We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



11

Functional Context Network of T2DM

Anja Thormann and Axel Rasche
Max-Planck-Institute for Molecular Genetics, Department of Vertebrate Genomics
Germany

1. Introduction

Type-2 diabetes mellitus (T2DM) is a complex disease with multiple causes covering several functional entities of the metabolism. Environmental factors contribute to the pathogenesis of the disease – most notably nutrition and weight of the organism. The identification of disease genes is the driving power of many research projects. In a previous paper (Rasche et al. 2008) we presented a method that integrates results from different T2DM related studies and identifies candidate genes with high disease relevance. This chapter is designated to elaborate on our work from a network based perspective. Network biology is a promising field that can shed light on interrelations between disease genes and from disease genes to their functional neighborhood. We use network-based tools to advance from a single-gene analysis towards a subnet, a functional module, of disease genes.

Proteins are gene products that are associated with particular molecular functions. Molecular functions are interpreted as activities that can be performed by individual proteins following the definitions introduced by the Gene Ontology Consortium (Ashburner et al. 2000). Examples of molecular functions are catalytic activity, transporter activity or binding. Additionally, a biological process is accomplished by one or more ordered assemblies of molecular functions (Ashburner et al. 2000).

Proteins physically interact with each other in order to carry out a biological function. A biological function is related to the term *biological process*. A signal transduction cascade whose biological function is to transmit information from a receptor to a transcription factor is a succession of protein-protein interactions (PPIs). Both the molecular function of a protein and the biological function in which it is involved are best deduced by studying the environment where it operates in.

To this end, scientists pursue the ambitious goal of assembling all PPIs in an organism – the *interactome* – to elucidate how proteins work together and promote individual biological processes and eventually the complete cellular machinery. Today, mainly two methods are used to detect PPIs: Yeast two-hybrid screens (Fields & Sternglanz 1994) and affinity purification (Pandey & Mann 2000). These large-scale technologies provide vast numbers of interactions but have high false positive rates. Additionally, such experiments only reflect one environmental condition and not the dynamics of interactions between different phyiological states leading to high false negative rates.

Regarding the current size of the human interactome, we have only a draft of the complete set of interactions. However, looking at the course of construction (fig. 1) so far and bearing in mind new quality standards we are continuously moving towards the completion of a

comprehensive human PPI network. For now we have to take into account that the network is incomplete and noisy.

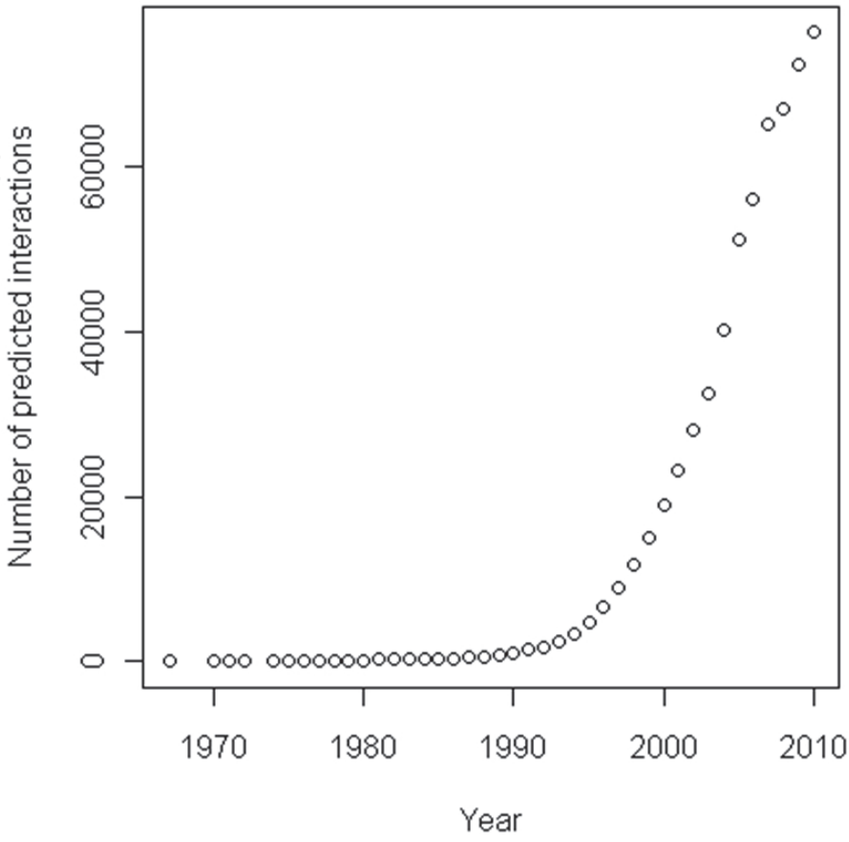


Fig. 1. Cumulative number of detected PPIs of the last years. The data is taken from the ConsensusPathDB website. The data may contain false positive interactions.

Interactions are consolidated in many different databases. For further analysis we take advantage of ConsensusPathDB (Kamburov et al. 2009; Kamburov et al. 2011), a resource joining various human molecular interaction networks including protein-protein, metabolic, signaling and gene regulatory interaction networks. ConsensusPathDB integrates interaction data from many interaction databases, consequently providing us with a comprehensive resource of the currently known interactome.

2. Meta-analysis

T2DM is a polygenic disease subject approached by diverse studies using a variety of experimental methods to dissect the molecular basis of T2DM. In Rasche et al. (2008) we conducted a meta-analysis approach merging different heterogeneous data sources for the identification of disease candidate genes. The analysis included transcriptome studies from multiple tissues in mouse and human, genetic information using knock-out mice, text mining as well as signaling protein data.

We computed scores for all genes in each individual study and summarized the scores across the different studies. Thus a basic disease relevance score was established.

Comparing the aggregated scores against a bootstrap background sample defined a cut-off score. Using this threshold, a list of 213 candidate genes was identified. The set of candidate genes was related to different T2DM gene predictions, monogenic mouse models for T2DM and major association studies with considerable overlap. These overlaps showed clearly that gene lists can be generated relying on a single aspect or technology but our meta-analysis rather encompasses a broad range of biomolecular aspects of T2DM. Functional enrichment analyses for KEGG pathways revealed a tight connection with diabetes-specific pathways. However, some genes exhibit a higher interconnection and contribute to an extensive crosstalk between *Insulin signaling*, *Type II diabetes mellitus* and *PPAR signaling*. Several candidate genes in particular are hubs in the protein interaction networks with many interactions and linking several of the pathways.

With the set of candidate genes we identified biological networks on different layers of cellular information: Signaling and metabolic pathways, gene regulatory networks and protein-protein interaction networks. However, we only provided parts of different networks as separated results. In this study the 213 candidate genes and their respective gene scores are used to identify a subnetwork of the human interactome provided over several functional levels by the ConsensusPathDB.

3. PPI networks

From a mathematical point of view proteins can be described as nodes (vertices) and interactions can be described as undirected links (edges) between interacting proteins. This abstraction allows us to characterize PPI networks by mathematical means. It helps to uncover underlying organizing principles of biological networks, describing the role of proteins in terms of topological parameters. Although computational methods are impaired by incomplete data sets they could be used to point out crucial proteins and structures.

Local topological properties characterize single proteins in a PPI network and may be averaged over all proteins. We give short definitions for the most common topological properties. More detailed descriptions can be found on the website introducing the Network Analyzer plug-in (Assenov et al. 2008). The defined topological parameters are computed in the Cytoscape (Cline et al. 2007) environment using the Network Analyzer plug-in and summary distributions are visualized in fig. 2 and 3.

Degree: The node degree of a node n is equal to the number of nodes that interact with node n.

Neighborhood connectivity: The connectivity of a node n is equal to its node degree. The neighborhood connectivity of a node n is defined as the average node degree of all neighbors of n.

Clustering coefficient: The clustering coefficient is a ratio between the number of edges between the neighbors of n, and the maximum number of edges that could possibly exist between the neighbors of n.

Betweenness centrality: The betweenness centrality of a node n equals the fraction of shortest paths (excluding paths starting or finishing in n) in a network that pass through the node n. A shortest path between two nodes corresponds to the minimal number of edges that has to be traversed in the graph to get from one node to the other.

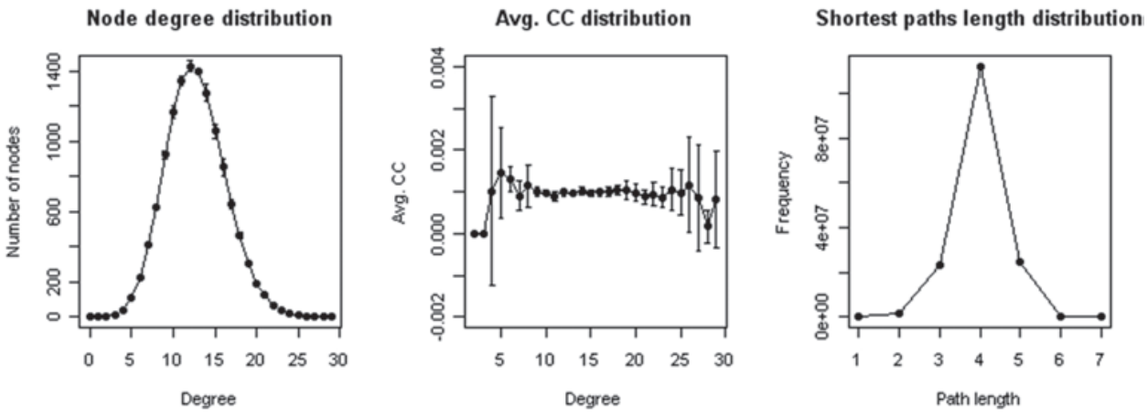


Fig. 2. Topological parameters are computed for five random PPI networks. Initial number of nodes is 12733. The probability for a node being part of the network is 0.01. The computation was done with igraph (Csardi & Nepusz 2006). Abbrv.: CC, clustering coefficient.

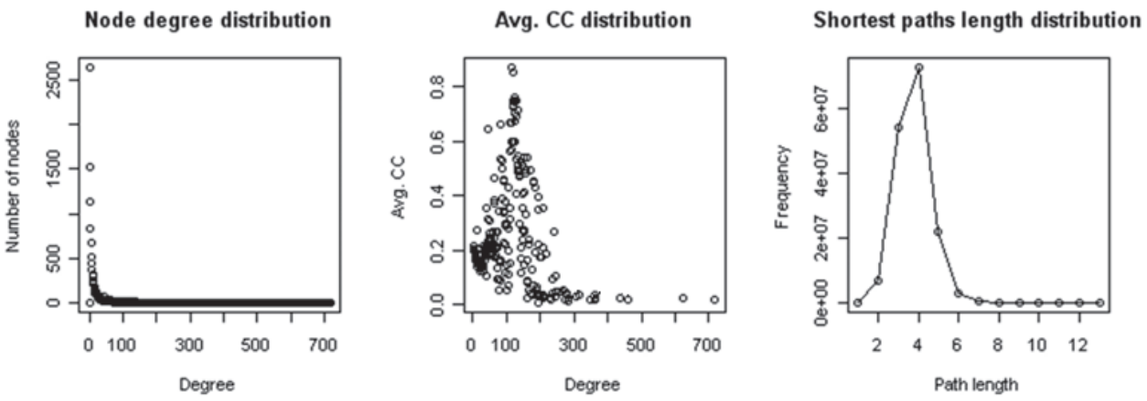


Fig. 3. Topological parameters for the ConsensusPathDB PPI network with 12733 nodes and 101613 undirected interactions

Global network properties emerge from the sum of all local topological properties and follow well-defined organizing principles (Barabási & Oltvai 2004):
Degree distribution: The degree distribution returns the probability that a randomly selected node is connected to *k* other nodes.
Average clustering coefficient distribution: The average clustering coefficient distribution returns the average over the clustering coefficients of all nodes with the same node degree *k*.
Shortest paths distribution: Considering all possible shortest paths in the network, the shortest paths distribution gives for each attained shortest path length the number of node pairs having such a path length.
These graph-theoretical criteria are important to show that biological networks are not comparable with random graphs following the well established Erdős–Rényi model (Erdős & Rényi 1960) since it does not sufficiently capture the wiring principles of PPI networks. In random graphs most nodes have approximately the same number of neighbors. In PPI networks there are only a few highly connected nodes called hubs. Most nodes only have a

few neighbors. This property is described by scale-free networks (Barabasi & Albert 1999) whose node degree distribution follows a power-law. Additionally, PPI networks have properties of "small-world" networks (Watts & Strogatz 1998): PPI networks exhibit a high degree of clustering and small path lengths between nodes. Modularity, a high degree of clustering and a degree distribution following a power law account for a hierarchical organization of the PPI network (Ravasz & Barabási 2003).

We build a PPI network from the set of PPIs in the ConsensusPathDB. We map genes to their respective protein identifiers and draw the parameter distributions for all candidate genes as well as for the total set of genes which are part of the PPI network (control). We want to quantify to which extent candidate genes separate from the whole network. Following Xu & Li (2006) we computed:

1N index: The 1N index is the ratio between the number of interactions with candidate genes and the number of all interactions for a given node *n*.

2N index: The 2N index is the average over all 1N indexes for interaction partners of node *n*.

Average distance to candidate genes: The average distance to candidate genes is the average over the shortest paths from a given node *n* to all candidate genes.

Positive topological coefficient: The positive topological coefficient is the average over the number of shared neighbors with any candidate genes.

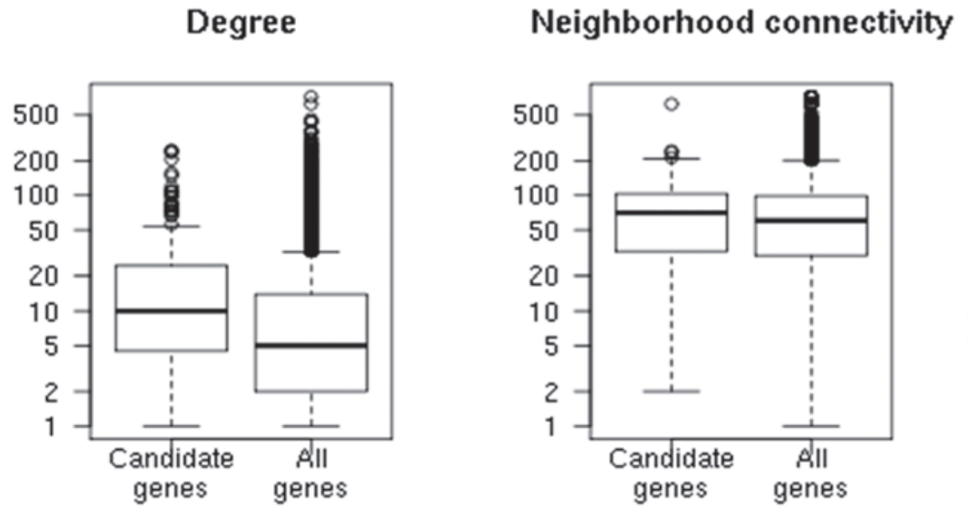


Fig. 4. Degree distributions and neighborhood connectivity distributions for candidate genes and all genes displayed on a log scale.

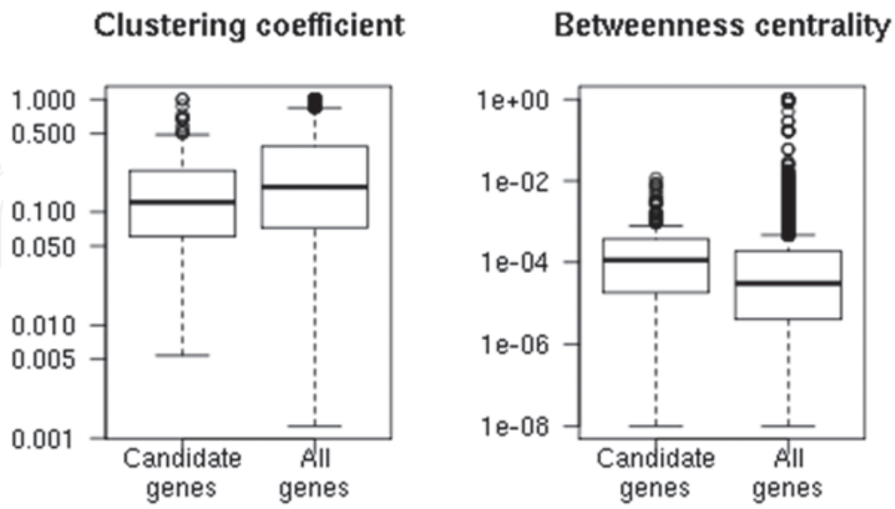


Fig. 5. Clustering coefficient distributions and betweenness centrality distributions for candidate genes and all genes displayed on a log scale.

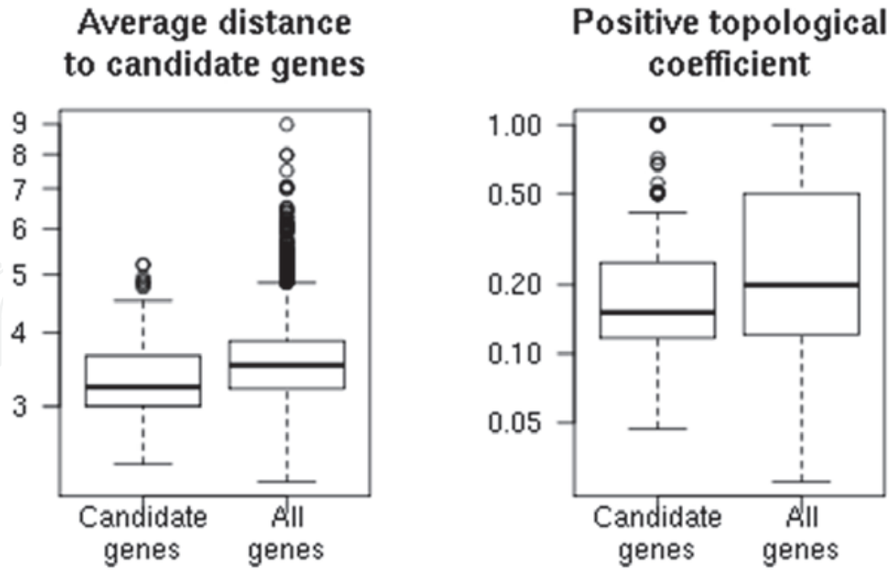


Fig. 6. Average distance to candidate genes distributions and positive topological coefficient distributions for candidate genes and all genes displayed on a log scale.

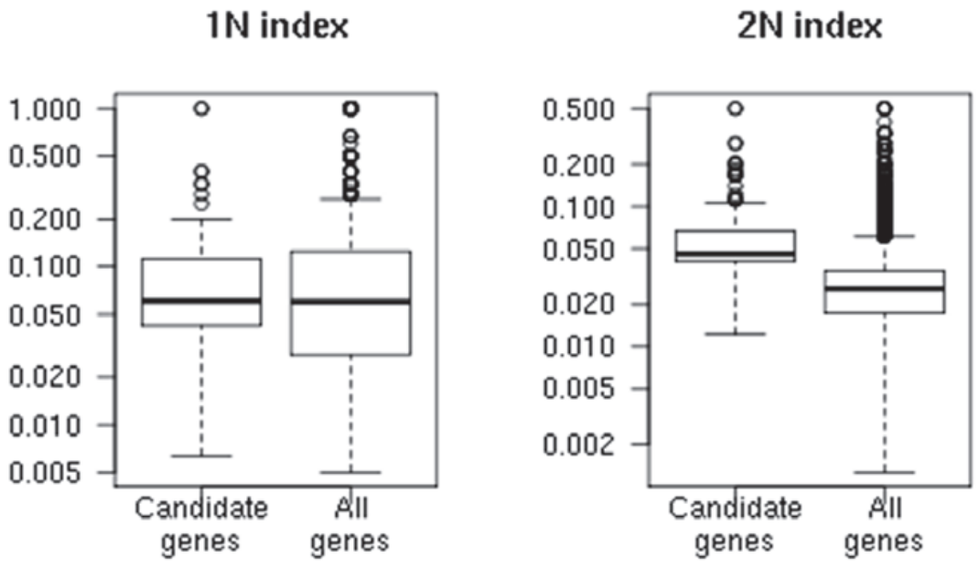


Fig. 7. 1N index distributions and 2N index distributions for candidate genes and all genes displayed on a log scale.

Distributions of the parameters are on display in fig. 4-7. In order to assess the significance of the distributions difference in the means parameter distributions for candidate genes and control we use the Wilcoxon rank sum test; resulting *p*-values are listed in table 1. For the degree, betweenness centrality, 2N index and average distance to candidate genes a significant deviation from the complete PPI network is ascertained.

Parameter	**H0**	**H1**	***p*-value**
Degree	A = B	A > B	2.038e-10
Neighborhood Connectivity	A = B	A > B	0.1103
Clustering coefficient	A = B	A > B	1
Betweenness Centrality	A = B	A > B	2.563e-08
1N index	A = B	A > B	0.1439
2N index	A = B	A > B	< 2.2e-16
Positive topological coefficient	A = B	A > B	0.9999
Average distance to candidate genes	A = B	A < B	3.038e-12

Table 1. Results for Wilcoxon rank sum test: A – distribution for candidate genes, B – distribution for all genes, H0 – null hypothesis, H1 – alternative hypothesis.

4. Functional modules

Proteins that form a local neighbourhood, topological module, and share a biological function can be summarized in a functional module. Following the interpretation that a disease results

as a consequence of a disrupted or disturbed functional module, such a module represents the fingerprint of a disease – the disease module (Barabási et al. 2011). The close relationship between topology, functionality and disease relevance demands for algorithms which can decompose the PPI network into distinct subnetworks. We want to identify a subnetwork (module) with high disease relevance.

As interaction data encodes only topological information we need to incorporate biological data which provides information on genes that are for example differentially expressed in the course of a disease and points to irregularities in biological function. Additionally, expression data provides temporal and spatial information. With the set of measured genes or proteins we build a node induced network containing the measured proteins and their interactions.

Finding a subnetwork of high disease relevance was first addressed by Ideker et al. (2002). The solution to the raised problem involves the following two steps: First, nodes in the network are weighted according to some criteria, usually according to their degree of differential expression. Highly differentially expressed nodes are assigned a positive value. Remaining nodes are assigned a negative value. Second, a maximally scoring network is computed.

Mathematically this is equivalent to finding a maximum-weight connected subnetwork. If the graph contains positive and negative weighted nodes finding such MWCS is an NP-hard problem (cannot be computed efficiently) (Ideker et al. 2002). NP-hard problems are often solved with heuristic algorithms. However, heuristic methods cannot guarantee optimal solutions and are highly sensitive to parameter settings. A review over the progress in computational methods for finding functional modules is given by Wu et al. (2009). A major progress was introduced with an algorithm (Dittrich et al. 2008) that computes exact solutions for the MWCS problem in reasonable time. They reformulate the MWCS problem and solve it with techniques from linear programming. Beforehand, a scoring function allows to aggregate *p*-values from several studies. The *p*-value distribution is decomposed into signal and noise modeled by different distributions. A likelihood ratio test computes positive values for highly differentially expressed genes and negative values for moderately or not differentially expressed genes belonging to the background noise. The score functions are provided as an R package (Beisser et al. 2010).

Here, we deviate from the presented approach. Our main objective is to present a functional module whose computation is based on the knowledge from the meta-analysis. Therefore we consider the complete PPI network and assign all candidate genes its scores from the meta-analysis. Non-candidate genes are assigned a negative value. With the algorithm introduced by Dittrich et al. (2008) we compute a functional module. The method reduces the complexity of large networks to biologically relevant modules of interpretable size. Induced by the weighted candidate genes we compute a functional module which points to biological functions that are impaired in T2DM.

The relevance of a module can be checked with gene set enrichments. Here we use an overrepresentation analysis (ORA) with the hypergeometric test as provided for all gene sets in the ConsensusPathDB (Kamburov et al. 2011). Reducing the list to Reactome pathways results in table 2 with an emphasis on inflammation and pyruvate metabolism pathways. Table 2 also shows nicely how candidate genes are complemented by closely related but non-significant genes. This modified set of module genes dissects the Reactome root pathways to closer defined metabolic or signaling entities. The ORA is also applied to the gene ontology (GO) database in

table 3. GO is only analysed on level 3 of its hierarchical biological process structure and highlights links between the functional module and several regulatory elements in metabolism. In fig.8 selected overrepresented pathways are highlighted in the functional module.

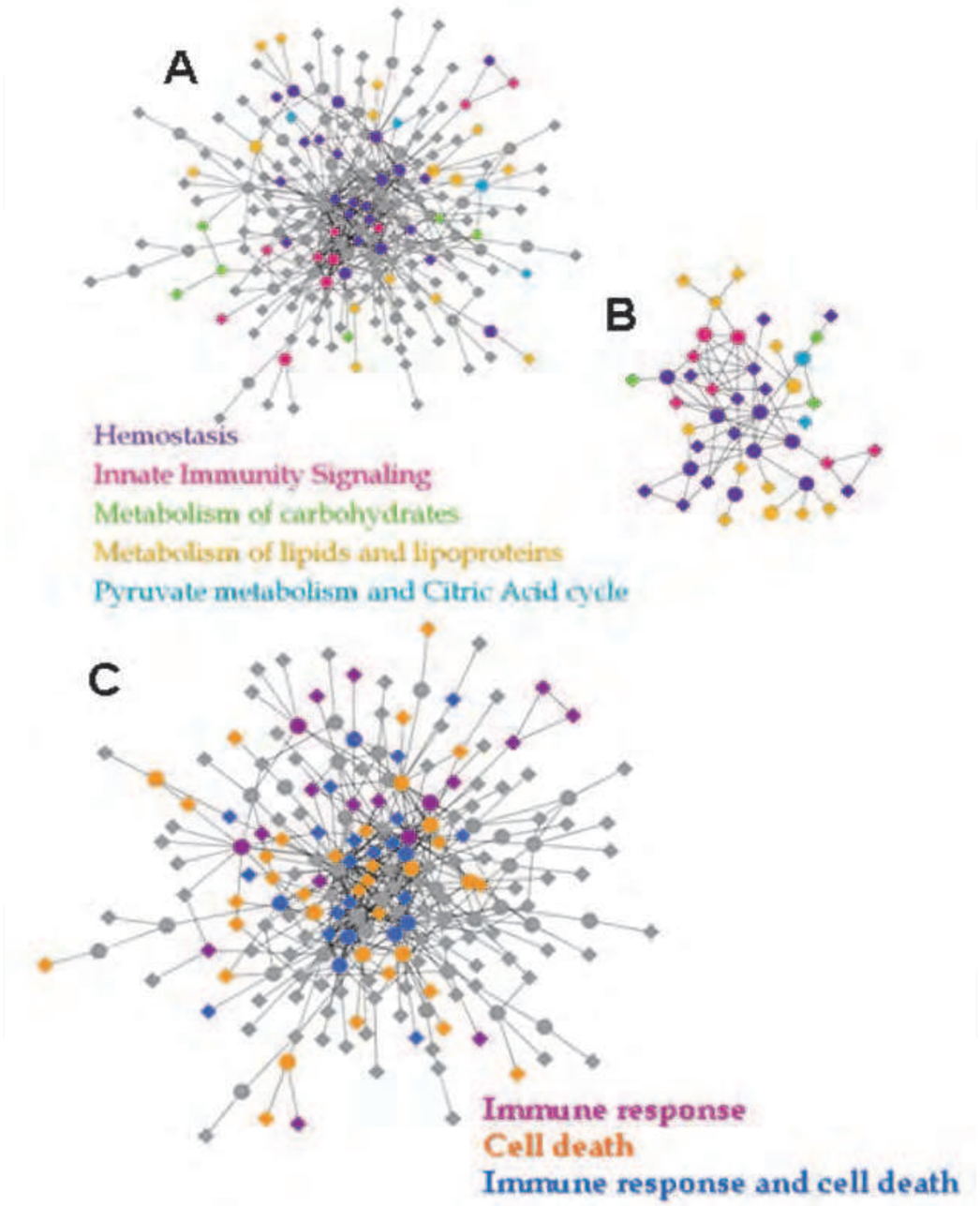


Fig. 8. Reactome interaction network. The concept of a reaction where reactant A is transformed to product B in reaction r1 and reactant B is transformed to product C in reaction r2 is reformulated into a relation where reaction r1 interacts with reaction r2. The relation is directed towards r2 because r2 precedes r1. The resulting reaction network consists of several sub-networks.

In the following we reinterpret the Reactome (Matthews et al. 2009) pathway information and characterize the module by impairment of reactions. Reactome is an expert-authored, peer-reviewed knowledge base. Reactome contains metabolic and signaling pathways. In metabolic pathways, proteins act as enzymes and in signaling pathways proteins are the main components that transfer information through interactions. We identify all reactions whose reactants, products or modifiers (enzymes) are part of the functional module and address them as covered. With the pathway information from Reactome we built a network (fig.9) where nodes represent reactions and edges represent relations between reactions: There is a directed out-going edge from a reaction to all its following reactions annotated in Reactome and there are directed in-coming edges to a reaction from all its preceding reactions. This interpretation may in mathematical terms be seen as a dual graph of the Reactome network. We compute shortest paths between all covered reactions and visualize the results in fig. 10. Nodes (covered reactions and non-covered) lying on these paths are included in the final set of reactions. The initial Reactome network is reduced to those reactions which are impaired in the course of T2DM and those reactions that link impaired reactions. Such a network can guide future research: Which pathways interfere with the proper functioning of other pathways? What is the link between proteins that interact with each other but are involved in different pathways?

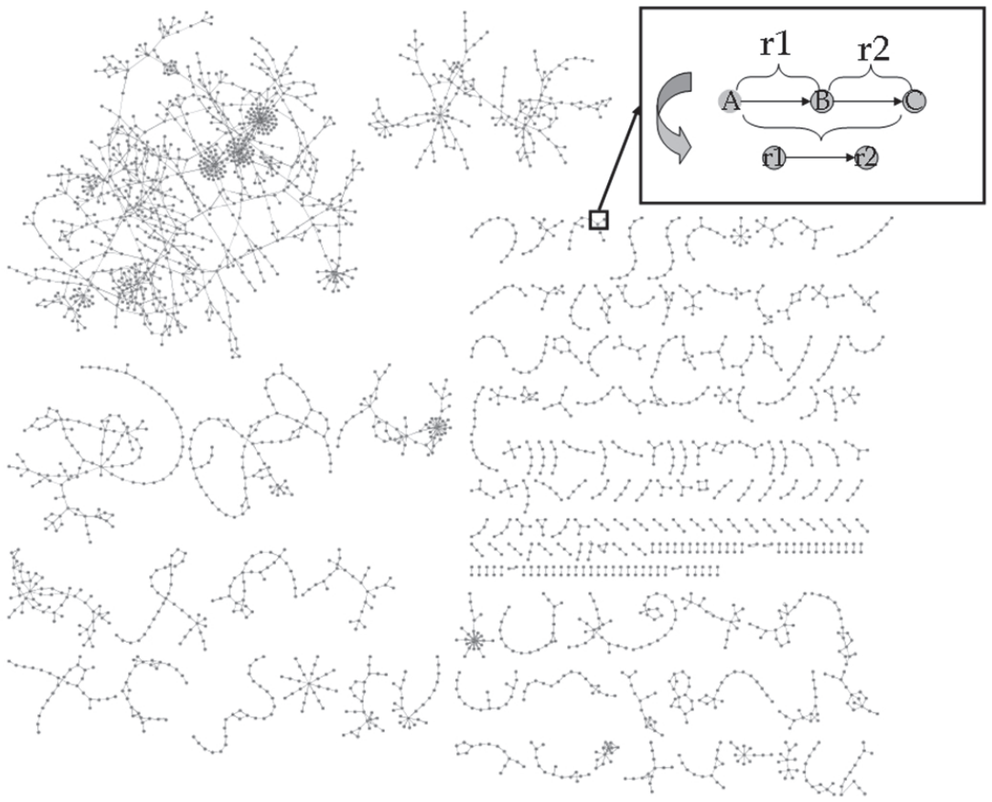


Fig. 9. A) The network shows the functional module where candidate genes are marked with diamonds and over-represented (OR) pathways are colored. B) Nodes in functional module involved in OR pathways tend to interact with each other. (C) Similarly OR GO-terms are colored in the functional module.

p-value	*q*-value	Pathway	Root	All	FM	Gene set
3.3e-5	0.0022	Alternative complement activation	IIS	3	3	C3*, CFB*, CFD*
2.0e-4	0.0087	TRAF6 Mediated Induction of proinflammatory cytokines	IIS	64	9	APP, ATF1, IKBKG, MAPK1*, MAPK9*, NFKB2, NFKBIA*, PPP2R1B, RELA*
3.0e-4	0.0087	NFkB and MAP kinases activation mediated by TLR4 signaling repertoire	IIS	67	9	
4.0e-4	0.0089	TLR3 Cascade	IIS	70	9	APP, ATF1, IKBKG, MAPK1*, MAPK9*, NFKB2, NFKBIA*, PPP2R1B, RELA*
5.0e-4	0.0089	MyD88-independent cascade initiated on plasma membrane	IIS	71	9	
6.0e-4	0.0089	TRAF6 mediated NF-kB activation	IIS	22	5	APP, IKBKG, NFKB2, NFKBIA*, RELA*
6.0e-4	0.0089	TRAF6 mediated induction of NFkB and MAP kinases upon TLR7/8 or 9 activation	IIS	74	9	APP, ATF1, IKBKG, MAPK1*, MAPK9*, NFKB2, NFKBIA*, PPP2R1B, RELA*
7.0e-4	0.0089	MyD88 dependent cascade initiated on endosome	IIS	75	9	APP, ATF1, IKBKG, MAPK1*, MAPK9*, NFKB2, NFKBIA*, PPP2R1B, RELA*
7.0e-4	0.0089	TLR7/8 Cascade	IIS	75	9	
8.0e-4	0.0091	TLR4 Cascade	IIS	90	10	APP, ATF1, CD14*, IKBKG, MAPK1*, MAPK9*, NFKB2, NFKBIA*, PPP2R1B, RELA*
9.0e-4	0.0091	Mitochondrial Fatty Acid Beta-Oxidation	MLL	14	4	ACADL*, HADHB, MCEE*, PCCB
9.0e-4	0.0091	human TAK1 activates NFkB by phosphorylation and activation of IKKs complex	IIS	24	5	APP, IKBKG, NFKB2, NFKBIA*, RELA*
0.0010	0.0091	TLR1, 2, 6, 9 Cascade	IIS	79	9	APP, ATF1, IKBKG, MAPK1*, MAPK9*, NFKB2, NFKBIA*, PPP2R1B, RELA*
0.0018	0.0129	TLR Cascades	IIS	100	10	APP, ATF1, CD14*, IKBKG, MAPK1*, MAPK9*, NFKB2, NFKBIA*, PPP2R1B, RELA*
0.0019	0.0129	Viral dsRNA:TLR3:TRIF Complex Activates RIP1	IIS	28	5	APP, IKBKG, NFKB2, NFKBIA*, RELA*
0.0019	0.0129	Chylomicron-mediated lipid transport	MLL	17	4	APOA1, LDLR*, LPL*, P4HB

p-value	*q*-value	Pathway	Root	All	FM	Gene set
0.0021	0.0141	Activated TLR4 signalling	IIS	86	9	APP, ATF1, IKBKG, MAPK1*, MAPK9*, NFKB2, NFKBIA*, PPP2R1B, RELA*
0.0022	0.0141	Lipoprotein metabolism	MLL	29	5	APOA1, LDLR*, LPL*, P4HB, PLTP*
0.0043	0.0253	Lipid digestion, mobilization, and transport	MLL	48	6	APOA1, FABP4*, LDLR*, LPL*, P4HB, PLTP*
0.0049	0.0277	Fatty acid, triacylglycerol, and ketone body metabolism	MLL	82	8	ACADL*, ACLY*, ACSL1*, FASN*, HADHB, MCEE*, MED1*, PCCB
0.0060	0.0327	Propionyl-CoA catabolism	MLL	4	2	MCEE*, PCCB
0.0075	0.0377	Advanced glycosylation endproduct receptor signaling	IIS	13	3	APP, LGALS3*, MAPK1*
0.0089	0.0434	Pyruvate metabolism and Citric Acid cycle	PM	40	5	BSG*, DLD, FH*, NNT*, PDK4*
0.0098	0.0458	Beta oxidation of lauroyl-CoA to decanoyl-CoA-CoA	MLL	5	2	ACADL*, HADHB

Table 2. Pathway over-representation analysis for proteins contained in the functional module, restricted to root level pathways IIS, MLL and PM. The table lists Reactome pathways in which proteins of our functional module are enriched. All: Number of proteins represented in the specified pathway. FM: Number of proteins represented in the specified pathway and contained in the functional module. Root: Specifies the root pathway as defined by Reactome for the given pathway (IIS – Innate Immunity Signaling, MLL – Metabolism of lipids and lipoproteins and PM – Pyruvate metabolism and Citric Acid (TCA) cycle). Candidate genes are succeded by a star in the last column.

5. Discussion

Many applications have been developed based on the analyses of topological network properties which provide insights into the evolution, function, stability and dynamic responses of PPI networks (Albert 2005). Deciphering the wiring scheme and determining topological properties of individual nodes could help to derive protein function and formulate predictions about disease involvement. Special attention is drawn towards highly connected nodes whose removal has serious, or even lethal, consequence for the network. Highly connected nodes are probably evolutionarily conserved or encoded by essential genes (Goh et al. 2007). There is evidence that in literature-curated PPI networks disease genes share common topological characteristics which differ from non-disease genes: Hereditary disease genes selected from OMIM (Hamosh et al. 2005) have a larger degree, the tendency to interact with each other, more common neighbors and fast communication to other disease genes (Xu & Li 2006).

The tendency of proteins involved in the same disease to interact with each other can be traced to the chromosome level (Oti et al. 2006). Genes that interact with known disease genes have a higher likelihood of being also disease relevant. In summary, network analysis reveals properties of potential disease genes. There are good reasons to assume that disease genes are not randomly placed in the interactome.

The meta-analysis is a valuable method for ranking genes according to their disease relevance. In a follow-up step we put the candidate genes from Rasche et al. into a functional context. We took advantage of PPI data in two ways: First, we characterized disease genes with respect to their topological parameters. Second, we applied an algorithm that channels all available PPIs into a sub-network. This subnetwork seems to contain relevant information about the underlying biological functions impaired in T2DM. The topological characterization of candidate genes reveals properties which distinguish them from the complete set genes: Compared to the complete set of candidate genes have higher node degrees (fig. 4), higher betweenness centrality coefficients (fig. 5), higher 2N indices (fig. 7) and shorter average distances to other candidate genes (fig. 6). The ten candidate genes with highest degree are: PIK3R1 (246), ACTB (244), RELA (236), MAPK1 (206), EIF4A2 (157), YBX1 (148), NFKBIA (119), TNFRSF1B (110) and B2M (108). These genes are well described in the literature. They are associated with different diseases. Although there is a relation between node degree and disease relevance we have to consider a bias towards genes where disease relevance and connectivity is established. The meta-analysis also identifies genes with a small node degree as relevant for T2DM: ACSL1, AKR1B10, AOX1, CCNI, GATM, GPD2, GPX2, LGMN, LRP10, NNT, P4HA1, RETN, SLC38A2, TMSB4X, YIPF5 and ZSCAN21 (all with node degree of one).

Candidate genes exhibit higher betweenness centrality coefficients. Candidate genes with highest betweenness centrality coefficients are: ACTB (0.011), PIK3R1 (0.009), MAPK1 (0.008), RELA (0.007), B2M (0.005), HSPA5 (0.004), DYNLL1 (0.003), C1QBP (0.003), TNFRSF1B (0.003) and NFKBIA (0.003). Nodes with a high betweenness centrality coefficient are termed bottlenecks (Yu et al. 2007). Many shortest paths pass through a node with high betweenness centrality coefficient; a perturbation in a node with high betweenness centrality coefficient easily deranges the rest of the network. Betweenness centrality better accounts for the prediction of node's essentiality in the network than the node degree. A perturbation in a node with a high degree which lies in the outer part of the network probably has less severe consequences than a node which lies more central in the network. Candidate genes do not differ in clustering coefficient and neighbourhood connectivity from the set of all genes. Direct neighbors of candidate genes are not more likely also candidate genes (1N index). However, the 2N index for candidate genes is higher than for non-candidate genes: Neighbors of neighbors of candidate genes are more likely also candidate genes. These results indicate that T2DM involves several impaired biological functions. A higher 1N index for candidate genes would suggest that a single biological function is perturbed. Related to the higher 2N index for candidate genes is the smaller average shortest paths length from a candidate gene to all other candidate genes. Topological parameters may not isolate disease genes if they are individually considered. But in this study they indicate that candidate genes link several biological processes as shown by the high betweenness centrality and the high 2N index.

p-value	*q*-value	Term name	All	FM
1.7e-39	1.1e-36	response to organic substance	1072	65
5.4e-30	1.7e-27	regulation of cell death	1115	49
3.2e-29	6.7e-27	positive regulation of biological process	2465	79
9.4e-29	1.5e-26	regulation of response to stimulus	648	40
1.4e-27	1.7e-25	regulation of immune system process	545	38
3.0e-27	3.1e-25	regulation of immune response	319	30
2.1e-26	1.9e-24	response to inorganic substance	316	33
3.3e-24	2.6e-22	programmed cell death	1321	59
4.8e-24	3.3e-22	response to drug	341	33
1.7e-23	1.1e-21	positive regulation of immune response	214	22
2.7e-23	1.6e-21	response to molecule of bacterial origin	187	22
7.9e-23	4.1e-21	response to hormone stimulus	580	43
2.5e-22	1.2e-20	negative regulation of cell death	548	39
4.4e-22	2.0e-20	cell differentiation	1899	67
1.7e-21	7.0 e-20	positive regulation of macromolecule metabolic process	1109	45
1.8e-21	7.0e-20	negative regulation of biological process	2193	76
2.5e-21	9.2e-20	regulation of developmental process	942	44
3.4e-21	1.2e-19	organ development	2026	71
7.2e-21	2.4e-19	system development	2622	81
1.0e-20	3.2e-19	antigen processing and presentation via MHC class Ib	16	3
1.9e-20	5.7e-19	adaptive immune response	166	20
4.7e-19	1.3e-17	antigen processing and presentation of exogenous antigen	19	4
8.3e-19	2.3e-17	regulation of cellular process	5922	126
1.2e-18	3.2e-17	positive regulation of cell death	613	27
3.1e-18	7.7e-17	positive regulation of cellular metabolic process	1141	44
1.8e-17	4.4e-16	protein complex assembly	681	42
2.0e-17	4.7e-16	T cell mediated cytotoxicity	29	7
2.7e-17	5.9e-16	immune effector process	269	27
3.4e-16	7.4e-15	positive regulation of biosynthetic process	875	35
3.9e-16	8.0e-15	cellular response to chemical stimulus	547	36
5.0e-16	9.8e-15	macromolecular complex assembly	853	45
5.0e-16	9.8e-15	positive regulation of immune effector process	72	14

Table 3. GO term over-representation analysis (terms downstream to term *biological process* level 3) for genes in the functional module.

Next, we identified a sub-network in the complete PPI network with enrichment in candidate genes. The algorithm used was proposed by Dittrich et al. for the computation of functional modules. Candidate genes were weighted with their meta-analysis score and the remaining nodes in the network with a negative score. A pathway over-representation analysis points to the pathways *Hemostasis, Innate immunity signaling, Pyruvate metabolism and citric acid cycle, Metabolism of lipids and lipoproteins and Metabolism of carbohydrates.* Our results confirm a known relation between inflammation and metabolic disorders (Hotamisligil 2006). The link between metabolism and immune response pathways can be retraced to common ancestral structures (Hotamisligil 2006). The Toll-like receptor (TLR) pathway comprises elements which regulate metabolic and immune functions. TLR4, receptor for bacterial LPS and component of innate immune system acts as a sensor for free fatty acids (Shi et al. 2006). Free fatty acids are increased in obesity and are a probable link to lipid-induced insulin resistance. The functional module contains genes (RELA, NFKBIA, NFKB2, ATF1, MAPK1, IKBKG, MAPK9) which are activated downstream to TLR4 (Akira & Takeda 2004). Analysis of the functional module also reveals a link to platelet dysfunction (Vinik et al. 2001).

Over-representation analysis for GO terms with root node biological function reveals terms lying downstream to cell death and immune response. Pathway and GO terms analysis suggests and supports the strong link between inflammation and T2DM. We extended this knowledge by annotated pathways, e.g. by introducing the notion of covered reactions. A covered reaction involves a protein from the functional module, either as enzyme, reactant or product. We suppose that an impaired covered reaction may have a negative influence on the network.

Using the PPI network, a list of candidate genes could be characterized according to distributions of topological parameters, especially in comparison to the full set of PPI. At the current stage of knowledge we can only use a static PPI graph, since the complete graph is unknown. We assume that we already have a representative subset of PPIs in the databases. We pointed out that known PPIs reflect only static, sometimes artificial, settings. In these settings interactions depend on many factors and thus proteins may only interact under certain circumstances. To overcome some of these constraints the candidate genes are extended to a functional module using the MWCS method. Genes lacking interaction information are skipped and only non-candidate genes which are directly linked and are in direct proximity to candidates are included in the module. The functional module genes are related to functional entities by applying ORA to Reactome and GO gene sets. These databases cover far less genes than PPI networks but with much more detailed descriptions about the purpose of the genes within a biological context. Module genes are related to the discussed functional entities which shows that current knowledge is well incorporated in the functional module. Furthermore, Reactome was also the basis for a modified description of its functional content with the notation of covered reactions. This is a possible way of identifying several pathways which interact in a direct or indirect manner. In the case of the functional module it elucidates how over-represented pathways are linked in T2DM and which module genes possibly modulate this link.

6. Conclusion

Results of a single-gene meta-analysis are combined with methods from network biology. We have to keep in mind that PPI networks are not static but are modified for changing cellular

states. In the long term it does not suffice to consider topological properties alone. We have to elaborate on an understanding of the dynamics of PPIs. Different conditions influence structural rearrangements in the cell which we need to measure and depict. Computation of functional modules is an attempt of including additional levels to the interaction data. We see overlapping functions rather than a clear division in single biological sections.

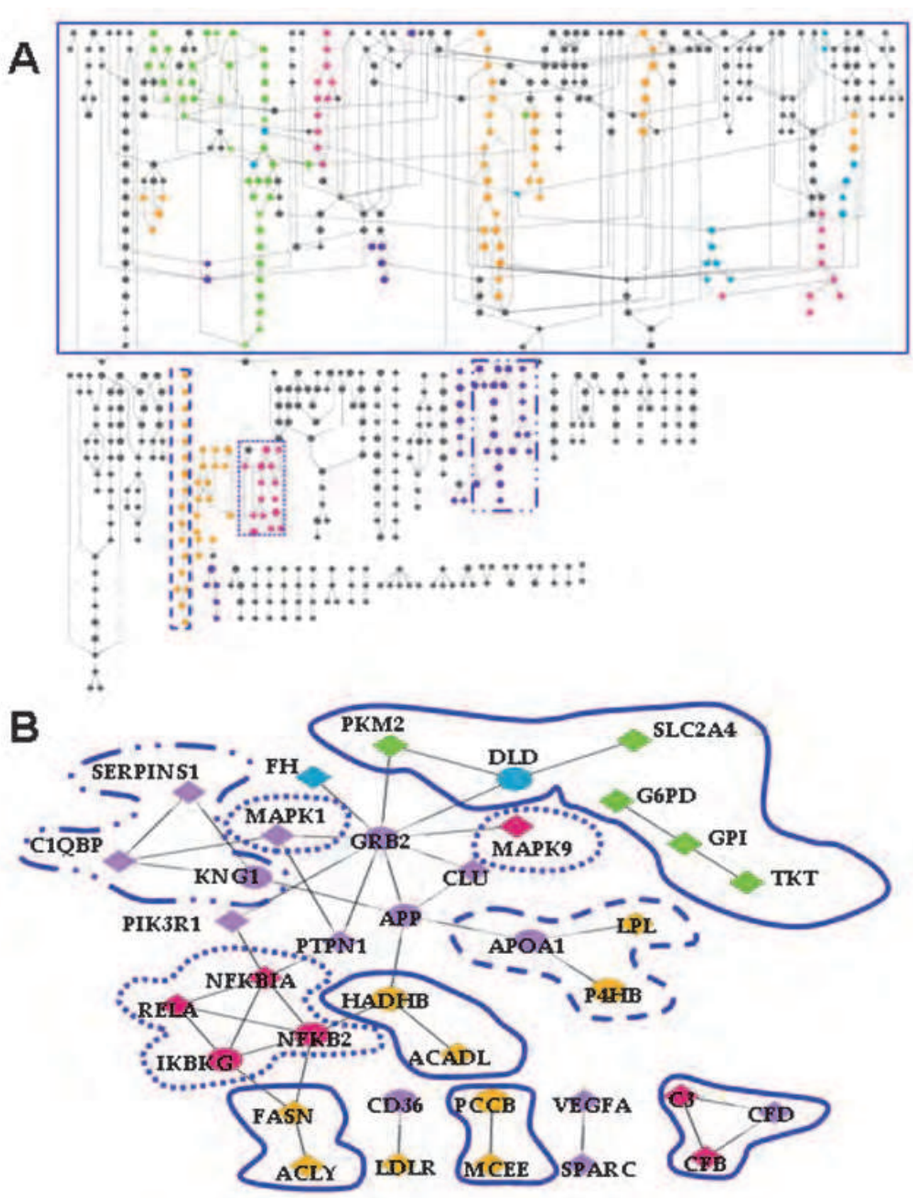


Fig. 10. A) Hierarchical view of shortest paths between all covered reactions. A covered reaction is shown as diamond, non-covered reaction as circle, non-covered reactions connect covered reactions. Over-represented pathways in the functional module are highlighted according to the color scheme in fig. 9. B) Genes from the functional module which are involved in a covered reaction. Frames with different line types in A) and B) elucidate how the functional module connects different pathways.

7. Acknowledgements

We want to acknowledge Atanas Kamburov who is the lead developer of the Consensus-PathDB and Dr. Ralf Herwig for initiating the topic, study design and funding. The work was partly funded by the European Union under its 6th Framework Programme with the grant SysProt (LSHG-CT-2006-037457) and the BMBF NGFN-transfer project (01GR0809).

8. References

Akira, S. & K. Takeda (2004). Toll-like receptor signalling. Nat Rev Immunol 4(7): 499-511.

Albert, R. (2005). Scale-free networks in cell biology. J Cell Sci 118(Pt 21): 4947-4957.

Ashburner, M., C. A. Ball, J. A. Blake, et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25(1): 25-29.

Assenov, Y., F. Ramírez, S.-E. Schelhorn, et al. (2008). Computing topological parameters of biological networks. Bioinformatics 24(2): 282-284.

Barabasi & Albert (1999). Emergence of scaling in random networks. Science 286(5439): 509-512.

Barabási, A.-L., N. Gulbahce & J. Loscalzo (2011). Network medicine: a network-based approach to human disease. Nat Rev Genet 12(1): 56-68.

Barabási, A.-L. & Z. N. Oltvai (2004). Network biology: understanding the cell's functional organization. Nat Rev Genet 5(2): 101-113.

Beisser, D., G. W. Klau, T. Dandekar, et al. (2010). BioNet: an R-Package for the functional analysis of biological networks. Bioinformatics 26(8): 1129-1130.

Cline, M. S., M. Smoot, E. Cerami, et al. (2007). Integration of biological networks and gene expression data using Cytoscape. Nat Protoc 2(10): 2366-2382.

Csardi, G. & T. Nepusz (2006). The igraph Software Package for Complex Network Research. InterJournal Complex Systems: 1695.

Dittrich, M. T., G. W. Klau, A. Rosenwald, et al. (2008). Identifying functional modules in protein-protein interaction networks: an integrated exact approach. Bioinformatics 24(13): i223-i231.

Erdos, P. & A. Renyi (1960). On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci 5: 17-61.

Fields, S. & R. Sternglanz (1994). The two-hybrid system: an assay for protein-protein interactions. Trends Genet 10(8): 286-292.

Goh, K.-I., M. E. Cusick, D. Valle, et al. (2007). The human disease network. Proc Natl Acad Sci U S A 104(21): 8685-8690.

Hamosh, A., A. F. Scott, J. S. Amberger, et al. (2005). Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res 33(Database issue): D514-D517.

Hotamisligil, G. S. (2006). Inflammation and metabolic disorders. Nature 444(7121): 860-867.

Ideker, T., O. Ozier, B. Schwikowski, et al. (2002). Discovering regulatory and signalling circuits in molecular interaction networks. Bioinformatics 18 Suppl 1: S233-S240.

Kamburov, A., K. Pentchev, H. Galicka, et al. (2011). ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res 39(Database issue): D712-D717.

Kamburov, A., C. Wierling, H. Lehrach, et al. (2009). ConsensusPathDB--a database for integrating human functional interaction networks. Nucleic Acids Res 37(Database issue): D623-D628.

Matthews, L., G. Gopinath, M. Gillespie, et al. (2009). Reactome knowledgebase of human biological pathways and processes. Nucleic Acids Res 37(Database issue): D619-D622.

Oti, M., B. Snel, M. A. Huynen, et al. (2006). Predicting disease genes using protein-protein interactions. J Med Genet 43(8): 691-698.

Pandey, A. & M. Mann (2000). Proteomics to study genes and genomes. Nature 405(6788): 837-846.

Rasche, A., H. Al-Hasani & R. Herwig (2008). Meta-analysis approach identifies candidate genes and associated molecular networks for type-2 diabetes mellitus. BMC Genomics 9: 310.

Ravasz, E. & A.-L. Barabási (2003). Hierarchical organization in complex networks. Phys Rev E Stat Nonlin Soft Matter Phys 67(2 Pt 2): 026112.

Shi, H., M. V. Kokoeva, K. Inouye, et al. (2006). TLR4 links innate immunity and fatty acid-induced insulin resistance. J Clin Invest 116(11): 3015-3025.

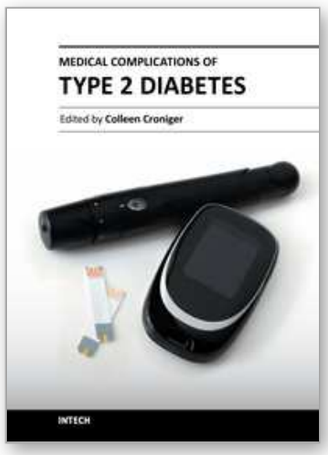
Vinik, A. I., T. Erbas, T. S. Park, et al. (2001). Platelet dysfunction in type 2 diabetes. Diabetes Care 24(8): 1476-1485.

Watts, D. J. & S. H. Strogatz (1998). Collective dynamics of 'small-world' networks. Nature 393(6684): 440-442.

Wu, Z., X. Zhao & L. Chen (2009). Identifying responsive functional modules from protein-protein interaction network. Molecules and Cells 27(3): 271-277.

Xu, J. & Y. Li (2006). Discovering disease-genes by topological features in human protein-protein interaction network. Bioinformatics 22(22): 2800-2805.

Yu, H., P. M. Kim, E. Sprecher, et al. (2007). The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. PLoS Comput Biol 3(4): e59.



Medical Complications of Type 2 Diabetes

Edited by Dr. Colleen Croniger

ISBN 978-953-307-363-7
Hard cover, 412 pages
Publisher InTech
Published online 12, September, 2011
Published in print edition September, 2011

Obesity and type 2 diabetes are increasing worldwide problems. In this book we reviewed insulin secretion in both healthy individuals and in patients with type 2 diabetes. Because of the risk associated with progression from insulin resistance to diabetes and cardiovascular complications increases along a continuum, we included several chapters on the damage of endothelial cells in type 2 diabetes and genetic influences on endothelial cell dysfunction. Cardiovascular complications occur at a much lower glucose levels, thus a review on the oral glucose tolerance test compared to other methods was included. The medical conditions associated with type 2 diabetes such as pancreatic cancer, sarcopenia and sleep disordered breathing with diabetes were also discussed. The book concludes with several chapters on the treatments for this disease offering us hope in prevention and successful alleviation of the co-morbidities associated with obesity and type 2 diabetes.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Anja Thormann and Axel Rasche (2011). Functional Context Network of T2DM, Medical Complications of Type 2 Diabetes, Dr. Colleen Croniger (Ed.), ISBN: 978-953-307-363-7, InTech, Available from: http://www.intechopen.com/books/medical-complications-of-type-2-diabetes/functional-context-network-of-t2dm

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen