kidney is inhibited by rising levels of FGF23 and by the progression of renal insufficiency *per se*, which finally leads to decreased plasma levels of active vitamin D_3 [14]. If glomerular filtration rate falls below 30 ml/min (i.e. stage IV and V of chronic kidney disease) the excretion of phosphate cannot be enhanced any further and hyperphosphatemia develops.

sHPT leads to severe changes in bone mineralization and structure, and the term chronic kidney disease-mineral bone disorder (CKD-MBD) was coined [16, 17]. However, PTH receptors are not only found in kidney, bone and enteric mucosal cells but also in the cardiovascular-system. Therefore, sHPT and adjacent vitamin D therapy not only lead to CKD-MBD but is also associated with the development of vascular, valvular and extravascular calcifications, all increasing mortality [18].

The degree of bone formation rate (BFR) can be somehow estimated by determining plasma levels of several marker proteins. While the plasma levels of bone-specific alkaline phosphatase (bAP), osteocalcin (OC) and procollagen type I carboxy-terminal extension peptide (PICP) stand for the degree of bone formation, the bone resorption rate is represented e.g. by procollagen type I crosslinked carboxy-terminal telopeptide (ICTP), plasma deoxypyridinoline (DPD), bone-specific tartrate-resistant acid phosphatase (TRAP), and some of the multiple products resulting from the degradation of type I collagen [19, 20]. Other circulating molecules are of growing interest as they may also be indicative for the bone turnover rate, namely osteoprotegrin (OPG), bone sialoprotein, β_2-microglobulin, cathepsins, nitric oxide, advanced oxidation protein products (AOPPs), advanced glycation products (AGEs), cytokines as interleukines (mostly IL-1, IL-6 and IL-11), soluble IL-6 receptor, tumor necrosis factor-α (TNF-α), transforming growth factor-β (TGF-β), bone morphogenetic proteins (BMPs) and their soluble receptors, growth factors such as insulin growth factor-I (IGF-I), macrophage colony stimulating factor (MCS-F), and granulocyte-macrophage colony stimulating factor (GMCF-F) [17, 20].

About thirty percent of patients with end stage renal failure exhibit coronary heart disease (USRDS registry 'Annual Data Report 2007 http://www.usrds.org/). At the same time, almost all patients with advanced renal impairment show a multifactorial bone disease [17, 21]. Progression of each of the three entities is strongest when the other two organ systems are malfunctioning. Furthermore, it has recently been shown by a worldwide multicenter trial, that traditional markers and risk factors for cardiovascular disease in the general population such as hypercholesterolemia, arterial hypertension or elevated body mass index exhibit a U-shaped association with cardiac events in patients with end stage renal disease [22, 23].

Thus based on this evidence the review sought to elucidate the current knowledge of molecular markers to uncover and correctly classify the individual risk for this dangerous triad. By identifying subjects at risk, potential prophylactic and/or therapeutic measure might be taken in time before end organ failure is clinically evident.

Data integration and systems biology analyses

Data preparation

Peer reviewed publications (PubMed, http://www.ncbi.nlm.nih.gov/, status as of December 2006) were screened for genes or proteins associated with cardiovascular diseases or bone metabolism disorders in chronic kidney patients. The following keywords were used during the literature search: "biomarker(s)", "risk factor(s)", "chronic kidney disease", "renal disease", "cardiovascular disease", "cardiovascular risk", and "bone metabolism disorder". Thirteen publications from 1999 to 2006 covered a non-redundant set of in total 73 genes associated with either cardiovascular disease

(n=31) or bone metabolism disorders (n=46) in patients with CKD, as summarized in table 1 and 2, respectively. Both serum as well as tissue markers were included without any restrictions in respect to the method of detection.

Table 1: CVD markers in CKD

Gene Symbol	Gene Name	Gene ID	References
ADIPOQ	adiponectin, C1Q and collagen domain containing	9370	Roberts 2006 [11]
AHSG	alpha-2-HS-glycoprotein	197	Roberts 2006 [11]
CCL2	chemokine (C-C motif) ligand	6347	Roberts 2006 [11]
CD40LG	CD40 ligand (TNF superfamily, member 5, hyper-IgM syndrome)	959	Roberts 2006 [11]
CRP	C-reactive protein, pentraxin-related	1401	Roberts 2006 [11]
CST3	cystatin C (amyloid angiopathy and cerebral hemorrhage)	1471	Shlipak 2006 [24]
EDN1	endothelin 1	1906	Dhaun 2006 [25]
FGA	Fibrinogen, alpha chain	2243	Roberts 2006 [11]
FGB	Fibrinogen, beta chain	2244	Roberts 2006 [11]
FGG	Fibrinogen, gamma chain	2266	Roberts 2006 [11]
ICAM1	intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	3383	Roberts 2006 [11]
IL6	Interleukin 6 (interferon, beta 2)	3569	Liu 2006 [26], Honda 2006 [27]
IL8	interleukin 8	3576	Roberts 2006 [11]
LEP	leptin (obesity homolog, mouse)	3952	Mallamaci 2005 [28]
LPA	Lipoprotein, Lp(a)	4018	Pernod 2006 [29]
MTHFR	5,10-methylenetetrahydrofolate reductase (NADPH)	4524	Pernod 2006 [29]
NPPB	natriuretic peptide precursor B	4879	Roberts 2006 [11]
NPY	Neuropeptide Y	4852	Vanholder 2005 [30]
PAPPA	pregnancy-associated plasma protein A, pappalysin 1	5069	Roberts 2006 [11]
PTH	parathyroid hormone	5741	Vanholder 2005 [30]
RLN1	relaxin 1	6013	Roberts 2006 [11]
RLN2	relaxin 2	6019	Roberts 2006 [11]
RLN3	relaxin 3	117579	Roberts 2006 [11]
SAA1	serum amyloid A1	6288	Roberts 2006 [11]
SAA2	serum amyloid A2	6289	Roberts 2006 [11]
SELE	selectin E (endothelial adhesion molecule 1)	6401	Roberts 2006 [11]
SELP	selectin P (granule membrane protein 140kDa, antigen CD62)	6403	Roberts 2006 [11]
TNF	tumor necrosis factor (TNF superfamily, member 2)	7124	Roberts 2006 [11]
TNNI3	troponin I type 3 (cardiac)	7137	Roberts 2006 [11]
TNNT2	troponin T type 2 (cardiac)	7139	Roberts 2006 [11]
VCAM1	vascular cell adhesion molecule 1	7412	Roberts 2006 [11]

Table 2: Bone markers in CKD

Gene Symbol	Gene Name	Gene ID	References
ACP5	acid phosphatase 5, tartrate resistant	54	Schwarz2006 [17]
ALPL	alkaline phosphatase, liver/bone/kidney	249	Rix1999 [19]
B2M	beta-2-microglobulin	567	Schwarz2006 [17]
BGLAP	bone gamma-carboxyglutamate (gla) protein (osteocalcin)	632	Schwarz2006 [17]
BMP1	bone morphogenetic protein 1	649	Urena1999 [20]
BMP10	bone morphogenetic protein 10	27302	Urena1999 [20]
BMP15	bone morphogenetic protein 15	9210	Urena1999 [20]
BMP2	bone morphogenetic protein 2	650	Urena1999 [20]
BMP3	bone morphogenetic protein 3 (osteogenic)	651	Urena1999 [20]
BMP4	bone morphogenetic protein 4	652	Urena1999 [20]
BMP5	bone morphogenetic protein 5	653	Urena1999 [20]
BMP6	bone morphogenetic protein 6	654	Urena1999 [20]
BMP7	bone morphogenetic protein 7 (osteogenic protein 1)	655	Schwarz2006 [17]
BMP8A	bone morphogenetic protein 8a	353500	Urena1999 [20]
BMP8B	bone morphogenetic protein 8b (osteogenic protein 2)	656	Urena1999 [20]
BMPR1A	bone morphogenetic protein receptor, type IA	657	Urena1999 [20]
BMPR1B	bone morphogenetic protein receptor, type IB	658	Urena1999 [20]
BMPR2	bone morphogenetic protein receptor, type II (serine/threonine kinase)	659	Urena1999 [20]
COL1A1	collagen, type I, alpha 1	1277	Schwarz2006 [17]
COL1A2	collagen, type I, alpha 2	1278	Schwarz2006 [17]
CSF1	colony stimulating factor 1 (macrophage)	1435	Urena1999 [20]
CSF2	colony stimulating factor 2 (granulocyte-macrophage)	1437	Urena1999 [20]
CTSL	cathepsin L	1514	Schwarz2006 [17]
FGF23	fibroblast growth factor 23	8074	Fukagawa2006 [16]
FN1	fibronectin 1	2335	Urena1999 [20]
GDF5	growth differentiation factor 5	8200	Reddi2000 [31]
GDF6	growth differentiation factor 6	392255	Reddi2000 [31]
GDF7	growth differentiation factor 7	151449	Reddi2000 [31]
IFNG	interferon, gamma	3458	Urena1999 [20]
IGF1	insulin-like growth factor 1 (somatomedin C)	3479	Schwarz2006 [17]
IL11	interleukin 11	3589	Urena1999 [20]
IL1A	interleukin 1, alpha	3552	Urena1999 [20]
IL1B	interleukin 1, beta	3553	Urena1999 [20]
IL6	interleukin 6 (interferon, beta 2)	3569	Urena1999 [20]
LEP	leptin (obesity homolog, mouse)	3952	Mallamaci 2005 [28]
MEPE	matrix, extracellular phosphoglycoprotein with ASARM motif (bone)	56955	Schwarz2006 [17]
PHEX	phosphate regulating endopeptidase homolog, X-linked (hypophosphatemia, vitamin D resistant rickets)	5251	Schwarz2006 [17]
PLAU	plasminogen activator, urokinase	5328	Urena1999 [20]
PTGES2	prostaglandin E synthase 2	80142	Urena1999 [20]
PTH	parathyroid hormone	5741	Fukagawa2006 [16]
SERPINE1	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	5054	Urena1999 [20]
SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	6678	Urena1999 [20]
SPP1	secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	6696	Urena1999 [20]
TGFB1	transforming growth factor, beta 1 (Camurati-Engelmann disease)	7040	Urena1999 [20]
TNF	tumor necrosis factor (TNF superfamily, member 2)	7124	Urena1999 [20]
TNFRSF11B	tumor necrosis factor receptor superfamily, member 11b (osteoprotegerin)	4982	Schwarz2006 [17]

Four genes are proposed as markers for cardiovascular as well as bone metabolism disorders, namely the parathyroid hormone (PTH), the tumor necrosis factor (TNF), leptin (LEP), as well as interleukin 6 (IL6).

Functional annotation

Functional categories as well as molecular pathways holding a significant number of genes were identified using the Gene Expression Data Analysis Tool of the PANTHER (Protein Analysis THrough Evolutionary Relationships) Classification System [32, 33], and are listed in table 3 and 4. In PANTHER proteins are assigned to families and subfamilies of shared function with two main categories, namely molecular function and biological process. Biological processes and molecular functions of our candidate genes were compared with the PANTHER-internal reference dataset holding all 25,431 currently annotated human genes. A chi-squared test including Bonferroni correction to account for multiple testing was applied to compare the ratio of expected to observed frequency of genes assigned to certain ontology categories. This procedure identifies if certain ontologies are over- or underrepresented on the basis of the given gene lists.

For both diseases, CVD and bone metabolism disorders, genes involved in the category 'signal transduction' were predominant. 20 out of the 46 bone metabolism disorder biomarker candidates and 14 out of the 31 cardiovascular disease marker candidates were assigned to this functional category. The most significantly enriched biological processes in CVD have been identified as immunity and defense (14 genes), blood circulation and gas exchange (5 genes), as well as cell communication (11 genes). Due to the fact that

several bone morphogenetic proteins are in the list of bone metabolism markers, the most significantly enriched biological processes in bone metabolism disorders are skeletal development (15 genes), mesoderm development (17 genes), and developmental processes (20 genes). The complete listing of all significant biological processes, molecular functions and biological pathways of the 77 biomarker candidates is given in tables 3 and 4 for cardiovascular and bone metabolism disorders, respectively.

Table 3: Functional classification of CVD markers

Biological Process	**REFLIST (25431)**	**CVD markers (31)**	**p-value**
Immunity and defense	1318	14	3.58E-09
Blood circulation and gas exchange	89	5	2.56E-06
Cell communication	1213	11	1.46E-05
Signal transduction	3406	14	5.14E-04
Blood clotting	92	4	7.23E-04
Ligand-mediated signaling	421	6	2.10E-03
Cytokine and chemokine mediated signaling pathway	252	5	2.59E-03
Cell proliferation and differentiation	1028	7	6.10E-03
Apoptosis	531	5	1.33E-02
Cell surface receptor mediated signal transduction	1638	9	1.51E-02

Molecular Function	**REFLIST (25431)**	**CVD markers (31)**	**p-value**
Signaling molecule	795	18	3.29E-18
Peptide hormone	102	9	8.05E-13
Extracellular matrix	384	4	3.42E-02
Cytokine	97	3	3.71E-02
Cell adhesion molecule	395	4	3.80E-02

Pathway	REFLIST (25431)	CVD markers (31)	p-value
Plasminogen activating cascade	21	4	1.88E-06
Blood coagulation	55	4	8.61E-05

Table 4: Functional classification of Bone markers

Biological Process	REFLIST (25431)	Bone markers (46)	p-value
Skeletal development	123	15	1.63E-21
Mesoderm development	551	17	7.12E-15
Developmental processes	2152	20	6.98E-09
Cell communication	1213	16	2.61E-08
Cytokine and chemokine mediated signaling pathway	252	9	1.44E-07
Ligand-mediated signaling	421	9	1.17E-05
Cell surface receptor mediated signal transduction	1638	15	1.48E-05
Signal transduction	3406	20	1.76E-05
Other receptor mediated signaling pathway	210	5	7.86E-03
Immunity and defense	1318	9	1.60E-02
Macrophage-mediated immunity	140	4	1.81E-02

Molecular function	REFLIST (25431)	Bone markers (46)	p-value
Signaling molecule	795	26	6.55E-26
Growth factor	125	9	2.52E-10
Cytokine	97	8	1.65E-09
Other signaling molecule	259	7	6.91E-05
Interleukin	34	3	6.95E-03
Extracellular matrix	384	5	1.86E-02

Pathway	REFLIST (25431)	Bone markers (46)	p-value
TGF-beta signaling pathway	154	16	3.58E-22

Protein-protein interaction network analysis

Next to identifying joint functional categories we used human protein-protein interaction data to determine the connectivity of the 77 biomarker candidates on the level of cellular protein networks. Human protein-protein interaction data from OPHID (Online Predicted Human Interaction Database) were used for the analysis (OPHID Version 2007-02-17) [34]. The generation of interaction networks followed the next neighbor expansion method as proposed by Chen *et al.* [35]. OPHID represents protein interactions as protein A interacts with protein B. If A and B are members of the list of 77 candidates a positive interaction is identified. The next neighbor expansion includes also interactions of the type A-X-B, where X represents a protein not given in the initial candidate list. All interacting partners of the initial set of 77 proteins were extracted from the OPHID database and the protein interaction network was generated. At least one interacting partner was found for 29 of the 31 CVD, and for 38 of the 46 bone metabolism markers. The resulting graph, composed of one large sub-graph and a number of smaller, disconnected sub-graphs, consisted of 353 protein nodes and 440 protein interaction edges, as depicted in figure 1a.

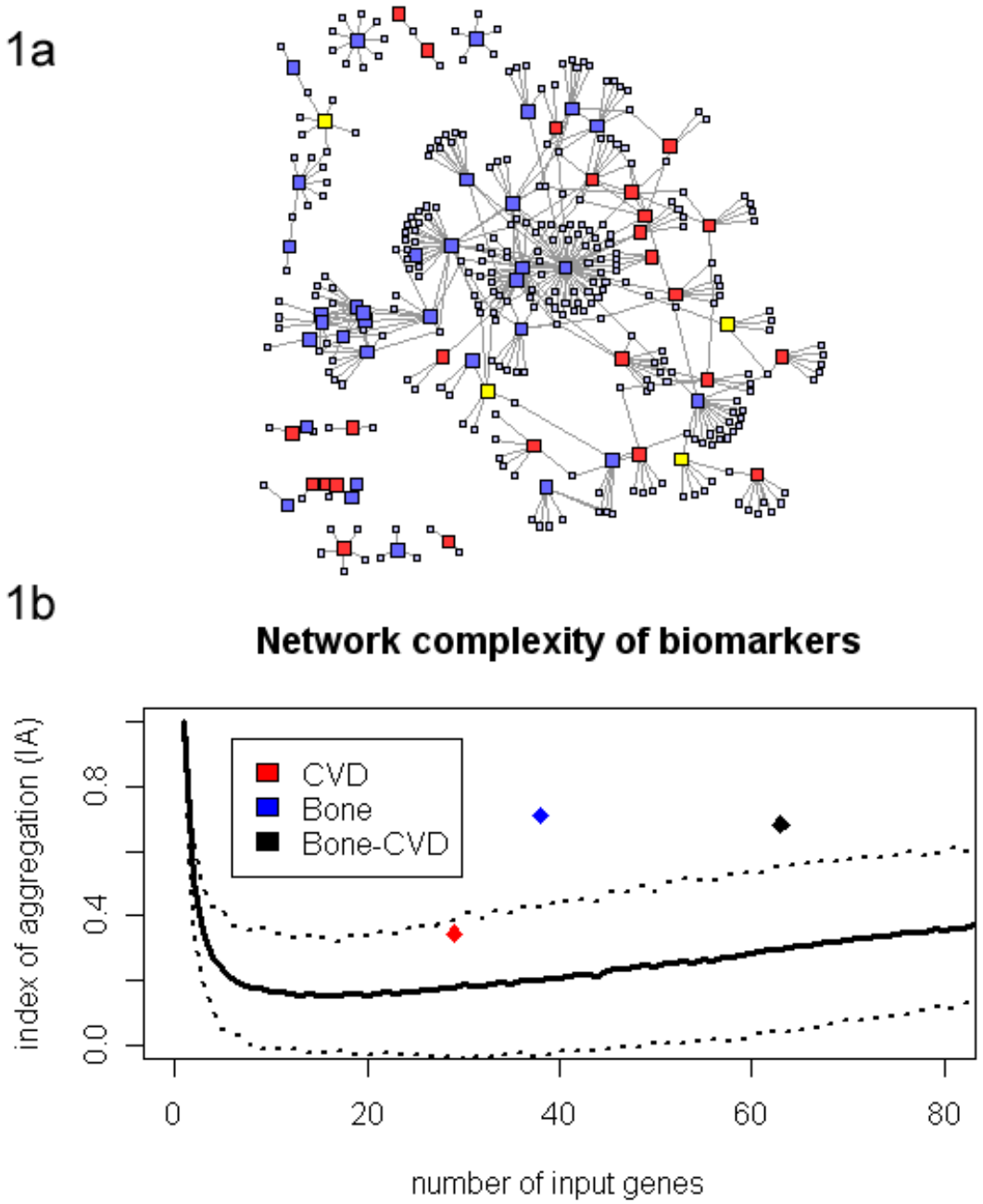


Figure 1: (a) Protein interaction networks after next neighbor expansion as given by the candidate biomarker lists of 77 proteins: Protein nodes given in blue depict bone markers, protein nodes given in red denote markers of cardiovascular disease, and protein nodes given in yellow represent proteins reported in both diseases. (b) Network complexity of protein interaction networks: Given is the index of aggregation (IA; y-axis) in relation to

the number of proteins used for constructing protein interaction networks (x-axis). The IA of protein interaction networks derived on the basis of randomly generated protein lists is given as reference (solid line, dashed lines gives the standard deviation). The IA for networks based on the list of cardiovascular marker candidates alone does not exceed values also derived for randomly generated protein lists. The IA for networks derived for the given bone metabolism disorder markers, but in particular the combined markers significantly exceed reference values as found for randomly generated lists with the same number of proteins involved.

The Index of Aggregation (IA) serves as aggregation and complexity measure of interaction networks for evaluating if the interaction characteristics differ with respect to networks derived on the basis of random protein lists. This measure therefore gives an indication if the connectivity for a given protein list is higher than statistically expected. The IA is given as percentage of protein nodes in the largest sub-graph with respect to all protein nodes in the network including all sub-graphs. The IA of the biomarker candidates' network given in figure 1 was compared to respective values of randomly generated protein lists. 43 of the 63 proteins which actually have interaction entries in the OPHID database were connected in a single sub-graph when including next neighbor expansion. The resulting index of aggregation of 0.68 for the combined list of potential biomarkers (CVD and Bone) is more than 2 standard deviations above the expected IA for randomly generated networks of equivalent size. Figure 1b shows the IA of the given biomarker lists in comparison to the distribution of the IA for randomly generated protein lists. Genes associated with cardiovascular diseases and those associated with bone metabolism disorders are highly interlinked on the level of protein-protein interactions.

Both, functional categories as well as protein interactions indicate the interrelation of biomarker candidates for CVD and bone metabolism disorders.

Integrated analysis

For further characterizing the interrelation between all 353 members of the interaction network represented by the largest sub-graph as given in figure 1 we extracted the following information for each single gene: The gene expression profile as found in chronic kidney disease biopsy material published by Rudnicki *et al*. [36], as well as gene ontology terms on molecular process and function as provided by the gene ontology consortium [37, 38]. Additionally, we computed the transcription factor binding site profiles for each of the genes following *in-silico* predictions as provided by the oPOSSUM tool [39, 40]. This procedure provides a list of transcription factors for each gene which appear to be involved in its differential regulation. Genes sharing transcription factors might be under similar expression control.

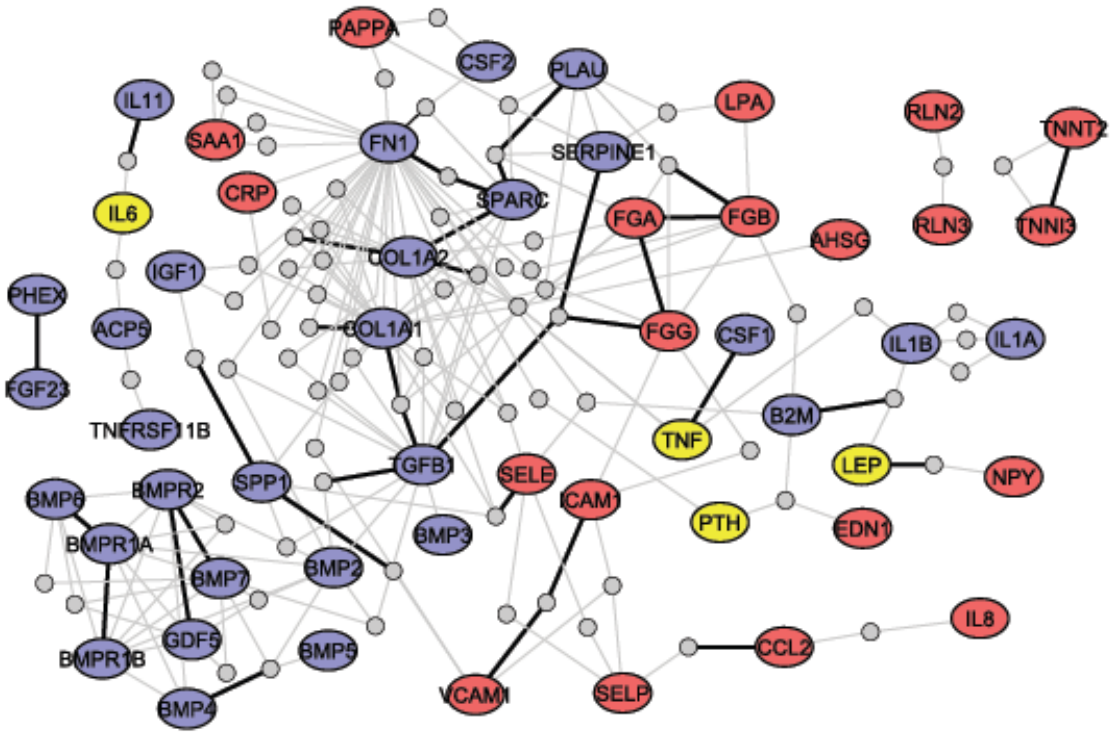
After assembling this set of properties for each of the 353 genes we computed pair-wise correlations including the parameters gene expression, functional category, and transcription factor modules. Rationale of this approach is the assumption that genes showing similarities on the level of these features might exhibit an increased likelihood for functional dependency in the context of cellular processes.

For characterizing co-expression of two genes we used the Pearson correlation coefficient. Two genes exhibiting a high correlation coefficient of their expression profile are co-expressed on the level of differential gene expression. For expressing the pair-wise similarity of two genes based on their gene ontology classification patterns the Dice

coefficient for bit-strings was calculated. This string comparison measure determines the ratio of joint annotation within given categories and the total number of annotations in categories. High values of the Dice coefficient found for a given biomarker candidate pair indicate similarity on the level of functional categorization. The same measure was used for identifying the ratio of joint transcription factors indicating co-regulation between two genes. A meta-correlation based on the three single parameters was finally calculated for expressing functional dependency between elements of our biomarker candidate list.

Applying this procedure provides correlation values for each interaction pair of the interaction graph given in figure 1. For subsequent analysis we focused on 'strong' pair-wise interactions, defined as meta-correlation values which were found as at least one standard deviation above the mean value of all meta-correlation values for all pairs analyzed. Figure 2 identifies these strong interactions as thick interaction lines.

Figure 2: Detailed representation of the largest sub-graphs of the protein interaction network derived on the basis of 77 candidate markers. Only nearest neighbors having two edges to the CVD and bone marker candidates are shown. Protein nodes given in blue depict markers for bone metabolism disorders, protein nodes given in red hold markers of cardiovascular disease, and nodes given in yellow represent proteins reported in both diseases.



Based on the protein interaction networks, and following the dependency measure expressed by the meta-correlation we identified two sub-networks, each connecting at least four of the reported biomarkers given in the initial list of 77 candidate biomarkers. The first sub-network holds the proteins fibronectin 1 (FN1), collagen, type I, alpha 2 (COL1A2), the plasminogen activator, urokinase (PLAU), and osteonectin (the secreted protein, acidic, cysteine-rich; SPARC). They were all reported to play a role in bone mineral disorders of patients with chronic kidney disease. FN1 is involved in various processes like cell adhesion and blood clotting, and has also been proposed as risk factor for arterial thrombosis [41]. SPARC regulates cell interactions with components of the extracellular matrix and is often found at sites of injury [42]. COL1A2 is mostly found in connective tissues and mutations in this gene regions were reported to lead to a variety of bone metabolism disorders including idiopathic osteoporosis, ostoegenesis imperfecta, or

the Ehlers-Danlos syndrome [43, 44]. Besides its function in hemostasis PLAU is also involved in cell attachment and deformation of the extracellular matrix [45].

Members of the second sub-network are collagen, type I, alpha 1 (COL1A1), the transforming growth factor beta 1 (TGFB1), the plasminogen activator inhibitor, also known as serpin peptidase inhibitor, clade E (SERPINE1), and the alpha, beta, and gamma chains of fibrinogen (FGA, FGB, FGG). COL1A1 is like COL1A2 found in most connective tissues. TGFB1 is a multifunctional protein involved in proliferation, differentiation, apoptotic processes, cell adhesion, and tissue remodeling [46]. SERPINE1 plasma concentrations are elevated in patients with increased risk of ischemic cardiovascular events [47]. All three chains of fibrinogen are part of the network. After cleavage by thrombin, fibrin fibers form blood clots after vascular injury.

Conclusion and Outlook

We provide an interactome analysis approach to characterize the interplay of reported biomarker candidates for cardiovascular diseases and bone metabolism disorders in chronic kidney disease patients. 46 potential biomarkers for bone metabolism disorders and 31 potential biomarkers for cardiovascular disease were identified in the literature and characterized with respect to biological function, gene expression in chronic kidney disease, and known protein-protein interactions.

A majority of marker candidates for cardiovascular diseases could be assigned to the functional category ‘immunity and defense’, whereas most of the bone metabolism genes were involved in skeletal and mesoderm development according to the PANTHER

classification scheme. A category significantly enriched in both diseases was ‘signal transduction’ with various secreted signaling molecules being proposed as potential biomarkers. On the level of protein-protein interactions proteins involved in bone metabolism disorders were highly interlinked. The resulting Index of Aggregation was significantly higher than one would expect from randomly drawn gene lists. Biomarker candidates of cardiovascular diseases were also closer connected as randomly generated gene lists although the statistical significance was not reached. The combined list of marker candidates from both diseases on the other hand was highly significant with around 68% of biomarkers forming the largest sub-graph of the overall protein-protein interaction network. Functional links of biomarkers proposed for CVD and bone metabolism disorders appears evident at least on this given level of data interpretation.

Of special note are the four potential biomarkers reported in both diseases, namely IL6, PTH, LEP, and TNF, as well as the three components of fibrinogen (FGA, FGB, and FGG) building a major link between the two diseases as indicated by strong interactions based on the meta-correlation as depicted in figure 2. Although causal inference cannot be drawn form our data, the coincidence of features in both disease entities may potentially suggest choreographed action via a common pathway.

Integration of data from various sources for characterizing diseases has the potential to unravel novel pathophysiological mechanisms. As more and more tools become available for predicting protein-protein interactions based on protein domain information, the in silico validation of given protein candidates, but also identification of novel proteins playing a role in a given disease will become feasible [48, 49]. This development allows

the analysis of the functional interplay between biomarker candidates, clearly providing routes towards identifying improved candidate markers.

Acknowledgements

Sources of support: Austrian Science Fund (P-18325-B13 to R.O. and P-18130-B13 to B.M.). We thank Gert Mayer, Arno Lukas and Alexander Kainz for their fruitful comments during the preparation of this manuscript.

Authors' contributions

The literature searches for CVD and bone metabolism biomarkers in CKD patients were performed by Julia Wilflingseder, Barbara Wimmer and Paul Perco. Bernd Mayer and Paul Perco designed the concept of the integrated analysis schema which was applied to the dataset by Martin Wiesinger and Andreas Bernthaler. Rainer Oberbauer and Michael Rudnicki made significant changes to the sections of the manuscript addressing the clinical background of the three diseases. All authors contributed to data interpretation and writing of the manuscript.

References

1. **Raggi P, Giachelli C, Bellasi A**. Interaction of vascular and bone disease in patients with normal renal function and patients undergoing dialysis. *Nat Clin Pract Cardiovasc Med*. 2007; 4:26-33.

2. **Moe SM**. Vascular calcification and renal osteodystrophy relationship in chronic kidney disease. *Eur J Clin Invest*. 2006; 36:51-62.

3. **Andress DL**. Vitamin D in chronic kidney disease: a systemic role for selective vitamin D receptor activation. *Kidney Int*. 2006; 69:33-43.

4. **Perco P, Pleban C, Kainz A, Lukas A, Mayer G, Mayer B, Oberbauer R**. Protein biomarkers associated with acute renal failure and chronic kidney disease. *Eur J Clin Invest*. 2006; 36:753-63.

5. **Dove A**. Screening for content--the evolution of high throughput. *Nat Biotechnol*. 2003; 21:859-64.

6. **Perco P, Rapberger R, Siehs C, Lukas A, Oberbauer R, Mayer G, Mayer B**. Transforming omics data into context: bioinformatics on genomics and proteomics raw data. *Electrophoresis*. 2006; 27:2659-75.

7. **Mondry A, Wang Z, Dhar PK**. Bone and the kidney: a systems biology approach to the molecular mechanisms of renal osteodystrophy. *Curr Mol Med*. 2005; 5:489-96.

8. **Drake TA, Ping P**. Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Proteomics approaches to the systems biology of cardiovascular diseases. *J Lipid Res*. 2007; 48:1-8.

9. **Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY**. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N Engl J Med*. 2004; 351:1296-305.

10. **Sarnak MJ, Levey AS, Schoolwerth AC, Coresh J, Culleton B, Hamm LL, McCullough PA, Kasiske BL, Kelepouris E, Klag MJ, Parfrey P, Pfeffer M, Raij L, Spinosa DJ, Wilson PW**. Kidney disease as a risk factor for development of cardiovascular disease: a statement from the American Heart Association Councils on Kidney in Cardiovascular Disease, High Blood Pressure Research, Clinical Cardiology, and Epidemiology and Prevention. *Circulation*. 2003; 108:2154-69.

11. **Roberts MA, Hare DL, Ratnaike S, Ierino FL**. Cardiovascular biomarkers in CKD: pathophysiology and implications for clinical management of cardiac disease. *Am J Kidney Dis*. 2006; 48:341-60.

12. **Baigent C, Landray M**. Which cardiovascular risk factors matter in chronic kidney disease? *Nephrol Dial Transplant*. 2007; 22:9-11.

13. **Levin A, Bakris GL, Molitch M, Smulders M, Tian J, Williams LA, Andress DL**. Prevalence of abnormal serum vitamin D, PTH, calcium, and phosphorus in patients with chronic kidney disease: results of the study to evaluate early kidney disease. *Kidney Int*. 2007; 71:31-8.

14. **Gutierrez O, Isakova T, Rhee E, Shah A, Holmes J, Collerone G, Juppner H, Wolf M**. Fibroblast growth factor-23 mitigates hyperphosphatemia but accentuates calcitriol deficiency in chronic kidney disease. *J Am Soc Nephrol*. 2005; 16:2205-15.

15. **LaClair RE, Hellman RN, Karp SL, Kraus M, Ofner S, Li Q, Graves KL, Moe SM**. Prevalence of calcidiol deficiency in CKD: a cross-sectional study across latitudes in the United States. *Am J Kidney Dis*. 2005; 45:1026-33.

16. **Fukagawa M, Hamada Y, Nakanishi S, Tanaka M**. The kidney and bone metabolism: Nephrologists' point of view. *J Bone Miner Metab*. 2006; 24:434-8.

17. **Schwarz C, Sulzbacher I, Oberbauer R**. Diagnosis of renal osteodystrophy. *Eur J Clin Invest*. 2006; 36:13-22.

18. **Lopez I, Mendoza FJ, Aguilera-Tejero E, Perez J, Guerrero F, Martin D, Rodriguez M**. The effect of calcitriol, paricalcitol, and a calcimimetic on extraosseous calcifications in uremic rats. *Kidney Int*. 2008; 73:300-7.

19. **Rix M, Andreassen H, Eskildsen P, Langdahl B, Olgaard K**. Bone mineral density and biochemical markers of bone turnover in patients with predialysis chronic renal failure. *Kidney Int*. 1999; 56:1084-93.

20. **Urena P, De Vernejoul MC**. Circulating biochemical markers of bone remodeling in uremic patients. *Kidney Int*. 1999; 55:2141-56.

21. **Ferreira A**. Development of renal bone disease. *Eur J Clin Invest*. 2006; 36:2-12.

22. **Kalantar-Zadeh K, Block G, Humphreys MH, Kopple JD**. Reverse epidemiology of cardiovascular risk factors in maintenance dialysis patients. *Kidney Int*. 2003; 63:793-808.

23. **Goodkin DA, Bragg-Gresham JL, Koenig KG, Wolfe RA, Akiba T, Andreucci VE, Saito A, Rayner HC, Kurokawa K, Port FK, Held PJ, Young EW**. Association of comorbid conditions and mortality in hemodialysis patients in Europe, Japan, and the United States: the Dialysis Outcomes and Practice Patterns Study (DOPPS). *J Am Soc Nephrol*. 2003; 14:3270-7.

24. **Shlipak MG, Katz R, Sarnak MJ, Fried LF, Newman AB, Stehman-Breen C, Seliger SL, Kestenbaum B, Psaty B, Tracy RP, Siscovick DS**. Cystatin C and prognosis for cardiovascular and kidney outcomes in elderly persons without chronic kidney disease. *Ann Intern Med*. 2006; 145:237-46.

25. **Dhaun N, Goddard J, Webb DJ**. The endothelin system and its antagonism in chronic kidney disease. *J Am Soc Nephrol*. 2006; 17:943-55.

26. **Liu Y, Berthier-Schaad Y, Fallin MD, Fink NE, Tracy RP, Klag MJ, Smith MW, Coresh J**. IL-6 haplotypes, inflammation, and risk for cardiovascular disease in a multiethnic dialysis cohort. *J Am Soc Nephrol*. 2006; 17:863-70.

27. **Honda H, Qureshi AR, Heimburger O, Barany P, Wang K, Pecoits-Filho R, Stenvinkel P, Lindholm B**. Serum albumin, C-reactive protein, interleukin 6, and

fetuin a as predictors of malnutrition, cardiovascular disease, and mortality in patients with ESRD. *Am J Kidney Dis*. 2006; 47:139-48.

28. **Mallamaci F, Tripepi G, Zoccali C**. Leptin in end stage renal disease (ESRD): a link between fat mass, bone and the cardiovascular system. *J Nephrol*. 2005; 18:464-8.

29. **Pernod G, Bosson JL, Golshayan D, Barro C, Forneris G, Martina G, Bonfant G, Hurot JM, Turc-Baron C, Jouet C, Theytaz J, Jeantet A, Wauters JP, Cordonnier D**. Phenotypic and genotypic risk factors for cardiovascular events in an incident dialysis cohort. *Kidney Int*. 2006; 69:1424-30.

30. **Vanholder R, Massy Z, Argiles A, Spasovski G, Verbeke F, Lameire N**. Chronic kidney disease as cause of cardiovascular morbidity and mortality. *Nephrol Dial Transplant*. 2005; 20:1048-56.

31. **Reddi AH**. Bone morphogenetic proteins and skeletal development: the kidney-bone connection. *Pediatr Nephrol*. 2000; 14:598-601.

32. **Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, Vandergriff JA, Doremieux O**. PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res*. 2003; 31:334-41.

33. **Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, Rabkin S, Guo N, Muruganujan A, Doremieux O, Campbell MJ, Kitano H, Thomas PD**. The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res*. 2005; 33:D284-8.

34. **Brown KR, Jurisica I**. Online predicted human interaction database. *Bioinformatics*. 2005; 21:2076-82.

35. **Chen JY, Shen C, Sivachenko AY**. Mining alzheimer disease relevant proteins from integrated protein interactome data. *Pac Symp Biocomput*. 2006; 367-78.

36. **Rudnicki M, Eder S, Perco P, Enrich J, Scheiber K, Koppelstatter C, Schratzberger G, Mayer B, Oberbauer R, Meyer TW, Mayer G**. Gene expression profiles of human proximal tubular epithelial cells in proteinuric nephropathies. *Kidney Int*. 2007; 71:325-35.

37. **Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G**. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000; 25:25-9.

38. **Gene-Ontology-Consortium**. Creating the gene ontology resource: design and implementation. *Genome Res*. 2001; 11:1425-33.

39. **Ho Sui SJ, Mortimer JR, Arenillas DJ, Brumm J, Walsh CJ, Kennedy BP, Wasserman WW**. oPOSSUM: identification of over-represented transcription factor binding sites in co-expressed genes. *Nucleic Acids Res*. 2005; 33:3154-64.

40. **Huang SS, Fulton D, Arenillas DJ, Perco P, Ho Sui SJ, Mortimer JR, Wasserman WW**. Identification of over-represented combinations of transcription factor binding sites in sets of co-expressed genes. *Advances in Bioinformatics and Computational Biology*. 2006; 3:247-56.

41. **Mosher DF**. Plasma fibronectin concentration: a risk factor for arterial thrombosis? *Arterioscler Thromb Vasc Biol*. 2006; 26:1193-5.

42. **Bradshaw AD, Sage EH**. SPARC, a matricellular protein that functions in cellular differentiation and tissue response to injury. *J Clin Invest*. 2001; 107:1049-54.

43. **Forlino A, Marini JC**. Osteogenesis imperfecta: prospects for molecular therapeutics. *Mol Genet Metab*. 2000; 71:225-32.

44. **Gajko-Galicka A**. Mutations in type I collagen genes resulting in osteogenesis imperfecta in humans. *Acta Biochim Pol*. 2002; 49:433-41.

45. **Czekay RP, Loskutoff DJ**. Unexpected role of plasminogen activator inhibitor 1 in cell adhesion and detachment. *Exp Biol Med (Maywood)*. 2004; 229:1090-6.

46. **Park SH**. Fine tuning and cross-talking of TGF-beta signal by inhibitory Smads. *J Biochem Mol Biol*. 2005; 38:9-16.

47. **Vaughan DE**. PAI-1 and atherothrombosis. *J Thromb Haemost*. 2005; 3:1879-83.

48. **Martin S, Roe D, Faulon JL**. Predicting protein-protein interactions using signature products. *Bioinformatics*. 2005; 21:218-26.

49. **Singhal M, Resat H**. A domain-based approach to predict protein-protein interactions. *BMC Bioinformatics*. 2007; 8:199.