V9 metabolische Netzwerke / Protein-Interaktionsnetzwerke

 Methoden der Systembiologie ermöglichen die integrierte, simultane Betrachtung von kompletten metabolischen Netzwerken.

wichtige Fragen:

- (1) wie soll man die Gesamtheit der metabolischen Flüsse im Netzwerk beschreiben?
- (2) wie können mathematische Techniken die gezielte Manipulation von Mikroorganismen im Hinblick auf die Produktion bestimmter Substanzen unterstützen?

Inhalt V9:

- (1) stöchiometrische Matrix
- (2) Flux Balance Analysis Methode
- (3) Elementary Flux Mode Analyse (anschaulich + Anwendung)
- (4) Aufbau und Visualisierung von Protein-Protein-Interaktionsnetzwerken

Analyse von Stoffwechselwegen: Beispiel E. coli

verwende Daten aus Datenbank **EcoCyc** (siehe auch **MetaCyc**: Stoffwechsel von > 150 Organismen)

EcoCyc enthält 905 Reaktionen für *E.coli* davon gehören 161 nicht zum Stoffwechsel kleiner Moleküle, z.B. DNA Replikation, von den verbleibenden 744 wurden 569 mindestens einem Pfad zugeordnet

Dagegen gibt es 607 Enzyme.

Es gibt also keine 1:1 Zuordnung zwischen Enzymen und Reaktionen, denn

- (1) manche Enzyme katalyiseren mehrere Reaktionen, und manche Reaktionen werden von mehreren Enzymen katalysiert
- (2) nicht zu allen Reaktionen sind die Enzyme bekannt, die sie katalysieren.

Ouzonis, Karp, Genome Research 10, 568 (2000)

Beispiel: Stoffwechsel von E. coli

Die 744 Reaktionen enthalten 791 verschiedene Substrate

Table 1. Most Frequently Used Metabolites in *E. coli* Central Metabolism

Occurrence	Name of metabolite	
205	H ₂ O	
152	ΑŤΡ	
101	ADP	
100	phosphate	
89	pyrophosphate	
66	NAD .	
60	NADH	
54	CO ₂	
53	H ⁺	
49	AMP	
48	NH ₃	
48	NADP	
45	NADPH	
44	Coenzyme A	
43	L-glutamate	
41	pyruvate	
29	acetyl-CoA	
26	O ₂	
24	2-oxoglutarate	
23	S-adenosyl-L-methionine	
18	S-adenosyl-homocysteine	
16	L-aspartate	
16	L-glutamine	
15	H ₂ O ₂	

Im Mittel enthält jede Reaktion 4 Substrate.

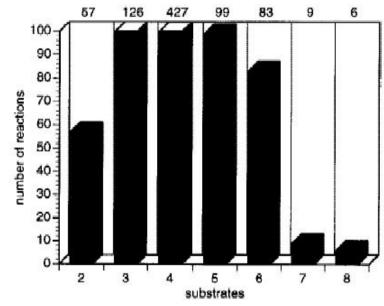


Figure 4 Diagram showing the number of reactions containing varying numbers of substrates (reactants plus products).

Ouzonis, Karp, Genome Research 10, 568 (2000)

Beispiel: Stoffwechsel von E. coli

EcoCyc enthält 131 Stoffwechsel-Pfade.

Die Länge der Pfade variiert von 1 bis 16. Der Mittelwert ist 5.4.

Von den 607 Enzymen sind 100 multifunktional.

Purin-Nukleosid-Phosphorylase und Nukleosid-Diphosphatkinase katalysieren 7 bzw. 9 Reaktionen.

483 Reaktionen gehören zu einem Pfad, 99 Reaktionen gehören zu mehreren Pfaden.

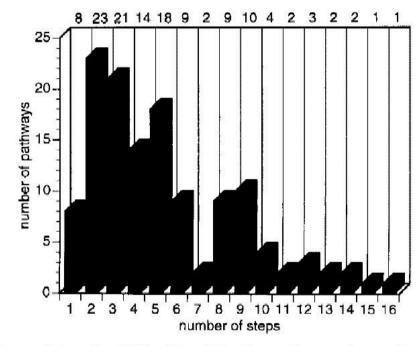


Figure 5 Length distribution of EcoCyc pathways; two pathways are not included because the number of steps is not known.

Ouzonis, Karp, Genome Research 10, 568 (2000)

Fazit

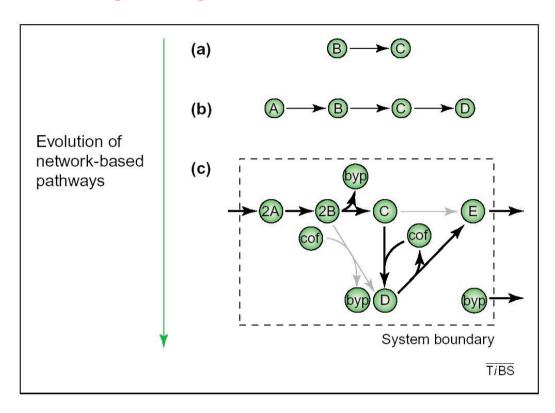
Stoffwechsel-Netzwerke von einfachen Organismen sind mittlerweile fast vollständig bekannt.

Ist die Beschreibung mit einzelnen Stoffwechsel-Wegen adäquat?

- Reaktionen, Enzyme und Substrate gehören oft zu mehreren Pfaden.
- Die Einteilung in einzelne Stoffwechsel-Pfade ist nicht immer eindeutig

Metabolische Pfade in der post-genomischen Ära

- (a) klassische Biochemie bestimmt Stöchiometrien einzelner **Reaktionen**
- (b) Katalogisierung vieler Reaktionen, Gruppierung nach gemeinsamen Metaboliten führt zu traditionellen **Pfaden** wie Glykolyse, Pentose-Phosphat-Pfad
- (c) Durch komplette Information können nun die kompletten metabolischen Pfade zu einem **Netzwerk** kombiniert werden.



Metabolische Pfade in der post-genomischen Ära

Traditionelle metabolische Pfade dienen als konzeptioneller Rahmen für Forschung und Lehre.

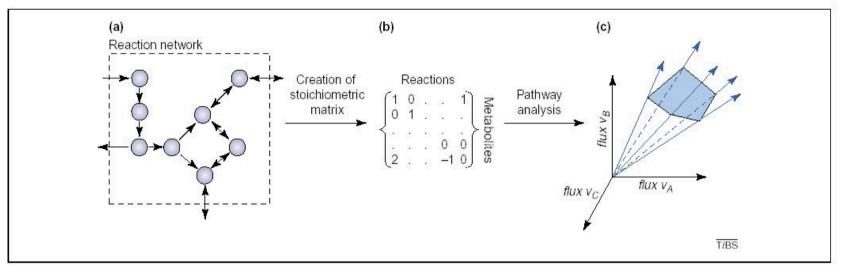
Man kann dadurch Metabolismen verschiedener Organismen vergleichen.

Jedoch sind sie nicht für quantitative, systemische Bewertungen biologischer Reaktionsnetzwerke geeignet, da sie nur Teile der Netzwerke darstellen.

Sie wurden oft in Zelltypen entdeckt, in denen sie wichtige metabolische Funktionen übernehmen (z.G. Glykolyse in Hefe).

Man kann diese Pfade jedoch nicht einfach auf andere Zelltypen mit anderen Enzym- und Metabolit-Konzentrationen übertragen.

Beschreibung vernetzter metabolischer Pfade



Papin et al. TIBS 28, 250 (2003)

- (a) aus genomischen, biochemischen, physiologischen Daten wird ein Reaktionsnetzwerk aufgestellt. Es gibt **interne Flüsse** innerhalb der Systemgrenzen und **externe Flüsse** zum Austausch mit der Umgebung.
- (b) Dieses Netzwerk wird durch eine $m \times n$ stöchiometrische Matrix dargestellt, in der m Metabolite durch n Reaktionen miteinander verbunden werden.
- (c) Mögliche Zustände der Zelle aufgrund dieser Matrix werden mit Techniken wie "elementary modes" oder "extreme pathways" identifiziert. Die möglichen Zustände liegen innerhalb eines Konus im durch die verschiedenen Flüsse aufgespannten Koordinatensystem.

Flux Balance Analyse

Jede chemische Reaktion erfüllt die Erhaltung der Massen.

Mit dieser grundlegenden Bedingung können metabolische Systeme untersucht werden.

Man muss lediglich die Stöchiometrie aller metabolischen Pfade und die metabolischen Anforderungen kennen. Für jeden Metabolit gilt:

$$v_{i} = \frac{dX_{i}}{dt} = V_{synthesized} - V_{degraded} - (V_{used} - V_{transported})$$

Im Gleichgewicht kann man die Massenerhaltung in einem metabolischen Netzwerk mathematisch durch eine Matrixgleichung darstellen:

$$S \cdot v = 0$$

Hierbei ist die Matrix S wiederum die $m \times n$ stöchiometrische Matrix, m = Anzahl der Metabolite, n = Anzahl der Reaktionen im Netzwerk.

Der Vektor v beinhaltet alle Flüsse des metabolischen Netzwerks, nämlich die internen Flüsse, die Transportflüsse, und das Zellwachstum.

Lösungsraum der metabolischen Flüsse

Da die Zahl an Metaboliten gewöhnlich kleiner ist als die Zahl an Reaktionen (m < n) ist die Flussgleichung üblicherweise **unterbestimmt**.

Daher gibt es im Allgemeinen eine Vielzahl an erlaubten Flussverteilungen, die die Massenerhaltung erfüllen. Dieser Satz von Lösungen entspricht dem Nullraum der Matrix S.

Um den "wahren" biologischen Fluss zu finden (\rightarrow e.g. Heinzle/UdS) benötigt man zusätzliche Informationen, z.B. aus Experimenten mit C¹³-Markierung.

Alternativ kann man Grenzwerte $\alpha_i \le v_i \le \beta_i$

für den Minimal- und Maximalbetrag jedes einzelnen metabolischen Flusses vorschreiben.

Der Durchschnitt des Nullraums und des durch diese linearen Ungleichungen eingeschränkten Raums definiert dann ein Volumen im hochdimensionalen Flussraum, das alle möglichen Flussverteilungen enthält.

E.coli in silico

Das am besten charakterisierte zelluläre System ist E. coli.

Im Jahr 2000 konstruierten Edwards & Palsson eine *in silico* Darstellung des *E.coli*- Metabolismus.

Dies erforderte eine Menge Handarbeit um die notwendigen Informationen

- aus der biochemischen Literatur,
- aus Genomannotationen und
- aus metabolischen Datenbanken wie EcoCyc und KEGG zusammenzustellen.



Bernhard Palsson, UC San Diego

Edwards & Palsson PNAS 97, 5528 (2000)

Gene im in silico Modell für E.coli Table 1. The genes Induded In the E. coli metabolic genotype (21) Zoom Out

potC, potD, potE, potE, potE, potE, potH, potI, proP, proV, proW, proX, pstA, pstB, pstC, pstS, ptsA, ptsG, ptsI,

ptsN, ptsP, purB, putP, rbsA, rbsB, rbsC, rbsD, rhaT, sapA, sapB, sapD, sbp, sdaC, srW_1, srW_2, srIB, tdcC, triaB, treA, treB, trkA, trkG, trkH, tsx, tyrP, ugpA, ugpB, ugpC, ugpE, uraA, xapB, xylE, xylF, xylG, xylH,

Table 1. The genes Induded In	the E. coll metabolic genotype (21)
Central metabolism (EMP, PPP, TCA cycle, electron transport)	aceA, aceB, aceE, aceF, ackA, acnA, acnB, acs, adhE, agp, appB, appC, atpA, atpB, atpC, atpD, atpE, atpF, atpF, atpF, atpF, atpH, atpI, cydA, cydB, cydC, cydD, cyoA, cyoB, cyoC, cyoD, dki, enc, fba, fbp, fdhF, fdnG, fdnH, fdrI, fdoG, fdoH, fdol, frdB, frdC, frdD, fumB, fumC, galM, gapA, gapC, I, gapC, Z, glcB, glgA, glgC, glgP, glk, glpA, glpB, glpC, glpD, gltA, gnd, gndB, gynB, hyaA, hyaB, hyaC, hybA, hybC, hycB, hycB, hycE, hycF, iycG, icdA, lctD, ldhA, lpcA, malP, mdh, nuoA, nuoB, nuoE, nuoF, nuoG, nuoH, nuoL, nuoL, nuoM, nuoM, nuoM, pckA, pfkA, pfkB, pflB, pflC, pflD, pgi, pgk, pntA, pntB, ppc, ppsA, pta, purT, pykA, pykF, rpe, rpiA, tpiB, sdhA, sdhB, sdhC, sdhD, sfcA, sucA, sucB, sucC, sucD, talB, tktA, tktB, tpiA, trxB, zwf, pgl (30), maeB (30)
Alternative carbon source	adhC, adhE, agaY, agaZ, aldA, aldB, aldH, araA, araB, araD, bglX, cpsG, deoB, fruK, fucA, fucA, fucK, fucO, galE, galK, galT, galU, gatD, gatY, glk, glpK, gntK, gntV, gpsA, lacZ, manA, melA, mtlD, nagA, nagB, nanA, pfkB, pqi, pqm, rbsK, rhaA, rhaB, rhaD, srlD, treC, xvlA, xvlB
Amino acid metabolism	adi, aktii, air, ansA, ansB, argA, argB, argC, argD, argE, argE, argE, argH, argI, arcA, aroB, aroC, aroD, aroE, aroF, aroG, aroH, aroK, aroI, asd, asnA, asnB, aspA, aspC, avtA, catA, catA, catB, cysC, cysD, cysE, cysH, cysI, cy
Purine & pyrimidine metabolism	add, adk, amn, apt, cdd, cmk, codA, dcd, deoA, deoD, dgt, dut, gmk, gpt, gsk, guaA, guaB, guaC, hpt, mutT, ndk, nrdA, nrdB, mdD, nrdE, mdF, purA, purB, purD, purE, purF, purH, purK, purL, purM, purN, pvrT, pyrB, pyrC, pyrD, pyrE, pyrE
Vitamin & cofactor metabolism	acpS, bioA, bioB, bioD, bioF, coaA, cyoE, cysG, entA, entB, entC, entD, entE, entF, epcl, folK, folK, folK, folK, gcvR, gcvR, gcvR, glxX, glyA, gor, gshA, gshB, hemA, hemB, hemC, hemD, hemE, hemF, hemH, hemK, hemL, hemX, hemX, hemX, ivC, lig. lpdA, menA, menB, menC, menD, menE, menF, menG, metF, mutT, nadA, nadB, nadC, nadE, ntpA, pabA, pabB, pabC, panB, panC, panD, pdxA, pdxB, pdxH, pdxI, pdxX, pncB, putC, ribA, ribB, ribD, ribE, ribH, serC, thiC, thiE, thiF, thiG, thiH, thrC, ubiA, ubiB, ubiC, ubiG, ubiH, ubiX, yaaC, ygiG, nadD (49), nadF (49), nadG (49), panE (50), pncA (49), pncC (49), thiB (51), thiD (51), thiD (51), thiK (51), thiM (51), thiM (51), ubiE (52), ubiF (52), arabinose-5-phosphate isomerase (22), phosphopentothenate-cysteine decarboxylase (90), phospho-pantetheine adenylyltransferase (90), dephosphoCoA kinase (50), NMN glycohydrolase (49)
Lipid metabolism	accA, accB, accD, atoB, cdh, cdsA, cls, dgkA, fabD, fabH, fadB, gpsA, ispA, ispB, pgpB, pgsA, pscl, pssA, pgpA (53)
Cell wall metabolism	dcl/A, ddlB, galF, galU, glmS, glmU, htrB, kdsA, kdsB, kdtA, lpxA, lpxB, lpxC, lpxD, mraY, msbB, murA, murB, murC, murD, murE, murF, murG, murl, rfaC, rfaD, rfaF, rfaG, rfaI, rfaI, rfaI, ushA, glmM (54), lpo4 (55), rfaE (55), tetraccyldisectharida 4' kinasa (55), 3-deoxy-p-manno-octulosonic-acid 8-phospheta phosphatasa (55)
Transport processes	araE, araF, araG, araH, argT, aroP, artl, artl, artM, artP, artQ, brnQ, cadB, chaA, chaB, chaC, cmtA, cmtB, codB, crr, cycA, cysA, cysP, cysU, cysW, cysZ, dctA, dcuA, dcuB, dppA, dppB, dppC, dppD, dppF, fadl, focA, fruA, fruB, fucP, gabP, galP, gatA, gatB, gatC, glnH, glnP, glnQ, glpF, glpT, gltU, gltK, gltL, gltP, gltS, gntT, gpt, hisI, hisM, hisP, hisQ, hpt, kdpA, kdpB, kdpC, kgtP, lacY, lamB, livF, livG, livH, livI, livK, livM, ltdP, lysP, malE, malF, malG, malK, malX, manX, manY, manZ, melB, mglA, mglB, mglC, mtlA, mtr, nagE, nanT, nhaA, nhaB, nupC, nupG, oppA, oppB, oppC, oppD, oppF, panF, pheP, pitA, pitB, pnuC, potA, potB,

Edwards & Palsson PNAS 97, 5528 (2000)

> 9. Vorlesung WS 2012/13 Softwarewerkzeuge 12

fruF (56), gntS (57), metD (43), pnuE (49), scr (56)

E.coli in silico

Begrenze $\alpha_i = 0$ für irreversible interne Flüsse, $\alpha_i = -\infty$ für reversible interne Flüsse (aufgrund der biochemischen Literatur)

Transportflüsse für PO₄²⁻, NH₃, CO₂, SO₄²⁻, K⁺, Na⁺ blieben unbeschränkt.

Für die anderen Metabolite wurden Obergrenzen verwendet $0 < v_i < v_i^{max}$ außer für die Metabolite, die das metabolische Netzwerk verlassen könne (d.h. Acetat, Äthanol, Laktat, Succinat, Format, Pyruvat etc.)

Aus dem Satz der möglichen Flussverteilungen wird mit **linearer Programmierung** eine bestimmte Lösung bestimmt, die eine bestimmte metabolische Zielfunktion Z maximiert (bzw. –Z minimiert). $Z = \sum_i c_i \cdot v_i = \langle \mathbf{c} \cdot \mathbf{v} \rangle$

Als Zielfunktion Z wird hier die Summe aller Einzelflüsse definiert, was der Bildung maximaler Biomasse entspricht.

Edwards & Palsson, PNAS 97, 5528 (2000)

E.coli in silico

Um die Korrektheit dieses Vorgehens zu testen, wurden Änderungen des metabolischen Netzwerkes simuliert, die sich durch hypothetische Gendeletionen ergeben.

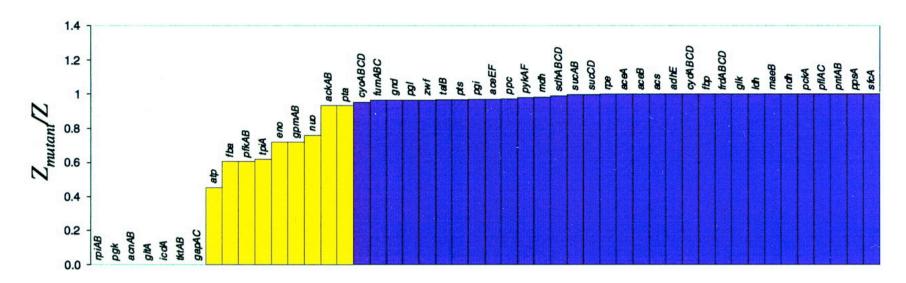
Dafür wird einfach der Fluss durch die entsprechende enzymatische Reaktion auf 0 gesetzt.

Dann wird der optimale Wert der Mutante (Z_{mutant}) mit dem des Wild-types verglichen (Z)

$$\frac{Z_{\it mutant}}{Z}$$

Edwards & Palsson PNAS 97, 5528 (2000)

Gendeletionen im zentralen E. coli MG1655 Metabolismus



Maximale Flüsse für virtuelles Wachstum auf Glucose für alle möglichen Deletionen einzelner Gene in Pfaden des zentralen Metabolismus (Glykolyse, Pentose Phosphat Pfad (PPP), Zitratzyklus, Atmung).

Gelbe Balken: Gendeletionen, die die maximale Biomasse (Fluss) auf weniger von 95% des Wert für den *in silico* Wildtyp drücken.

Edwards & Palsson PNAS 97, 5528 (2000)

Interpretation der Ergebnisse

Die essentiellen Gene gehörten zur 3-Kohlenstoff-Stufe der Glykolyse, 3 Reaktionen des Zitratzyklus und mehrere Mitglieder des PPP.

Alle anderen Gene des zentralen Metabolismus konnten entfernt werden ohne nenneswerten Effekt auf das *in silico*-Wachstum von *E.coli*.

Umlenken der metabolischen Flüsse

(Schwarz) Flussverteilung für Wildtyp.

(Rot) *zwf*- Mutante. Die Biomasse-Erzeugung ist 99% des Wildtyps.

(Blau) *zwf- pnt-* Doppelmutante. Biomasse-Erzeugung ist 92% des Wildtyps.

E.coli in silico umgeht die Deletion einer kritischen Reaktion (roter Pfeil) durch Erhöhung des Flusses durch die alternative G6P → P6P Reaktion.

Edwards & Palsson PNAS 97, 5528 (2000)

E.coli in silico

+ und – bedeuten Wachstum bzw.

Nichtwachstum.

± heisst, dass Suppressormutanten beobachtet wurden, die diesem Mutantenstamm Wachstum erlaubten.

glc: Glucose, gl: Glycerol,

succ: Succinat, ac: Acetat.

In 68 von 79 Fällen stimmte die Vorhersage mit dem Experiment überein.

Roten und gelbe Kreise kennzeichnen die vorhergesagten Mutationen, die Wachstum eliminieren bzw. reduzieren.

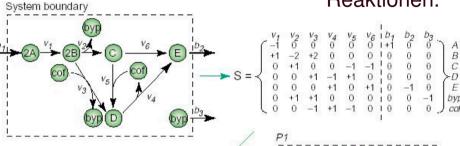
Edwards & Palsson PNAS 97, 5528 (2000)

Table 2. Comparison of the predicted mutant growth characteristics from the gene deletion study to published experimental results with single mutants

Gene	glc	gl	succ	ac
aceA	+/+		+/+	-/-
aceB				-/-
aceEf	* -/+			
ackA				+/+
acn	-/-			-/-
acs				+/+
cyd	+/+			
cyo	+/+			
eno†	-/+	-/+	-/-	-/-
fbal	-/+			
fbp	+/+	-/-	-/-	-/-
frd	+/+		+/+	+/+
gap	-/-	-/-	-/-	-/-
glk	+/+			
gltA.	-/-			-/-
gnd	+/+			
idh	-/-			-/-
mdh1	1 +/+	+/+	+/+	
ndh	+/+	+/+		
nuo	+/+	+/+		
pfk [†]	-/+			
pgi‡	+/+	+/-	+/-	
pgk	-/-	-/-	-/-	-/-
pgl	+/+			
pntA	B +/+	+/+	+/+	
ppc⁵	±/+	-/+	+/+	
pta				+/+
pts	+/+			
pyk	+/+			
- rpi	-/-	-/-	-/-	-/-
sdhA.			-/-	-/-
sucAl			-/+	-/+
tkt.A.E				
tpi**	-/+	-/-	-/-	-/-
unc	+/+		±/+	-/-
zwf	+/+	+/+	+/+	

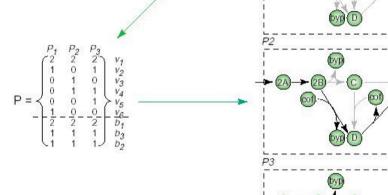
Berechnung von Elementaren Flussmoden aus der stöchiometrischen Matrix

Darstellung des Reaktionsnetzwerks mit stöchiometrischer Matrix S. stöchiometrische Koeffizienten der einzelnen Reaktionen.



Metabolite

Analyse der Matrix S → Pathway-Darstellung P. Deren Zeilen enthalten den Reaktionen entsprechende Flüsse und die Spalten die sich ergebenden Pfade.



Elementare Flussmoden für dies System.

Papin et al. TIBS 28, 250 (2003)

Elementare Flussmoden

Ein Pfad P(v) ist eine elementare Flussmode, wenn er Bedingungen C1 – C3 erfüllt.

(C1) **Pseudo Gleichgewicht**. $S \cdot e = 0$.

Dies garantiert, dass kein Metabolit angereichert oder abgereichert wird.

(C2) **Machbarkeit**: alle irreversiblen Reaktionen haben einen Rate $e_i \ge 0$. Dadurch können nur thermodynamisch realisierbare Flüsse in e enthalten sein.

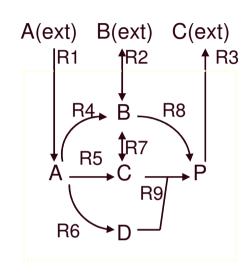
(C3) Nicht-Zerlegbarkeit:

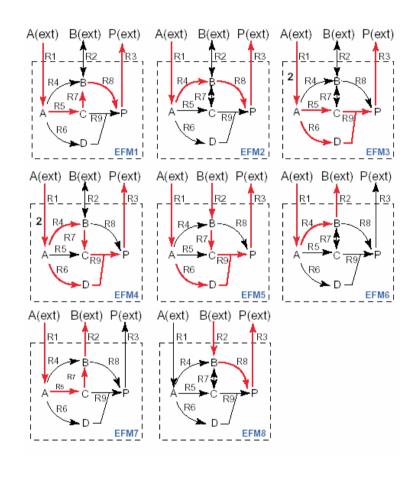
keine echte Teilmenge von Reaktionen aus P(v) ausser dem Nullvektor erfüllt ebenfalls Bedingungen C1 und C2.

Dadurch wird das Netzwerk in kleinste Einheiten P(v) zerlegt, die für sich jeweils ein Gleichgewicht im Netzwerk erzeugen.

Klamt & Stelling Trends Biotech 21, 64 (2003)

Elementare Flussmoden in Modell-Netzwerk





16. Lecture WS 2011/12

Klamt & Stelling Trends Biotech 21, 64 (2003)

21

Software: FluxAnalyzer, based on Matlab



Steffen Klamt

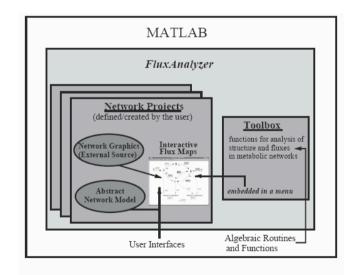


Fig. 2. Structural setup of the FluxAnalyzer.

9. Vorlesung WS 2012/13

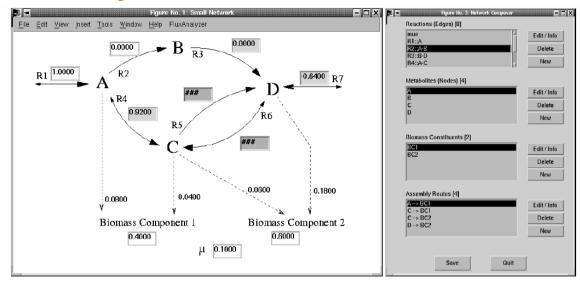


Fig. 1. The network project of 'SMALLNET' constructed by the FluxAnalyzer. Left: interactive flux map displaying a flux scenario (unknown rates are denoted by '###'). Right: network composer.

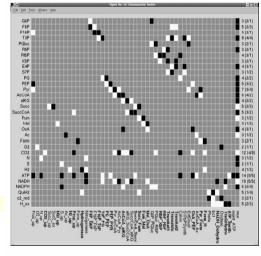


Fig. 3. Concise graphical representation of the stoichiometric matrix (here: catabolic part of the network studied in Klamt *et al.*, 2002)

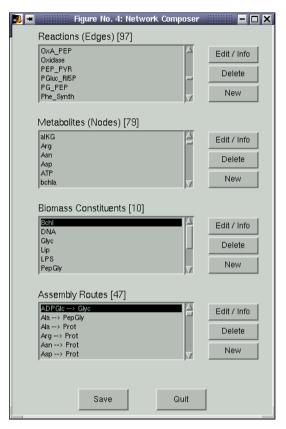
Softwarewerkzeuge

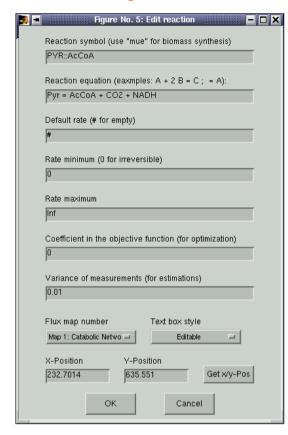
FluxAnalyzer has both EPs and EFMs implemented.

Allows convenient studies of metabolicsystems.

Klamt et al. Bioinformatics 19, 261 (2003)

Software: FluxAnalyzer, based on Matlab





Network composer .. Show all reaction equations Show names of network elements Show constraints Set constraints Set default scenario Reset last scenario Clear all values Save scenario ... Load scenario ... Copy values to flux clipboard Paste values from flux clipboard Arithmetic operations . Show stoichiometric matrix Basic topological properties Conservation relations Connectivity histogram Graph-theoretical path lengths ... Elementary flux modes and pathway analysis ... Minimal cut sets ... Check feasibility Classify rates (balanceability/calculability) Flux analysis: calculate/balance rates ... Sensitivity analysis ... Optimization (for underdetermined scenarios) Export stoichiometric matrix ... Show current values in bar chart Set original flux map size Info ...

Left: Network composer of the *FluxAnalyzer* facilitating the definition of the network structure.

Middle: Input mask for defining a new network element of type reaction Bioinformatics III

Right Mult-down menu of the Flux Analyzer

providing an interactive use of the various functions of the toolbest.//pingui

Stamm-Optimierung basierend auf EFM-Analyse

Metabolic Engineering 12 (2010) 112-122



Contents lists available at ScienceDirect

Metabolic Engineering





Rational design and construction of an efficient *E. coli* for production of diapolycopendioic acid

Pornkamol Unrean, Cong T. Trinh, Friedrich Srienc*

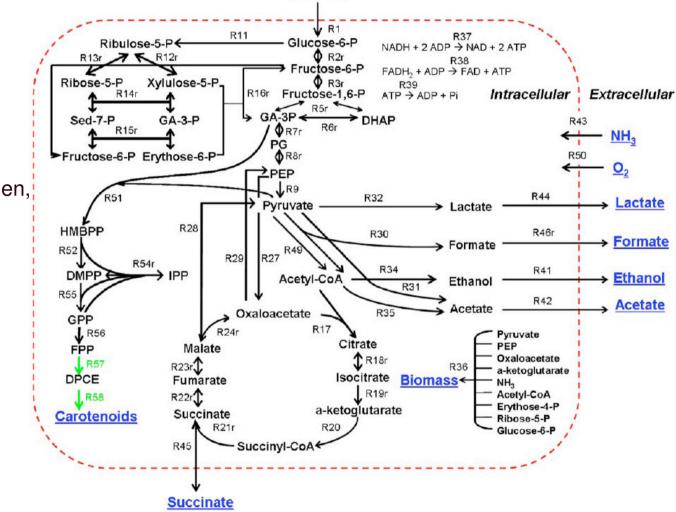
Department of Chemical Engineering and Materials Science, and BioTechnology Institute, University of Minnesota, 240 Gortner Laboratory, 1479 Gortner Ave, St. Paul. MN 55108. USA

Carotenoide (z.B. DPL und DPA) sind Licht-absorbierende Pigmente, schützen vor UV-Strahlung, regulieren die Fluidität von Membranen, und wirken antioxidierend.

Sie werden als Nahrungszusatzstoffe, Pharmazeutika und als Lebensmittelfarbstoffe verwendet.

Ziel des Projekts: erhöhe die Carotenoid-Synthese in E.coli

Metabolisches Netzwerk für rekombinantes *E.coli*



Glucose

58 metabolische Reaktionen,

22 reversibel

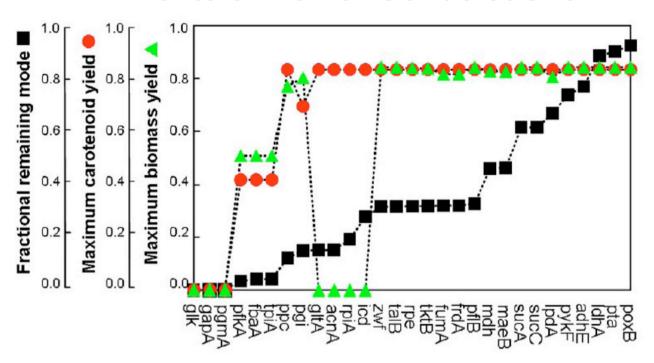
36 irreversibel

57 Metabolite

29532 EFMs

In 5923 EFMs ist die Produktion von Biomasse und DPA gekoppelt.

Effekte einzelner Gendeletionen



Resulte für virtuelle Gen-Kockouts (Abzählen der EFMs und Berechnung der CRT- und Biomasse-Produktion gemäß Stöchiometrien).

Wähle Targetgene, deren Knockouts noch eine möglichst hohe Ausbeute an CRT behalten sowie eine sinnvolle Biomasse-Produktion, jedoch möglichst viele EFMs eliminiert werden.

Effekte einzelner Gendeletionen

Strain	Total modes	Aerobic modes	Anaerobic modes	Predicted CRT yield ^a
Wild-type	29,532	24,155	5377	0.0-426
$\Delta ldhA$	15,662	13,405	2257	0.0-426
$\Delta ldhA\Delta frdA$	8573	7810	763	0.0-426
$\Delta ldhA\Delta frdA\Delta poxB$	7541	6861	680	0.0-426
$\Delta ldhA\Delta frdA\Delta poxB\Delta pta$	6171	5600	571	0.0-426
$\Delta ldhA\Delta frdA\Delta poxB\Delta pta\Delta adhE$	4099	4099	0	0.0-426
$\Delta IdhA\Delta frdA\Delta poxB\Delta pta\Delta adhE\Delta pykF$	2573	2573	0	0.0-426
$\Delta IdhA\Delta frdA\Delta poxB\Delta pta\Delta adhE\Delta pykF\Delta zwf$	375	375	0	0.0-426
$\Delta ldhA\Delta frdA\Delta poxB\Delta pta\Delta adhE\Delta pykF\Delta zwf\Delta maeB$	5	5	0	0.4-426

^a Yield is in mg-diapolycopendioic acid/g-glucose.

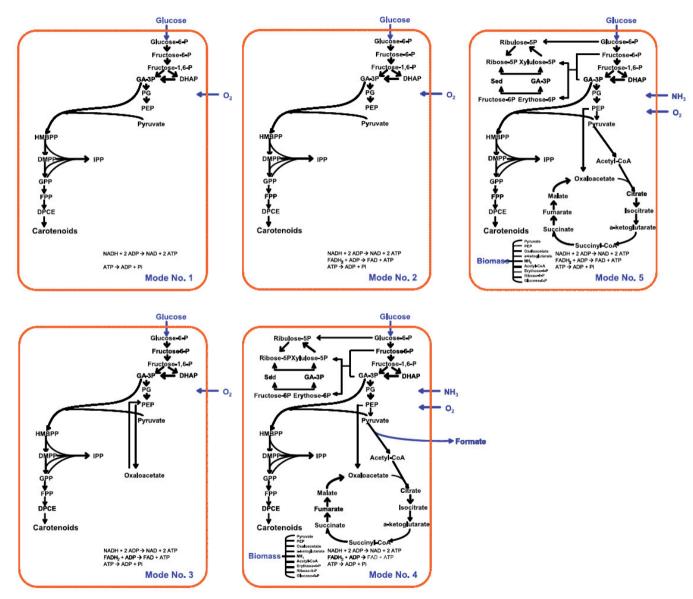
Deleted Reaction	leted Reaction Corresponding gene Enzyme		Pathway	
R9	pykF	Pyruvate kinase	Glycolysis	
R11	zwf	Glucose-6-phosphate-1-dehydrogenase	Pentose phosphate	
R22	frdA	Fumarate reductase	Fermentation	
R28	таеВ	Malate dehydrogenase	Anapleurotic	
R31	poxB	Pyruvate oxidase	Fermentation	
R32	ldhA	Lactate dehydrogenase	Fermentation	
R34	adhE	Alcohol dehydrogenase	Fermentation	
R35	pta	Phosphate acetyltransferase	Fermentation	

Als optimale Lösung vorhergesagt:

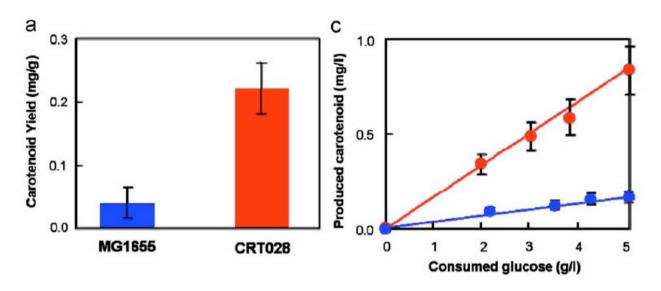
der Knockout von 8 Genen sollte zur Überproduktion von DPL und DPA führen.

nur 5 EFMs verbleiben für das restliche Netzwerk.

Verbleibende EFMs



Experimentelle Verifikation: erhöhte Carotenoid-Ausbeute



Mutante wächst langsamer, aber CRT-Produktion ist vierfach erhöht.

	MG1655/ pACMNOx	CRT028/ pACMNOx
Growth rate (/h) Carotenoid production (mg/l) Carotenoid yield (mg carotenoid/g glucose) Specific production (mg carotenoid/g cell dry weight-h)	$\begin{array}{c} 0.17 \pm 0.02 \\ 0.19 \pm 0.02 \\ 0.04 \pm 0.00 \\ 0.01 \pm 0.00 \end{array}$	$\begin{array}{c} 0.13 \pm 0.01 \\ 0.83 \pm 0.20 \\ 0.17 \pm 0.04 \\ 0.10 \pm 0.02 \end{array}$

Unrean et al. Metabol Eng 12, 112-122 (2010

9. Vorlesung WS 2012/13 Softwarewerkzeuge 29

Protein-Interaktionsnetzwerk in S. cerevisiae

Abb. zeigt das Protein–Protein Interaktionsnetzwerk in *Saccharomyces cerevisiae*, basierend auf **yeast two-hybrid** Experimenten.

Knoten: einzelne Proteine

CITIZOTTO I TOTOTTO

Verbindungen/Kanten: physikalische Interaktionen.

Das Netzwerk wird von einigen Knoten mit sehr vielen Verbindungen zusammengehalten. Diese nennt man **Hubs**.

Dieser größte Cluster enthält 78% aller Proteine.

Rot: Gendeletion ist tödlich

Grün: Gendeletion ist nicht tödlich

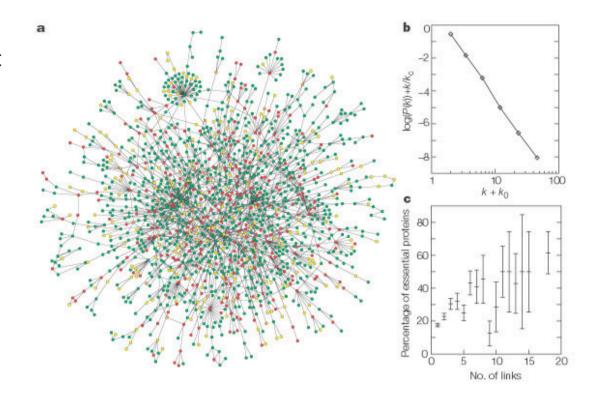
Gelb: Effekt der Gendeletion ist unbekannt.



Barabasi & Oltvai, Nature Rev Gen 5, 101 (2004)

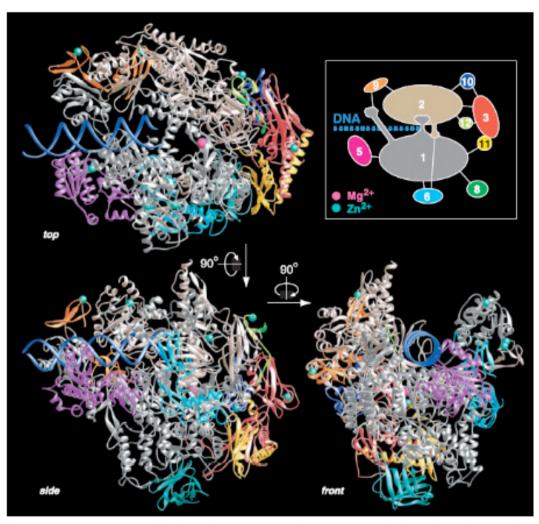
Welche Proteindeletionen sind tödlich?

- (b) Die Häufigkeit von Hub-Proteinen mit *k* Interaktionen nicht exponentiell ab (wie in einem Zufallsnetzwerk mit Poisson-Verteilung der Verknüpfungen erwartet), sondern gemäss einem Potentzgesetz.
- (c) Die Deletion von Genen, die für Hub-Proteine kodieren, die viele Links besitzen, ist eher tödlich als die Deletion von Proteinen, die mit wenigen anderen Proteinen interagieren.



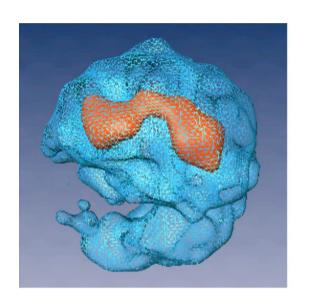
H. Jeong, S. P. Mason, A.-L. Barabási and Z. N. Oltvai Nature 411, 41-42

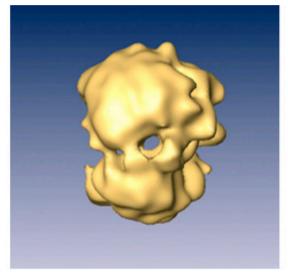
Proteinkomplexe: RNA Polymerase II

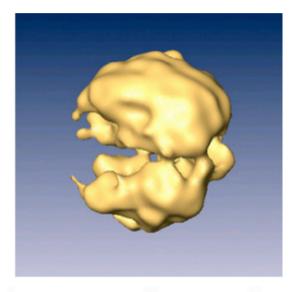


Cramer et al., Science 288, 640 (2000)

Proteinkomplexe: Spliceosom

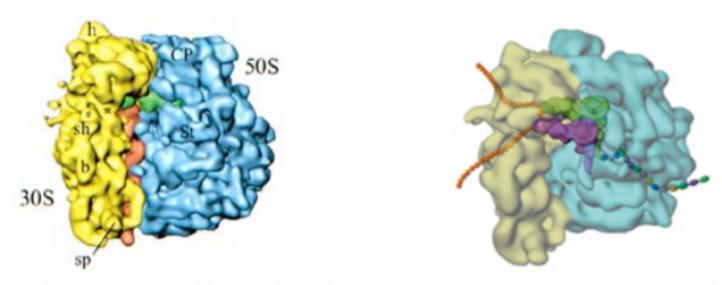






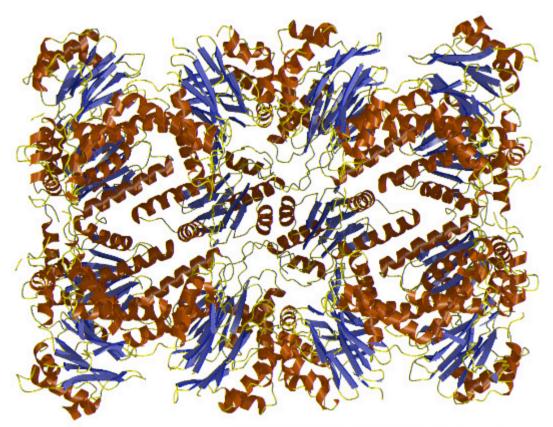
http://www.weizmann.ac.il/

Proteinkomplexe: Ribosom



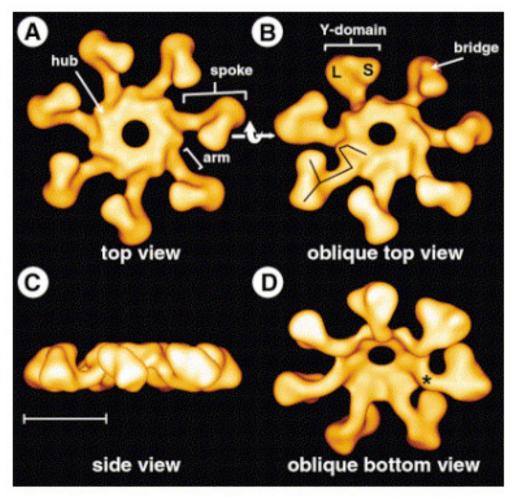
http://www.millerandlevine.com/chapter/12/cryo-em.html

Proteinkomplexe: Proteasom



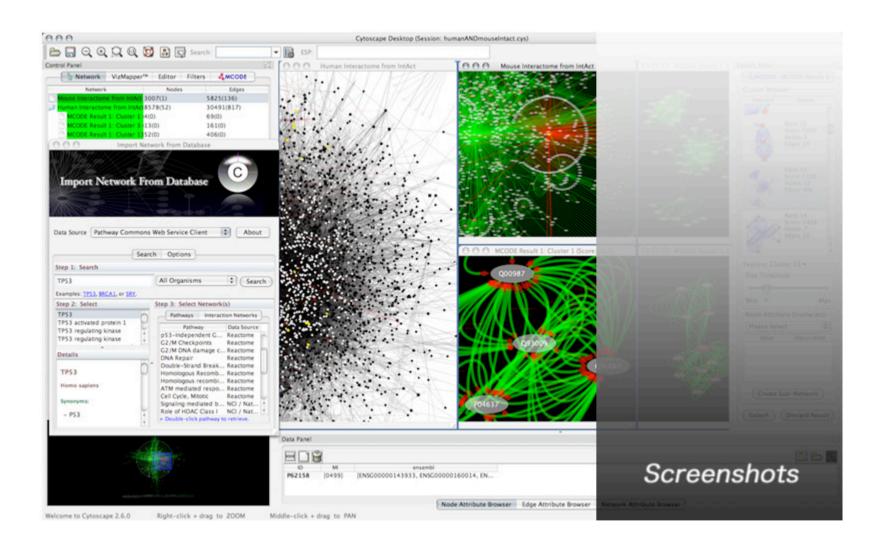
http://www.biochem.mpg.de

Proteinkomplexe: Apoptosom



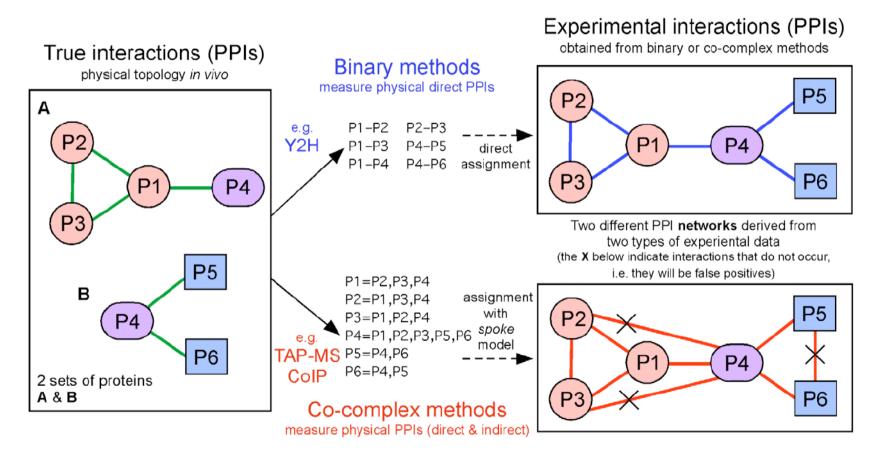
Acehan et al. Mol. Cell 9, 423 (2002)

Proteininteraktionsnetzwerke



Proteininteraktionsnetzwerke

Verschiedene experimentelle Techniken messen verschiedene Eigenschaften von Proteinkomplexen.



Datenbanken zu Proteininteraktionen

Table 1. Description of PPI databases and repositories.

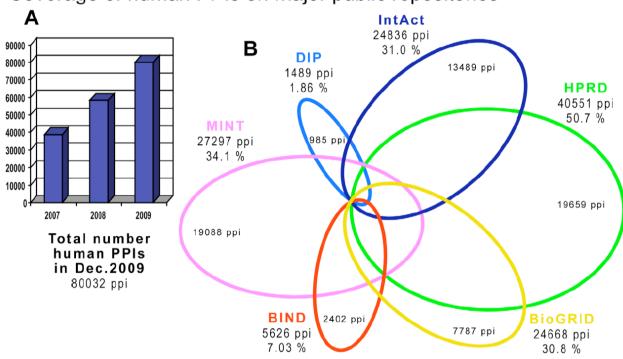
Acronym	Database Full Name and URL	PPI Sources	Type of MI	Species	n Proteins (Dec. 2009)	n Interactions (Dec. 2009)
Primary Datab	ases: PPI experimental data (curated from specific SSc & LSc publish	ed studies)				
BIND	Biomolecular Interaction Network Database, http://bond. unleashedinformatics.com/	Ssc & Lsc published studies (literature-curated)	PPIs & others	All	[31,972]	[58,266]
BioGRID	Biological General Repository for Interaction Datasets, http://www.thebiogrid.org/	Ssc & Lsc published studies (literature-curated)	PPIs & others	All	[28,717]	[108,691]
DIP	Database of Interacting Proteins, http://dip.doe-mbi.ucla.edu/dip/	Ssc & Lsc published studies (literature-curated)	Only PPIs	All	20,728	57,683
HPRD	Human Protein Reference Database, http://www.hprd.org/	Ssc & Lsc published studies (literature-curated)	Only PPIs	Human	27,081	38,806
IntAct	IntAct Molecular Interaction Database, http://www.ebi.ac.uk/intact/	Ssc & Lsc published studies (literature-curated)	PPIs & others	All	[60,504]	[202,826]
MINT	Molecular INTeraction database, http://mint.bio.uniroma2.it/mint/	Ssc & Lsc published studies (literature-curated)	Only PPIs	All	30,089	83,744
MIPS-MPact	MIPS protein interaction resource on yeast, http://mips.gsf.de/ genre/proj/mpact/	Derived from CYGD	Only PPIs	Yeast	1,500	4,300
MIPS-MPPI	MIPS Mammalian Protein-Protein Interaction Database, http://mips.gsf.de/proj/ppi	Ssc published studies (literature-curated)	Only PPIs	Mammalian	982	937
Meta-Database	s: PPI experimental data (integrated and unified from different pub	lic repositories)				
APID	Agile Protein Interaction DataAnalyzer, http://bioinfow.dep.usal.es/apid/	BIND, BioGRID, DIP, HPRD, IntAct, MINT	Only PPIs	All	56,460	322,579
MPIDB	The Microbial Protein Interaction Database, http://www.jcvi.org/mpidb/	BIND, DIP, IntAct, MINT, other sets (exp & litcurated)	Only PPIs	Microbial	7,810	24,295
PINA	Protein Interaction Network Analysis platform, http://csbi.ltdk. helsinki.fi/pina/	BioGRID, DIP, HPRD, IntAct, MINT, MPact	Only PPIs	All	[?]	188,823
Prediction Dat	abases: PPI experimental and predicted data ("functional interaction	ns", i.e., interactions lato sensu derived from different	types of data)			
MiMI	Michigan Molecular Interactions, http://mimi.ncibi.org/MimiWeb/	BIND, BioGRID, DIP, HPRD, IntAct, & nonPPI data	PPIs & others	All	[45,452]	[391,386]
PIPs	Human PPI Prediction database, http://www.compbio.dundee. ac.uk/www-pips/	BIND, DIP, HPRD, OPHID, & nonPPI data	PPIs & others	Human	[?]	[37,606]
OPHID	Online Predicted Human Interaction Database, http://ophid. utoronto.ca/	BIND, BioGRID, HPRD, IntAct, MINT, MPact, & nonPPI data	PPIs & others	Human	[?]	[424,066]
STRING	Known and Predicted Protein-Protein Interactions, http://string.embl.de/	BIND, BioGRID, DIP, HPRD, IntAct, MINT, & nonPPI data	PPIs & others	All	[2,590,259]	[88,633,860]
UniHI	Unified Human Interactome, http://www.mdc-berlin.de/unihi/	BIND, BioGRID, DIP, HPRD, IntAct, MINT, & nonPPI data	PPIs & others	Human	[22,307]	[200,473]

The table divided in three sections: **primary databases**, which include PPIs from large- and small-scale (Lsc & Ssc) experimental data that are usually obtained from curation of research articles (8 resources included: BIND, BioGRID, DIP, HPRD, IntAct, MINT, MIPS-MPACT, MIPS-MPACT, MIPS-MPPI); **meta-databases**, which include PPIs derived from integration and unification of several primary repositories (3 resources: APID, MPIDB, PINA); **prediction databases**, which include PPIs from experimental analyses together with predicted PPIs obtained from the analyses of heterogenous biological data (5 resources: MiMI, PIPs, OPHID, STRING, UniHI). The table shows the total number of proteins and interactions that were reported by each repository in December 2009 (as far as we could see in the respective Web site). The numbers are in brackets [] when the repository includes PPIs and other types of interactions (e.g., protein-ligand interactions or for the case of prediction databases nonPPI data). The question mark [?] indicates that the number of distinct proteins included is such repository could not be found in the Web. doi:10.1371/journal.pcbi.1000807.t001

Proteininteraktionsnetzwerke

Human Interactome

Coverage of human PPIs on major public repositories



Der Überlapp zwischen verschiedenen Datenbanken ist leider gering.

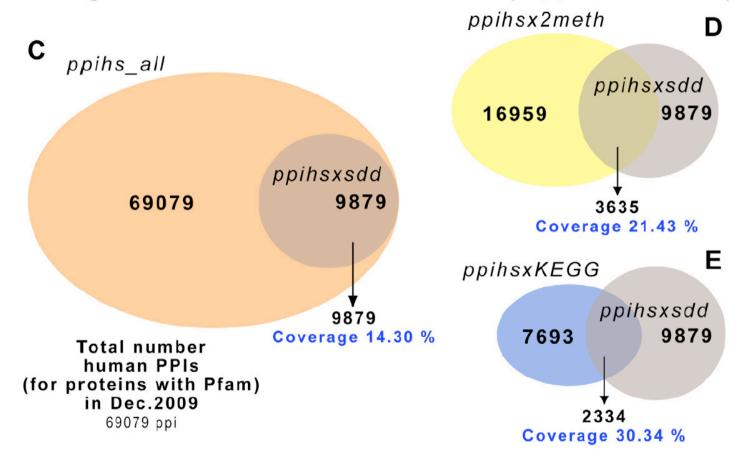
Man kennt bisher etwa 80 000 Interaktionen von menschlichen Proteinen.

Man schätzt, dass es in *S. cerevisae* (6000 Gene) ca. 20 000 – 30 000 Proteininteraktionen gibt.

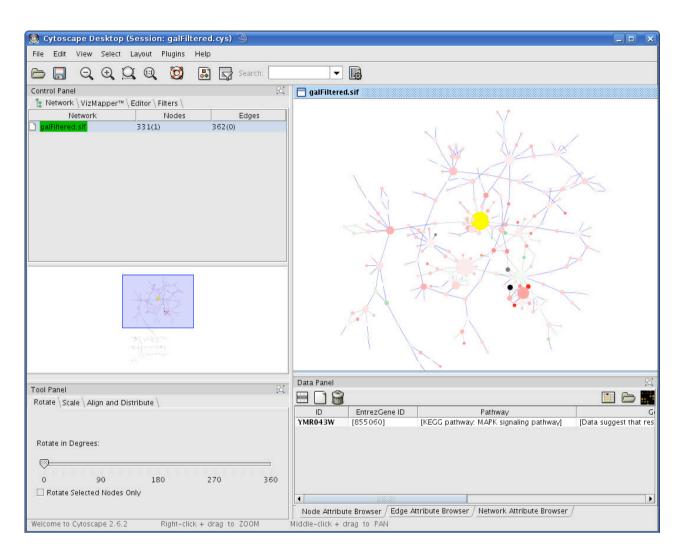
D.h. im Mensch wird eine erheblich höhere Anzahl erwartet.

Proteininteraktionsnetzwerke

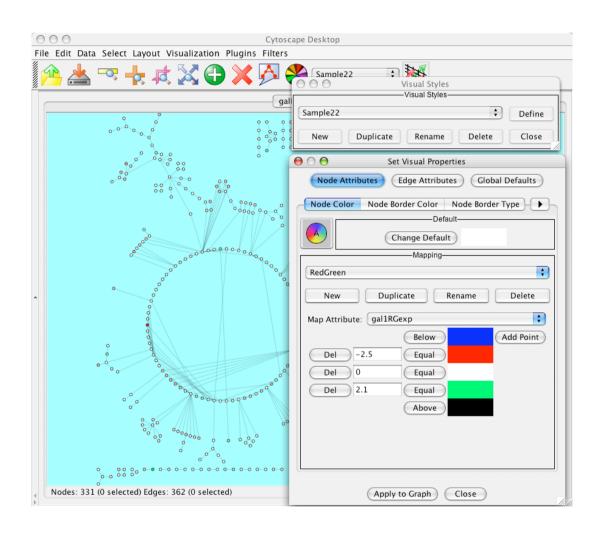
Coverage of human PPIs with 3D structure (only proteins with Pfam)



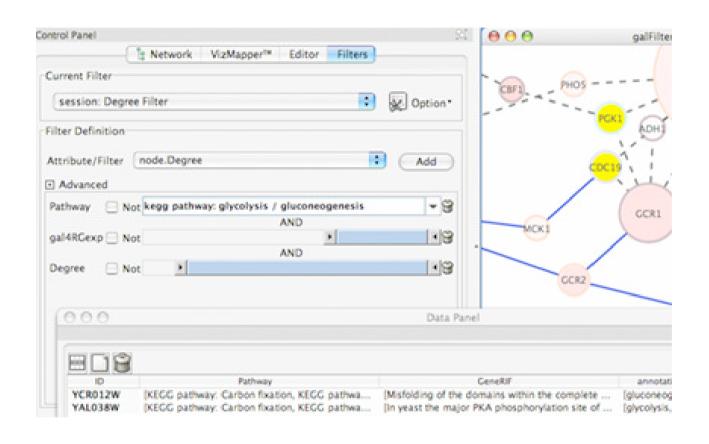
Proteininteraktionsnetzwerke - Visualisierung



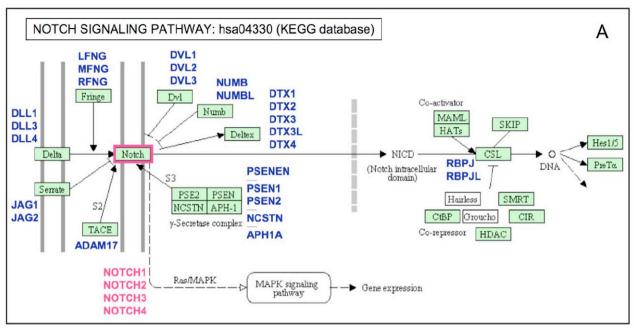
Proteininteraktionsnetzwerke - Visualisierung



Proteininteraktionsnetzwerke - Visualisierung

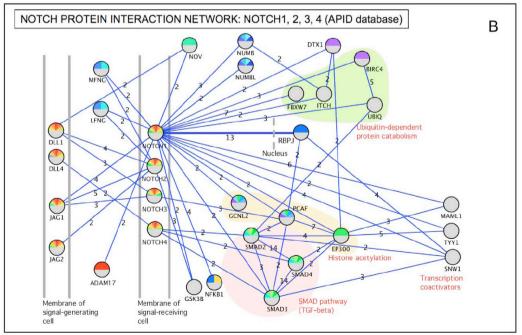


Proteininteraktionsnetzwerke



De Las Rivas, PLOS Comp Biol. 6, e1000807 (2010)

9. Vorlesung WS 2012/13



Auffinden aktiver Teilnetze/Pfade

