

Metabolomics in Systems Biology: Integrating Multi-omics Data

K.A. Aliferis, I.F. Kalambokis and S. Jabaji

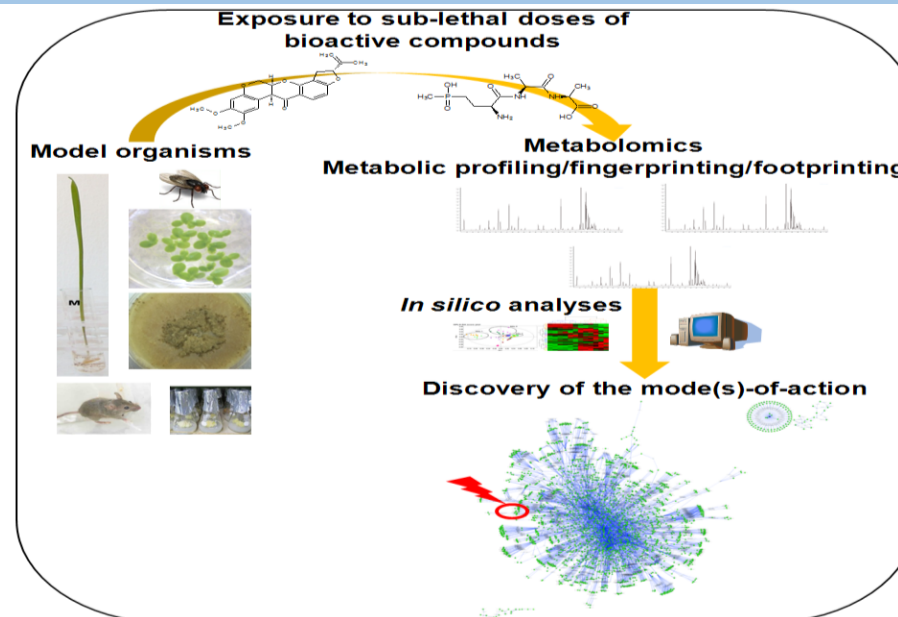
ioannis.kalambokis@aua.gr
konstantinos.aliferis@aua.gr
konstantinos.aliferis@mcgill.ca



Current & Future Projects Applying Metabolomics

➤ Mode of action of bioactive substances

Aliferis and Jabaji, 2011 PBP; Aliferis and Chrysayi 2011, Metabolomics; Aliferis and Chrysayi, 2006 JAFAC



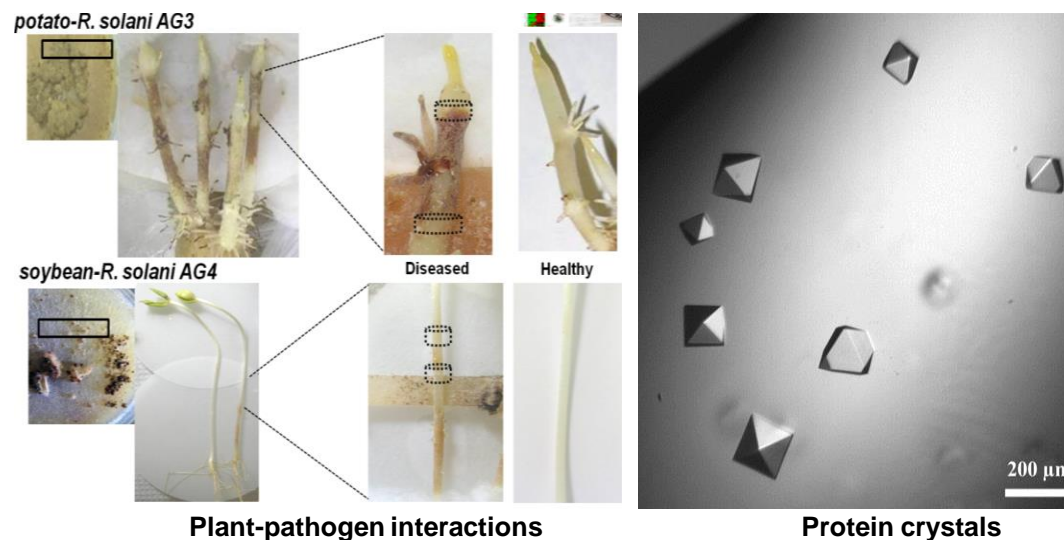
➤ Interactions in biological systems

FQRNT Equipe 2011-2014

➤ Plant and pathogen resistance

➤ Pathogenicity of microorganisms

Aliferis and Jabaji, 2011 Nat Biotechnol., Subm.



Current & Future Projects Applying Metabolomics

➤ Toxicology-ecotoxicology

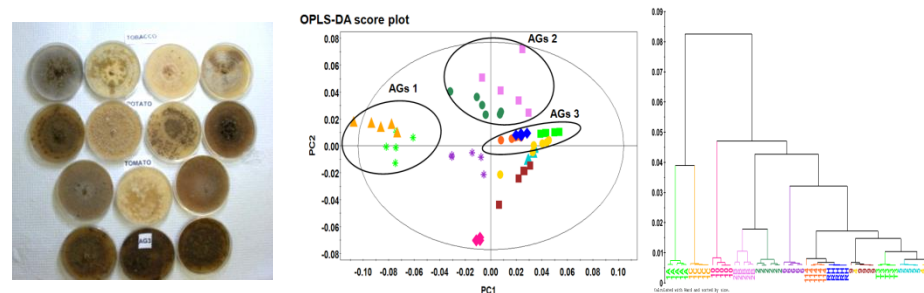
Aliferis et al., 2009 Chemosphere



Genome Canada Project-phytoremediation

➤ Taxonomy of microorganisms

Aliferis and Jabaji, 2011 Metabolomics, Subm.



➤ Pesticide R&D

Aliferis and Jabaji, 2011 PBP; Aliferis and Chrysayi 2011, Metabolomics; Aliferis and Chrysayi, 2006 JAFAC



➤ Risk assessment of GM organisms

➤ Food science

Aliferis et al., 2010, Food Chem.

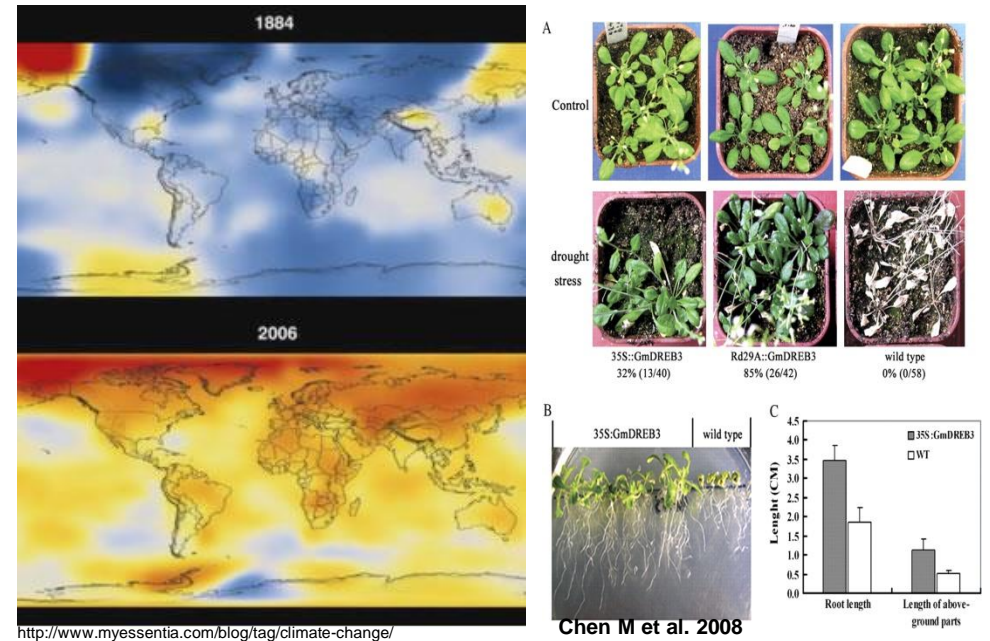
Substantial equivalence was introduced by OECD and FAO as a reliable indicator of GM food safety assessment



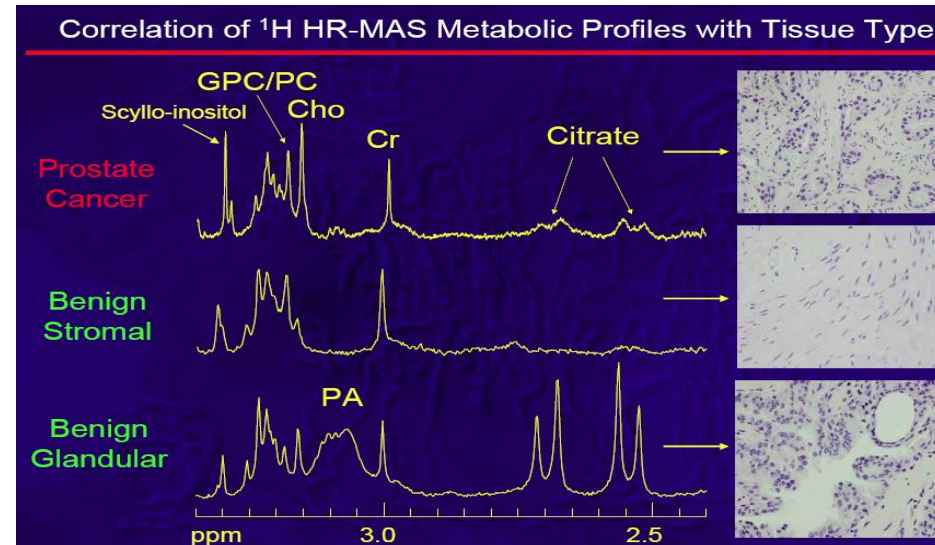
GM canola (MAPAQ project with Dr J. Singh)

Current & future projects applying metabolomics

➤ Climate change, drought/salt tolerance



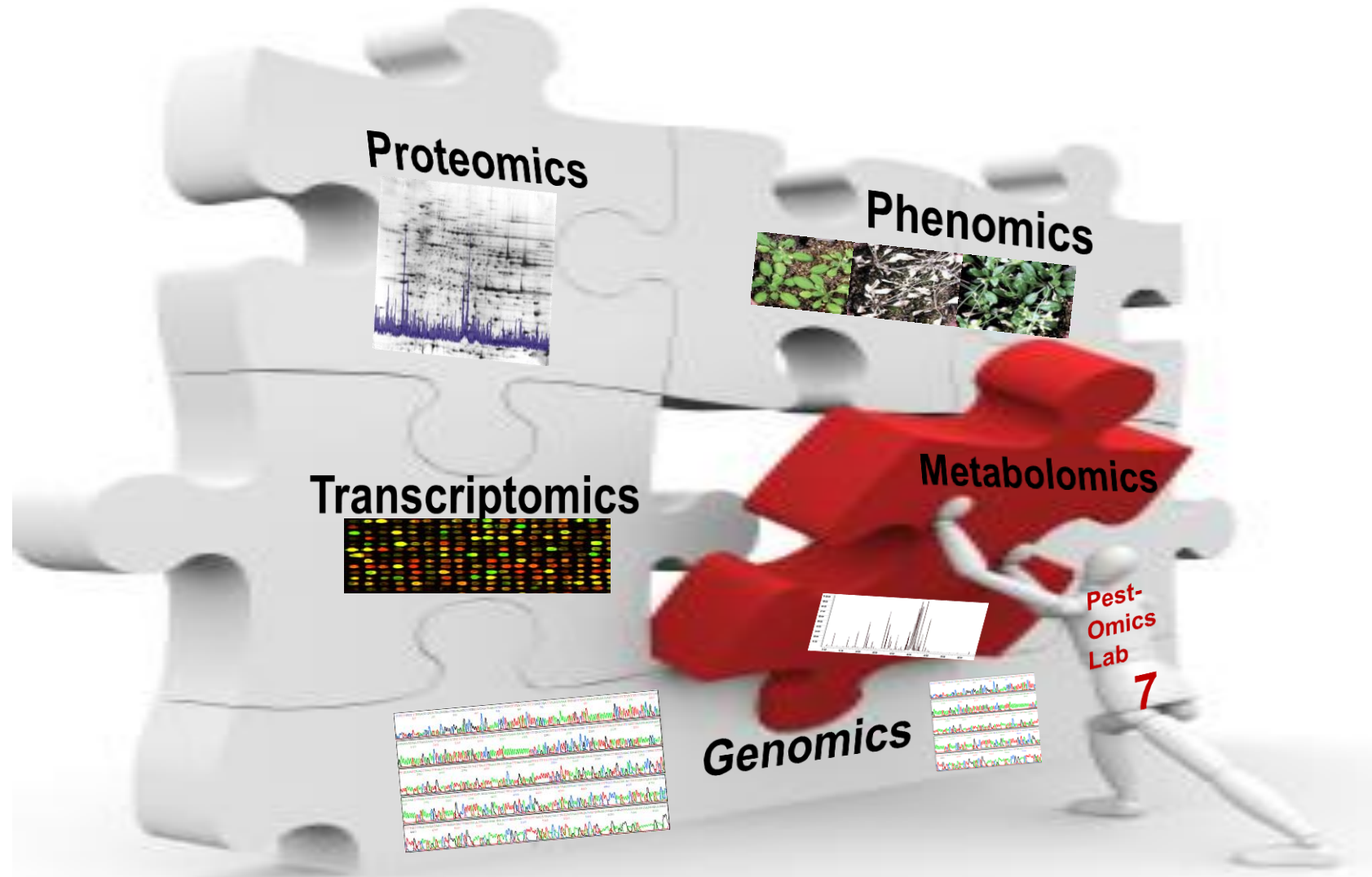
➤ Human-animal physiology and diseases



<http://www.iatmo.org/presentazioni/pdf/Conti.pdf>

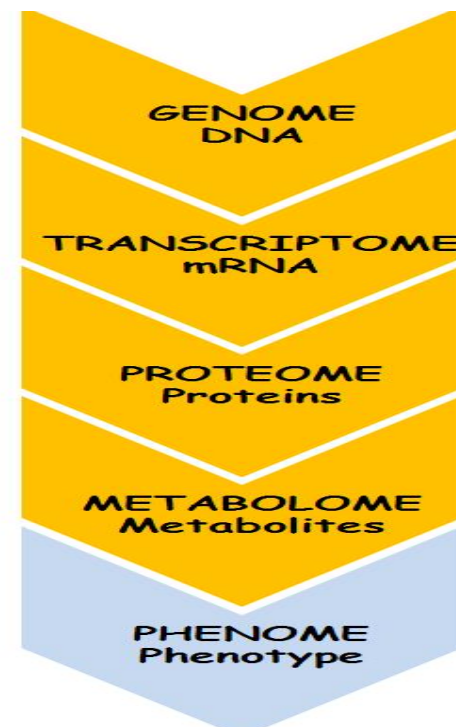
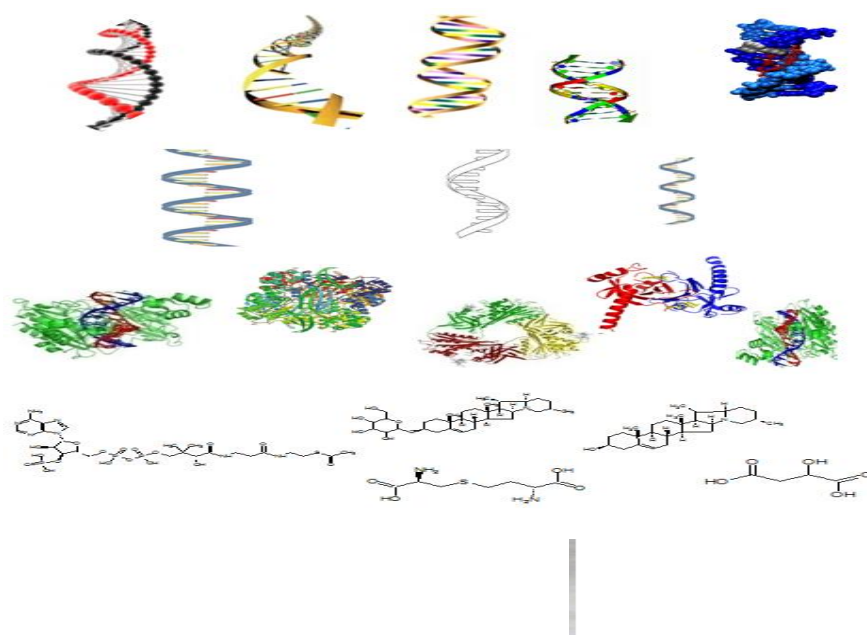
The Jigsaw Puzzle of Systems Biology

“the whole is not represented by the sum of its components”
Aristotle 384-322 BC



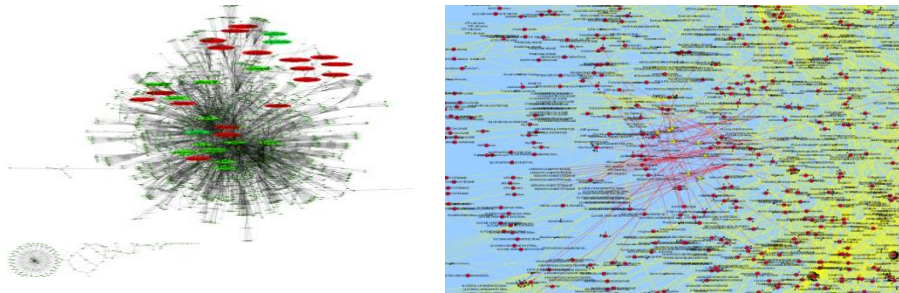
Why Integration of “omics”

- Facilitate the profiling and functional characterization of biomolecules in complex biological systems
- Retain all the advantages of more classical genomics/proteomics approaches
- Gain the capacity to elucidate endogenous biochemical activities for genes/proteins via metabolite profiling

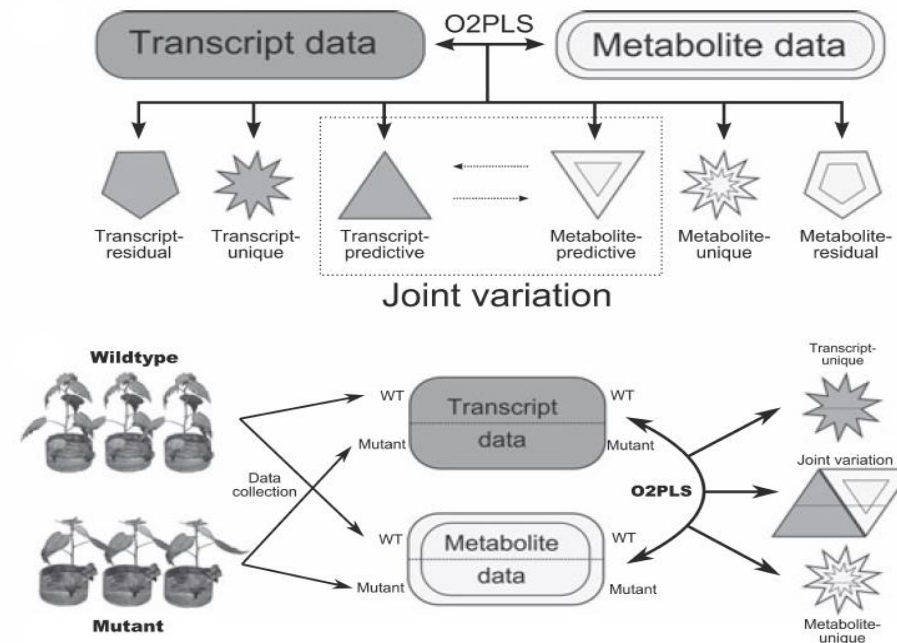
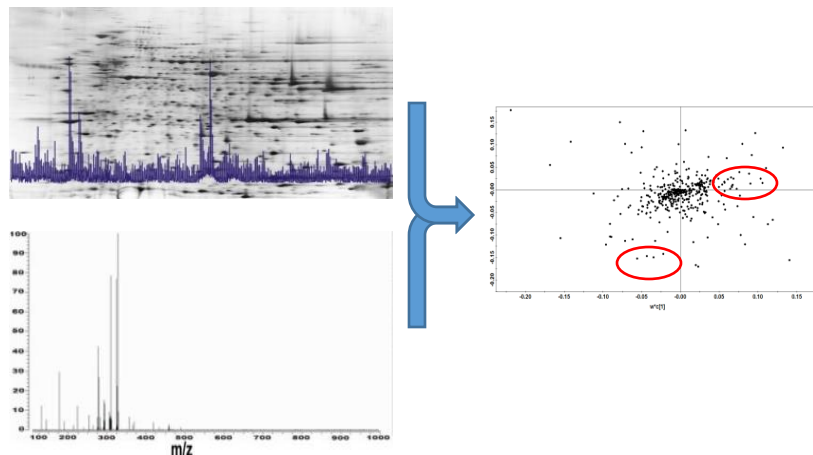


Integration of “omics”, but how?

- 1) Bioinformatics software (i.e., Simca P+, Cytoscape) and databases (i.e., BioCyc)
Correlation between metabolites, enzymes and genes



- 2) Multivariate analyses of combined metabolomics-proteomics (i.e., shotgun proteomics) or transcriptomics data: discovery of trends between metabolite level changes and gene expression

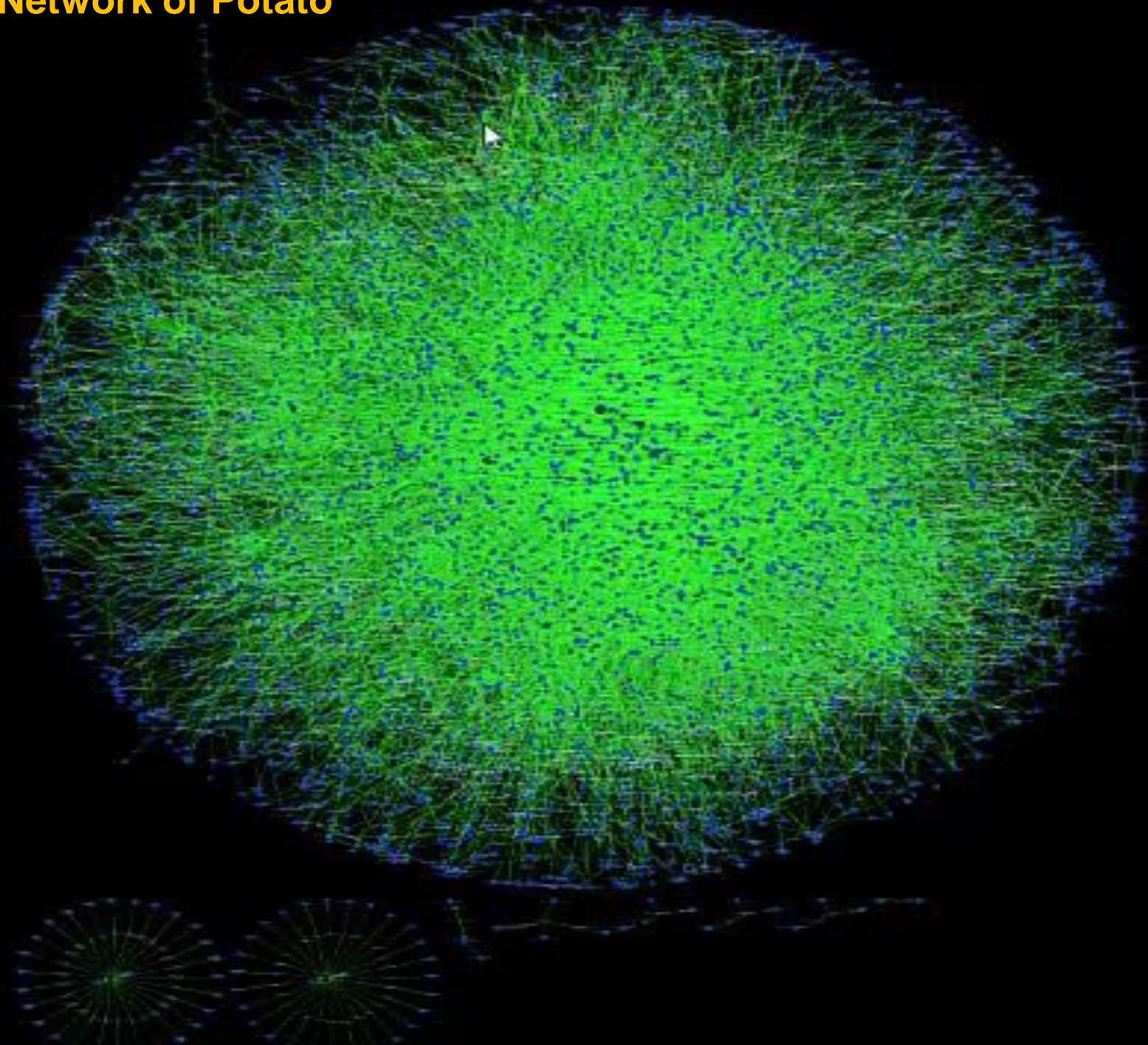


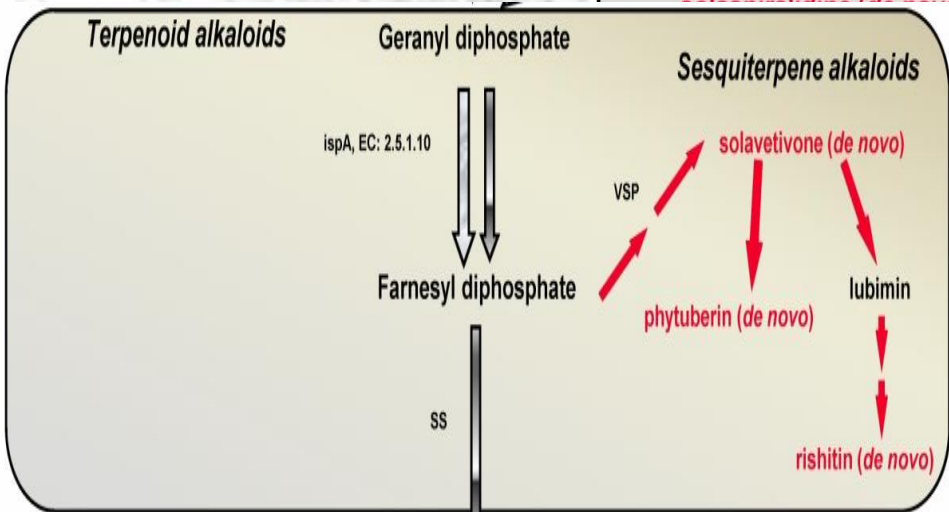
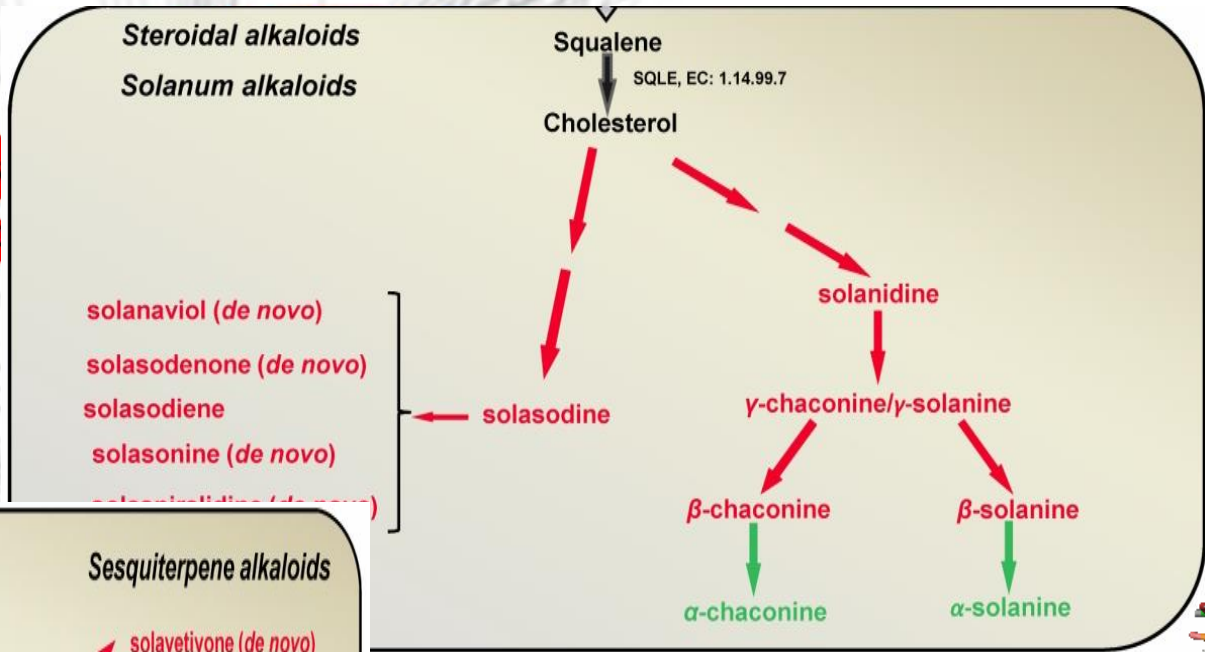
Bylesio et al., 2007

**Bioinformatics software (i.e., Simca P+, Cytoscape) and databases (i.e., BioCyc)
Correlation between metabolites, enzymes and genes**



Global Metabolic Network of Potato





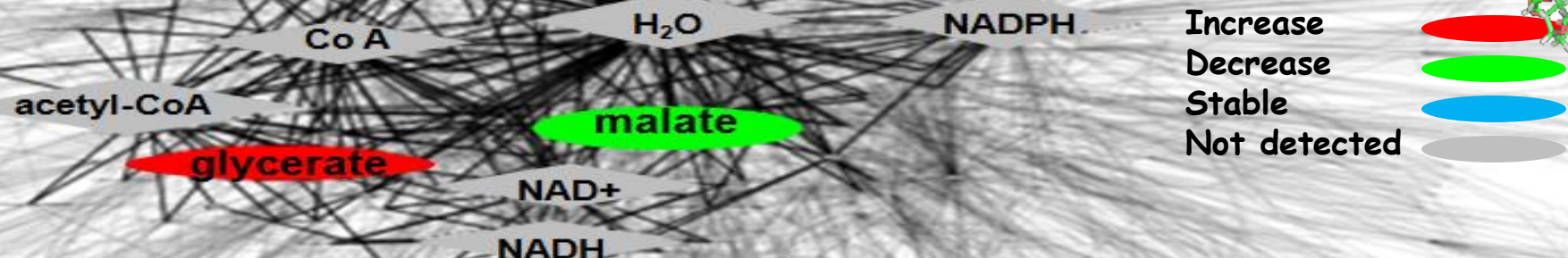
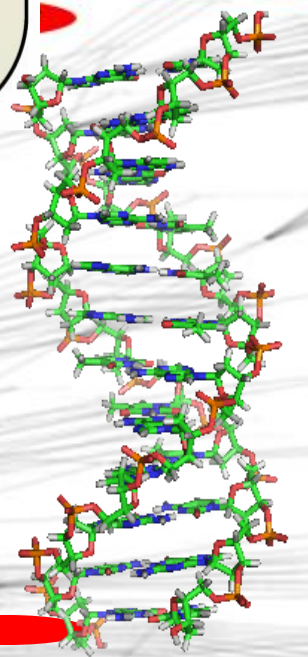
anaviol*

solaspiralidine*

solasodiene*

trehalose

[SGNU269148](#), [SGNU269149](#),
[SGNU287228](#), [SGNU287229](#)
[SGNU270961](#), [SGNU270962](#),
[SGNU270963](#), [SGNU269148](#)



2. Multivariate analyses of combined metabolomics, shotgun proteomics, or transcriptomics data

The Concept of Orthogonal Variation

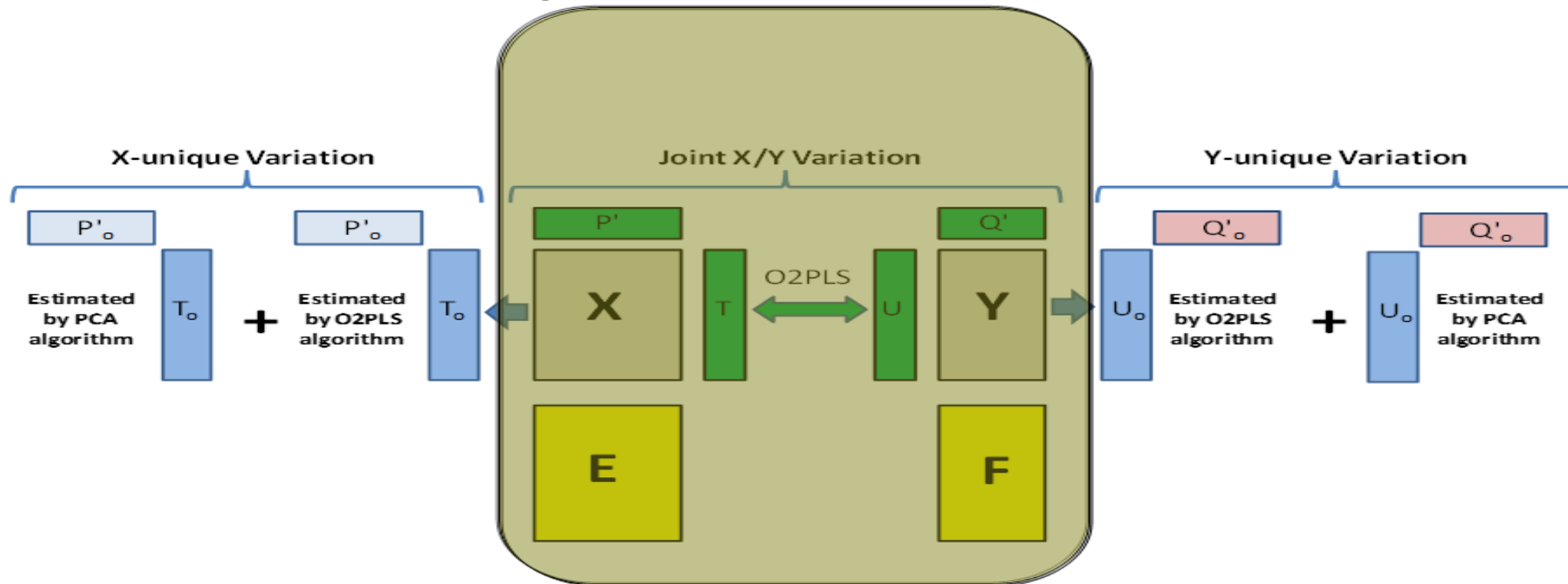
- Not all systematic variation in the X-block (e.g., metabolomics data set) is related to the Y-block (e.g., transcriptomics data set)
- The new 'O'-methods, OPLS (Orthogonal PLS) and O2PLS (Two-way Orthogonal PLS), are able to divide the systematic X-variation in two parts:
 - What in X is correlated to Y (**predictive variation**)
 - What in X is not correlated to Y (**orthogonal variation**)
 - What in Y is not correlated to X (**orthogonal variation**)
- The orthogonal variation is important information for the total understanding of the studied system or process

Introduction to OPLS/O2PLS

- Regression challenge
PLS and OPLS are **unidirectional**, i.e., $X \rightarrow Y$
- Integration challenge
O2PLS is **bi-directional**, i.e., $X \leftrightarrow Y$
- Differences in preferred terminology
OPLS: **Predictive & Orthogonal variabilities**
O2PLS: **Joint & Unique variabilities**

O2PLS

- The primary objective of O2PLS is the **integration of data in the two-block (X/Y) context** and understanding of which information overlaps between the two data sets and which information is unique to a specific data set (X or Y)
- O2PLS accomplishes this task by incorporating three types of components,
 - components expressing the **joint X/Y** information overlap,
 - components expressing what is **unique to X**, and
 - components expressing what is **unique to Y**

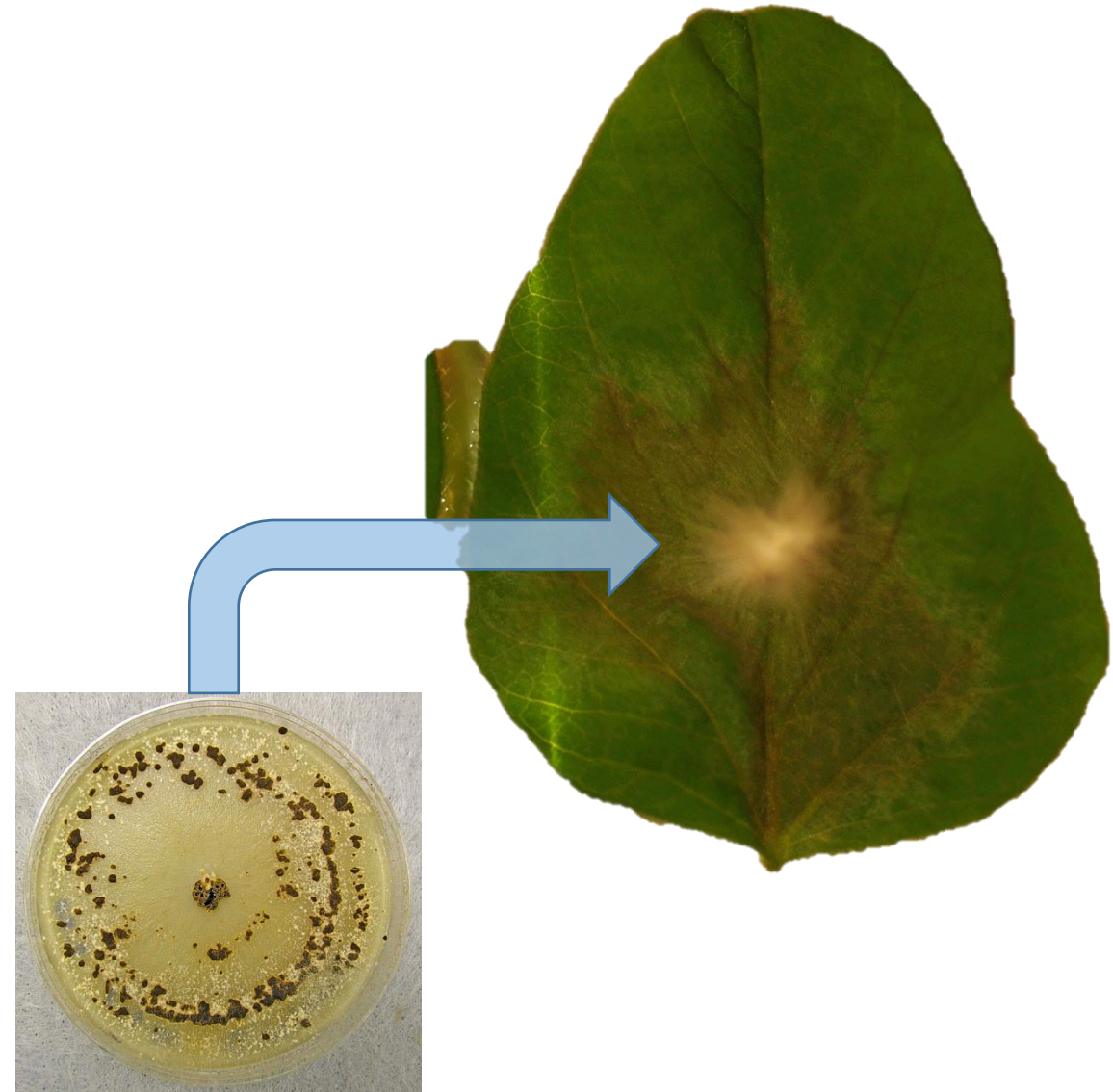


**Integration of RNAseq with ^1H NMR Data Sets Reveals
Differential Regulation of Soybean Primary Metabolism in
Response to Rhizoctonia Foliar Blight**

Aliferis et al. 2016. Plant Cell, Subm

Soybean-Rhizoctonia Foliar Blight Pathosystem

- *R. solani* AG-1 1A
- **Necrotic** fungal pathogen
- Up to **60%** annual yield losses on soybean
- Little knowledge on soybean defense mechanisms

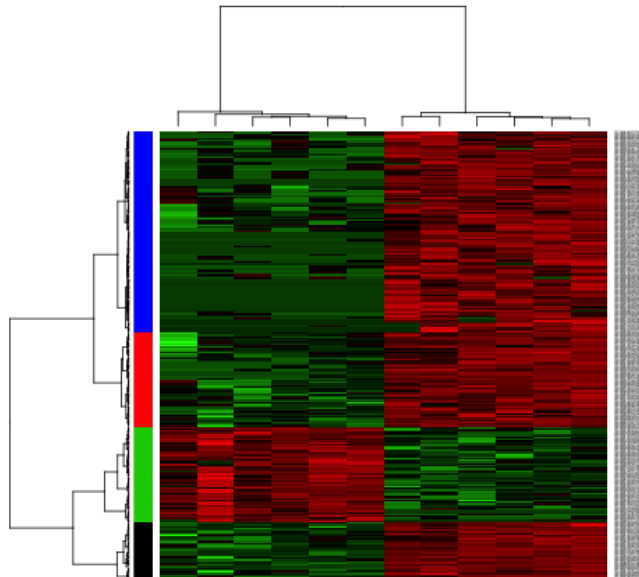


TRANSCRIPTOMICS

```
>AF025366#  
CCTGGCTCAGATTGAAAGTGTGGCGGCAGGCCTAACACATGCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTTTGCT  
>AF530475#  
GCTGGCGGCAGGCCTAACACATGCAAGTCGAACGGTAACAGGRATCAGCTTGCTGATTGCTGCTGACGAGTGGCGGACGGC  
>AJ508303#  
GGCGGCAGGCCTAACACATGCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTGCTGACGAGTGGCGGACGGGTGAC  
>AJ853889#  
ATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTGCTGCTGACGAGTGGC  
>AJ853890#  
ATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTGCTGCTGCTGCTGCTGCTG  
>AJ854063#  
AGAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  
>AY217654#  
AGAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  
>AY995561#  
GGCGCGGGAATTCGATTAGAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  
>EF059845#  
CTAACACATGCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  
>EF059872#  
GGCTCAGATTGAAACGGCGGCGGCAGGCCTAACACATGCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTGCTGCTGCTG  
>EF059884#  
ACATGCAAGTCGAACGGTAACAGGAAGCAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
```

INPUT

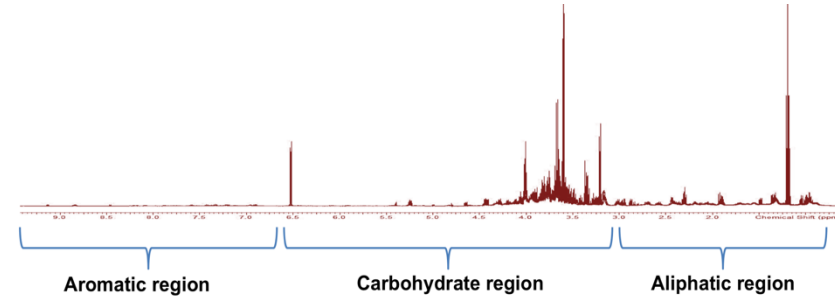
Fastq sequence files



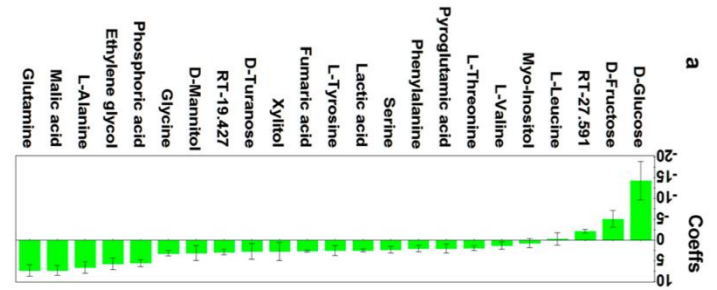
OUTPUT

Heat map

METABOLOMICS



NMR resonance signals



PLS coefficient plots

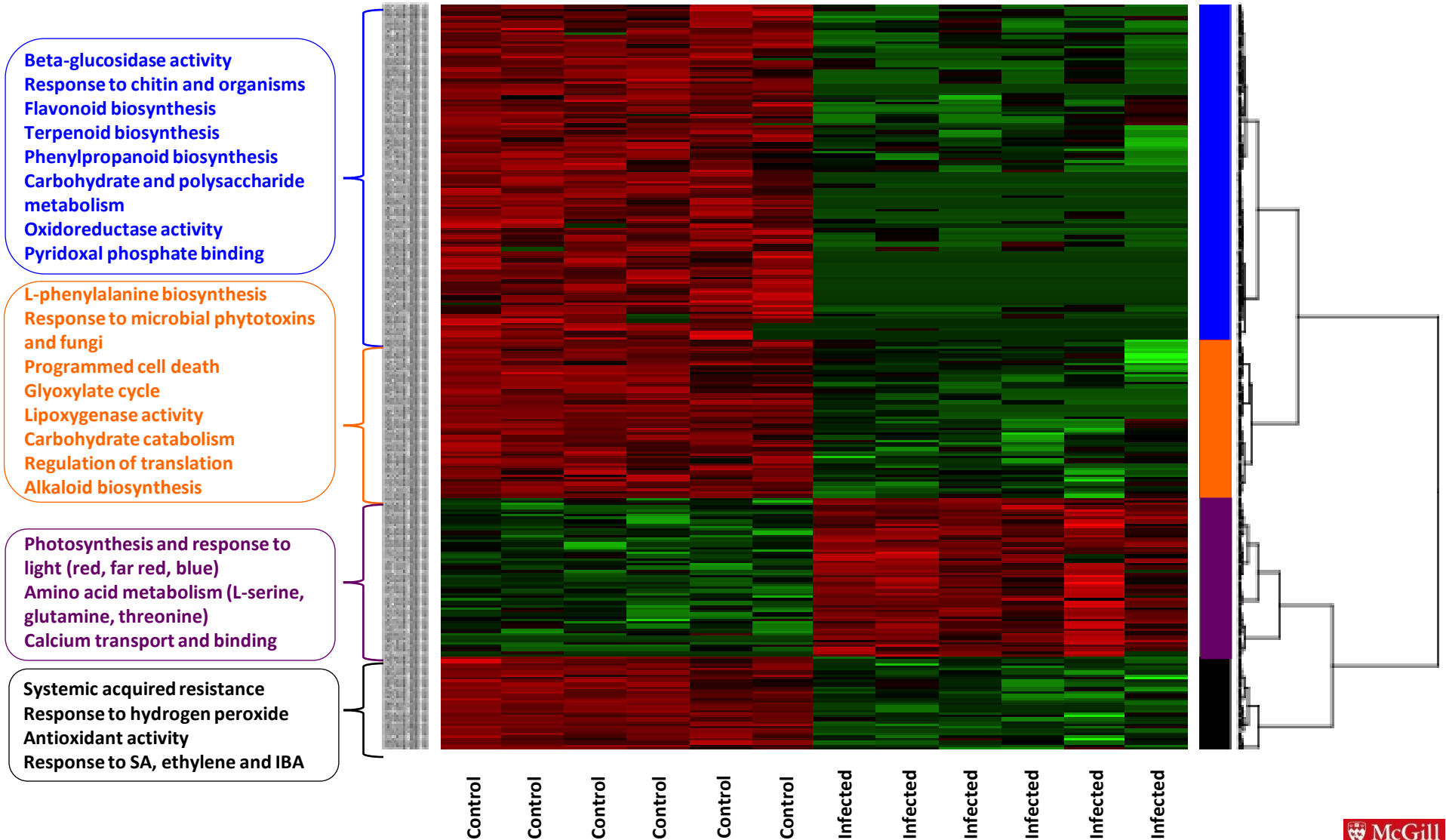


O2PLS

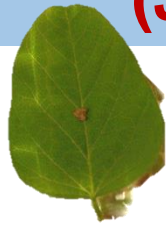
258 DE genes

55 down-regulated

203 up-regulated

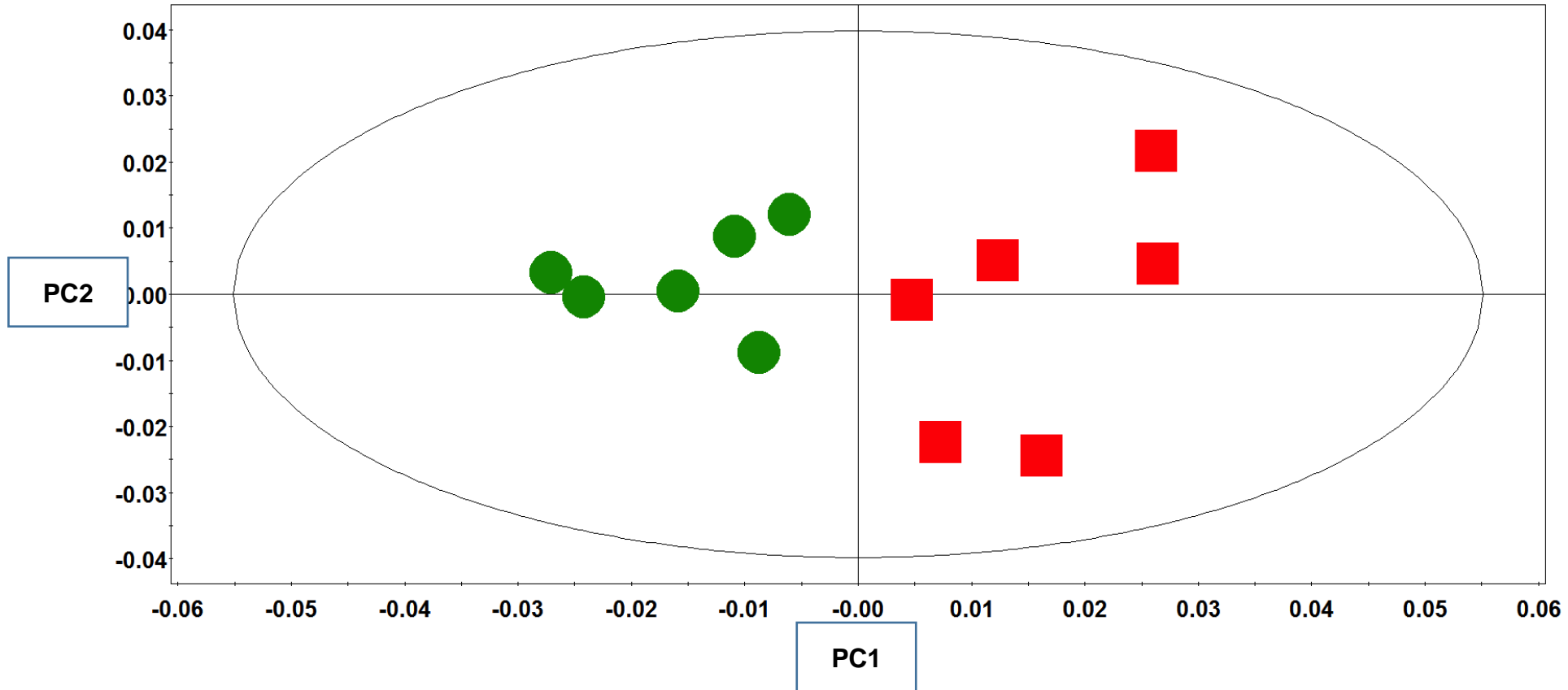


O2PLS Score Plot (JOINED METABOLOMICS-TRANSCRIPTOMICS)



← CONTROL LEAVES

→ INFECTED LEAVES



Conclusions

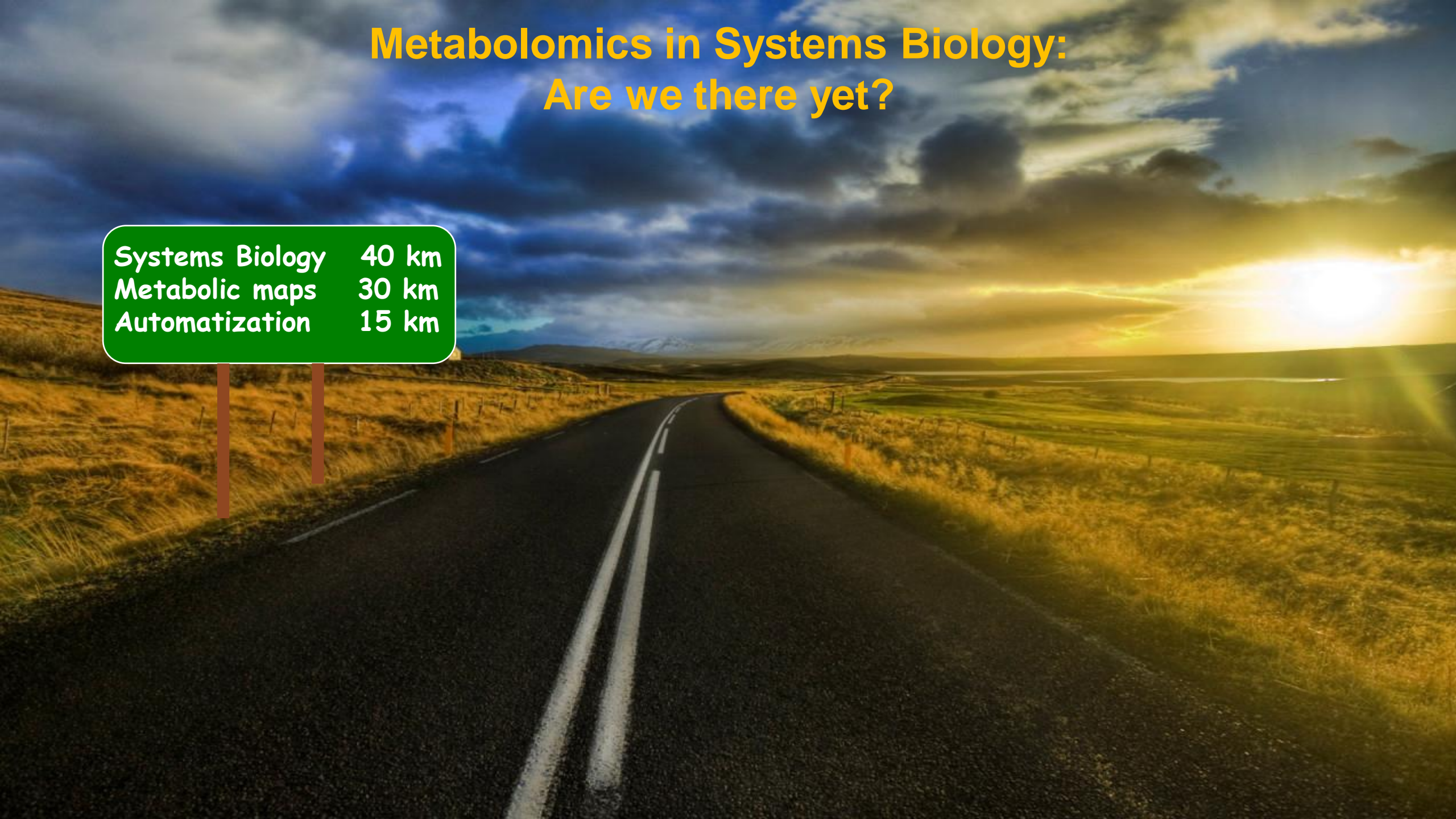
- **O2PLS is suitable for analyzing data from different platforms.**
- **O2PLS highlights potential mechanisms whereas OPLS highlights potential biomarkers**
- **The same approach described here can easily be applied to other varying data sources, for example, transcriptomics vs. proteomics data etc**

Freely available software for metabolomics data analyses and visualization

Software	Web address	Reference
COMSPARI	http://www.biomechanic.org/comspari/	Katz et al. (2004)
Cytoscape	http://www.cytoscape.org/	Shannon et al. (2003)
DrDMASS	http://kanaya.aist-nara.ac.jp/DrDMASS/	Oikawa et al. (2006)
KaPPA-View	http://kpv.kazusa.or.jp/kappa-view/	Tokimatsu et al. (2005)
KNApSAcK	http://kanaya.naist.jp/KNApSAcK/	Shinbo et al. (2006)
MapMan	http://mapman.gabipd.org/web/guest/mapman	Thimm et al. (2004)
MassTRIX	http://metabolomics.helmholtz-muenchen.de/masstrix2/	Suhre and Schmitt-Kopplin (2008)
MathDAMP	http://mathdamp.iab.keio.ac.jp/	Baran et al. (2006)
Metaboanalyst	http://www.metaboanalyst.ca	Xia et al. (2009)
MetHouse	http://msbi.ipb-halle.de/msbi/metware/methouse	Gaida and Neumann (2007)
MET-IDEA	http://met-idea.software.informer.com/download/	Broeckling et al. (2006)
MOLGEN	http://www.molgen.de/	Kerber et al. (2001)
MSFACTs	http://www.noble.org/PlantBio/Sumner/msfacts/index.html	Duran et al. (2003)
Mzmine	http://mzmine.sourceforge.net/	Katajamaa et al. (2006)
OmicSpace	http://omicSpace.riken.jp/gps/	Toyoda and Wada (2004)
PathVisio	http://www.pathvisio.org/	van Iersel et al. (2008)
PlantCyc	http://www.plantcyc.org/	
PRIME	http://prime.psc.riken.jp/	Akiyama et al. (2008)
SetupX	http://fiehnlab.ucdavis.edu:8080/m1/index.jsp	Scholz and Fiehn (2007)
Seven Golden Rules Software	http://fiehnlab.ucdavis.edu/projects/Seven_Golden_Rules/Software/	Kind and Fiehn (2007)
SIRIOUS	http://bio.informatik.uni-jena.de/sirius/download/index.html	Böcker et al. (2009)
TargetSearch	http://www.bioconductor.org/packages/2.5/bioc/html/TargetSearch.html	Cuadros-Inostroza et al. (2009)
VANTED	http://vanted.ipk-gatersleben.de/	Junker et al. (2006), Klukas and Schreiber (2007)
XCMS	http://metlin.scripps.edu/download/	Smith et al. (2006)

Metabolomics in Systems Biology: Are we there yet?

Systems Biology	40 km
Metabolic maps	30 km
Automatization	15 km



Acknowledgments

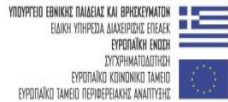
Dr S. Jabaji, McGill University
Dr B. F. Lang, Université de Montréal
Dr M. Labreque, Université de Montréal, IRBV
Dr D. Faubert, Institut de Recherches Clinique de Montréal (IRCM)
Dr T. Sprules, Quebec High Field NMR Facility, McGill University
Dr J. Sutton, Thermo Scientific, Waltham, MA, USA
Dr R. Chamoun, McGill University
Mrs M. Rani, McGill University
FunLab, McGill University



Genome Québec



Genome Canada



Fonds de recherche
sur la nature
et les technologies

