PROJECT TITLE

Microbial community modeling for the production of 'designer' yogurt

PROJECT DURATION	TOTAL REQUESTED FUNDING	TOTAL COSTS
36 Months (01 / 2018 to 12 / 2020)	946.000 €	1.023.000 €

CONSORTIUM

P 1	Prof Bas Teusink Vrije Universiteit Amsterdam Molecular Cell Biology	De Boelelaan 1085, NL-1081HV Amsterdam Netherlands	b.teusink@vu.nl Tel.: 0031 621835201 http://www.teusinkbruggemanlab.nl
P 2	Prof Ursula Kummer Heidelberg University Modeling of Biological Processes	Im Neuenheimer Feld 267, 69120 Heidelberg Germany	ursula.kummer@bioquant.uni-heidelberg.de Tel.: 0049 6221 5451278 Fax: 0049 6221 5451483 http://www.uni-heidelberg.de
P 3	Prof Ralf Takors University of Stuttgart Institute of Biochemical Engineering	Allmandring 31, 70569 Stuttgart Germany	takors@ibvt.uni-stuttgart.de Tel.: 0049 71168564535 Mobile: 0049 15119616717 http://www.ibvt.uni-stuttgart.de/
P 4	Dr Petri-Jaan Lahtvee University of Tartu Institute of Technology	Nooruse 1, 50411 Tartu Estonia	lahtvee@ut.ee Tel.: 00372 5088117 Mobile: 00372 5088117 http://www.ut.ee
P 5	Dr Ana Rute Neves Chr. Hansen A/S Bacterial Physiology and Improvement (BactoPhI)	Bøge Allé 10-12, 2970 Hørsholm Denmark	DKRUNS@chr-hansen.com Tel.: +45 60431012 http://www.chr-hansen.com

Yoghurt consortium
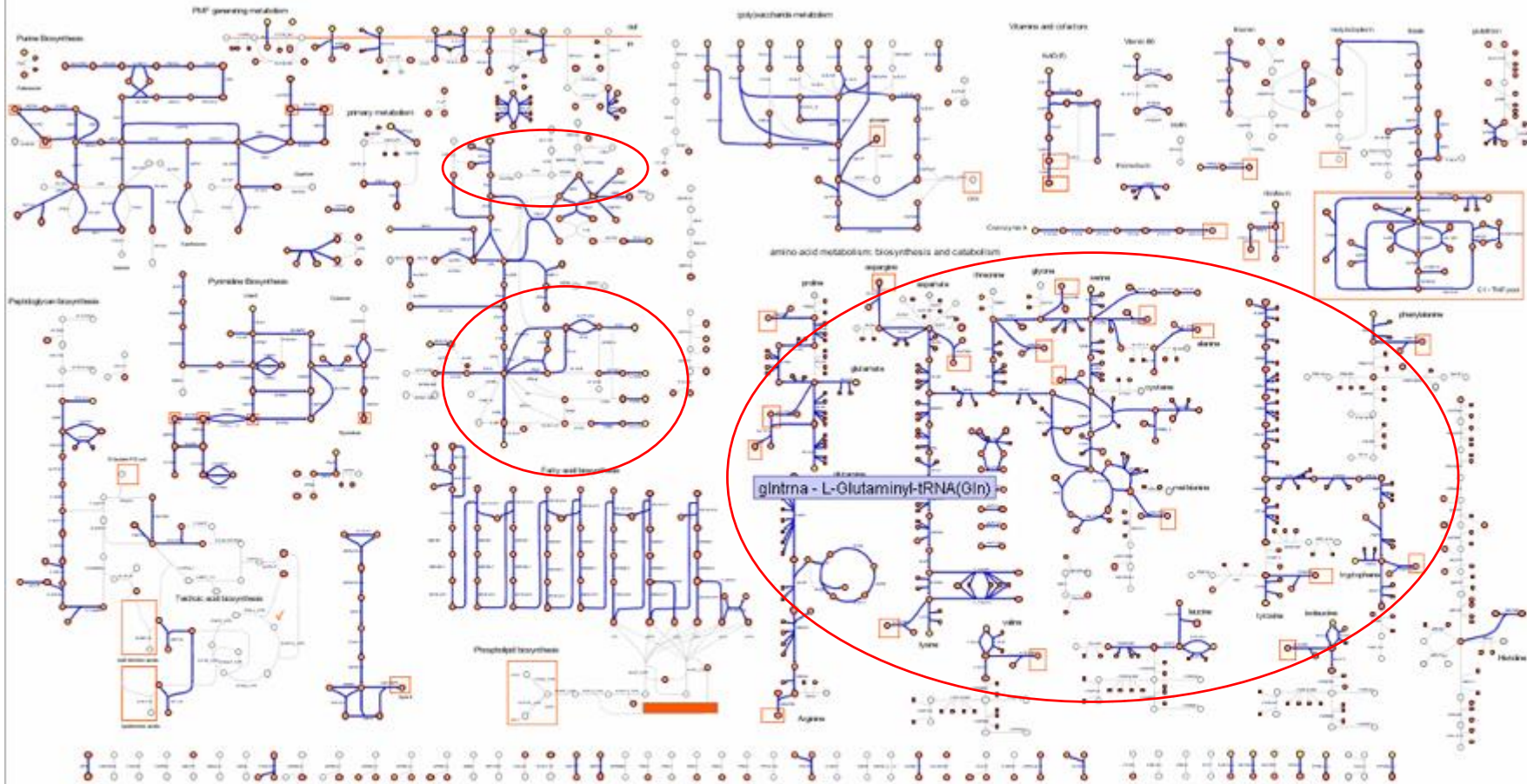

L. bulgaricus   S. thermophilus

Lactobacillus bulgaricus
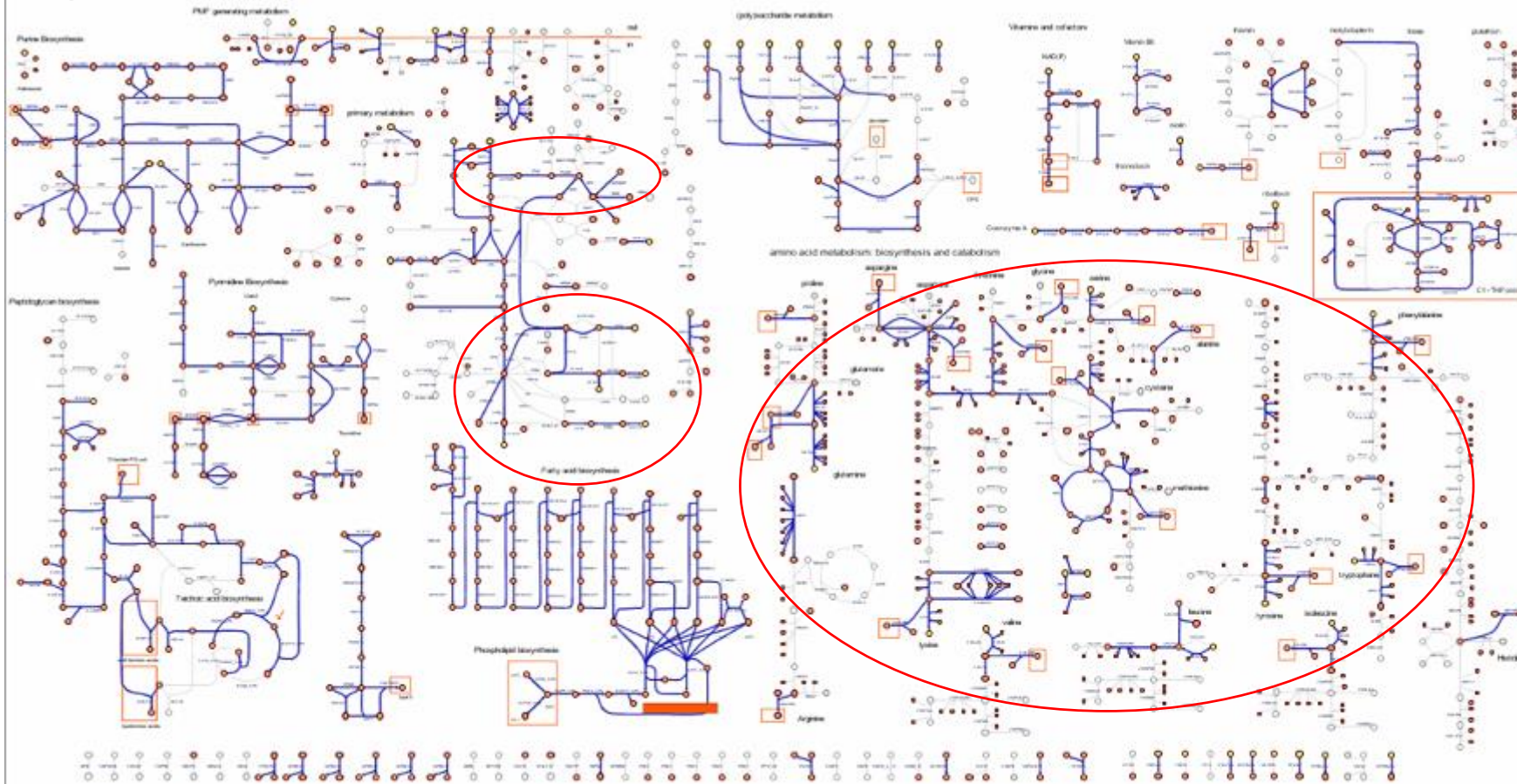*Streptococcus thermophilus*

Key attributes: acidification rate, acidity (yield), robustness, flavours (primary metabolism compounds: acetaldehyde, butyrate), texture (exopolysaccharides)

S. Thermophilus CNRZ1066

-reduced carbohydrate metabolism
-extensive AA metabolism

L. bulgaricus ATCC BAA-365

-carbohydrate metabolism more flexible
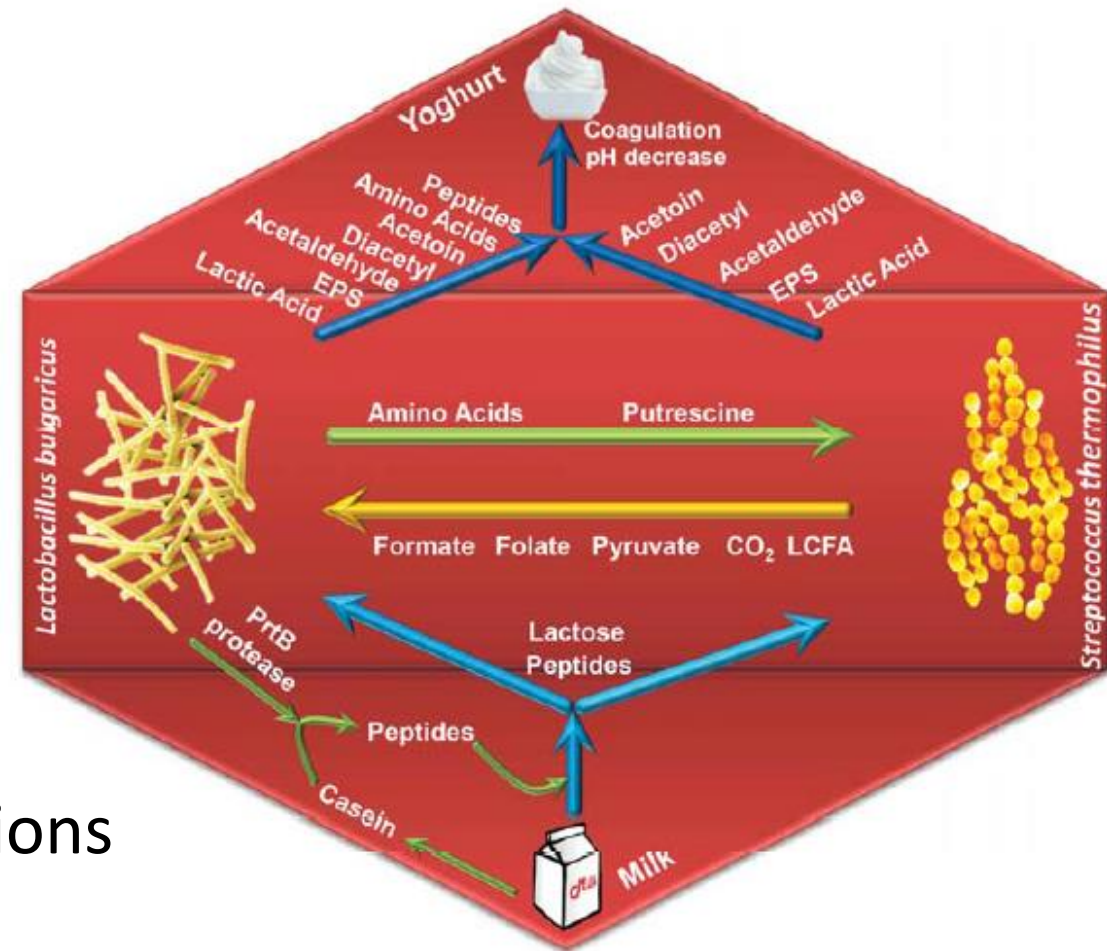-many gaps in AA metabolism

Yogurt... a model for metabolic interactions

The problem for industry:

100's of different strains

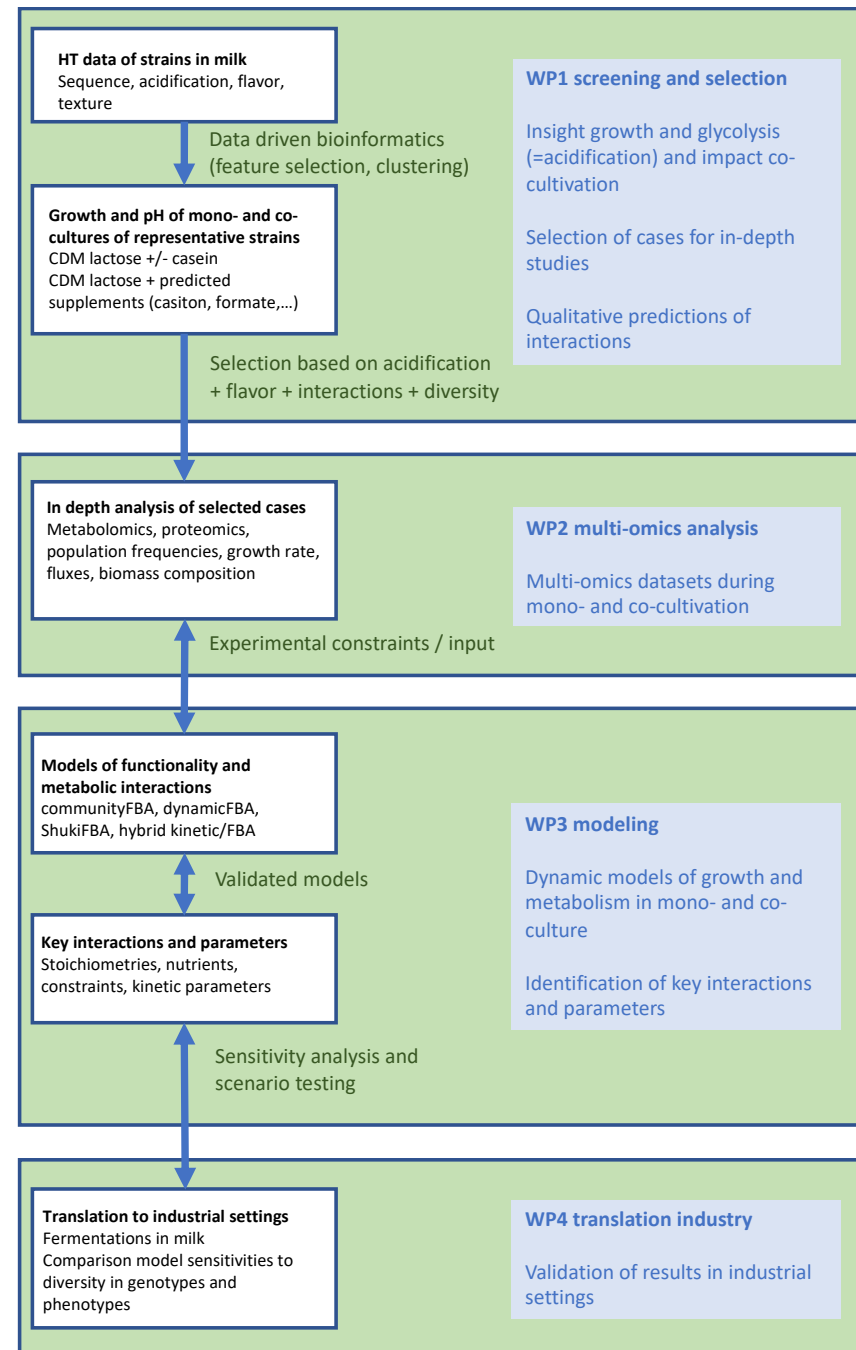=> Impossible to test all different combinations
=> Rational approach warranted

Key challenge:

- Can we predict community performance on the basis of properties of individual strains?

  - Data driven approach:
    - How much variation is there?
    - Which features can explain this?
    - Can we underpin this with molecular mechanisms?

  - Mechanistic approach for selected cases:
    - Genome-scale models
    - Community modeling
    - Dynamic modeling

The project plan:

HT data of strains in milk
Sequence, acidification, flavor, texture

Data driven bioinformatics (feature selection, clustering)

Growth and pH of mono- and co-cultures of representative strains
CDM lactose +/- casein
CDM lactose + predicted supplements (casiton, formate,…)

Selection based on acidification + flavor + interactions + diversity

WP1 screening and selection

Insight growth and glycolysis (=acidification) and impact co-cultivation

Selection of cases for in-depth studies

Qualitative predictions of interactions

In depth analysis of selected cases
Metabolomics, proteomics, population frequencies, growth rate, fluxes, biomass composition

Experimental constraints / input

WP2 multi-omics analysis

Multi-omics datasets during mono- and co-cultivation

Models of functionality and metabolic interactions
communityFBA, dynamicFBA, ShukiFBA, hybrid kinetic/FBA

Validated models

Key interactions and parameters
Stoichiometries, nutrients, constraints, kinetic parameters

Sensitivity analysis and scenario testing

WP3 modeling

Dynamic models of growth and metabolism in mono- and co-culture

Identification of key interactions and parameters

Translation to industrial settings
Fermentations in milk
Comparison model sensitivities to diversity in genotypes and phenotypes

WP4 translation industry

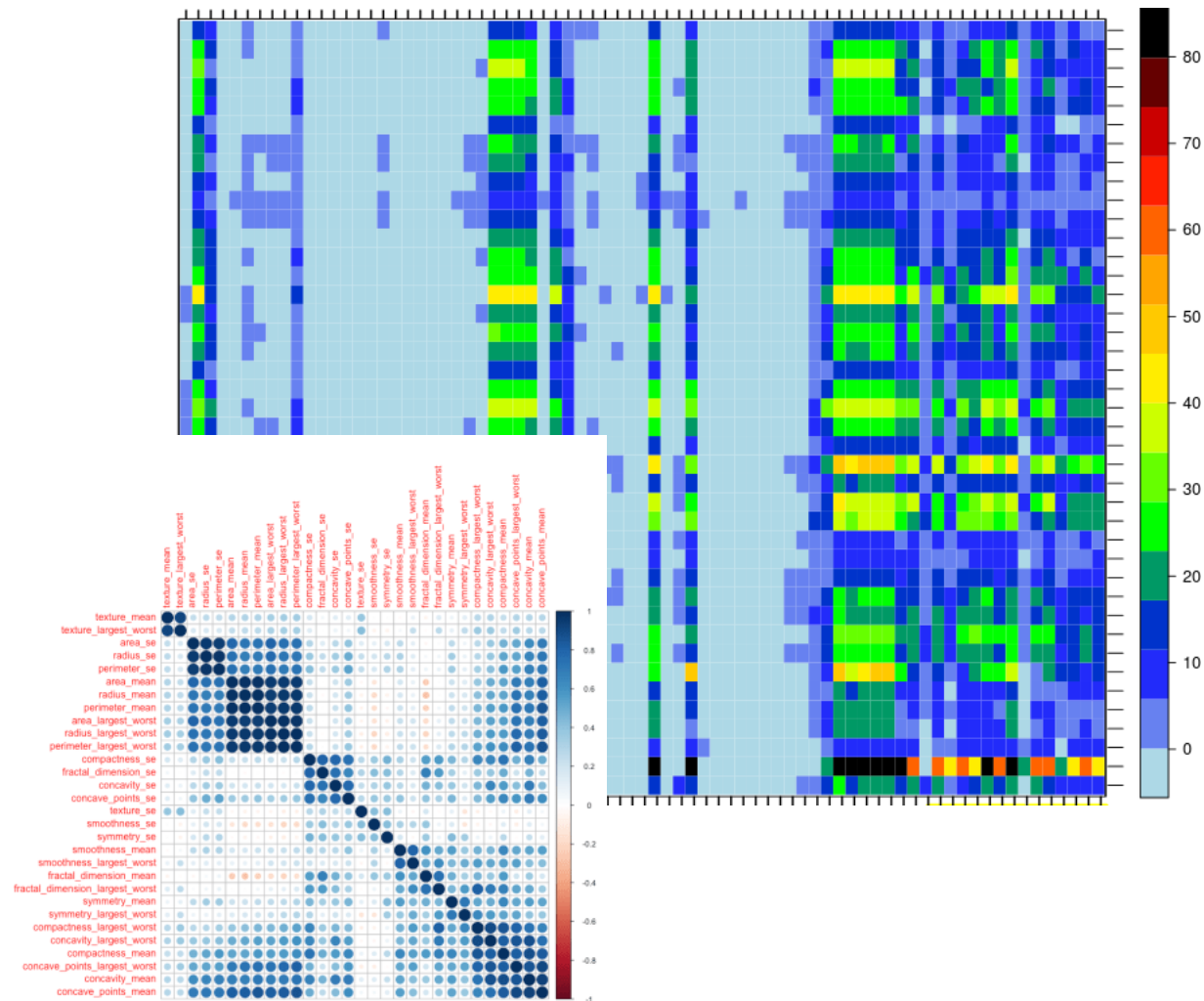Validation of results in industrial settings

WP1: high throughput data

- Data consolidation: what is available?

- New growth/acidification experiments: conditions, number of strains, what can be done?

- Growth media, determination of biomass

- Strain transfer to different labs
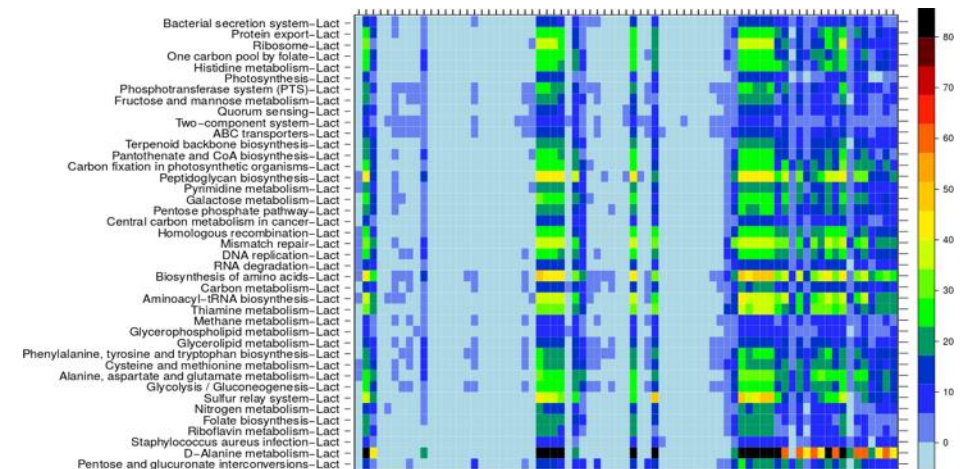
Feature selection from data

Which features explain which differences between samples (strain)?

- Presence/absence of genes (gene-trait matching)
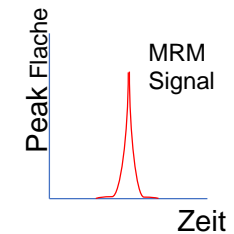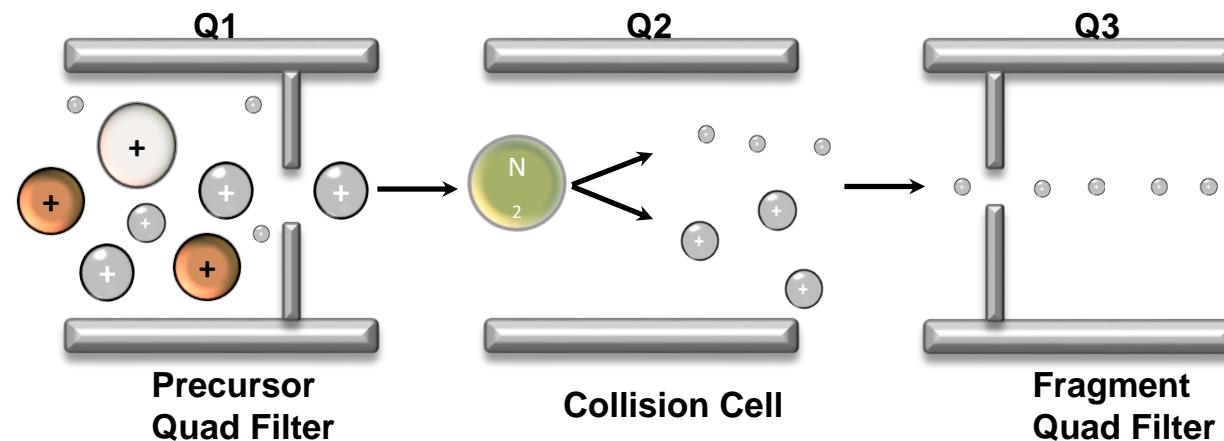- Promotor sequences, TF binding sites
- Codon usage
- …

WP2: omics analyses

For a selected number of strains with high diversity

- Proteomics

- Metabolomics

- Flux analysis in communities

LC/MS-based intracellular analytics

- **LC-ESI-TripleQuad-MS/MS-System**

- **HILIC** (*H*ydrophilic *I*nteraction *L*iquid *C*hrom.)

→ bicratic partition chromatography for polar analytes

→ MRM (multiple reaction monitoring)

→ **broad spectrum: >100 metabolites (CM, anabolic routes)**

→ **high sensitivity, LOQ (linearity range) > 10 nM**

Agilent 6410B

	Metabolite	Polarity mode	Linearity range a) Min (nM)	Max (nM)	R²	Linearity range b) Min (nM)	Max (nM)	R²	Method detection limit a) (nM)	b) (nM)	a) fmol (on column)	b) fmol (on column)	MS/MS 12C Precursor ion	Daughter ion	MS/MS 13C Precursor ion	Daughter ion	CE	Fragmentor
Citric acid cycle	aKB	(-)	20	100000	0.9938	20	5000	0.9991	20.54	4.94	0.1027	0.0247	145	101	150	105	4	70
	Fum	(-)	20	10000	0.9998	200	5000	0.9990	27.76	331.71	0.1388	1.6586	115	71	119	74	3	70
	Suc	(-)	500	10000	0.9983	200	5000	0.9980	17.62	129.03	0.0881	0.6452	117	73	121	76	9	82
	Mal	(-)	20	10000	0.9902	10	5000	0.9940	2.23	6.49	0.0112	0.0325	133	115	137	119	7	80
	aKG	(-)	20	10000	0.9978	20	5000	0.9915	7.6	24.06	0.0380	0.1203	145	101	150	105	4	70
	Cit	(-)	50	10000	0.9966	-	-	-	439	-	2.1950	-	191	111	197	116	10	90
	IsoCit	(-)	20	10000	0.9949	200	5000	0.9908	21.52	426.25	0.1076	2.1313	191	173	197	179	6	90
	cis-Aco	(-)	20	500	0.9999	10	50000	0.9985	3.91	18.83	0.0196	0.0942	173	85	179	89	10	60
Pentose phosphate pathway	Xu5P	(-)	50	3000	0.9933	-	-	-	30.1	67.5	0.1505	0.3375	229	97	234	97	4	70
	6PG	(-)	1000	5000	0.9903	500	20000	0.9963	744.31	1294.19	3.7216	6.4710	275	97	281	97	12	100
	E4P	(-)	-	-	-	2000	5000	0.9997	-	1713	-	8.5650	199	97	203	97	4	70
	R5P	(-)	20	200000	0.9931	-	-	-	19.73	-	0.0987	-	229	97	234	97	6	90
	Ru5P	(-)	20	200000	0.9918	-	-	-	30.33	-	0.1517	-	229	97	234	97	6	90
	R(u)5P Pool	(-)	-	-	-	50	20000	0.9956	-	67.5	-	0.3375	229	97	234	97	6	90
Emdenhof-Meyer-Parnas-Weg	F6P	(-)	20	200000	0.9958	50	5000	0.9977	55.46	88.22	0.2773	0.4411	259	97	265	97	10	100
	G6P	(-)	20	200000	0.9953	20	5000	0.9961	43.84	40.26	0.2192	0.2013	259	97	265	97	12	100
	FbP	(-)	10												345	97	15	100
	DHAP	(-)	10												172	97	4	60
	PEP	(-)	2(												170	79	14	70
	2(3)PG	(-)	10												-	-	-	-
	2PG	(-)	10												188	79	38	90
	3PG	(-)	10												188	97	14	80
Energy	NAD	(+)	2(												687	670	65	135
	NADP	(+)	2(												765	141	58	135
	AMP	(-)	2(												356	79	60	135
	ADP	(-)	5(												436	79	98	135
	ATP	(-)	10												516	79	108	135
	GMP	(-)	2(												372	79	60	135
	GDP	(-)	2(												452	79	108	135
	GTP	(-)	20												532	79	108	135
	CMP	(-)	5(												331	79	60	135
	CDP	(-)	2(												411	79	98	135
	CTP	(-)	10												491	79	108	135
	UMP	(-)	5(												332	79	60	135
	UDP	(-)	2(												412	79	98	135
	UTP	(-)	10												492	79	108	135
	cAMP	(-)	2(												338	139	22	135
Non proteinogenic amino acids	AIBA	(+)	5												-	-	10	75
	GABA	(+)	5												-	-	10	75
	L-Hom-Ser	(+)	2(												124	77	10	75
	O-Ac-Ser	(+)	2(												153	62	20	60
	OSHS	(+)	5												228	106	8	70
	LL-Cysta	(+)	5												230	91	27	85
Proteinogenic amino acids	Gly	(+)	10												78	31	8	70
	L-Ala	(+)	2(												93	46	10	70
	L-Ser	(+)	100	5000	0.9988	100	5000	0.9969	9.94	76.7	0.0497	0.3835	106	60	109	62	10	75
	L-Pro	(+)	10	5000	0.9992	10	5000	0.9969	9.1	31.09	0.0455	0.1555	116	70	121	74	17	75
	L-Val	(+)	20	5000	0.9991	100	5000	0.9964	26.4	49.74	0.1320	0.2487	118	72	123	76	10	75
	L-Thr	(+)	20	5000	0.9993	100	5000	0.9968	16.65	97.61	0.0833	0.4881	120	74	124	77	9	75
	L-Leu	(+)	20	5000	0.9993	10	5000	0.9966	7.91	9.71	0.0396	0.0486	132	44	138	47	20	75
	L-Ile	(+)	20	5000	0.9990	10	5000	0.9967	17.15	1.84	0.0858	0.0092	132	44	138	47	20	75
	L-Asn	(+)	20	5000	0.9991	20	5000	0.9967	11.5	23.95	0.0575	0.1198	133	74	137	76	15	70
	L-Asp	(+)	20	5000	0.9889	100	5000	0.9972	8.83	192.02	0.0442	0.9601	134	74	138	76	13	75
	L-Gln	(+)	20	5000	0.9987	20	5000	0.9959	3.39	37.53	0.0170	0.1877	147	84	152	88	18	70
	L-Lys	(+)	20	5000	0.9987	100	5000	0.9933	20.74	36.28	0.1037	0.1814	147	84	153	89	17	75
	L-Glu	(+)	20	5000	0.9990	50	5000	0.9965	7.75	34.62	0.0388	0.1731	148	84	153	88	17	75
	L-Met	(+)	10	5000	0.9990	10	5000	0.9966	16.83	5.99	0.0842	0.0300	150	56	155	59	17	75
	L-Phe	(+)	10	5000	0.9989	10	5000	0.9966	10.18	7.13	0.0509	0.0357	166	120	175	128	12	75
	L-Arg	(+)	20	3000	0.9936	50	2000	0.9946	6.11	31.63	0.0306	0.1582	175	70	181	74	30	75
	L-Trp	(+)	20	5000	0.9991	20	5000	0.9983	65.52	24.99	0.3276	0.1250	205	188	216	199	4	75
	L-His	(+)	-	-	-	50	3000	0.9975	-	157.82	-	0.7891	156	110	162	115	13	85

quantification of > 100 metabolites of central metabolism, nucleotides, amino acids in negative and positive polarisation mode > 10 nM.

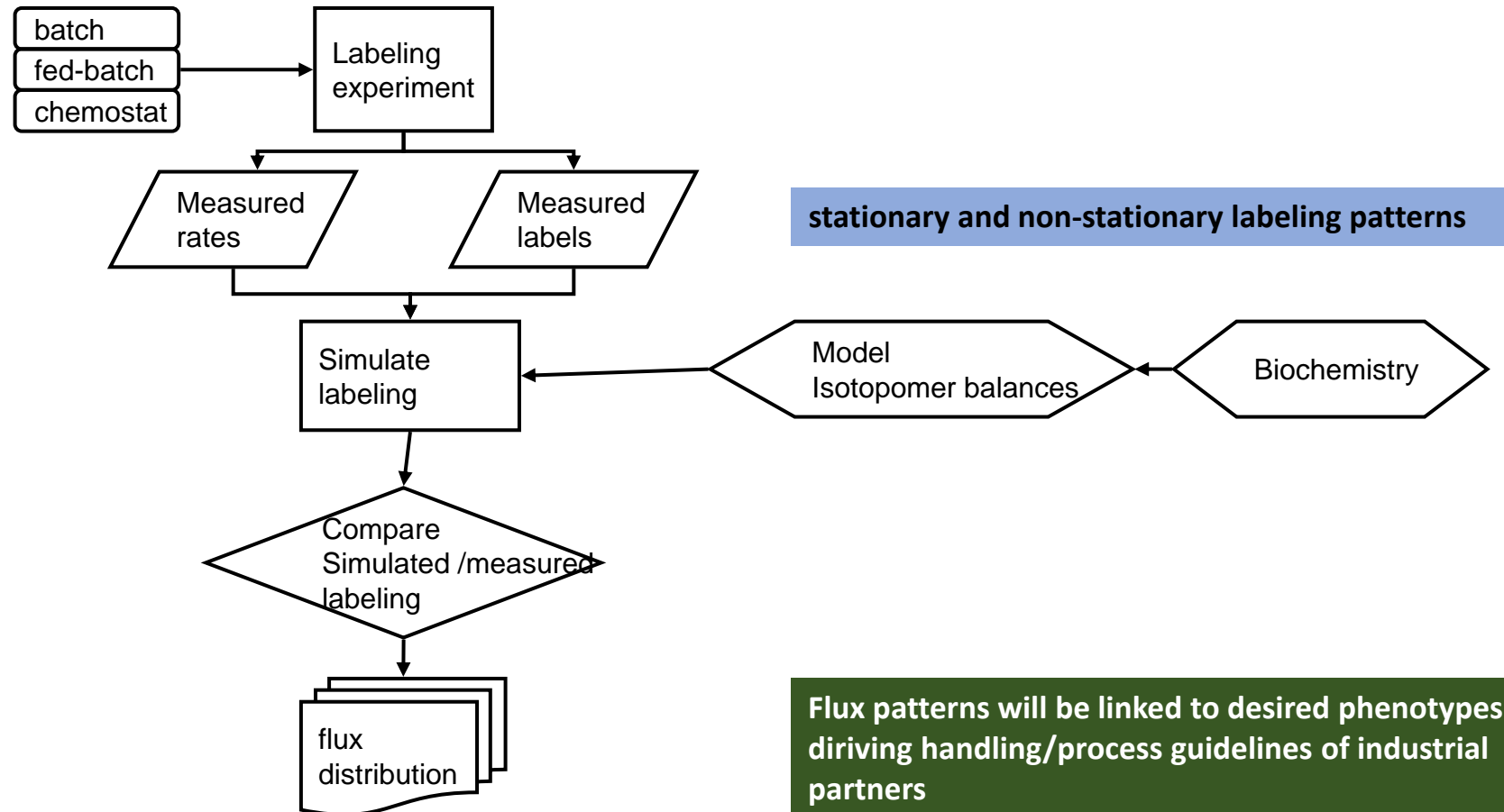**Metabolomics is a standard tool at IBVT today-**

Exploiting the Wealth of Metabolomics for Flux Analysis

1. mono cultures
2. mixed cultures

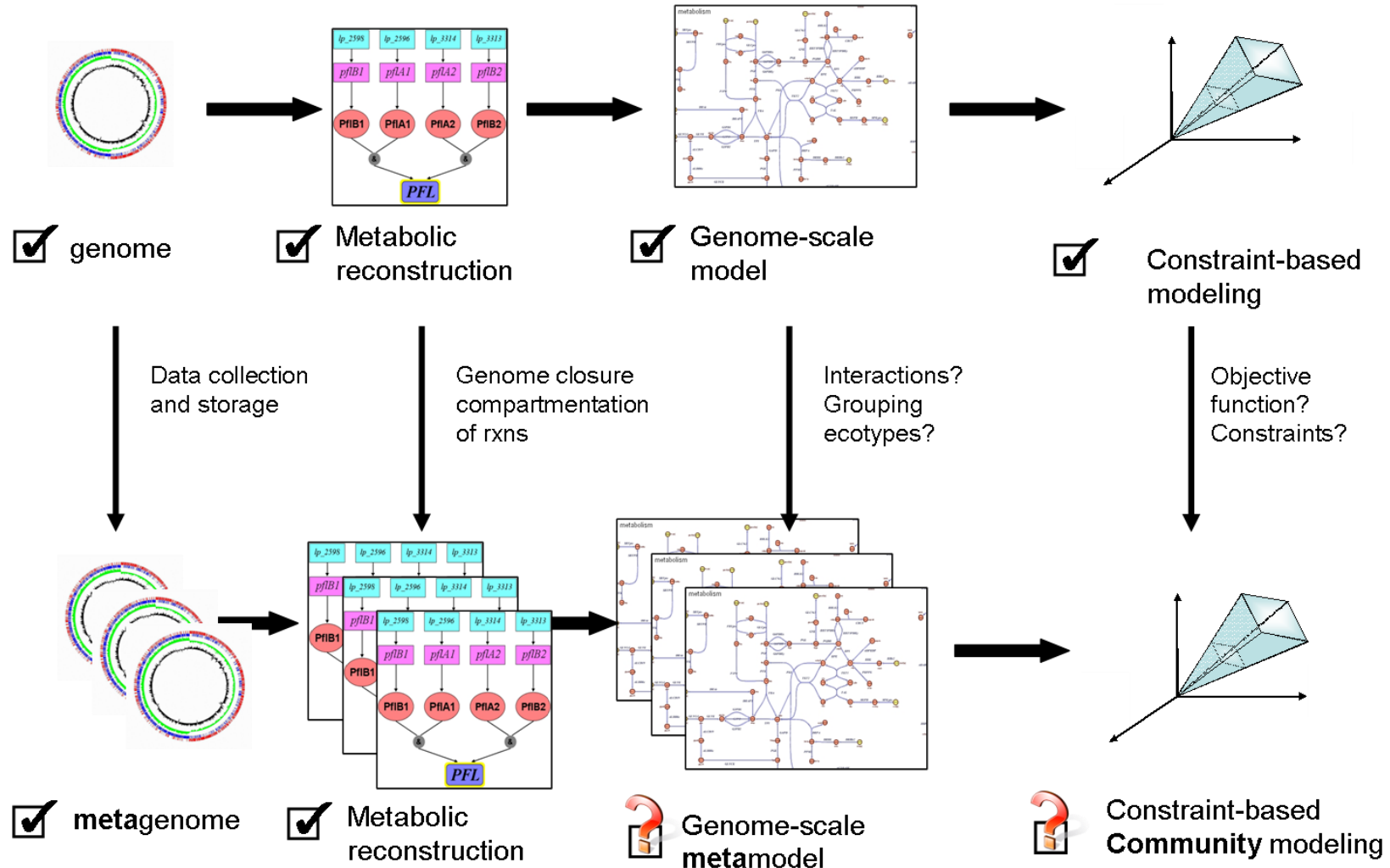balancing experimental data

Insilico Discovery™

batch
fed-batch
chemostat

Labeling experiment

Measured rates

Measured labels

stationary and non-stationary labeling patterns

Simulate labeling

Model Isotopomer balances

Biochemistry

Compare Simulated /measured labeling

flux distribution

Flux patterns will be linked to desired phenotypes for diriving handling/process guidelines of industrial partners

WP3: modeling approaches

- Community and dynamic FBA

- Constraints from proteomics

- Hybrid kinetic / stoichiometric models

- Predict return of investments - somehow

from monoculture to communities

☑ genome

☑ Metabolic reconstruction

☑ Genome-scale model

☑ Constraint-based modeling

Data collection and storage

Genome closure compartmentation of rxns

Interactions? Grouping ecotypes?

Objective function? Constraints?

☑ **meta**genome

☑ Metabolic reconstruction

❓ Genome-scale **meta**model

❓ Constraint-based **Community** modeling

Steady state: community FBA

Steady state only possible if:

$$\mu_i = \mu_j = \mu$$

FBA formulation:

$Max\ \mu$
$Subject\ to:$
$N_i\,q_i = 0$
$N_j\,q_j = 0$
$N_{exchange}\,J = 0$

$J:$ flux in mmol $L^{-1}\ h^{-1}$
$q:$ specific flux in mmol $h^{-1}\ gDW^{-1}$ $\Big\}$ $J = q \times X$
$X:$ biomass in gDW L^{-1}

Steady state: community FBA
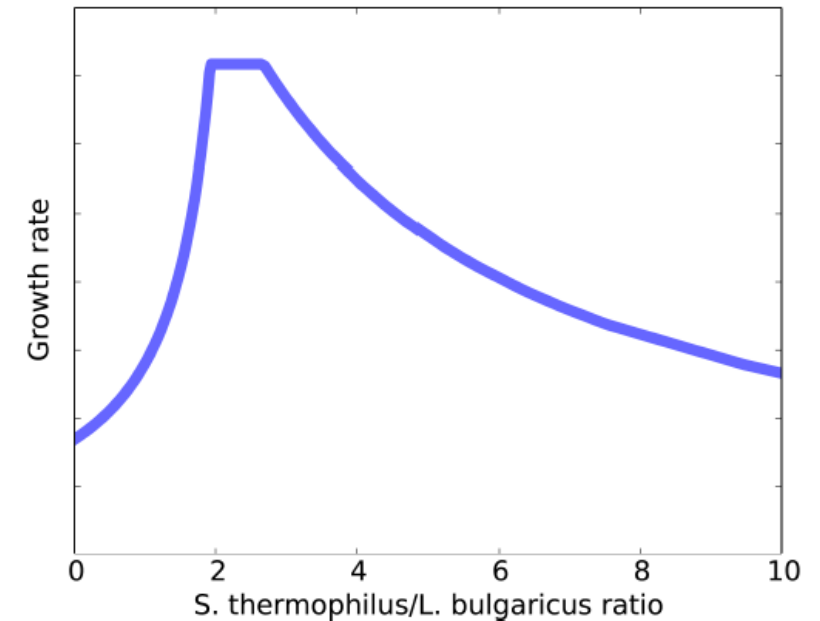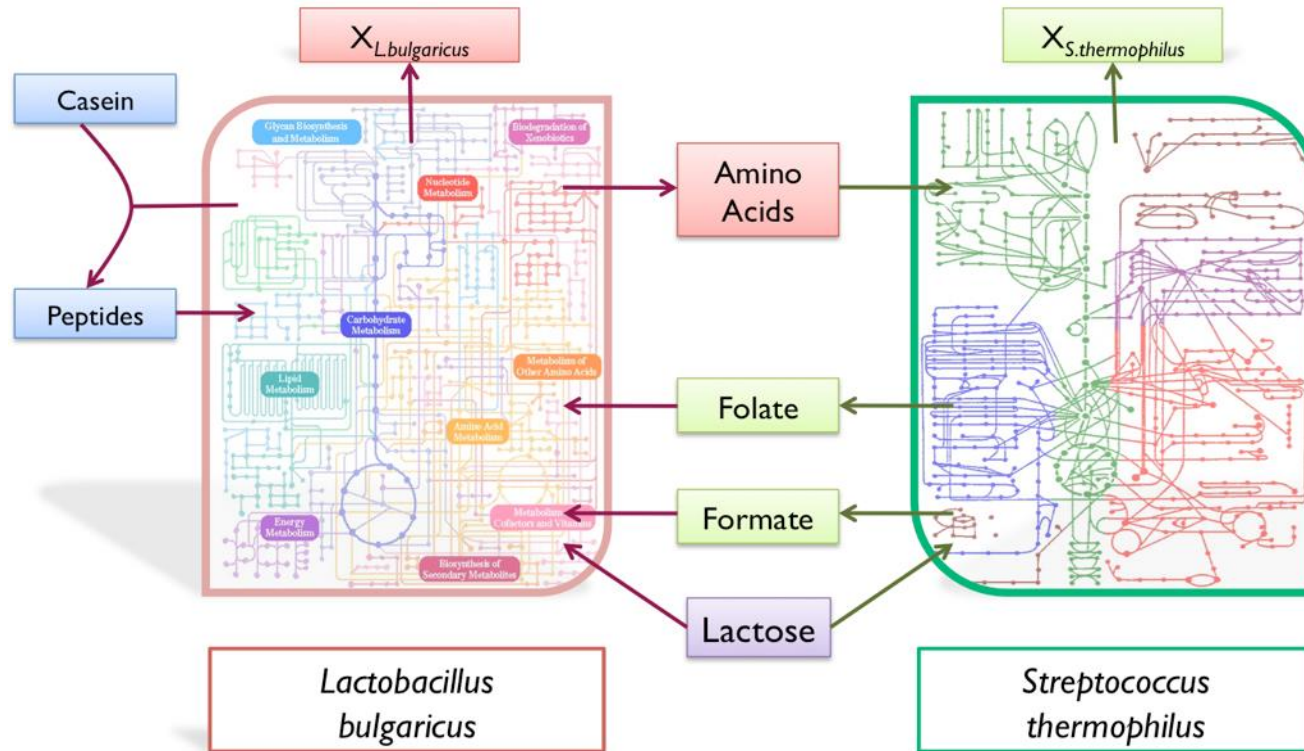
Toy example, preliminary models

=> Can predict ratio of biomasses that maximizes growth rate (and the fluxes)

Dynamic FBA

Substrate uptake kinetics

$$Vupt = \frac{Vuptmax * [S]}{[S] + Km}$$
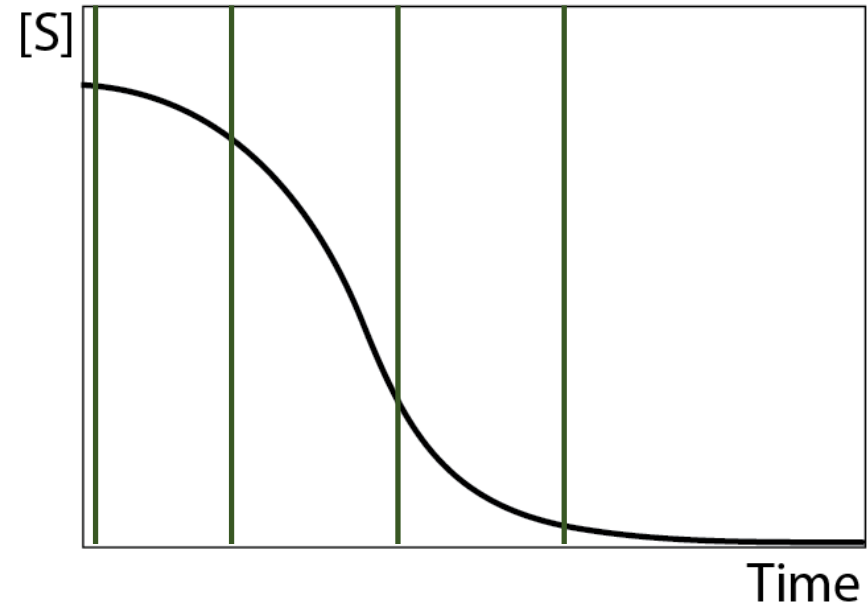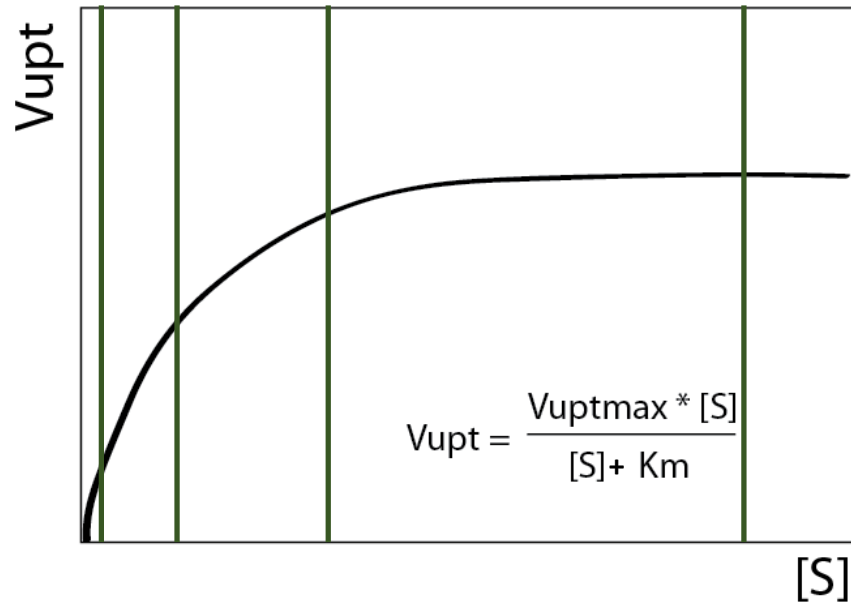
Dynamic FBA on monoculture

Approach 2:
2nd generation dynamic FBA to model acidification and flavor formation
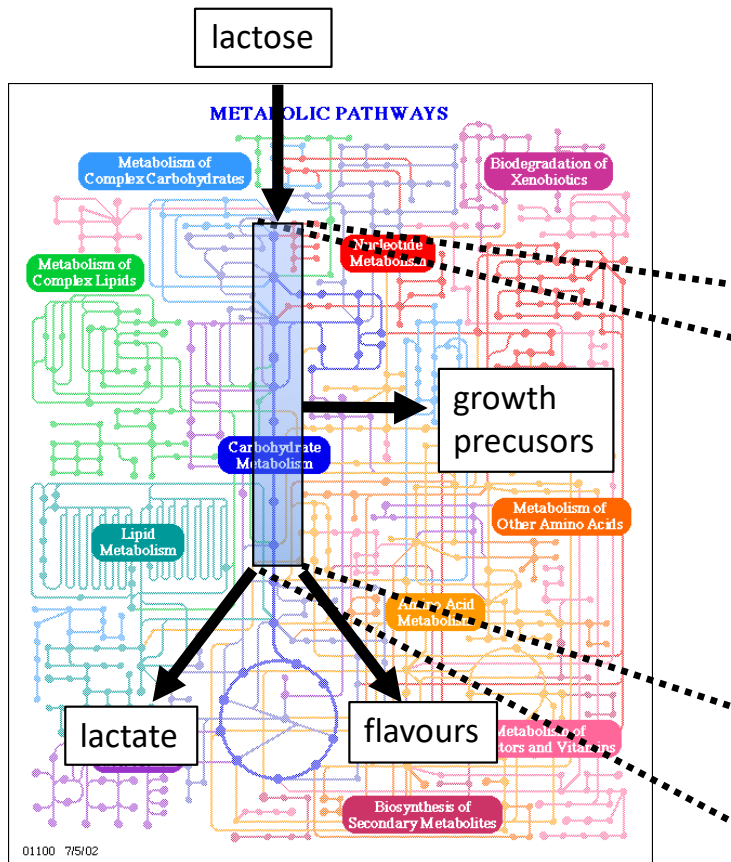
Growth: Genome-scale model

Acidification: a function of glycolytic flux and growth
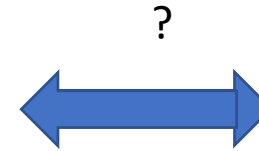Flavor formation: branches of glycolysis (acetaldehyde, diacetyl)

Levering FEBS J 2012
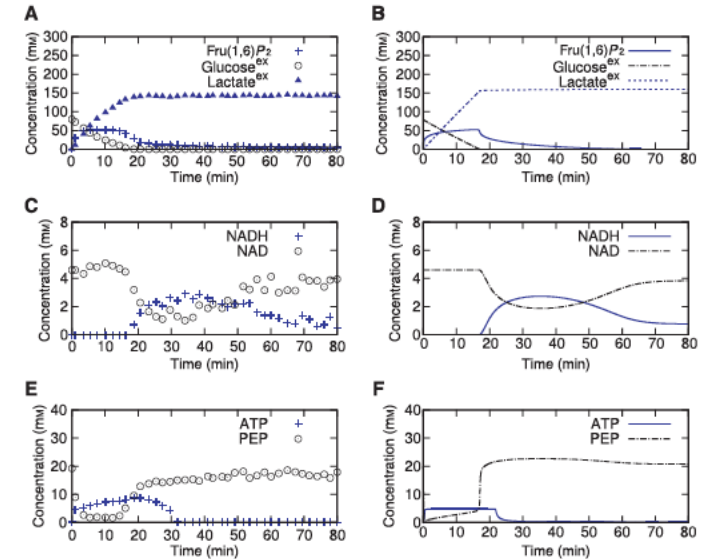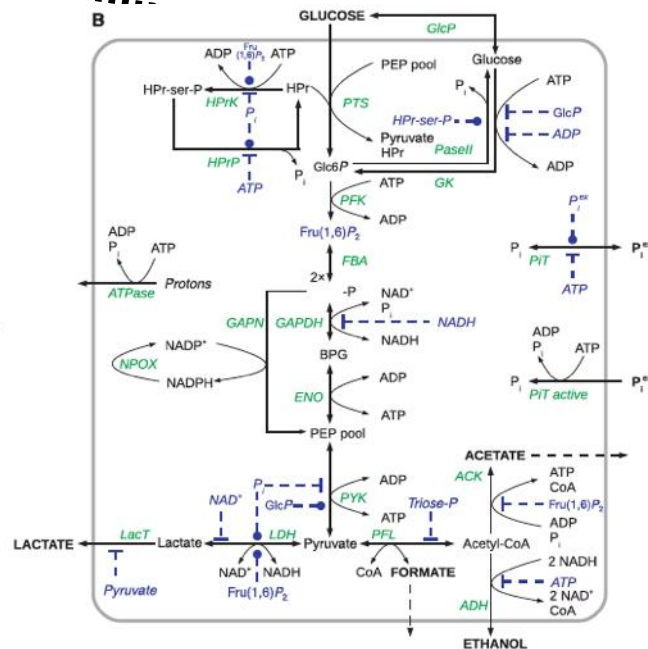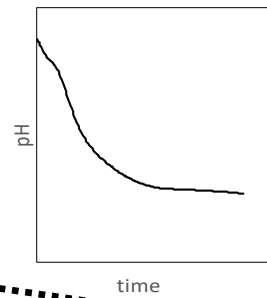
Add-on: From models to sensitivities to strain diversity

- Key idea: if parameters affect phenotype strongly, than variations is phenotype should be explainable by variations in that parameter
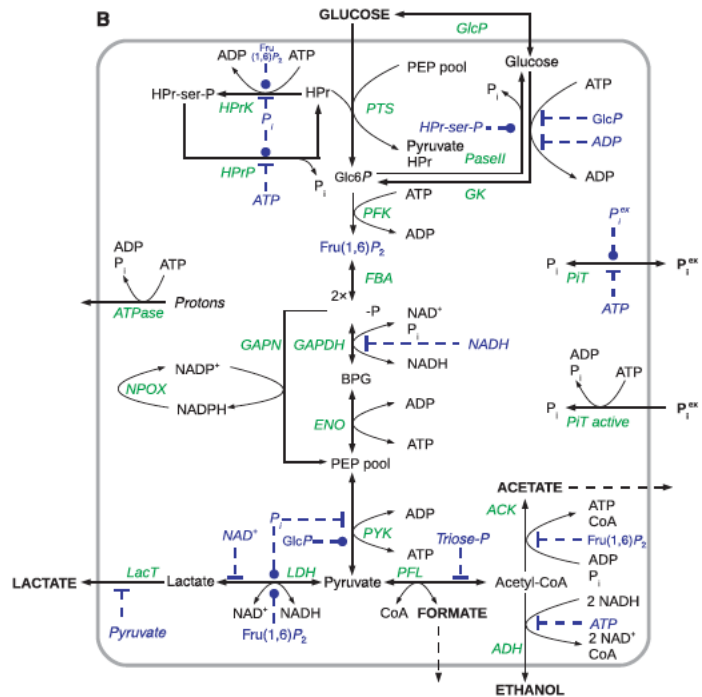
$$C_i^J = \frac{\partial \ln J}{\partial \ln e_i}$$

Identifies sensitive enzyme
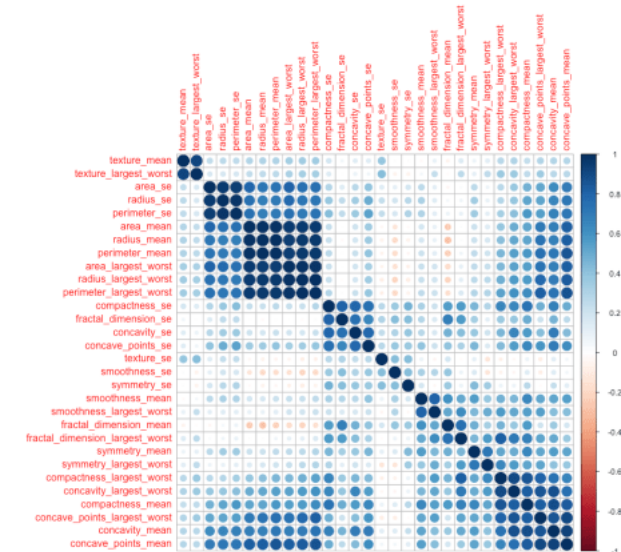
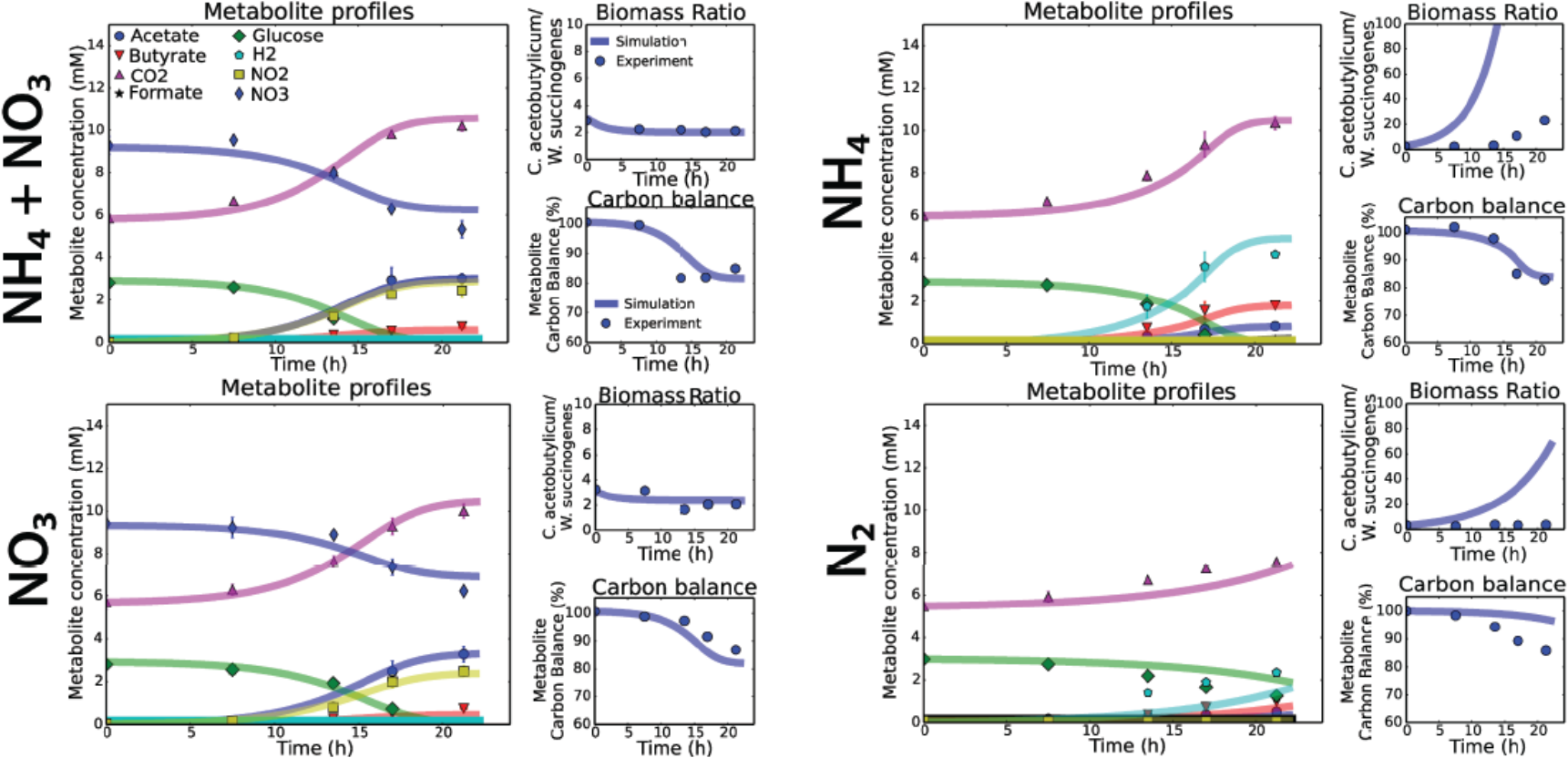# Dynamic FBA on co-culture: fit dynamic data…

Fig 3. The dFBA simulations agree mostly with the experimental data for the four different cultivation conditions and were used to infer the metabolic fluxes. The metabolite profile plots contain error bars, but the other subplots not. The biomass ratios are based on the gene-copy data and the carbon balance consisted of the measured metabolites that contained carbon. The remaining missing carbon is assumed to be incorporated into biomass.

Example from Hanemaaijer PLOS One 2015
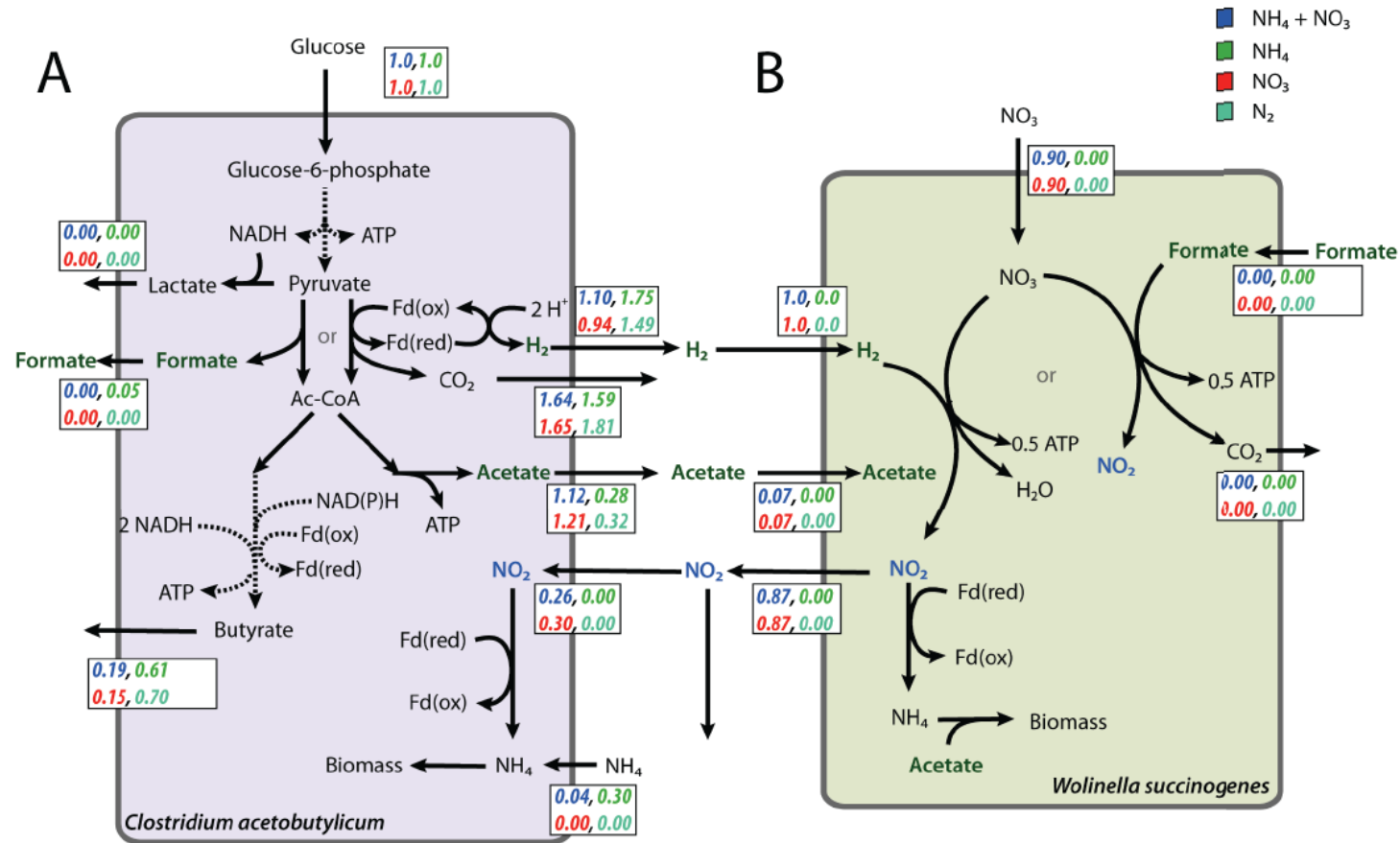
…to infer exchange fluxes…

Fig 5. Nitrogen-source also had an impact on the H₂ and the NO₂⁻ exchange rates. In **A** the calculated flux-values for *C. acetobutylicum* are normalized for the glucose uptake. **B** shows the calculated flux-values for *W. succinogenes* and are normalized for the H₂ uptake. Note that H₂ production by *C. acetobutylicum* is not equal to the specific H₂ uptake rate of *W. succinogenes* as the biomass abundances should be taken into account, which were not equal during the experiment.

Example from Hanemaaijer PLOS One 2015

dynamic cFBA: how to *predict* the future?
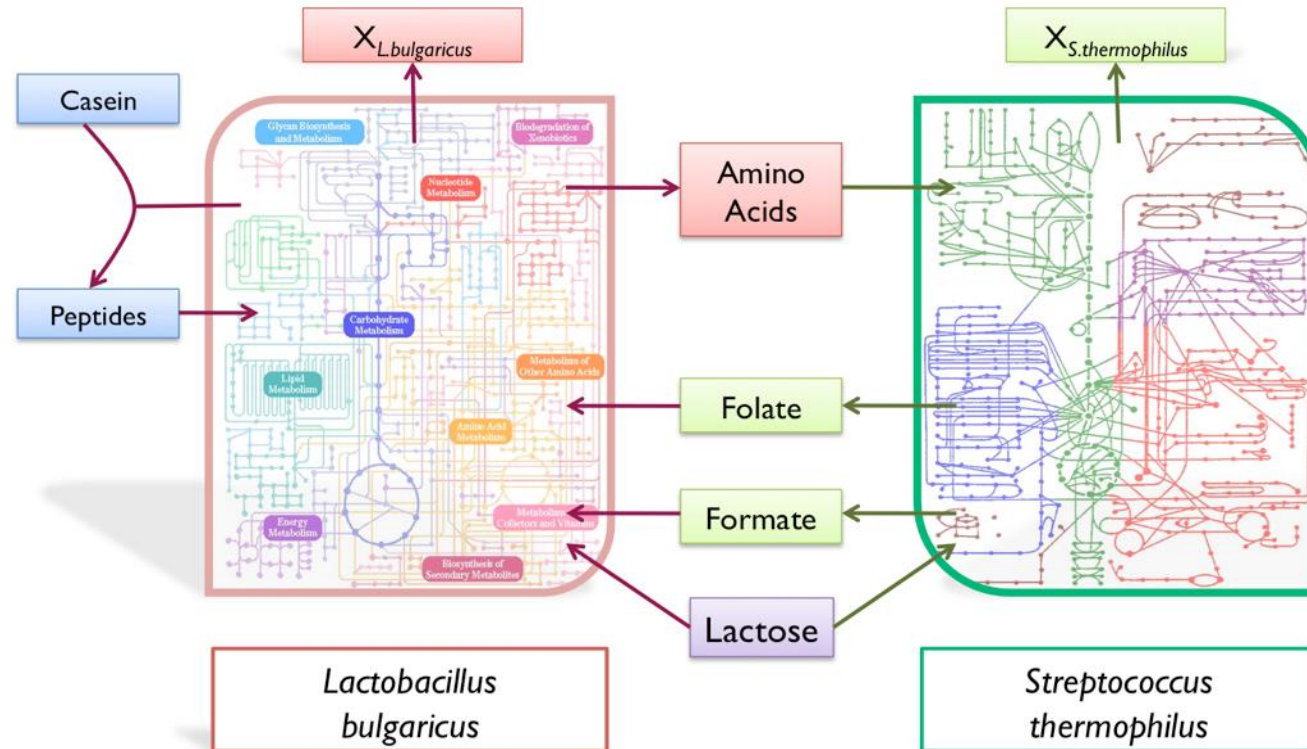
Dynamic FBA assumes instantaneous optimality, and cannot predict an investment at the cost of current growth for higher future returns

To be developed…

WP4: translation to applications

Test predictions of strain combinations in milk

PROJECT TITLE

Microbial community modeling for the production of 'designer' yogurt

PROJECT DURATION	TOTAL REQUESTED FUNDING	TOTAL COSTS
36 Months　(01 / 2018 to 12 / 2020)	946.000 €	1.023.000 €

CONSORTIUM

P 1	Prof Bas Teusink Vrije Universiteit Amsterdam Molecular Cell Biology	De Boelelaan 1085, NL-1081HV Amsterdam Netherlands	b.teusink@vu.nl Tel.: 0031 621835201 http://www.teusinkbruggemanlab.nl
P 2	Prof Ursula Kummer Heidelberg University Modeling of Biological Processes	Im Neuenheimer Feld 267, 69120 Heidelberg Germany	ursula.kummer@bioquant.uni-heidelberg.de Tel.: 0049 6221 5451278 Fax: 0049 6221 5451483 http://www.uni-heidelberg.de
P 3	Prof Ralf Takors University of Stuttgart Institute of Biochemical Engineering	Allmandring 31, 70569 Stuttgart Germany	takors@ibvt.uni-stuttgart.de Tel.: 0049 71168564535 Mobile: 0049 15119616717 http://www.ibvt.uni-stuttgart.de/
P 4	Dr Petri-Jaan Lahtvee University of Tartu Institute of Technology	Nooruse 1, 50411 Tartu Estonia	lahtvee@ut.ee Tel.: 00372 5088117 Mobile: 00372 5088117 http://www.ut.ee
P 5	Dr Ana Rute Neves Chr. Hansen A/S Bacterial Physiology and Improvement (BactoPhI)	Bøge Allé 10-12, 2970 Hørsholm Denmark	DKRUNS@chr-hansen.com Tel.: +45 60431012 http://www.chr-hansen.com