Biomathematics and Bioinformatics at Rothamsted Research

Professor Chris Rawlings

chris.rawlings@bbsrc.ac.uk
www.rothamsted.bbsrc.ac.uk/bab

Outline

- Overview of Rothamsted Research
- What we do in Biomathematics
- What we do in Bioinformatics

Rothamsted Research

- Largest agricultural and crop science research institute in UK
- Research started in 1853
- 400 Staff

New Facilities

Rothamsted's six scientific goals

- Enhanced crop quality and production efficiency (in the context of sustainability)
- Environmentally sensitive management practices for arable agriculture and associated habitats
- Protection of soils and the global environment
- New products from crops
- Conservation and exploitation of biodiversity
- Elevation of public confidence in science related to agriculture

ROTHAMSTED RESEARCH

Biomathematics and Bioinformatics

- Integrate data from multiple biological sources and develop tools to analyse and interpret results
- Exploit mathematics and computational sciences to develop methods for detection of subtle signals in complex and noisy datasets
- Develop predictive systems models of plants and their interactions with pathogens and the environment at a variety of scales
- Validate and apply the models to support the development of sustainable agricultural practises

ROTHAMSTED
RESEARCH

What we do in Biomathematics and Bioinformatics

- 30 Staff, 5 PhD Students
- Statistics
  - research into new statistical methods
  - consultancy with scientists (including quality assurance issues)
  - statistics training
- Modelling (3 groups)
  - development of population dynamic and simulation models of biological processes
- Bioinformatics
  - research on efficient storage, processing, integration and analysis of ‘omics’ data (genes, proteins, pathways etc)
  - Development of databases and software
  - Consultancy – internal collaborations applied bioinformatics

ROTHAMSTED RESEARCH

Applied Statistics Research

Robin Thompson, Sue Welham

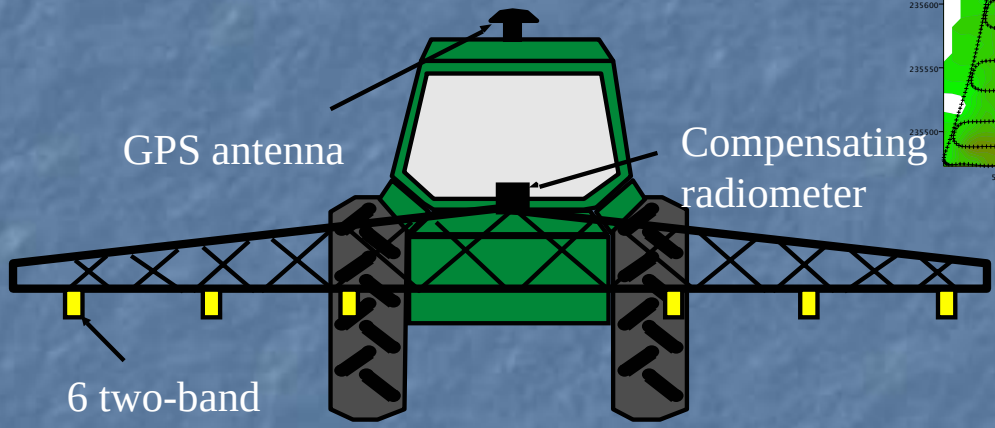
- General methods for dealing with noisy and complex datasets
- Novel approach to modelling correlated effects (errors) in data
  - Common in biology because of “hidden” relationships e.g. genetic, time, space
  - REML, AsREML
- More sensitive for detecting small signals in lots of noise
  - E.g gene effects in a dominating background of environment effects
  - Applications in analysis of gene expression data
- Contribute to GenStat software with VSN international

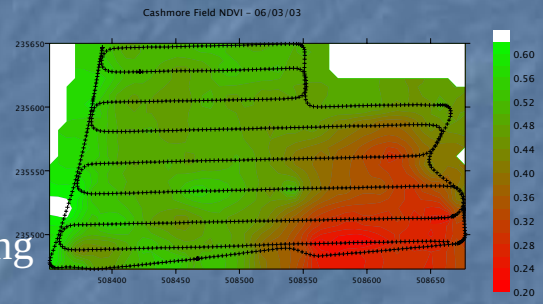
ROTHAMSTED RESEARCH

Environmetrics
Murray Lark

- Studying spatial variability in environment
  - Nutrient concentration
  - Water retention
  - *Heavy metal contamination*
- Combine geostatistics with other methodologies
  - Wavelets
  - Bayesian Networks
- Applications
  - Precision Farming
  - Optimal sampling technologies and telemetry
  - Environmental pollution monitoring/control

Telemetry & Adaptive Sampling





ROTHAMSTED RESEARCH

Mathematical Modelling

Frank van den Bosch

- Interactions between crop, pest and environment
  - Optimising crop performance, pesticide management
- Population Dynamics
  - Understanding how pests co-evolve with host plant e.g. in response to disease resistant varieties
- Mathematical Physics – *Andy Reynolds*
  - Fluid dynamic models of movements of particles in turbulent air flow – applications to insect behaviour and fungal spore dispersal

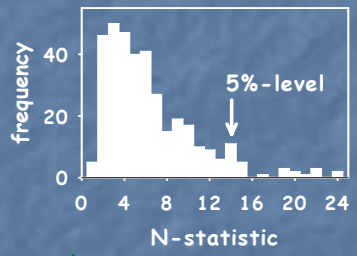
ROTHAMSTED RESEARCH

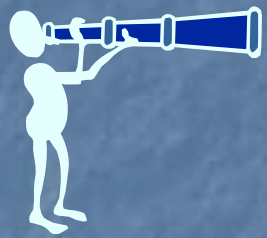
Plant pathogen population dynamics

Methods developed to study disease-weather relations.
Applied to light leaf spot, stem canker, Septoria

Future aim: To develop generic methodology for disease forecasting system and to apply these to foliar plant pathogens.

Delivery through decision support systems





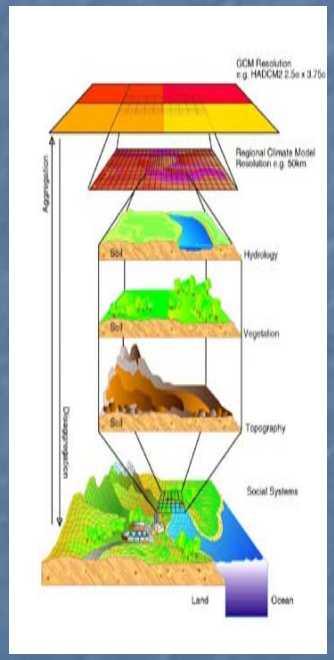
sowing

GS31

GS75

Climate and Crop Models

Mikhail Semenov



Meta-model

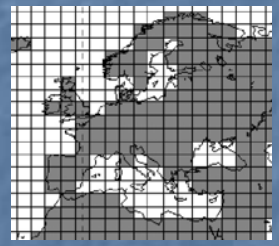
Upscaling

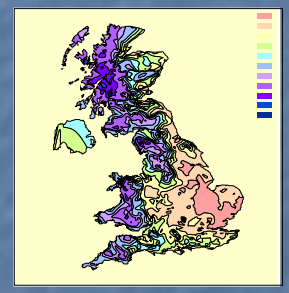
Crop Model

GCM

Downscaling

Weather Generator

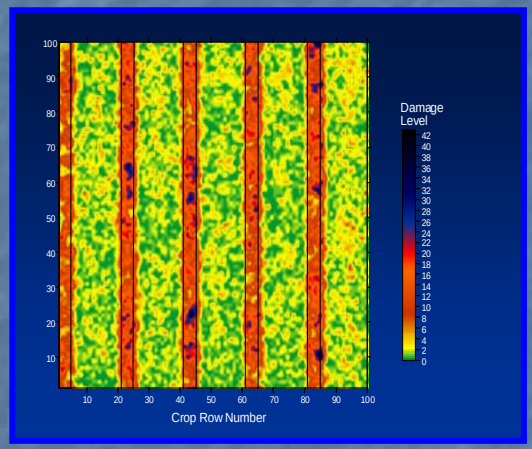




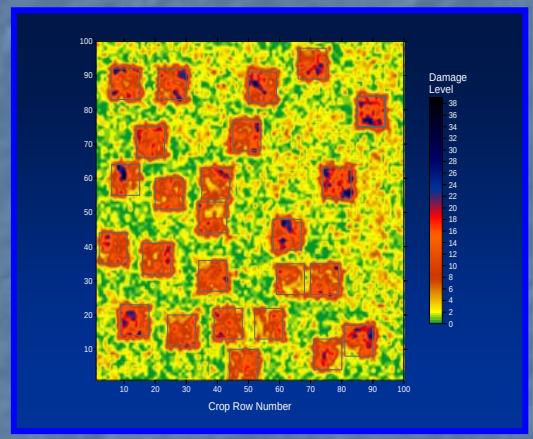
Modelling complex interactions: Individual Based Modelling

Mikhail Semenov

Intercrop



Patches



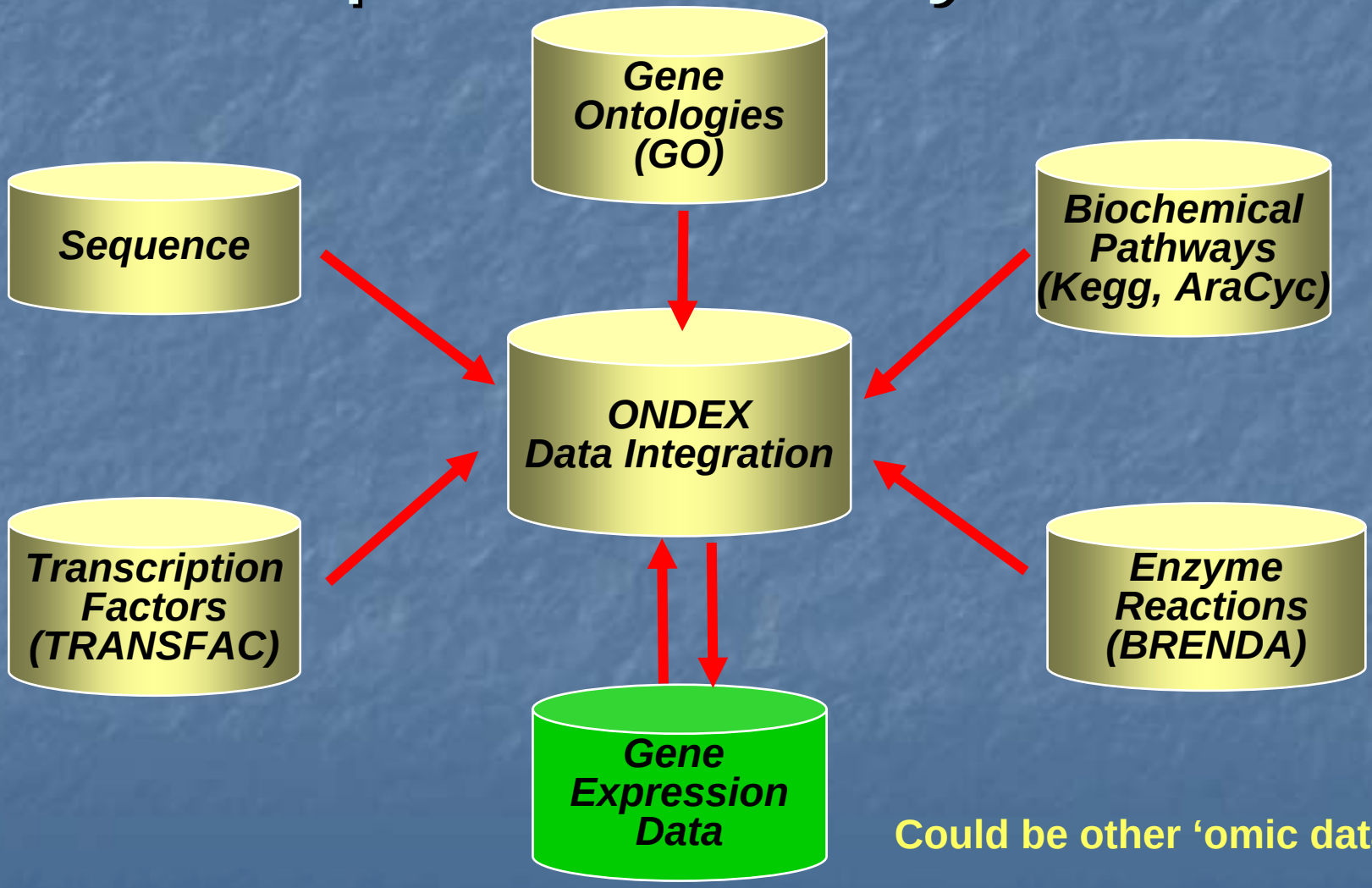
Slug-nematode interactions: optimal field management

Bioinformatics

Jacob Köhler, Paul Verrier

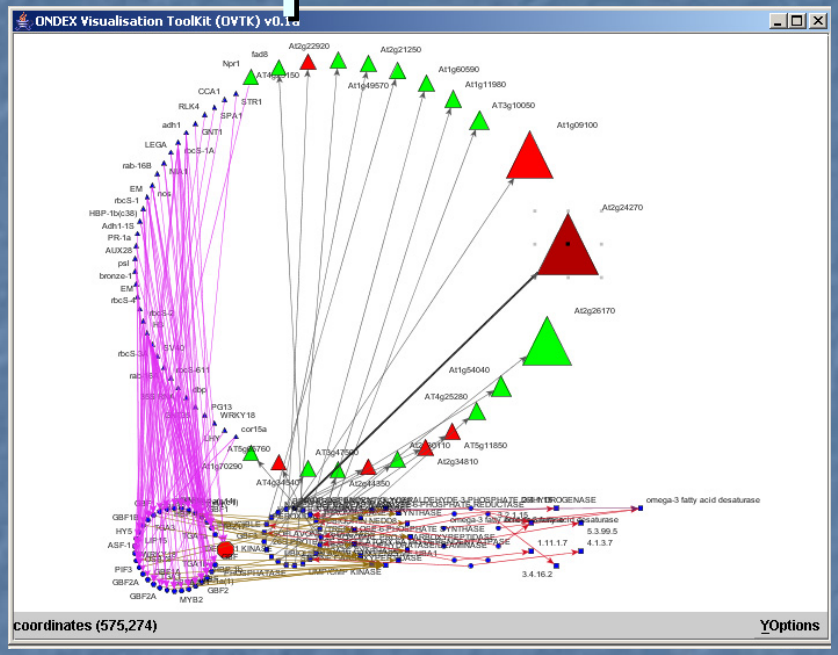
- Data integration
  - Integrating ‘omics datasets
  - Software platform using graph-based method – ONDEX
  - ondex.sourceforge.net
- Text mining
  - Use for supporting database curation for a database of pathogen host interactions
  - http://www.phi-base.org/
- Systems models of plant metabolic processes

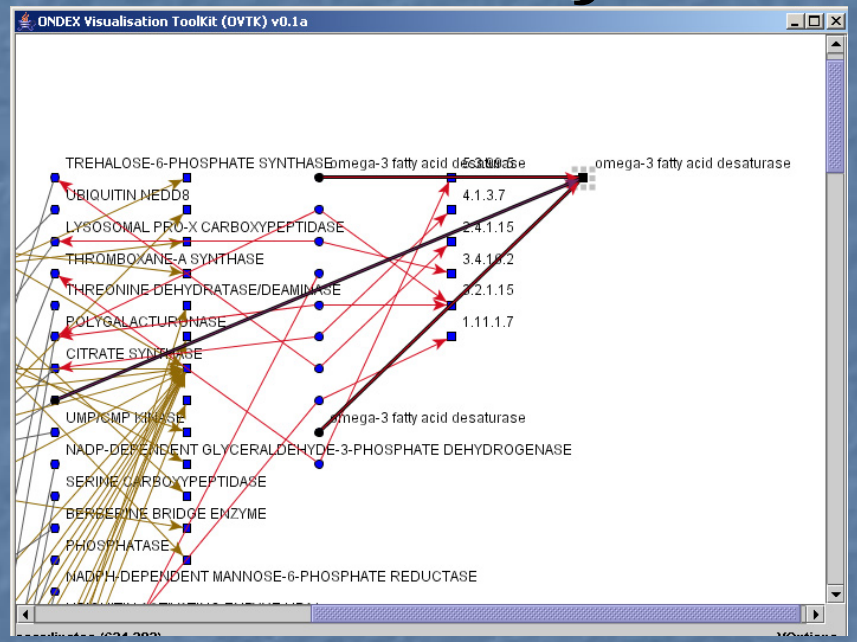
Data Integration – Gene Expression Analysis



Could be other 'omic data

Graph Visualisation & Analysis





Gene expression signal strength expressed as colour and size of glyph
Relationship between genes/proteins shown as lines
Circular layout designed to display maximum number of concepts/relations

ROTHAMSTED RESEARCH

Pilot Study

- microarray analysis – published study
- Parani, M., et al. (2004) *Microarray analysis of nitric oxide responsive transcripts in Arabidopsis.* Plant Biotechnology Journal*, 2,* 359-366.

Arabidopsis data with 120 “novel” genes

New observations not in original paper made because of access to integrated data:

→ provided annotation to 50 “novels”
→ an important “unspotted” gene (a TF)
→ drought stress
→ jasmonic acid biosynthesis

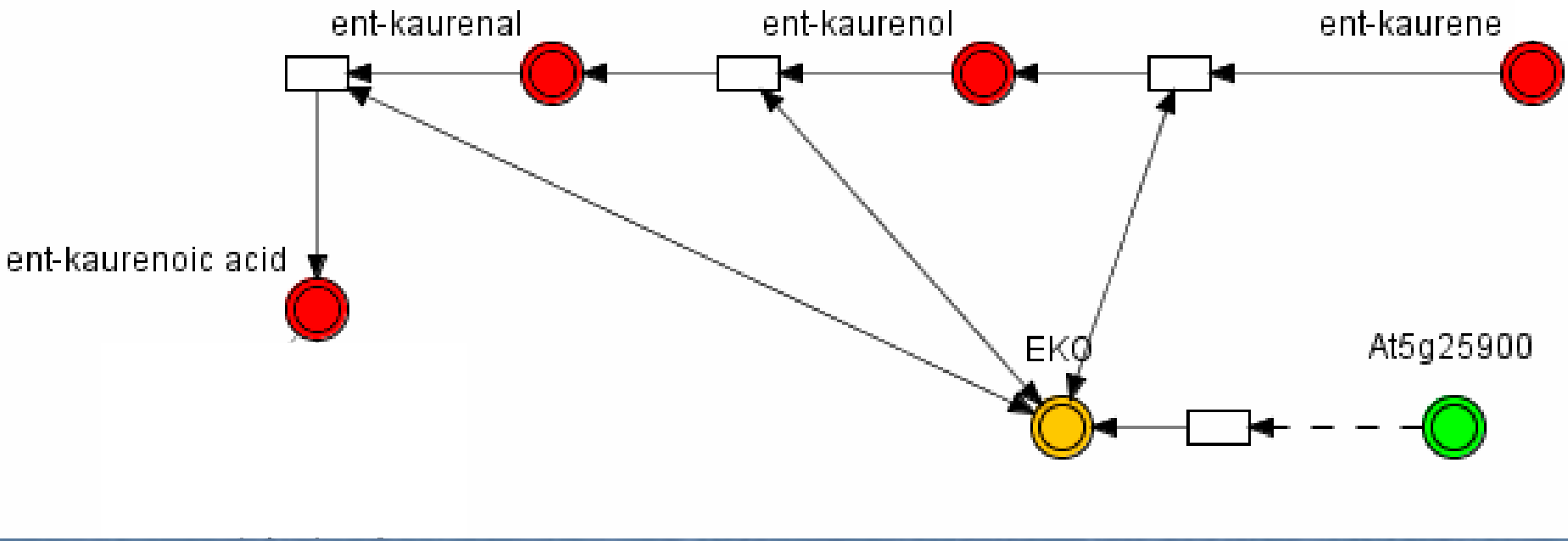
Other Applications of Data Integration

- General method – can be applied to many different types of data
- Potential underpinning for data mining and machine learning projects
- Integration of gene and trait ontologies for text mining
- Provide information infrastructure for pathway modelling projects

Petri net models of Gibberellin biosynthesis

ROTHAMSTED RESEARCH

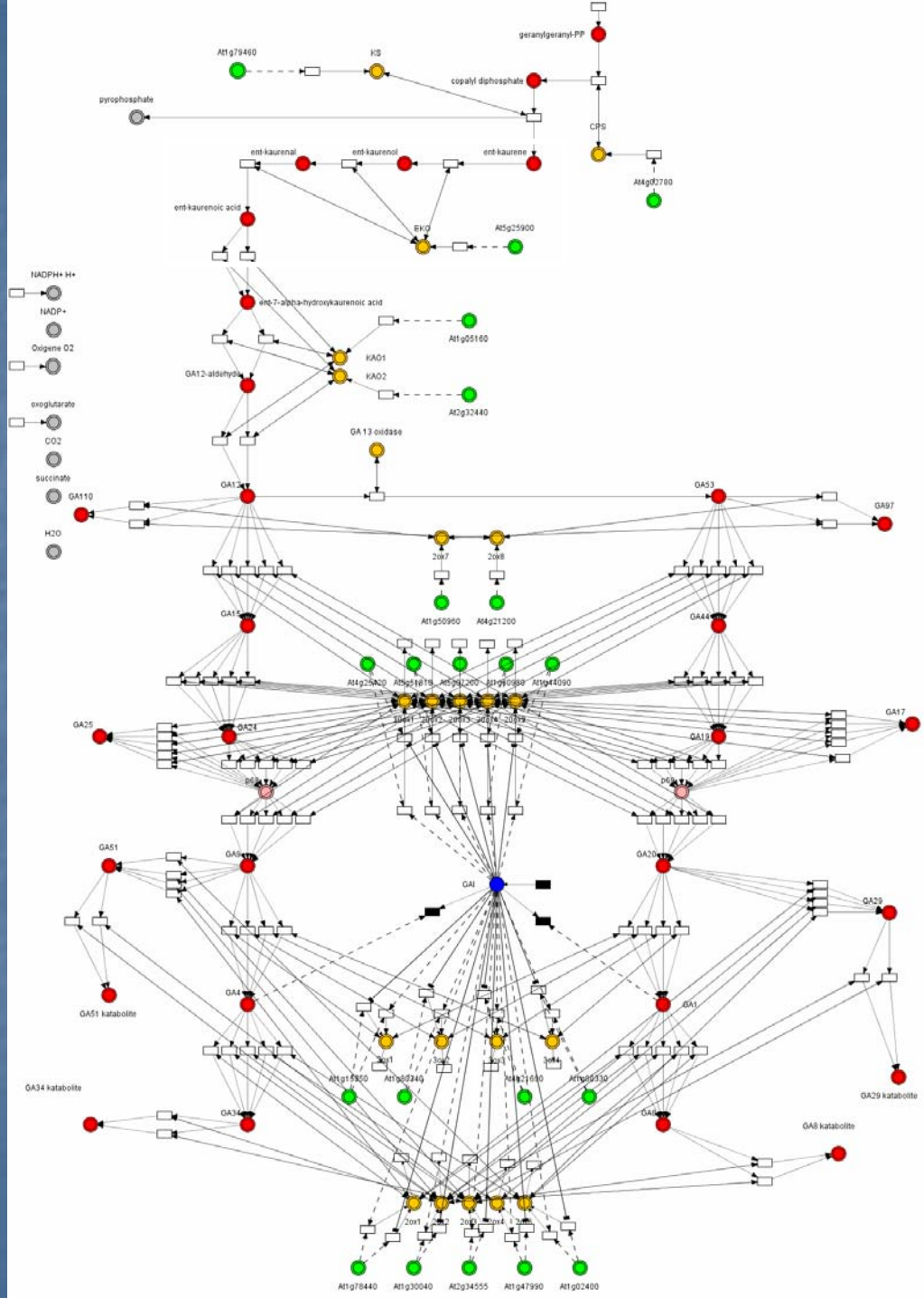
Biological pathways represented as Petri nets



Hybrid Petri-Net - Cell Illustrator

Investigating phenotype of mutants and knock-downs

ROTHAMSTED RESEARCH

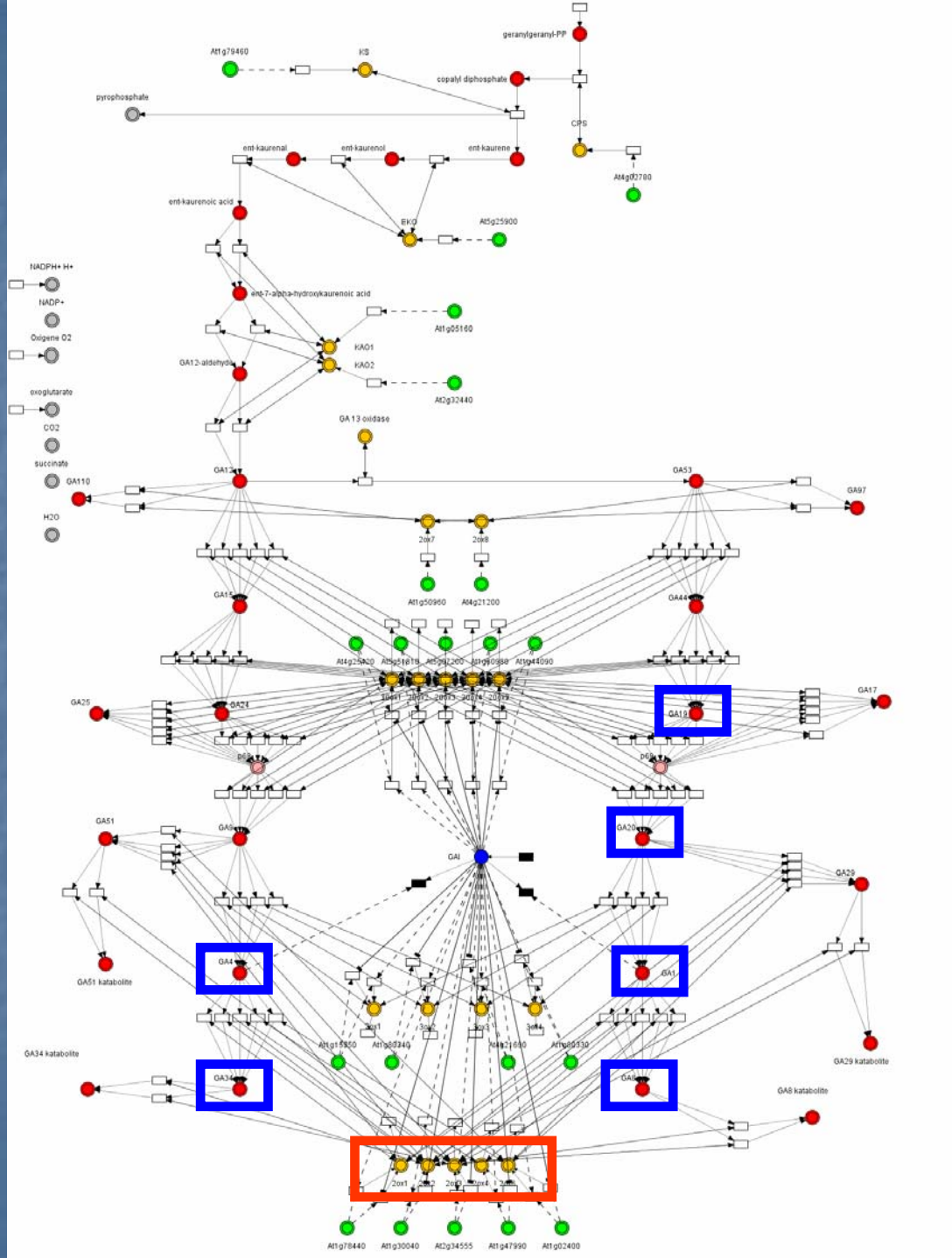


Gibberellin biosynthesis

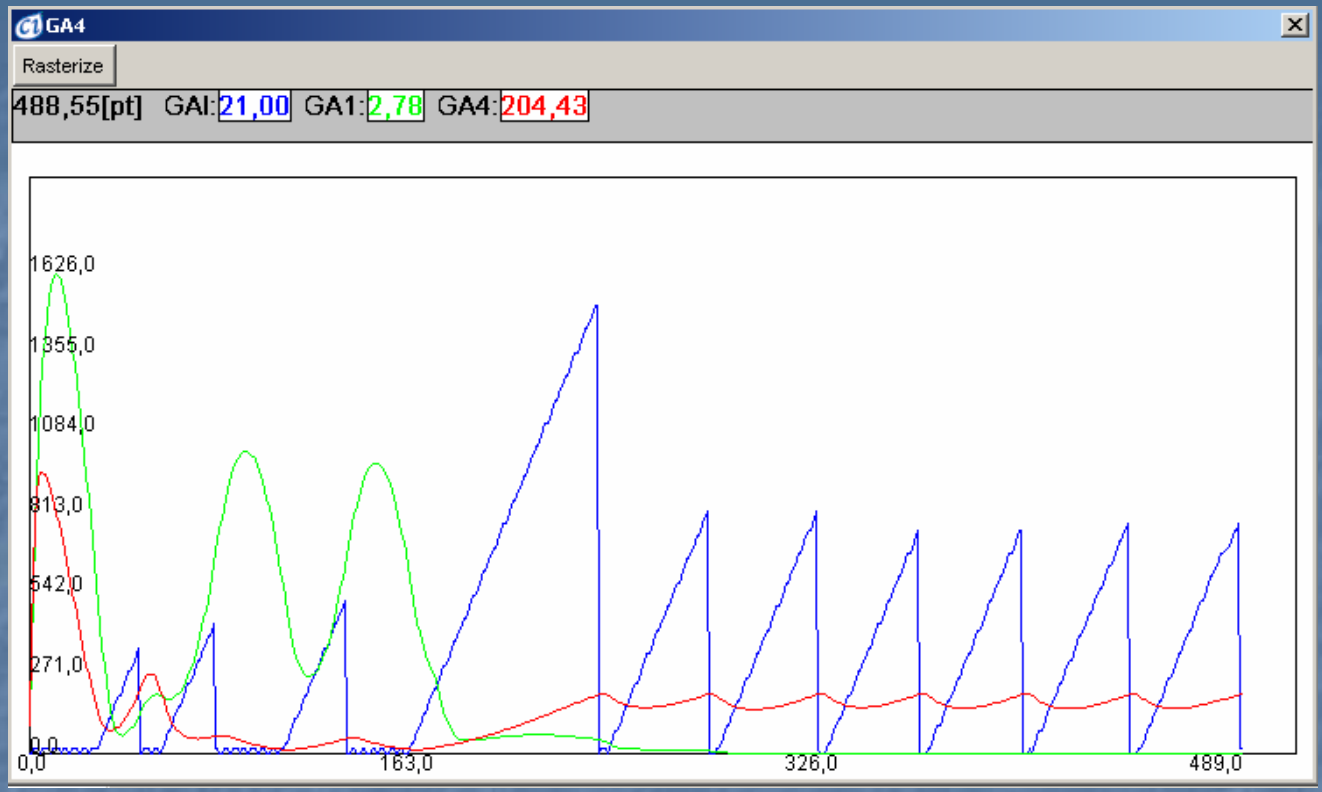
…according to Andy Philips and Peter Hedden and curated by Burkhard Steuernagel

ROTHAMSTED RESEARCH

Gibberellin biosynthesis

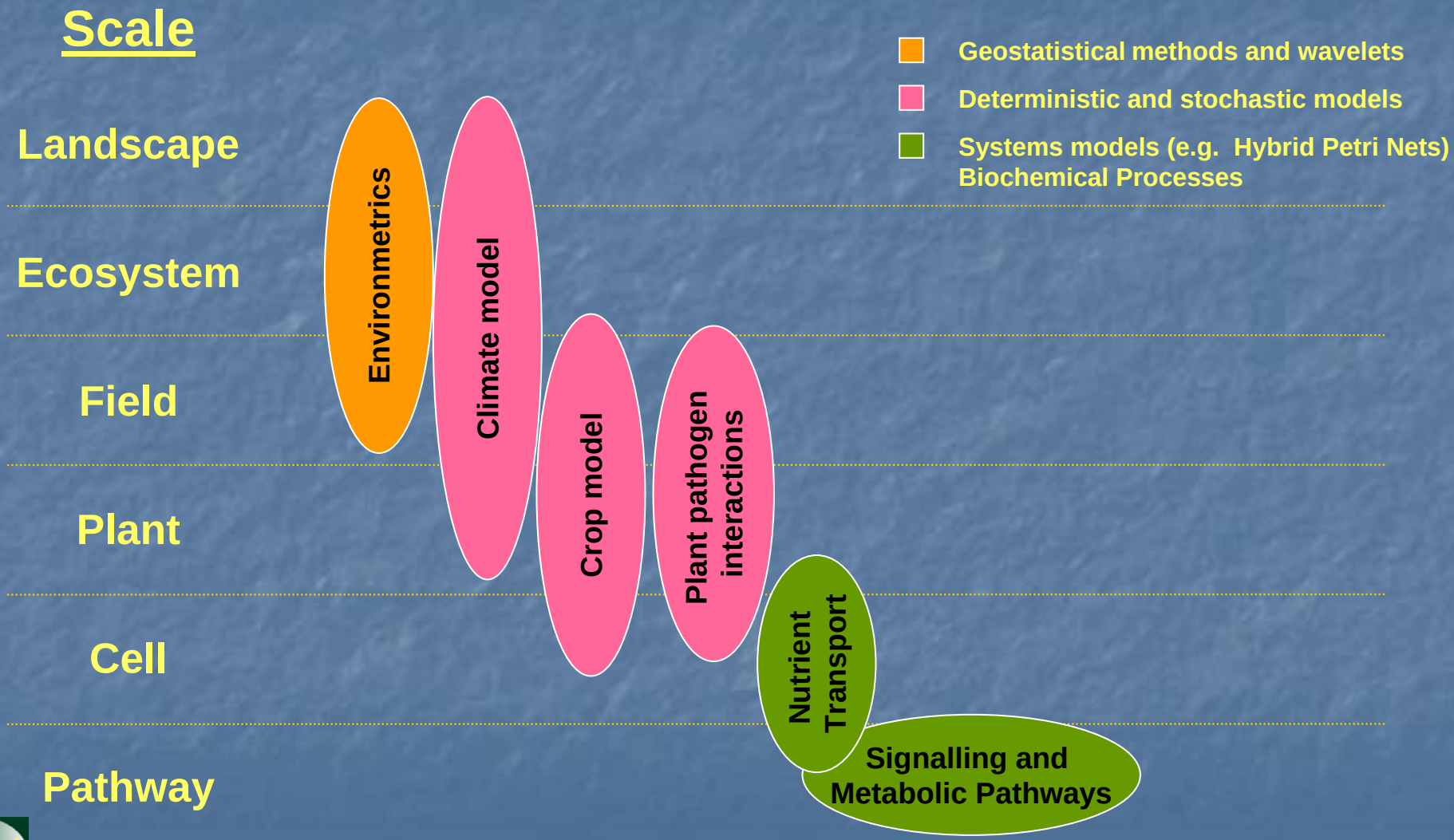


Gibberellin biosynthesis



		GA_4	GA_{34}	GA_1	GA_8	GA_{20}	GA_{19}
Experiment	Control	99	89	38	47	18	111
	GA 2-ox	16	3	3	6	2	2
Petri Net	Control	205	5861	3	4439	3	115
	GA 2-ox	9	453	0.3	0.8	3	102
ODE	Control	103-201	4061	17.8-30.7	7584.6	249.2	0.00052
	GA 2-ox	0.73	2.3	0.6	1.4	98	0.00052

Many Modelling Scales - Rothamsted



Plant Systems Biology

- Collaborator in two UK National systems biology centres:
  - Nottingham University – virtual root
  - Imperial College – plant pathogen interactions