

XomicsToModel: Multi-omics data integration and generation of thermodynamically consistent metabolic models.

X-omics festival 2023
Nijmegen, The Netherlands

17/04/23

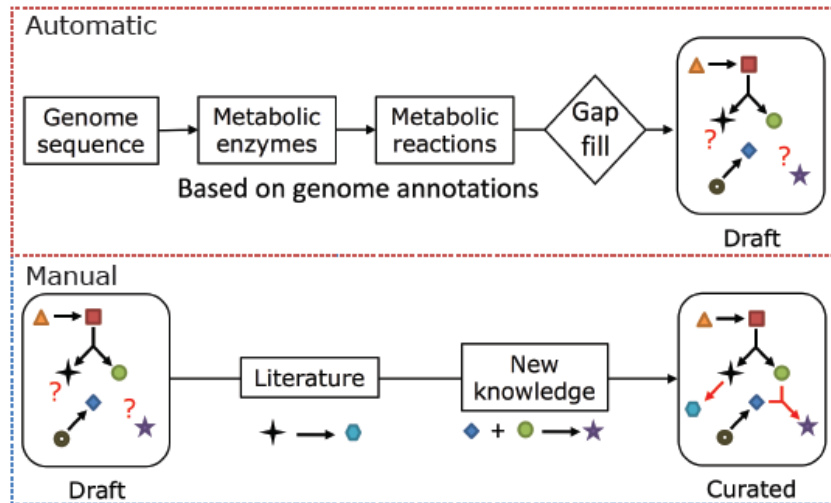
Ronan M.T. Fleming, B.V.M.S., Dip. Math., PhD.

Leiden Academic Centre for Drug Research, Leiden University, The Netherlands.
School of Medicine, National University of Ireland, Galway, Ireland.



Universiteit
Leiden

a) Genome-scale metabolic reconstruction



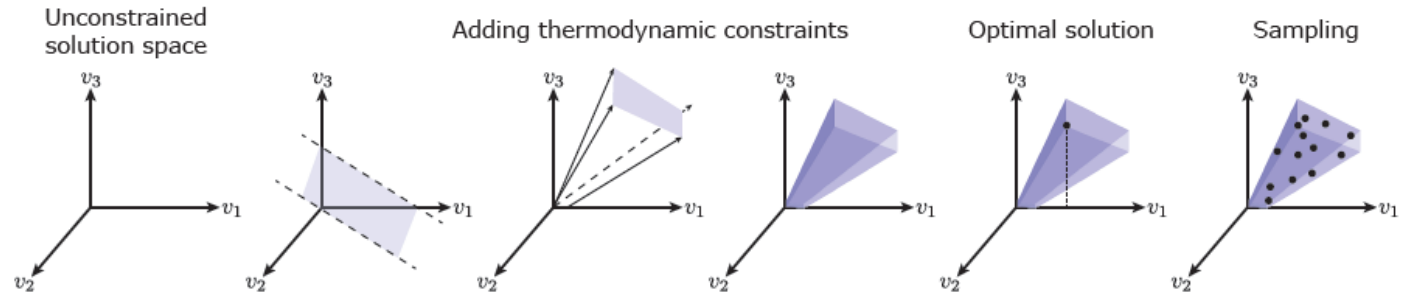
c) Flux balance analysis

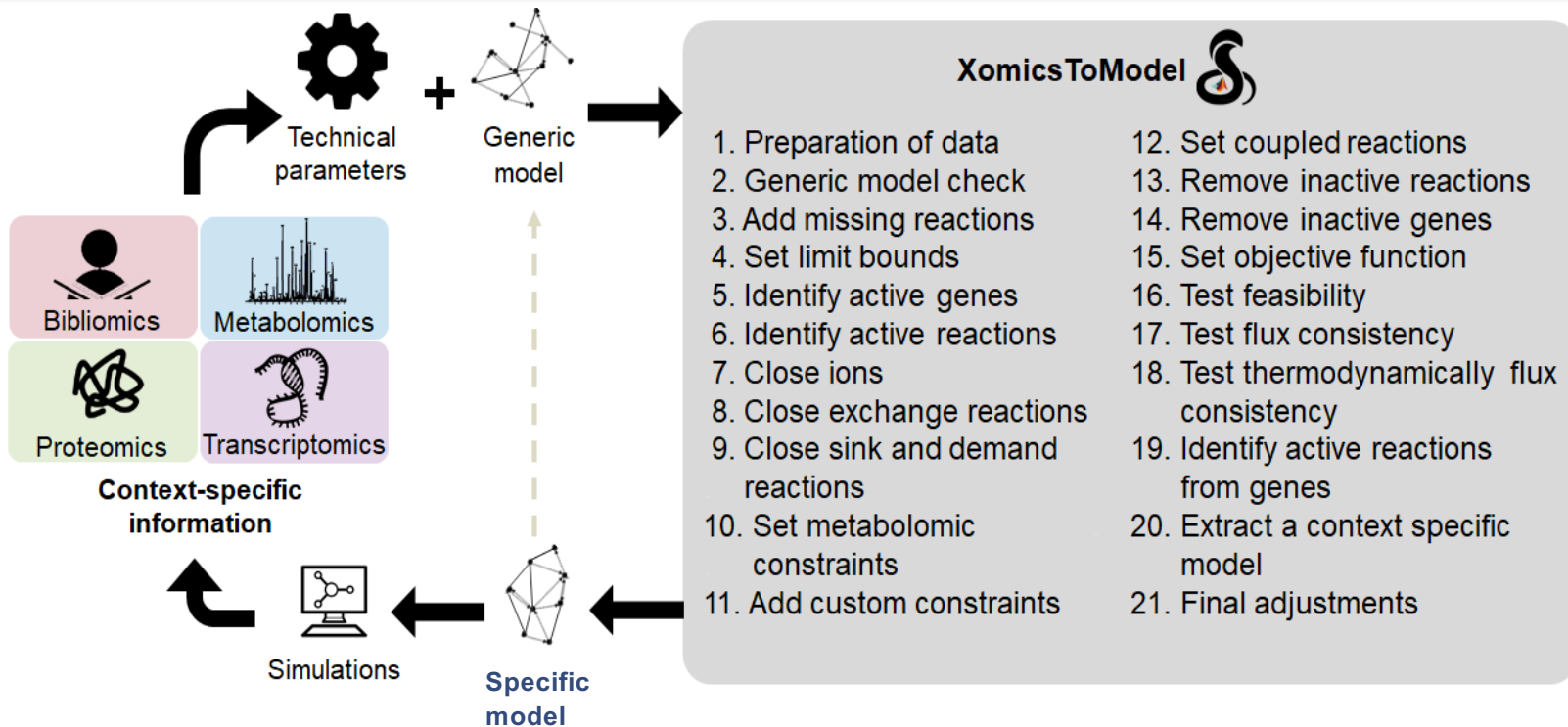
Maximize/minimize an objective function
 $\psi = c_1v_1 + c_2v_2 + \dots + c_5v_5$ such that:

		Reactions						
		R ₁	R ₂	R ₃	R ₄	R ₅		
Metabolites	▲	-1	0	0	0	0	⊗	=
	■	1	-1	0	0	0		
	✦	0	1	-1	0	0		
	●	0	1	0	0	-1		
	⬡	0	0	1	0	0		
	⊙	0	0	0	-1	0		
◆	0	0	0	1	-1	⊖	=	
★	0	0	0	0	1			
		S-matrix					Flux vector	

and for every reaction i : $lb_i \leq v_i \leq ub_i$

b) Solution spaces





Protocol paper + COBRA Toolbox extension + executable tutorials and examples:

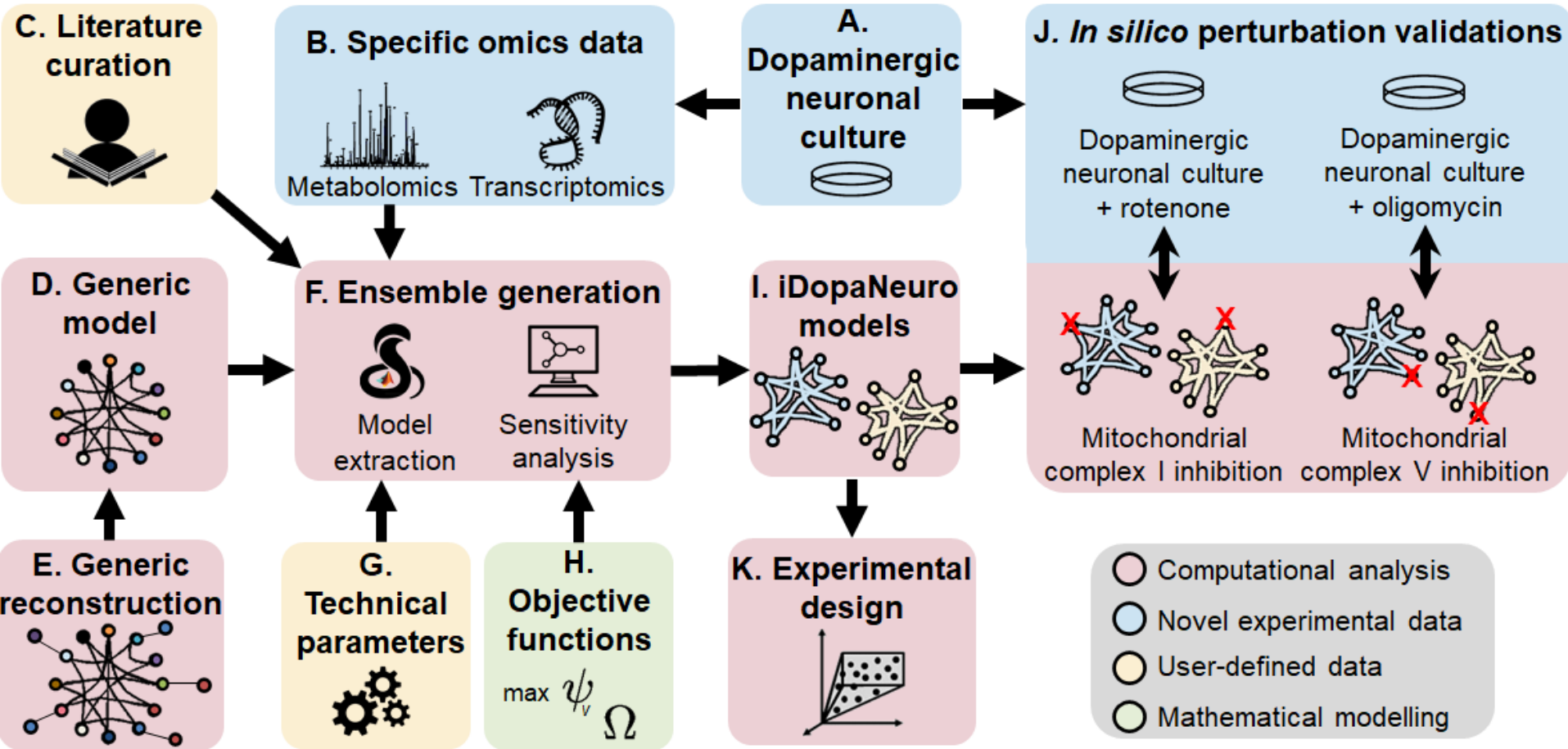
<https://www.biorxiv.org/content/10.1101/2021.11.08.467803v2>

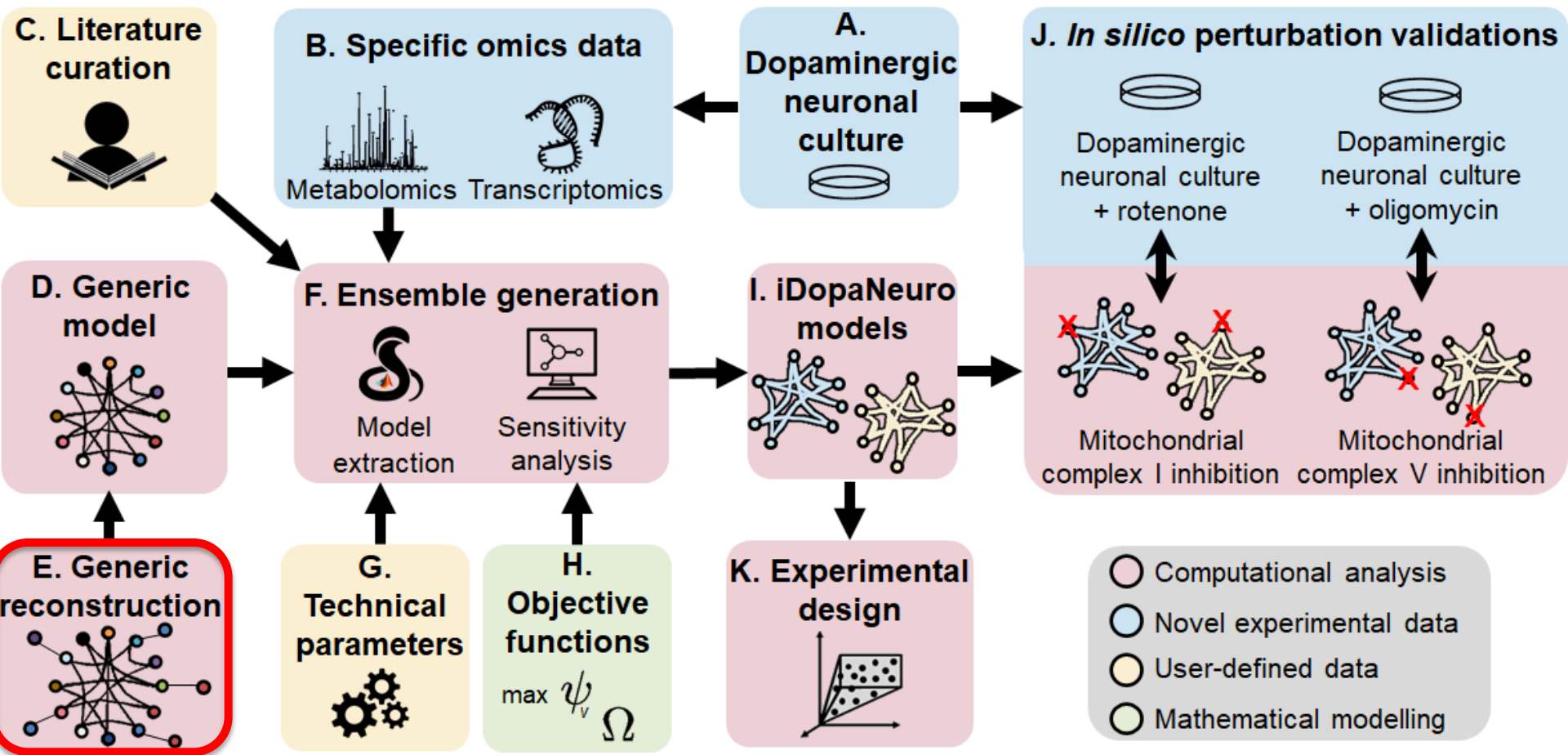
<https://github.com/opencobra/COBRA.tutorials/tree/master/dataIntegration/XomicsToModel>

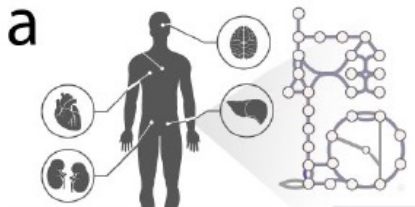
<https://github.com/opencobra/cobratoolbox/tree/master/src/dataIntegration/XomicsToModel>

https://github.com/opencobra/COBRA.papers/tree/master/2023_iDopaNeuro

Generic model + specific data + XomicsToModel \rightarrow Specific model

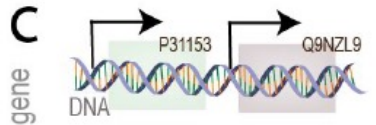
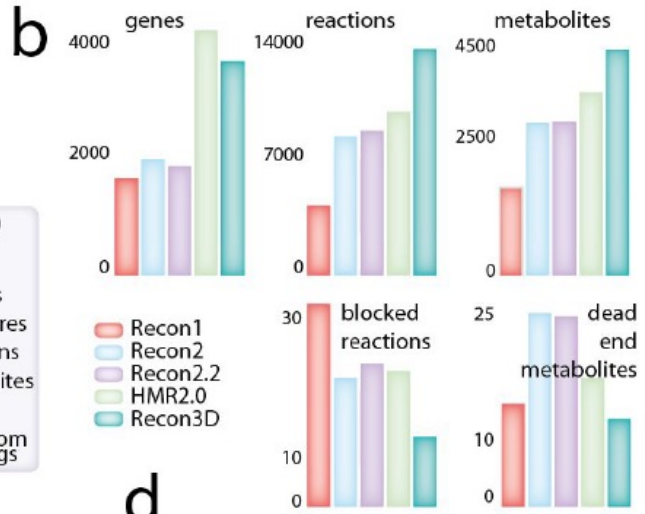




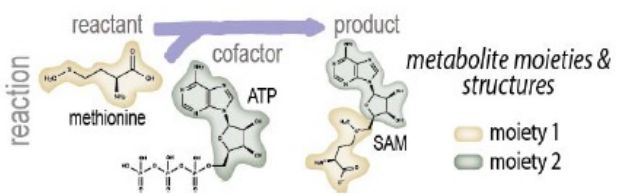
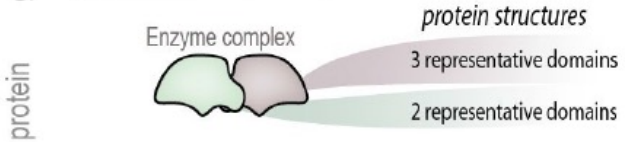


biochemical data
genomic data
tissue-specific localization
proteomic data
metabolomic data
protein structural data
pharmacogenomic data
atom-atom mappings

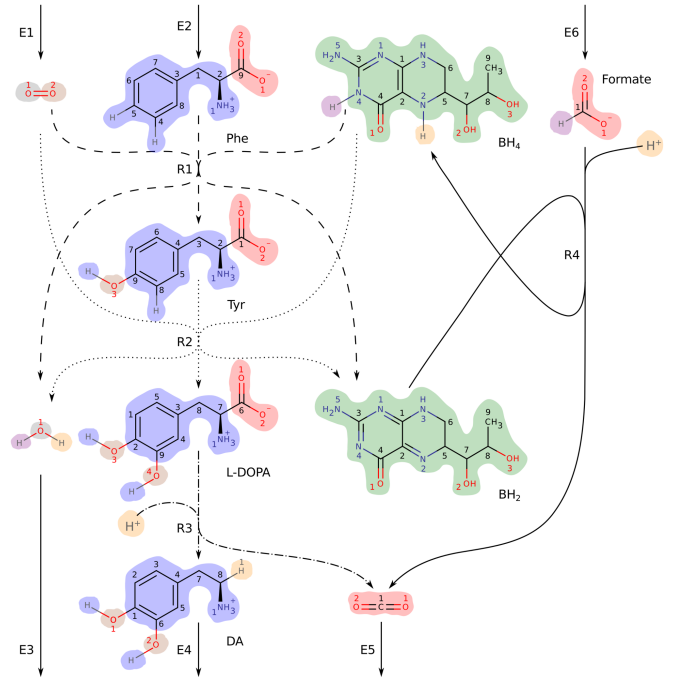
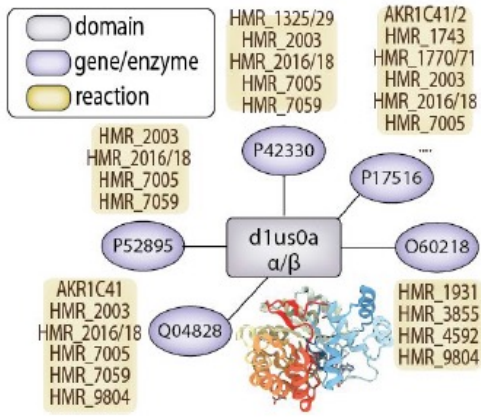
Recon 3D
3288 genes
2908 domains
12890 structures
13543 reactions
4140 metabolites
3536 SNVs
8315 atom-atom mappings



New datatypes

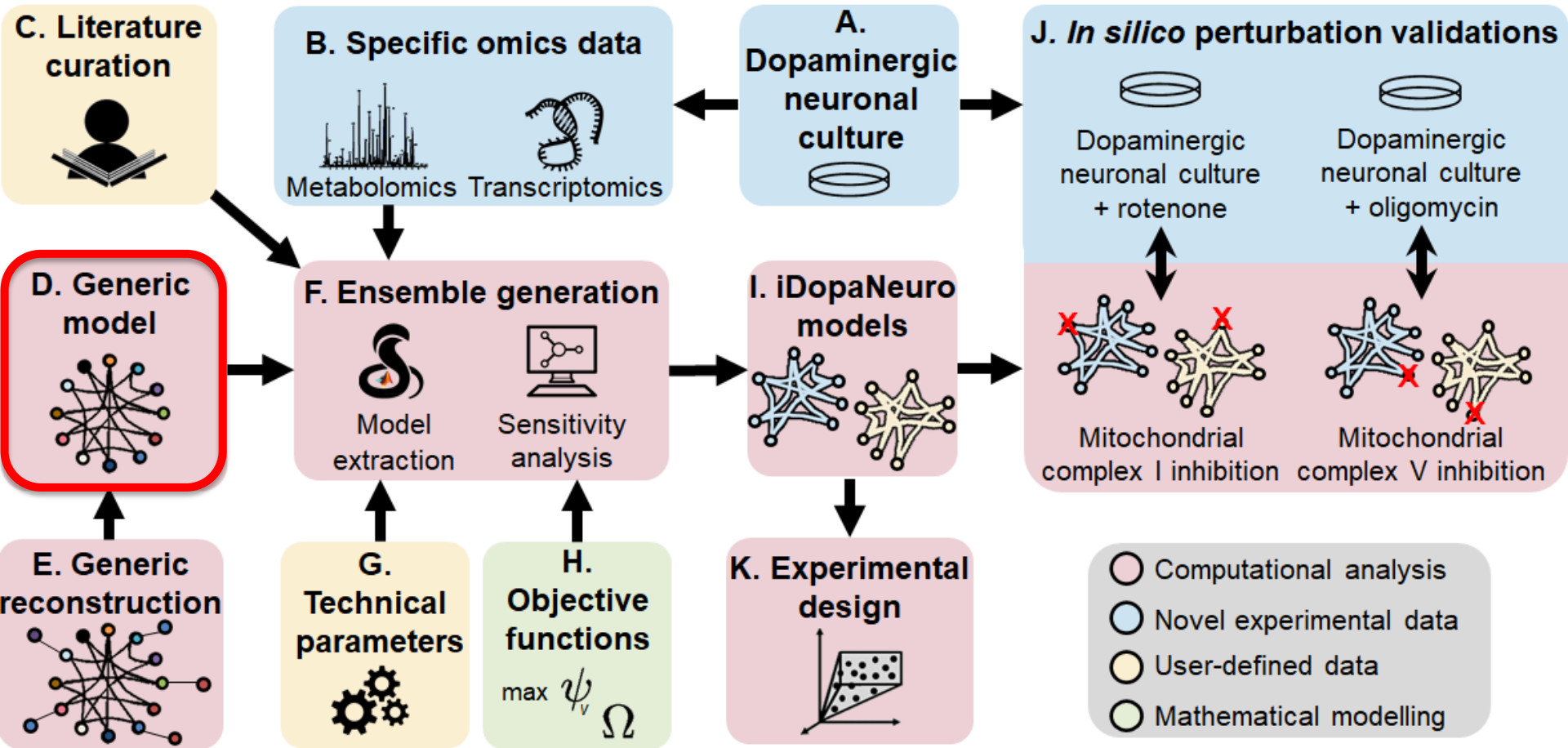


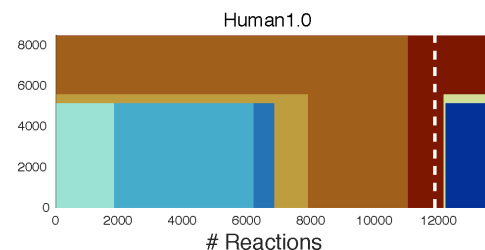
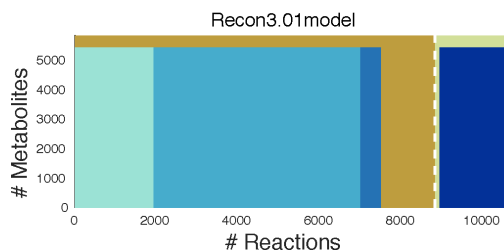
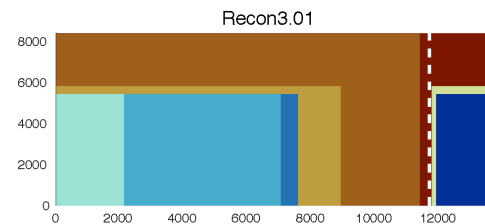
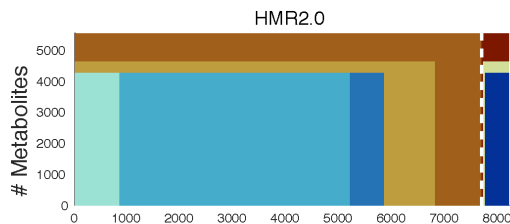
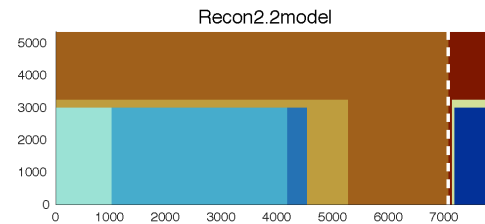
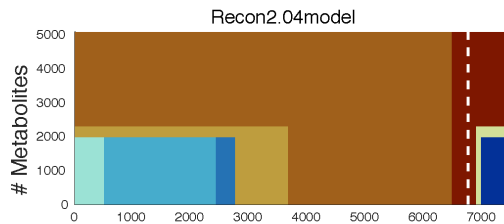
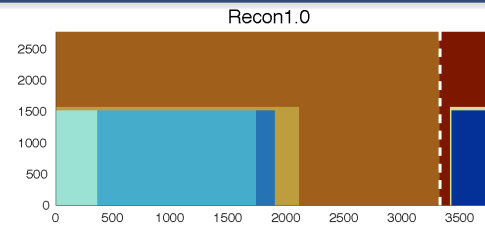
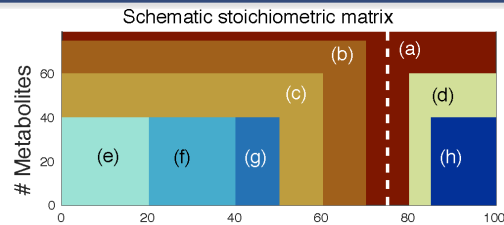
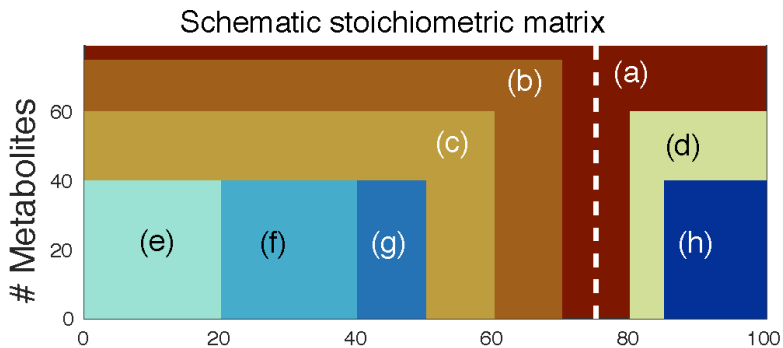
d



Haraldsdóttir HS, et al. (2016) Identification of Conserved Moieties in Metabolic Networks by Graph Theoretical Analysis of Atom Transition Networks. PLoS Comput Biol 12(11): e1004999.

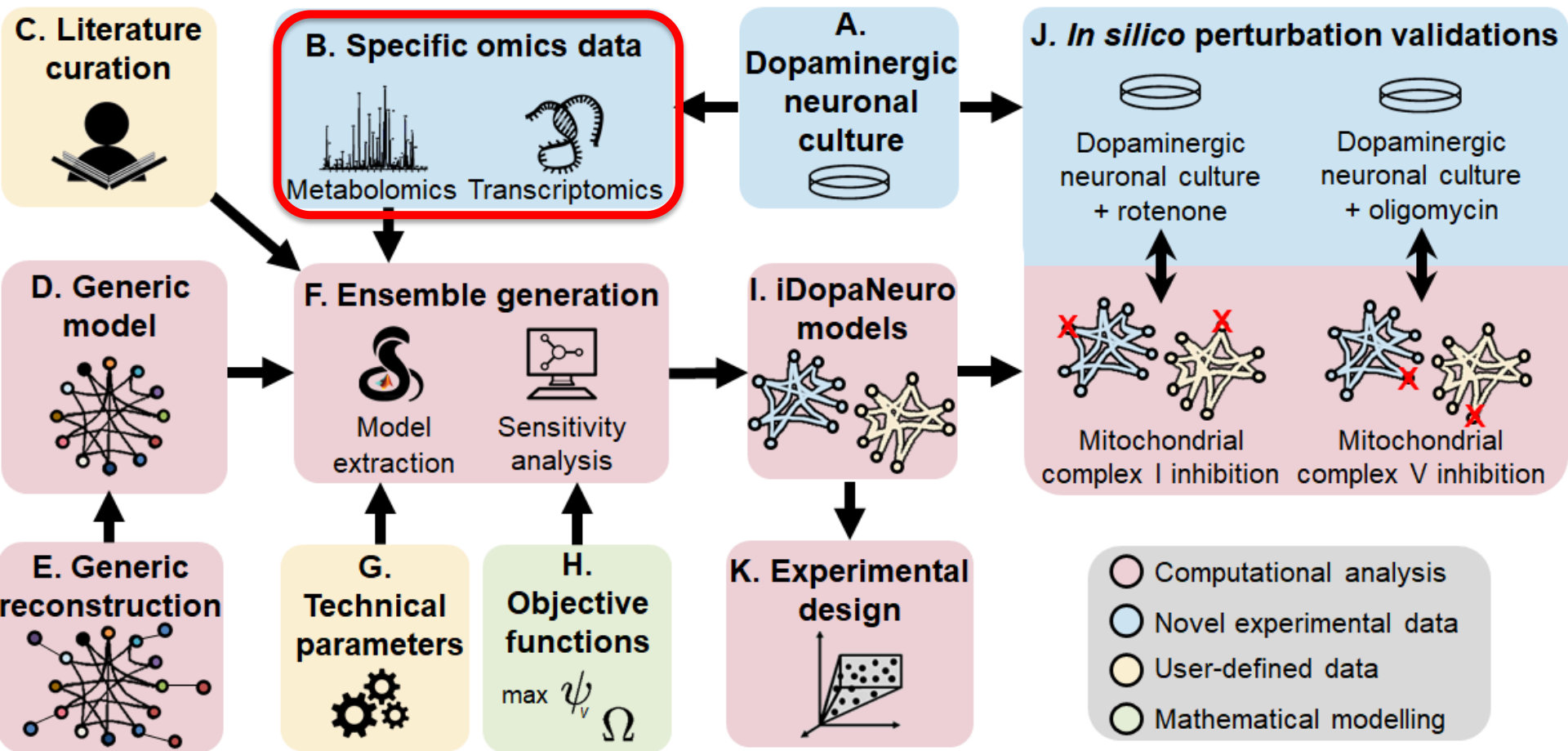
Brunk, et al (2017) Recon 3D: A resource enabling a three dimensional view of gene variation in human metabolism, Nature Biotech.





- (a) reconstruction
- (b) stoichiometrically consistent subset
- (c, d) + flux consistent
- (e, f, g, h) + thermodynamically flux consistent

Fleming et. al. Cardinality optimisation in constraint-based modelling: Application human metabolism (submitted)



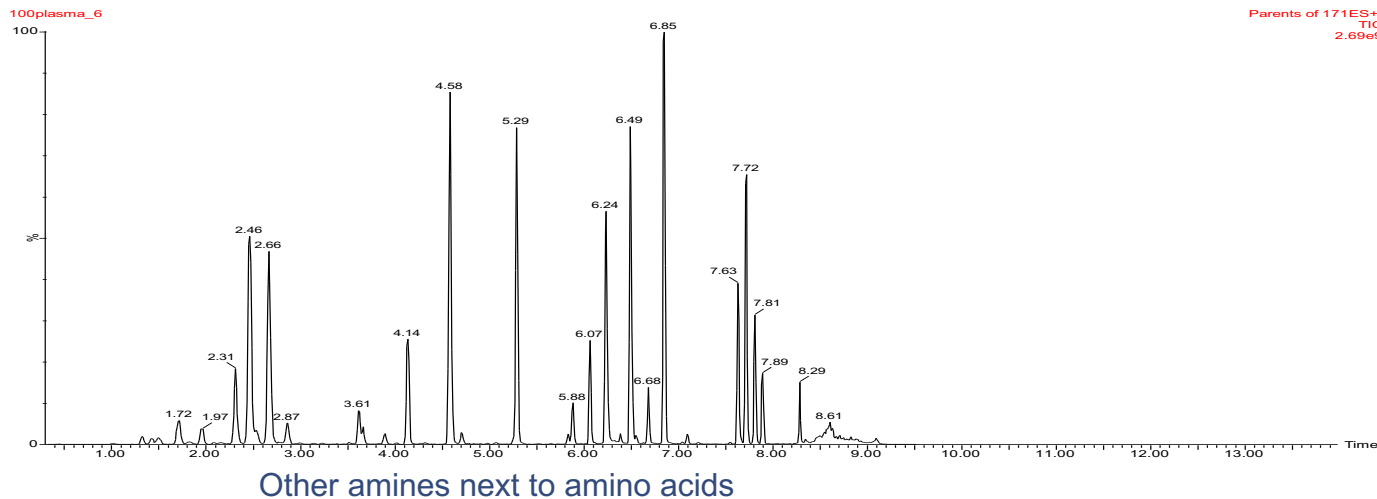


74 Amines: Targeted LC-MS/MS after AccQ-Tag derivatization

Marek J. Noga, et al. *Metabolomics of cerebrospinal fluid reveals changes in the central nervous system metabolism in a rat model of multiple sclerosis.* 8(2):253-263, 2012.

24 Organic acids: an established GC-MS method

Rodrigo D. A. M. Alves, et. al. *Global profiling of the muscle metabolome: method optimization, validation and application to determine exercise-induced metabolic effects.* 11(2):271-285, 2015.



Parents of 171ES+
TIC
2.69e9



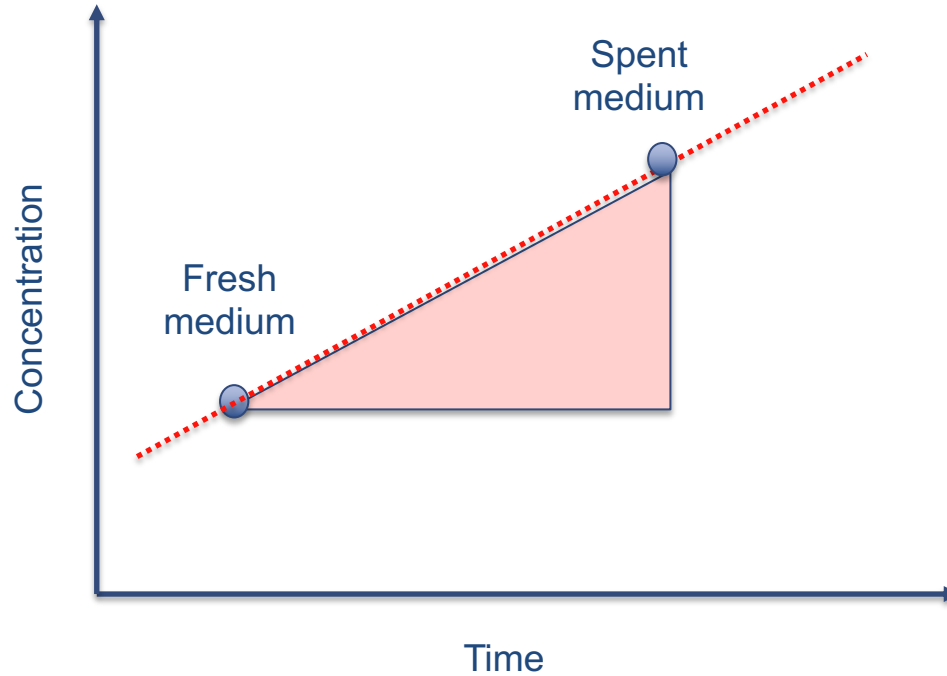
Thomas
Hankemeier,
Leiden University

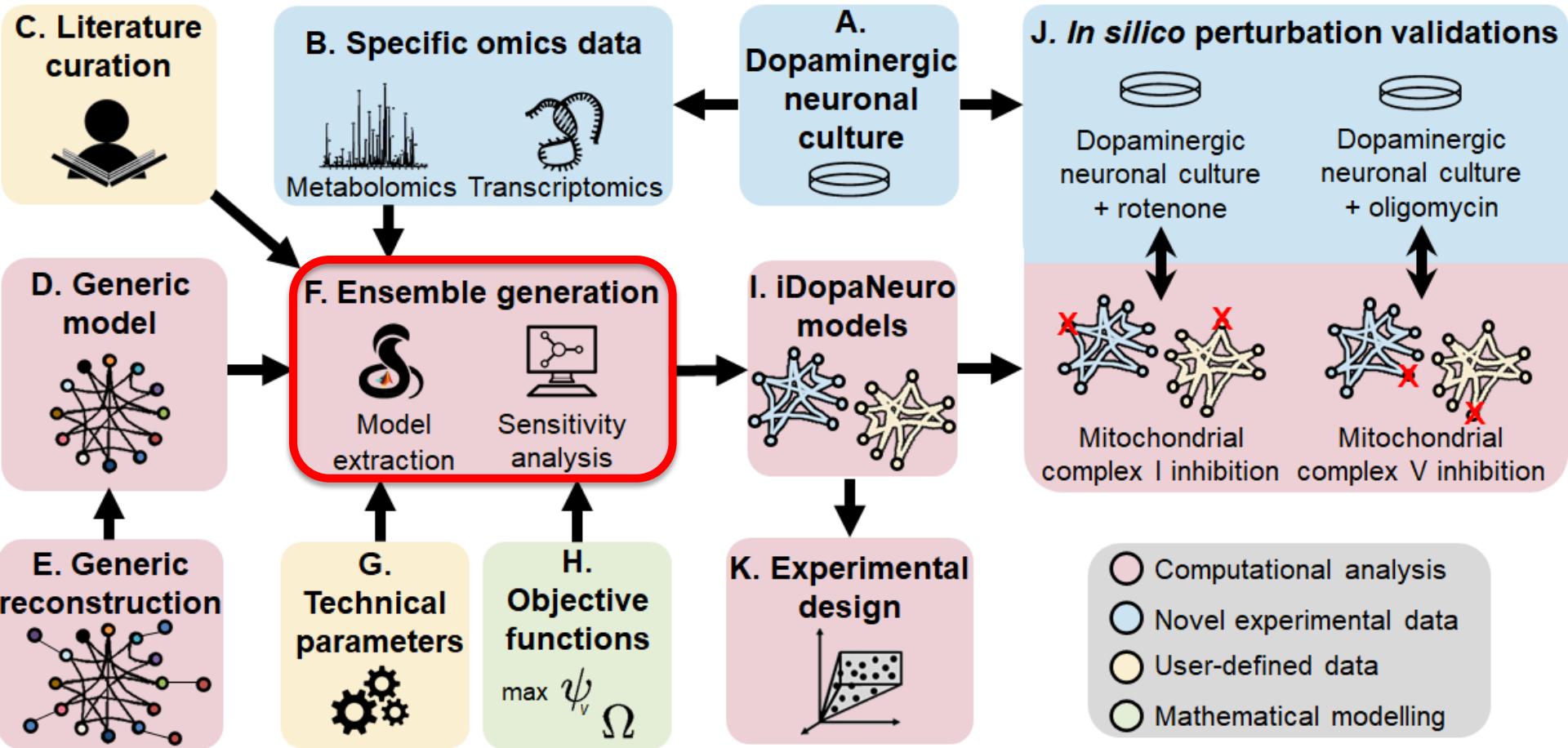


Amy Harms,
Leiden University



Conversion of extracellular concentration to exchange reaction rate





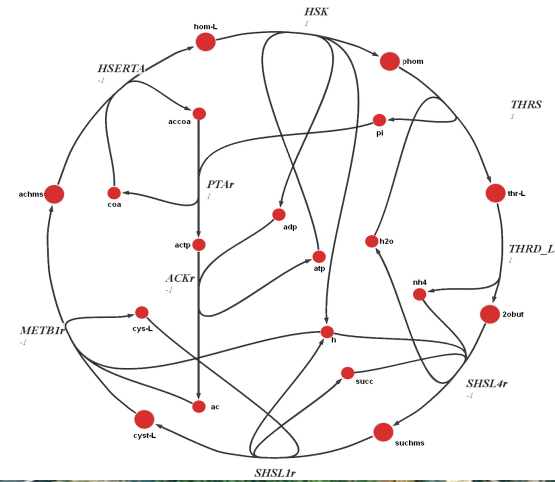
thermoKernel: a novel thermodynamically consistent model extraction algorithm

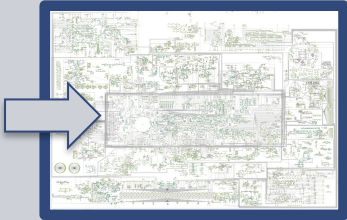
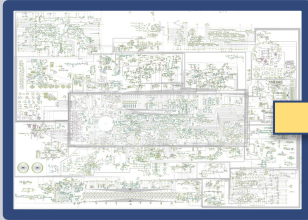
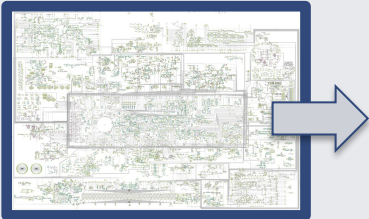
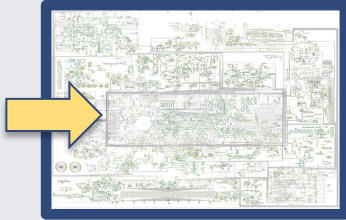
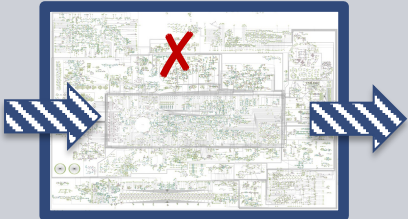
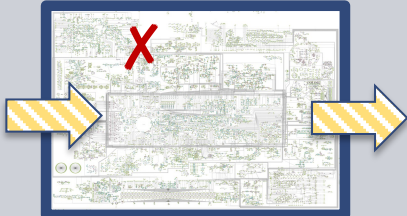
- **New options for input data**

- presence/absence of metabolites (not just reactions or genes) to be specified
- weights on metabolite/reaction/genes
 - e.g. transcript abundance

- **Improved context-specific model output**

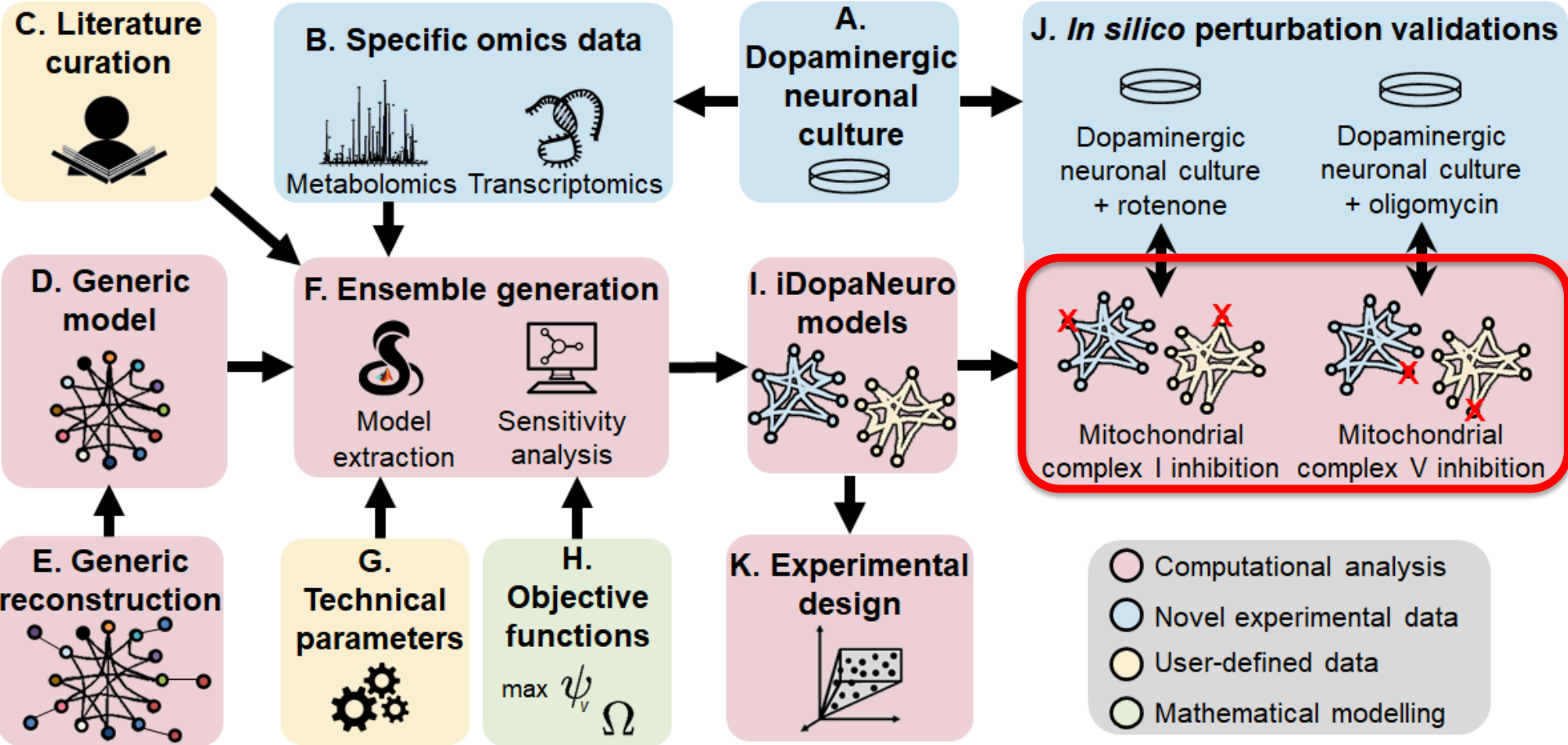
- all reactions are thermodynamically flux consistent
 - admits a flux satisfying energy conservation
 - admits a flux satisfying 2nd law of thermodynamics
 - internal reaction rates can be predicted (no artefactual flux around cycles)
- admits reconstruction directionality constraints
- minimal sized model
- scalable: algorithm based on a sequence of linear optimization problems



Model type	Data	Prediction
Uptake constrained (training set)	Uptake rates 	Secretion rates 
Secretion constrained (training set)	Secretion rates 	Uptake rates 
Optimal model (validation)	Uptake & secretion rates 	Metabolic perturbations 

Leave one (metabolite) out cross validation (LOOCV)

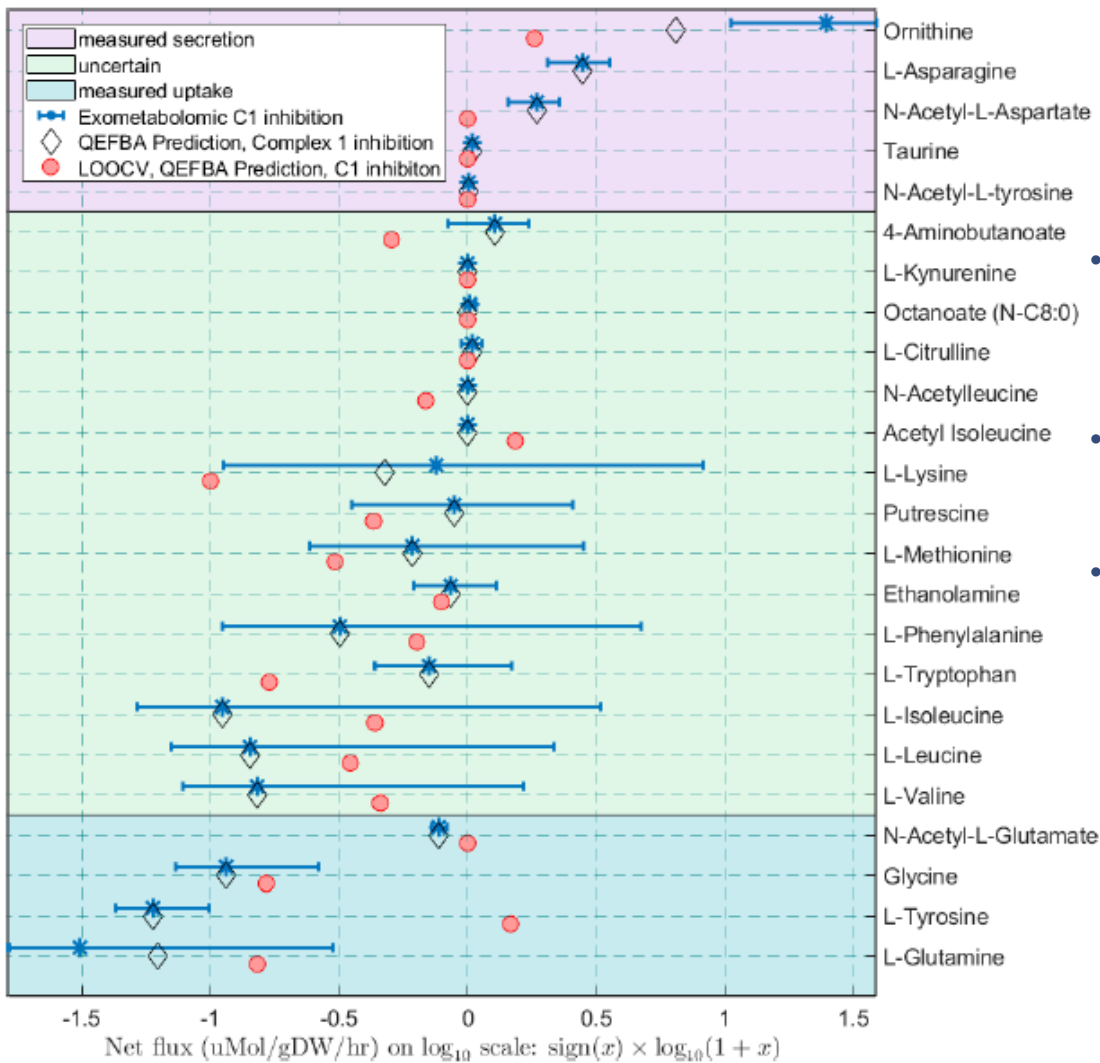
Human dopaminergic neuronal metabolic model validation



Human dopaminergic neuronal metabolic model validation

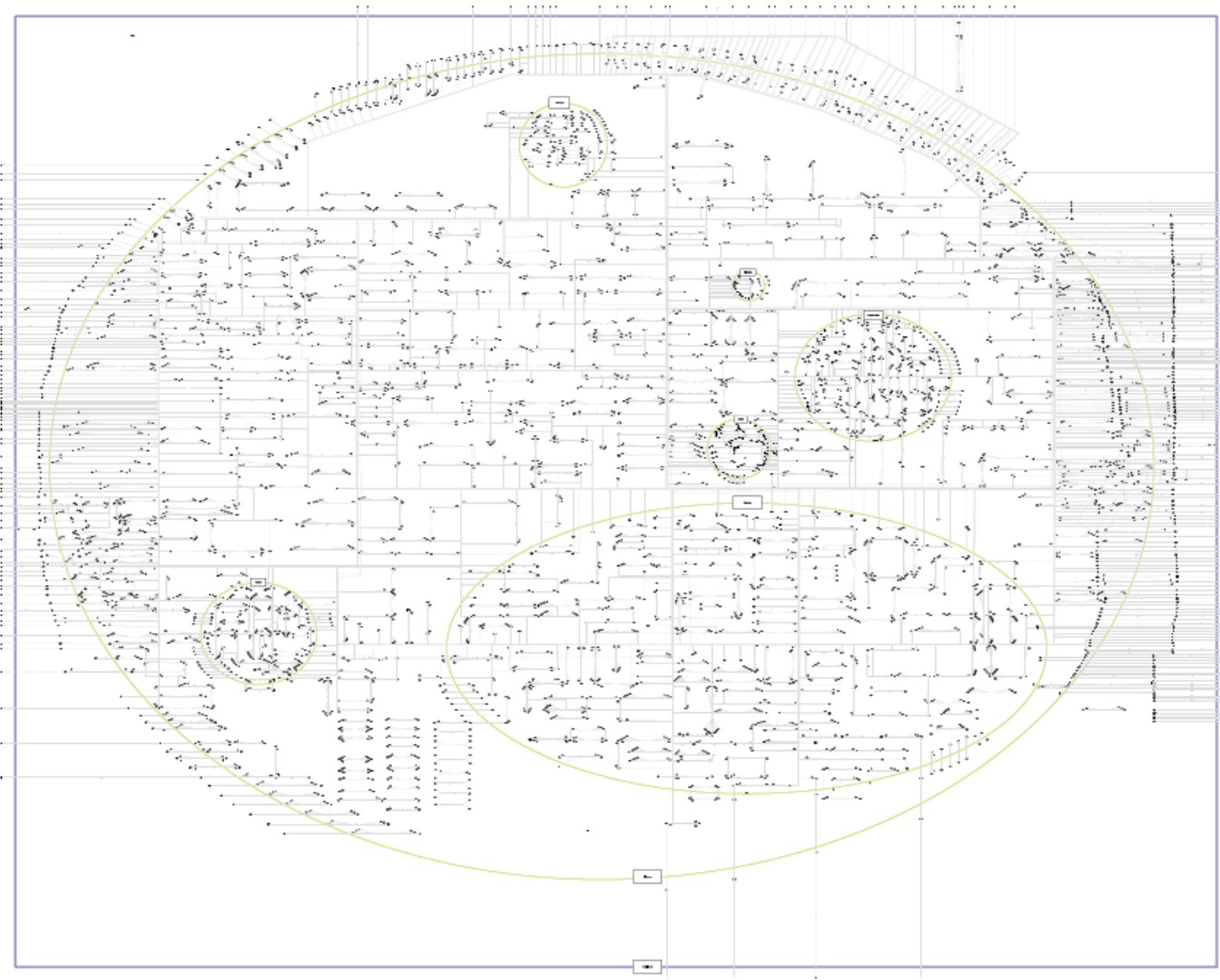
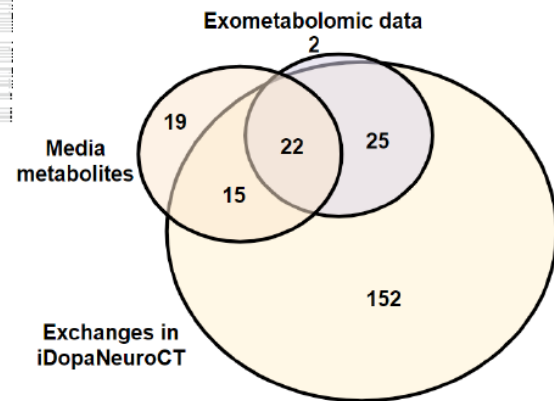
- In vitro complex I inhibition (rotenone) of iPSC derived dopaminergic neuronal culture from normal control subjects.
- In silico complex I inhibition (reduced upper bound on complex 1 metabolic reaction)
- Prediction of metabolite exchange
 - high qualitative accuracy
 - (correct/total = 0.78, n = 9)
 - moderate semi-quantitative accuracy (Spearman rho = 0.48, *pval* = 0.018).

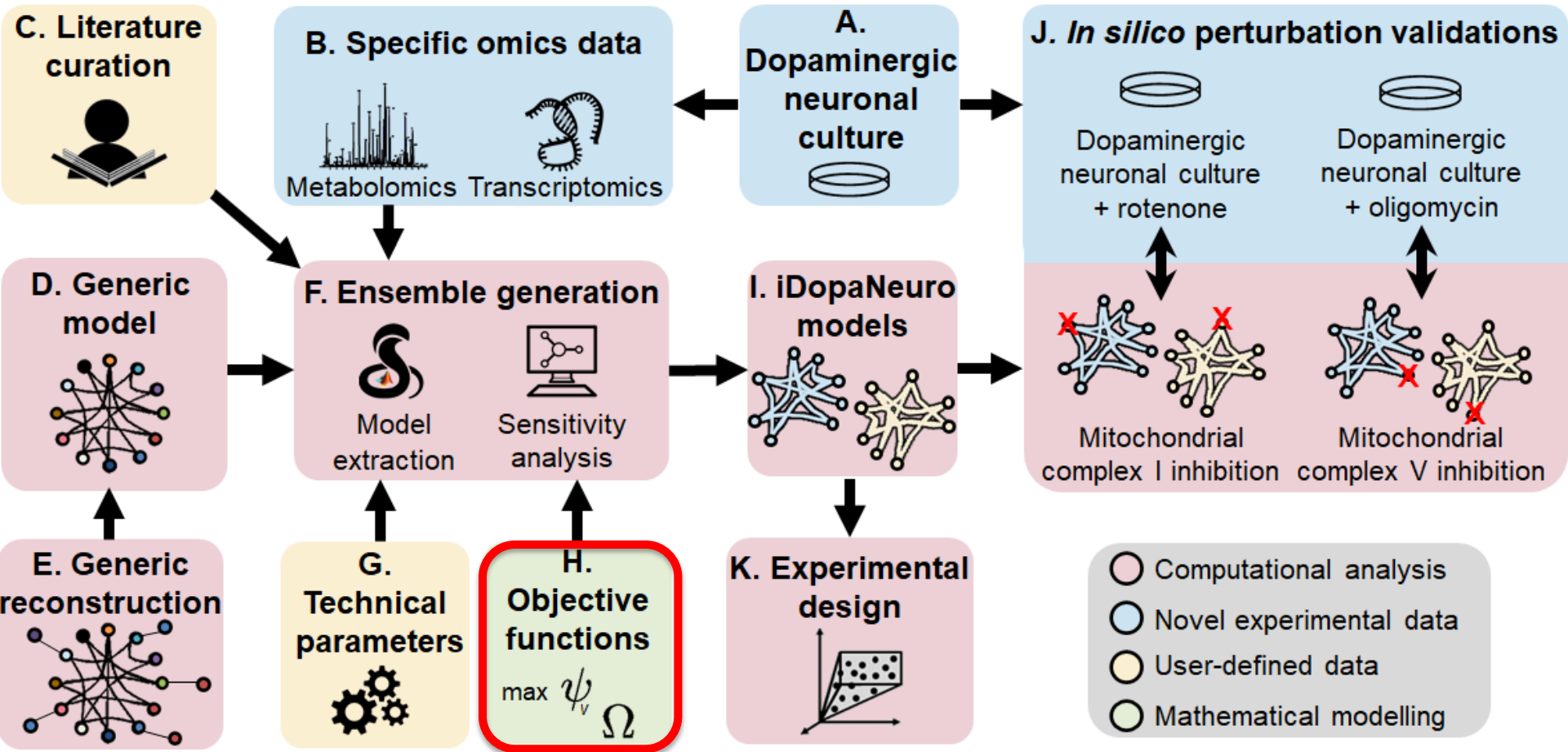
Preciat et. al. Mechanistic model-driven exometabolomic characterisation of human dopaminergic neuronal metabolism (submitted)



Cell-type specific dopaminergic neuronal model (iDopaNeuroCT)

- 1,229 genes
- 2,065 reactions
 - 238 metabolites exchanged with the environment
- 1,313 metabolites





Maximum entropy: least biased prediction given the data

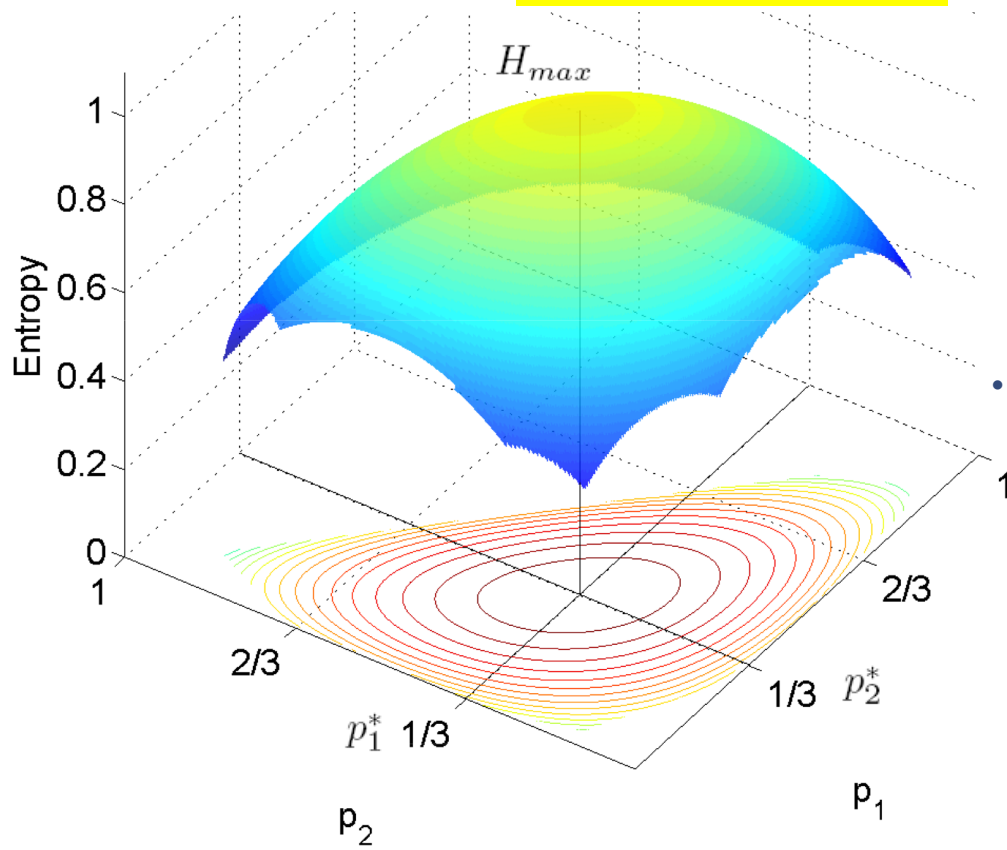
$$\text{maximize } H = -(p_1 \ln p_1 + p_2 \ln p_2 + p_3 \ln p_3)$$

$$\text{subject to } p_1 + p_2 + p_3 = 1$$

Maximum entropy subject to constraints

- Thermodynamically feasible model
 - admits at least one flux vector that satisfies energy conservation and the second law of thermodynamics

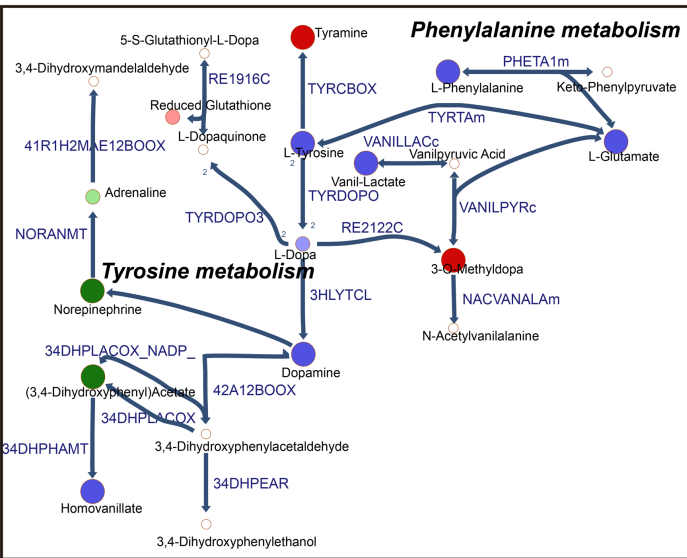
With no constraints other than the normalization of probability, then equiprobability is the least biased prediction.



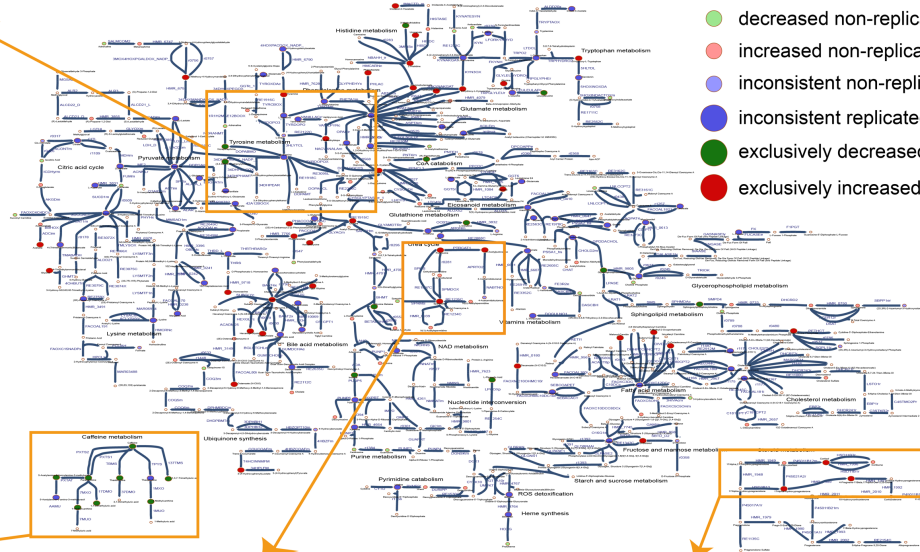
- Maximisation of unidirectional flux entropy
 - predicts a particular type of thermodynamically feasible flux vector

$$v_f^T \ln v_f + v_r^T \ln v_r$$

Identification of robust metabolic biomarkers for Parkinson's Disease diagnosis based on meta-analysis of clinical data and computational modelling



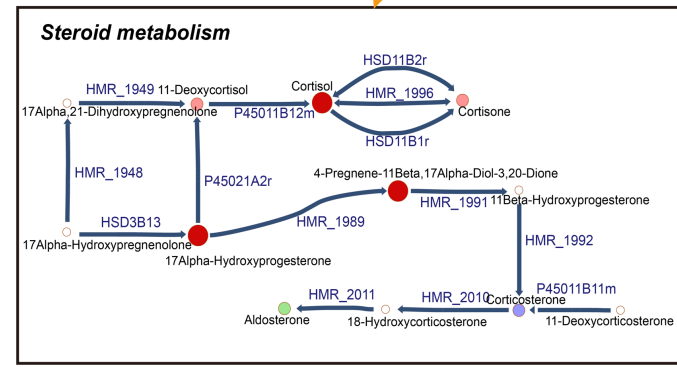
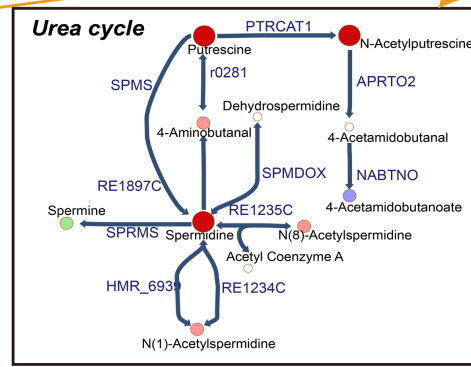
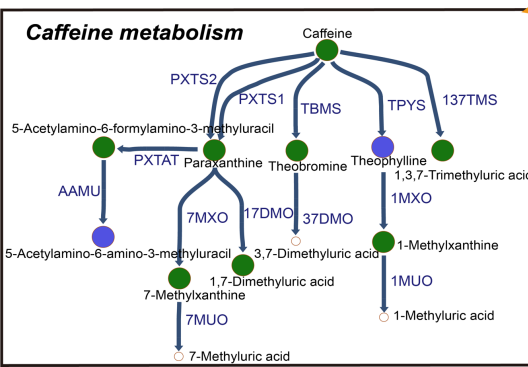
The map related to the replicated metabolites



Metabolites Legend:

- non diagnosis-related
- decreased non-replicated
- increased non-replicated
- inconsistent non-replicated
- inconsistent replicated
- exclusively decreased replicated
- exclusively increased replicated

74 studies
928 metabolites
190 replicated
60 exclusively increased
54 exclusively decreased

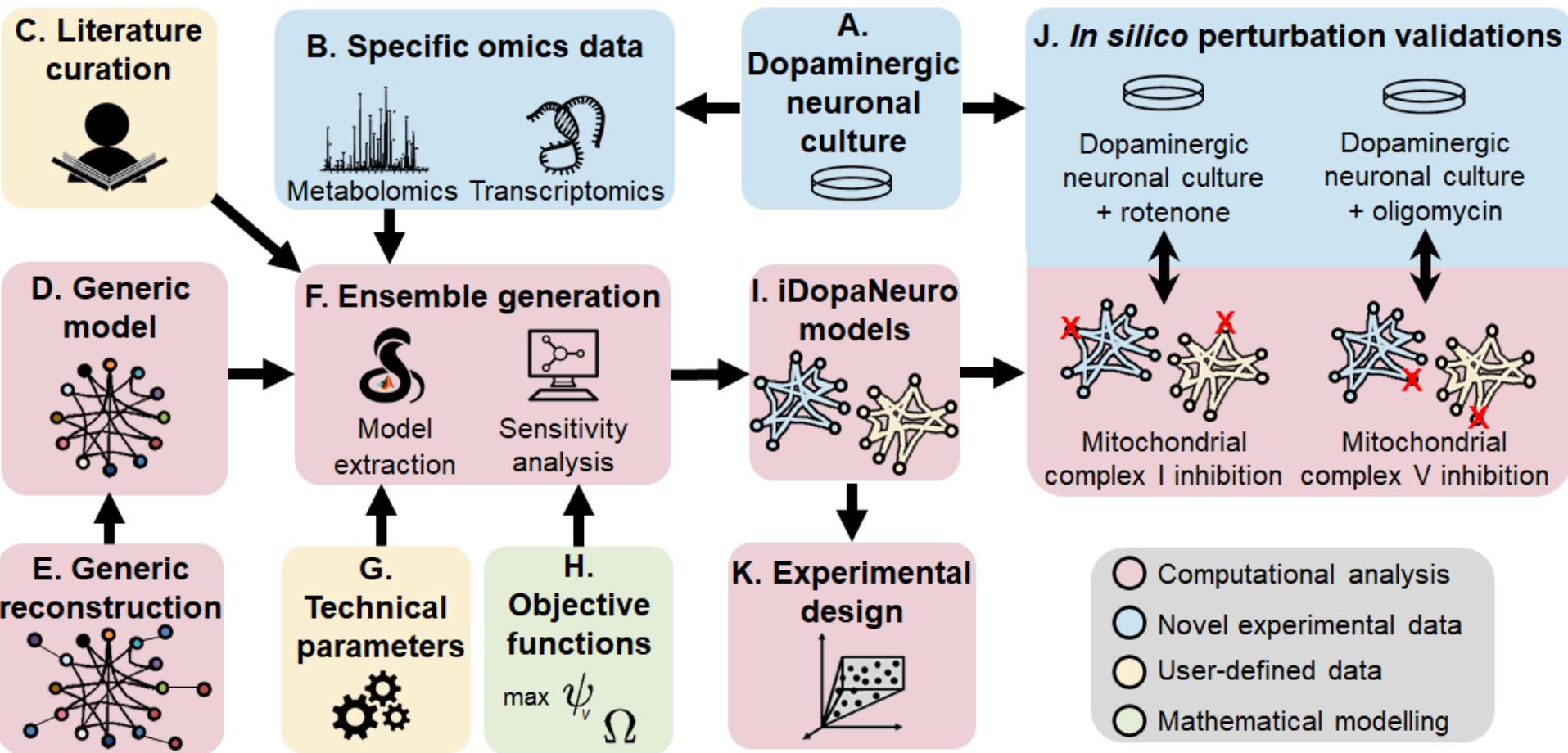


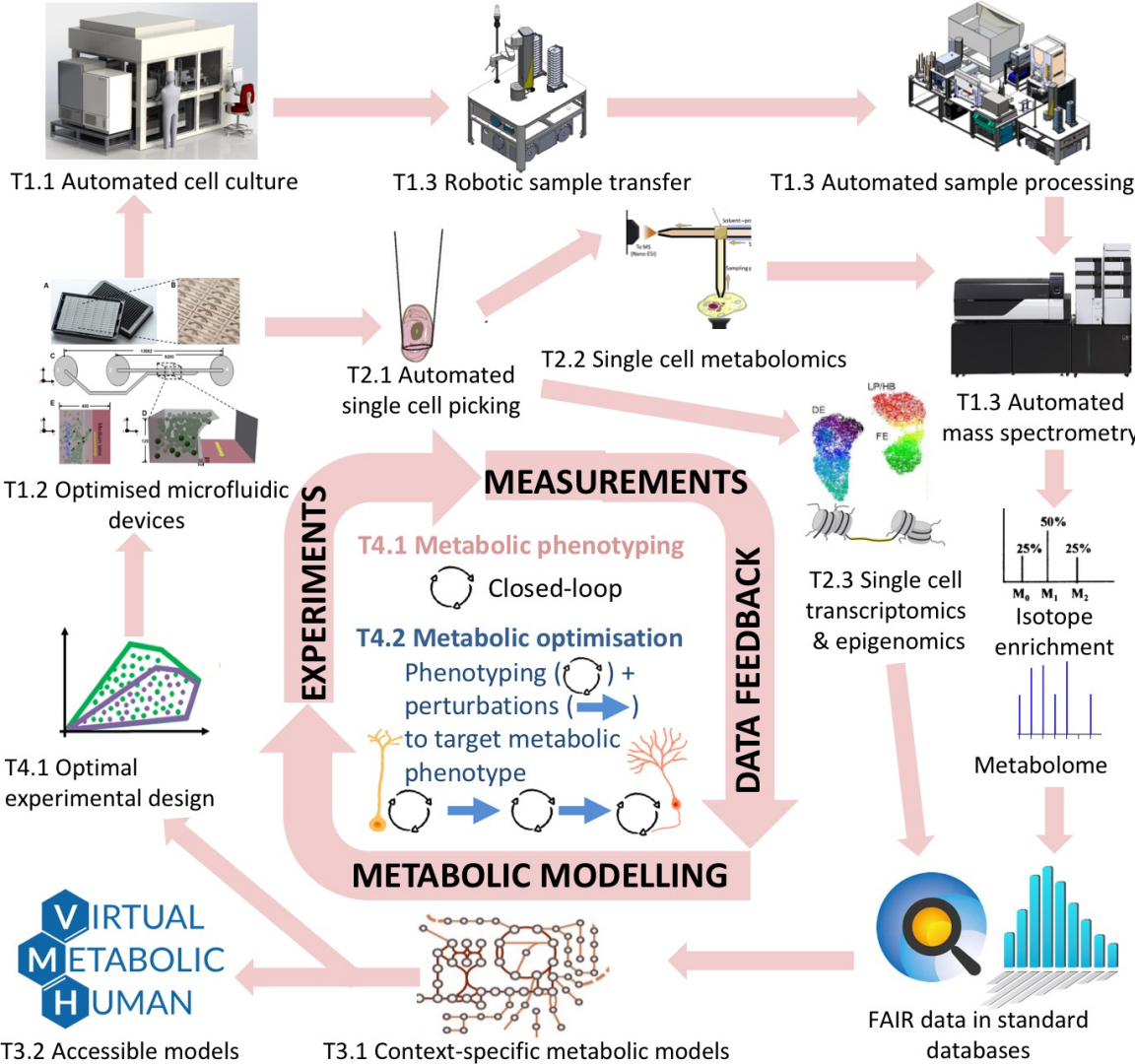


Open challenges

Which one of the 3 next research proposals should be funded?

Semi-automated generation of specific metabolic models

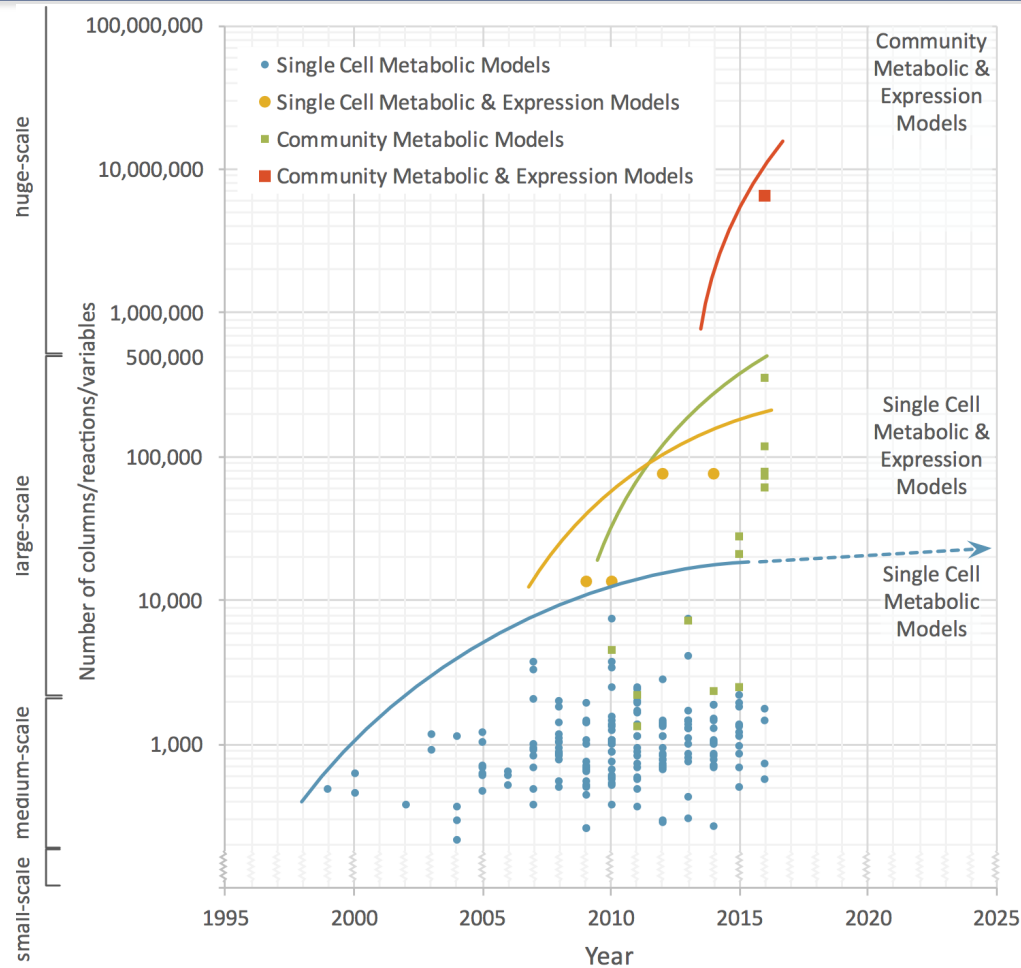


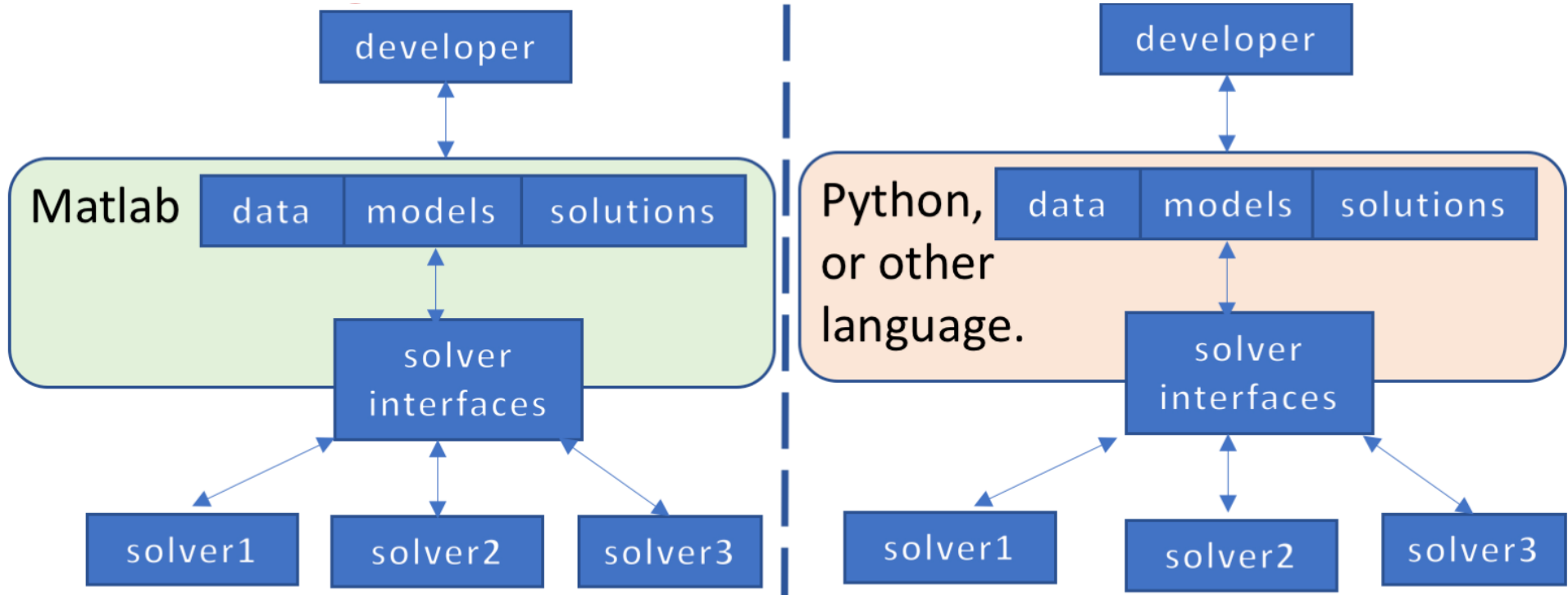


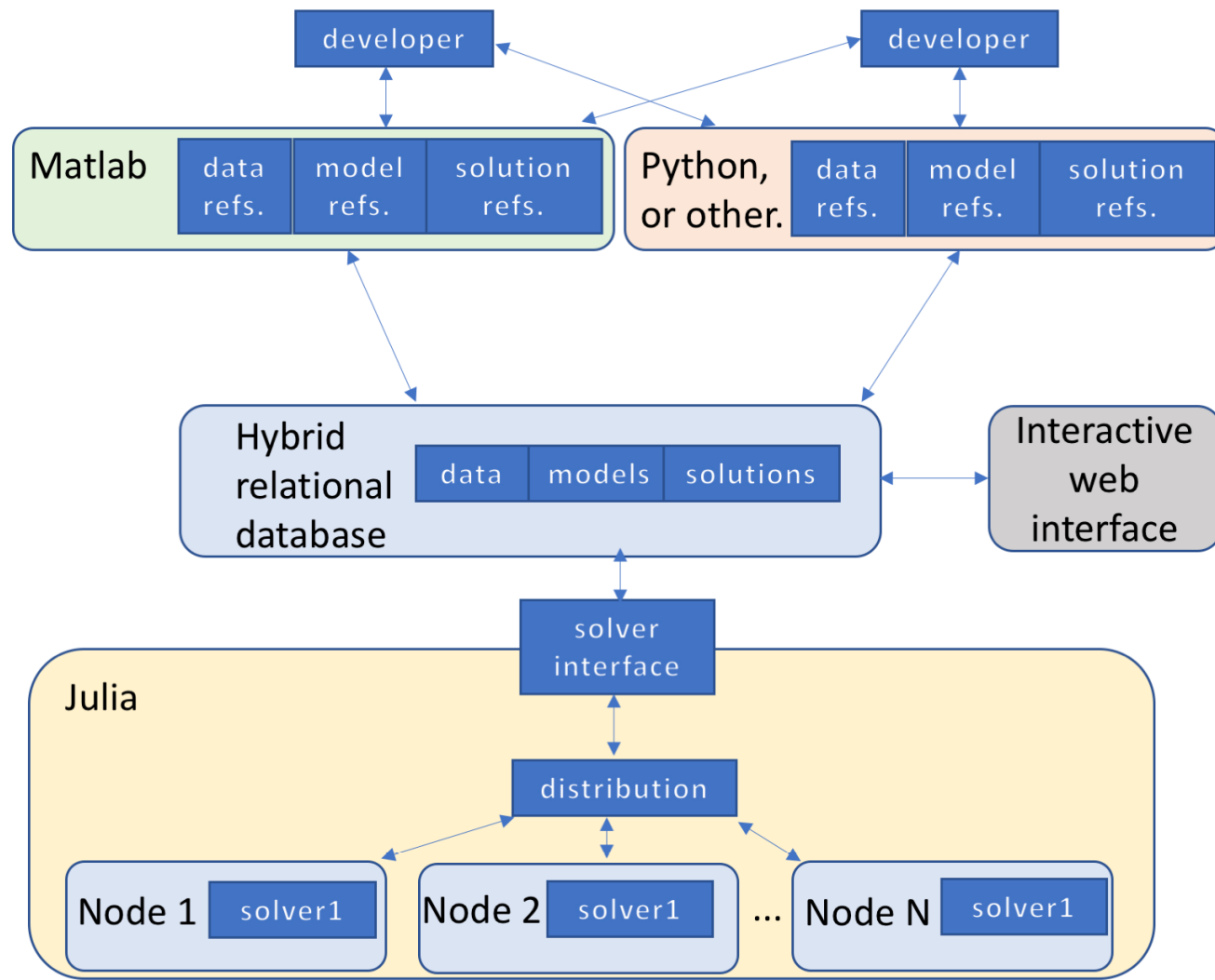
1. Automated in vitro metabolic phenotypic characterisation and optimisation

AutoMetabo (€4 Million)

Growing size of constraint-based models

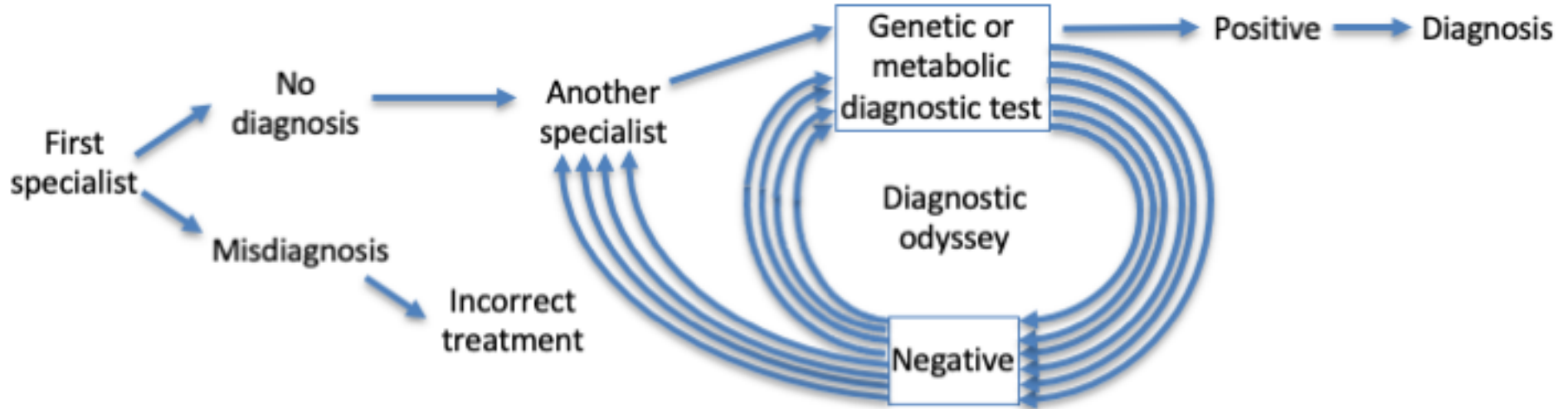




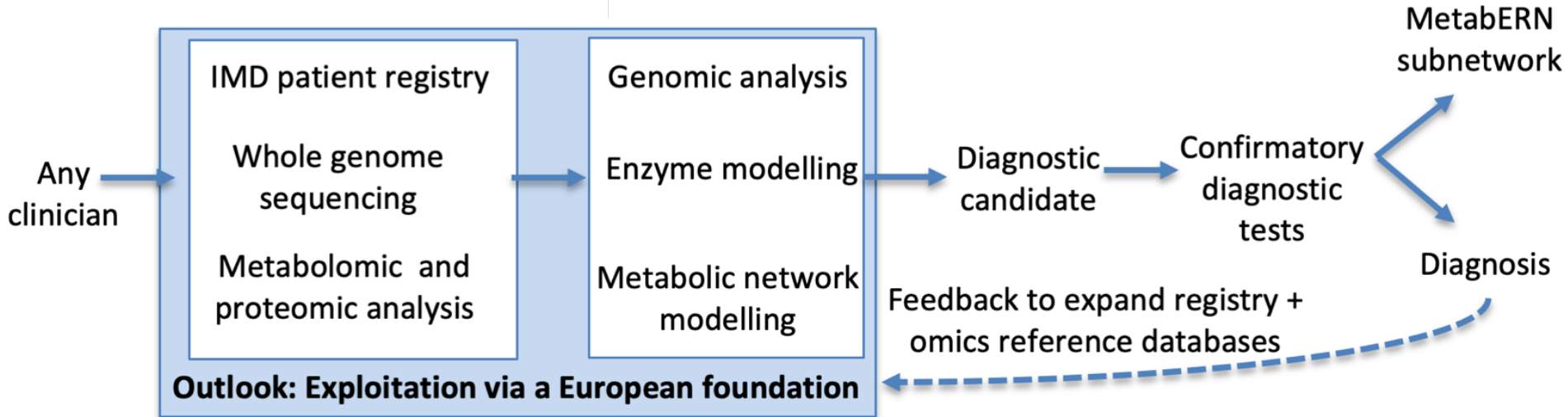


2. High-performance optimisation for molecular systems biology (OptSys, 1.5 Million)

Existing approach to diagnosis of inherited metabolic diseases



3. Reconstruction and Computational Modelling for Inherited Metabolic Diseases (Recon4IMD, € 11 Million)



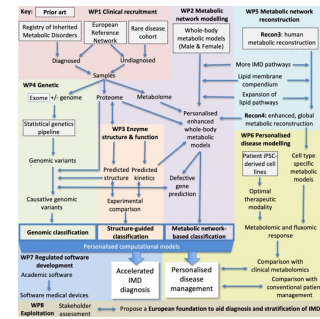
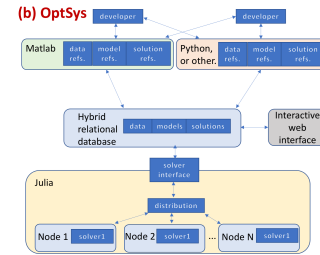
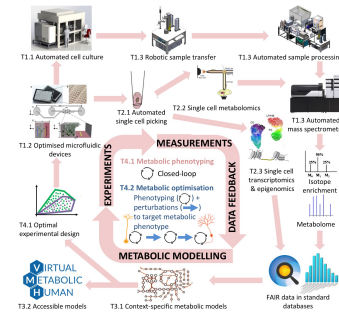
1. €4 Million for “Automated in vitro metabolic phenotypic characterisation and optimisation” (AutoMetabo)

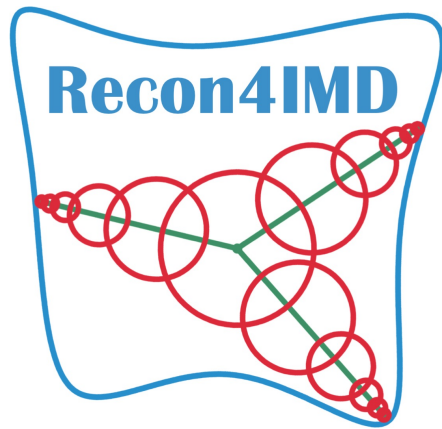
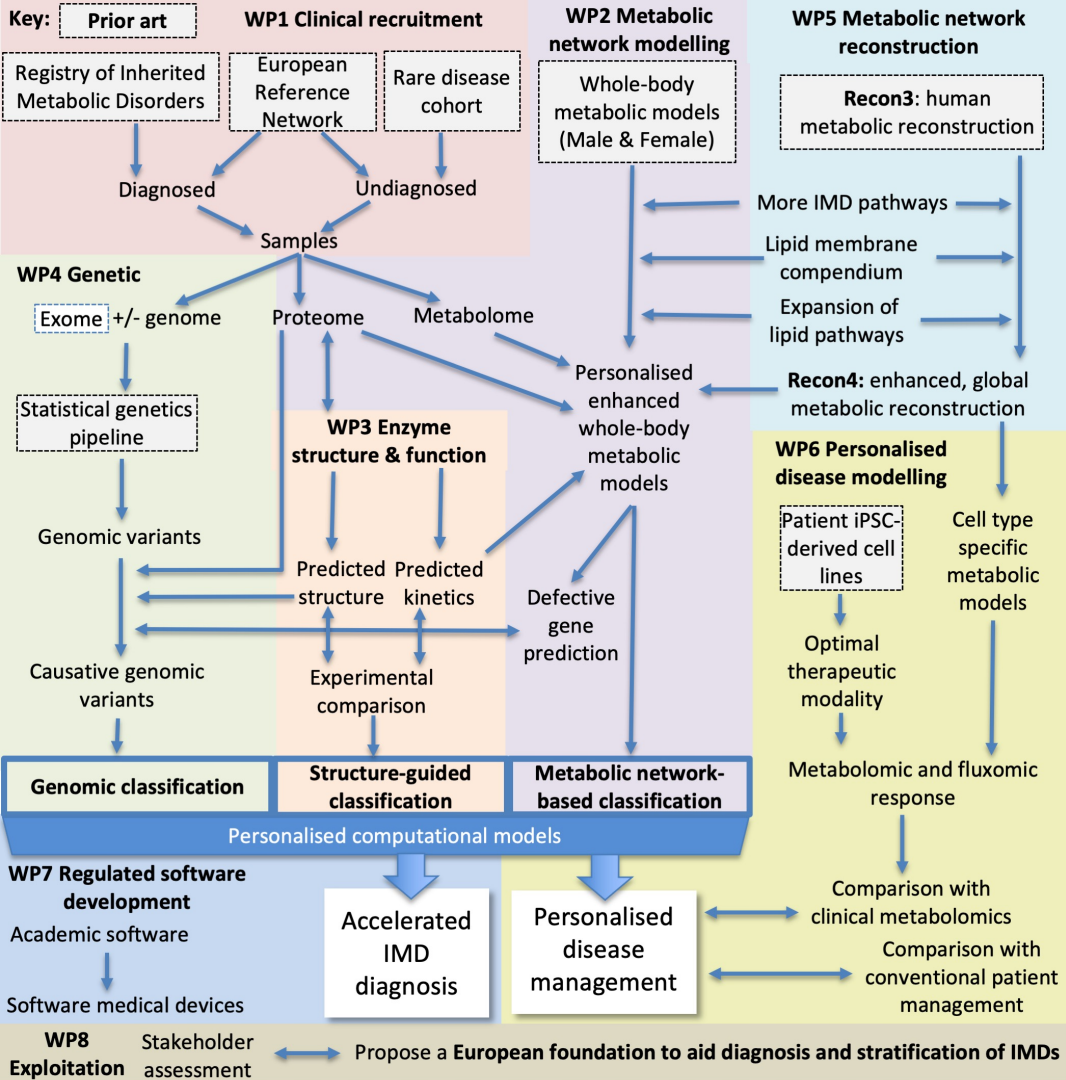
or

2. €1.5 Million for “High-performance optimisation for molecular systems biology” (OptSys)

or

3. €11 Million for “Reconstruction and Computational Modelling for Inherited Metabolic Diseases” (Recon4IMD)





www.recon4IMD.org

20+ open research assistant, PhD, postdoc, clinical fellow, project manager positions open now!

www.recon4imd.org/open-positions

recruit.recon4imd@gmail.com



Thanks for your attention





minimise $\psi(v_f, v_r, w) := \dots$

$$q^T \|v_f - v_r\|_p + \dots$$

$$a^T (v_f - v_r) + \dots$$

$$2v_f^T \ln v_f + \dots$$

$$2v_r^T \ln v_r + \dots$$

◇

◇

Objective function

Weighted p-norm

Linear objective on net flux

Forward flux entropy maximisation

Reverse flux entropy maximisation

(6)

(7)

(8)

(9)

subject to $\Omega := \dots$

$$N(v_f - v_r) + Bw = 0$$

$$C(v_f - v_r) \leq d$$

$$l_v \leq v_f - v_r \leq u_v$$

$$l_w \leq w \leq u_w$$

$$0 \leq v_f \leq u_f$$

$$0 \leq v_r \leq u_r$$

Constraint set

Stoichiometric mass balance

Flux coupling constraints

Net internal flux bounds

Net external flux bounds

Forward internal flux bounds

Reverse internal flux bounds

(10)

(11)

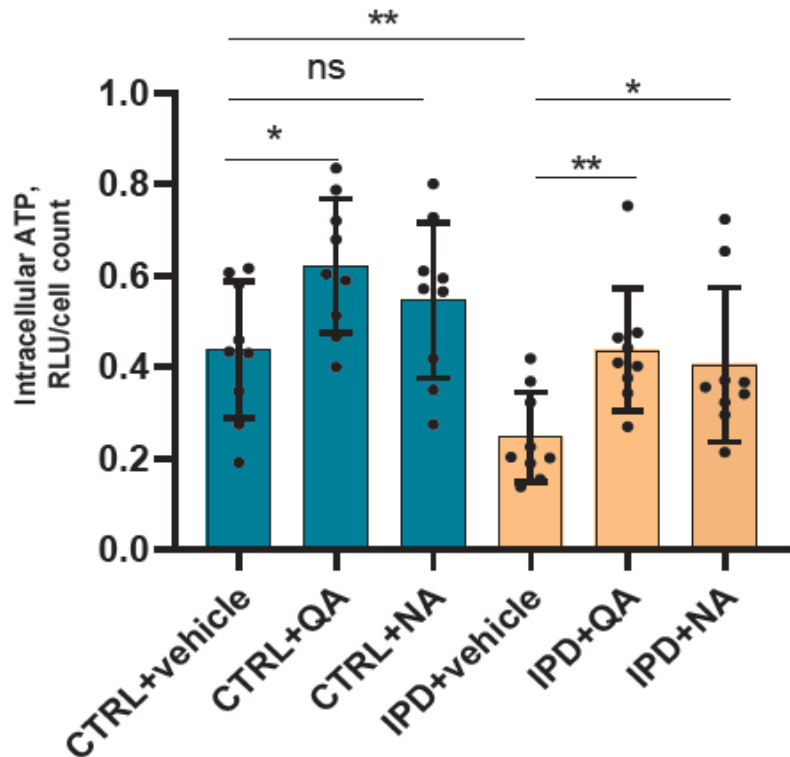
(12)

(13)

(14)

(15)

Neural stem cell metabolism: control vs Idiopathic Parkinson's disease



- In vitro cell culture of iPSC-derived neural stem cells from idiopathic Parkinson's disease versus controls
- Comparison of XomicsToModel generated control and idiopathic metabolic models
- Treatment with NAD precursors quinolinic acid (QA, 20nM) and nicotinic acid (NA, 5mM)
- Recovery of intracellular ATP levels measured in relative light units (RLU), normalized to the cell number in samples treated with vehicle
- Significance asterisks represents $P < 0.05$ * or $P < 0.001$ **.

Constraint-based modelling and data integration identifies a specific idiopathic Parkinson's disease metabolic signature, A. Zagare et al (submitted)