Big data in the agricultural and ecological sciences – a growing challenge

chris.rawlings@rothamsted.ac.uk

Head of Computational and Systems Biology

BBSRC bioscience for the future

ROTHAMSTED RESEARCH

Rothamsted Research

- Rothamsted is an independent scientific research institute
- Longest running agricultural research institute in the world (est. 1843)
- Delivering knowledge, innovation and new practices to increase crop productivity and quality
- Develop environmentally sustainable solutions for agriculture

Main funder

BBSRC bioscience for the future

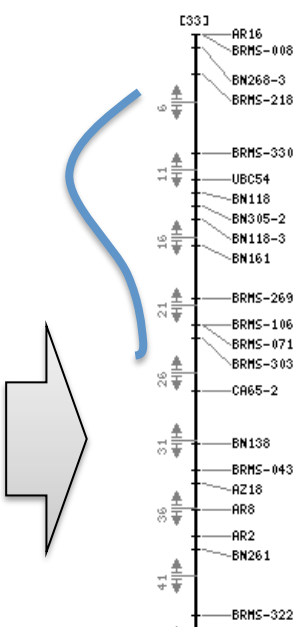
Relating genotype to phenotype

Genetics

- Studies of disease families
- Population studies (association genetics)
- Crop breeding populations
- Originally genetic markers
- Now genotyping by sequencing

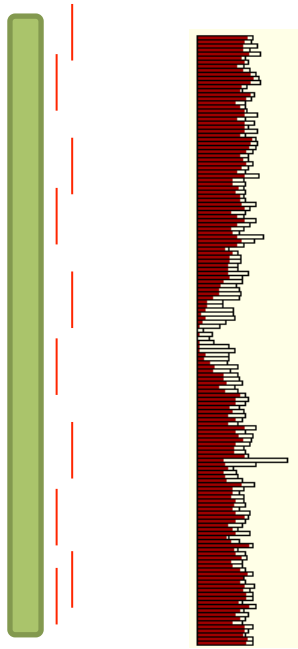
Phenotypic measurements

QTL **Linkage maps**

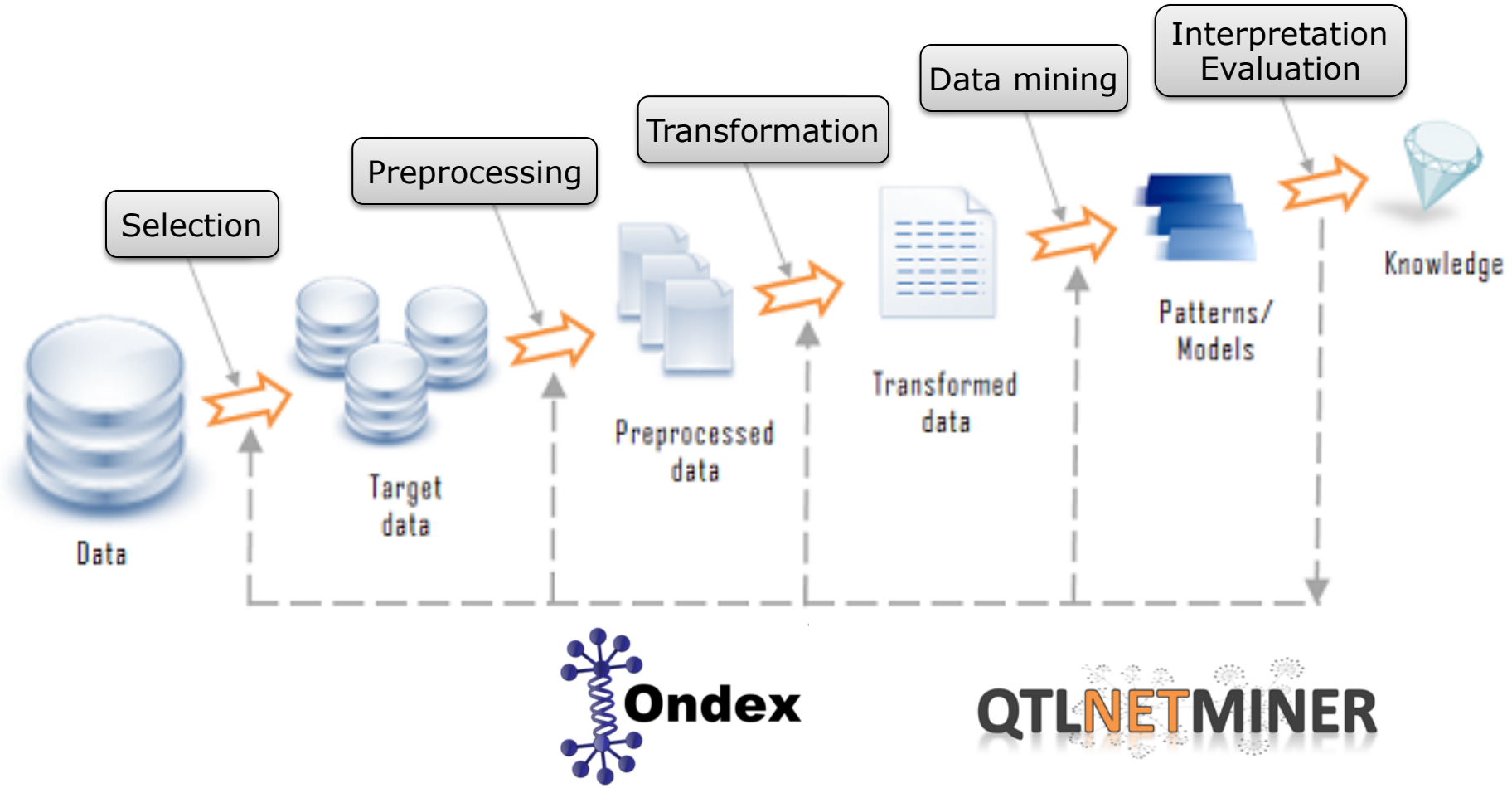


Genomic sequence

S.viminalis chromosome



Knowledge discovery process

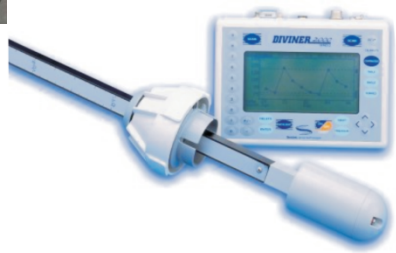


Crops - Phenotype

- Agronomic
  - Crop yield
  - Total biomass
- Plant Morphology
  - Architectural - macro-measurements
    - E.g. stem number, height, thickness, root structure
    - Light Interception – canopy
  - Non-destructive and destructive methods
- Biophysical and Biochemical
  - Photosynthesis
  - Tissue composition e.g.
    - Lipid content
    - Sugars – bioethanol
- **Environmental factors**
  - Weather
  - Soil moisture
  - Treatments (nutrients, pest management etc).

LiCOR Fluorometer

Delta T Sunscan



Sentek Diviner

New - Phenomics – HT Phenotyping

ROTHAMSTED RESEARCH

- Automated measurements
- Exploit high resolution, multi-spectral cameras
- Image processing, computer vision techniques
- Non-invasive
- Measurements e.g.
  - Plant growth and development
  - Photosynthetic activity
  - Stress measurements

BBSRC bioscience for the future

Field Based Phenotyping

Phenomobile

http://www.plantphenomics.org.au/

http://www.plantphenomics.org/hrppc/capabilities/fieldmodule

Agricultural Interactions with the Environment

Delivering sustainable intensification

BBSRC
bioscience for the future

ROTHAMSTED
RESEARCH

AnaEE

A European infrastructure for analysis and experimentation on agro ecosystems

http://www.anaee.com/

Data Rich Interactions in Agri-Ecology

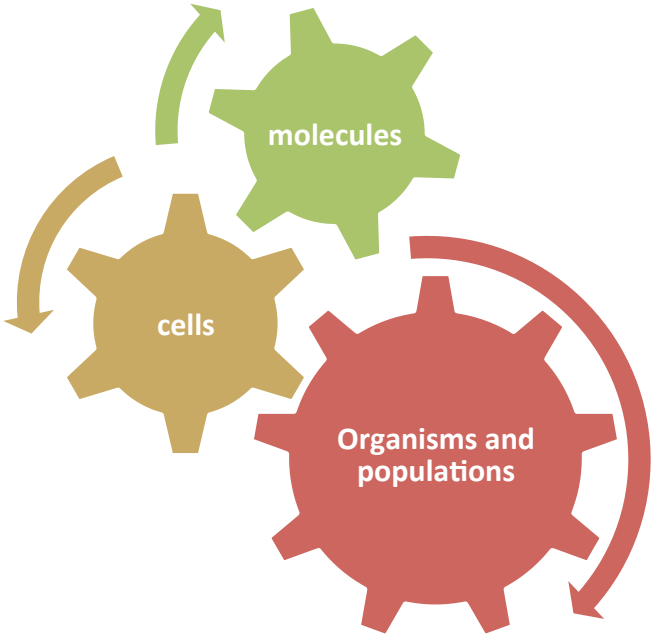
ROTHAMSTED RESEARCH

- Next generation sequencing
  - Genomes of host organisms
    - Large and complex
  - Genomes of pest and pathogen organisms
- Managing / integrating multi-omics datasets
  - Variety of data resources
  - Transcriptomics and metabolomics most important
  - Importance of model organisms
  - Range of data types

ELIXIR

- Measuring phenotypes and traits –
  - Until now - low throughput
  - Now moving to high throughput
  - Range of automation and imaging technologies
- Measuring the environmental factors
  - affecting agricultural production
  - Increasing sustainability of agriculture

AnaEE



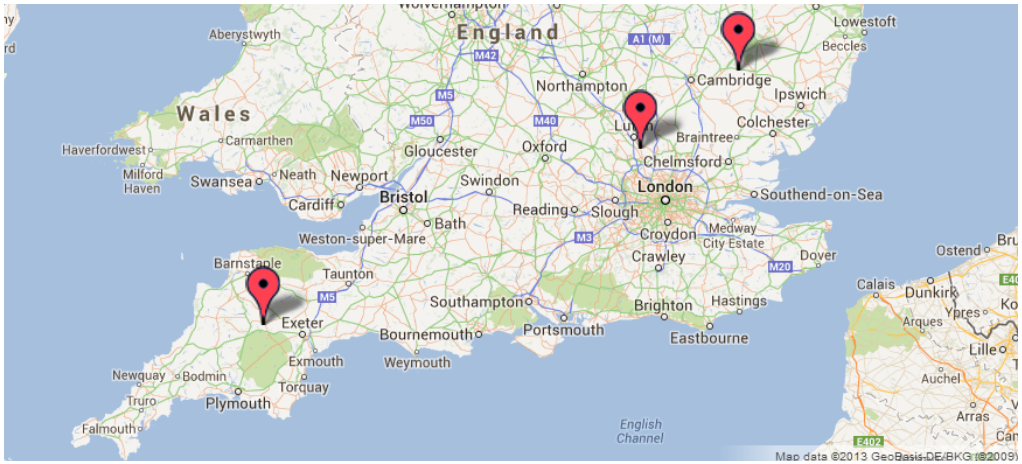
ISBE

Modelling interactions and dynamics of biological systems at all scales

BBSRC bioscience for the future

Rothamsted North Wyke Farm Platform

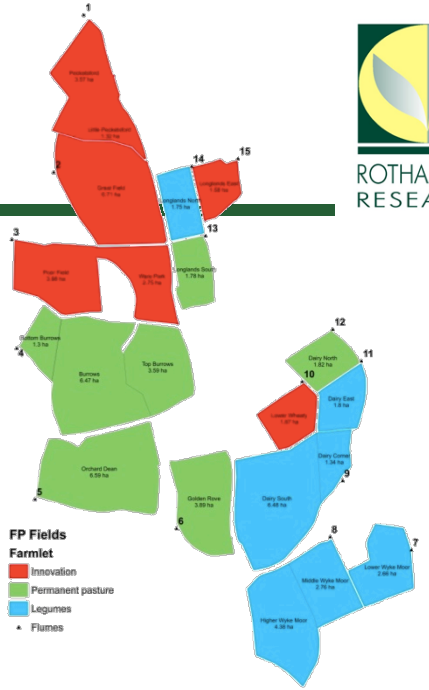
Example AnaEE Site



BBSRC bioscience for the future

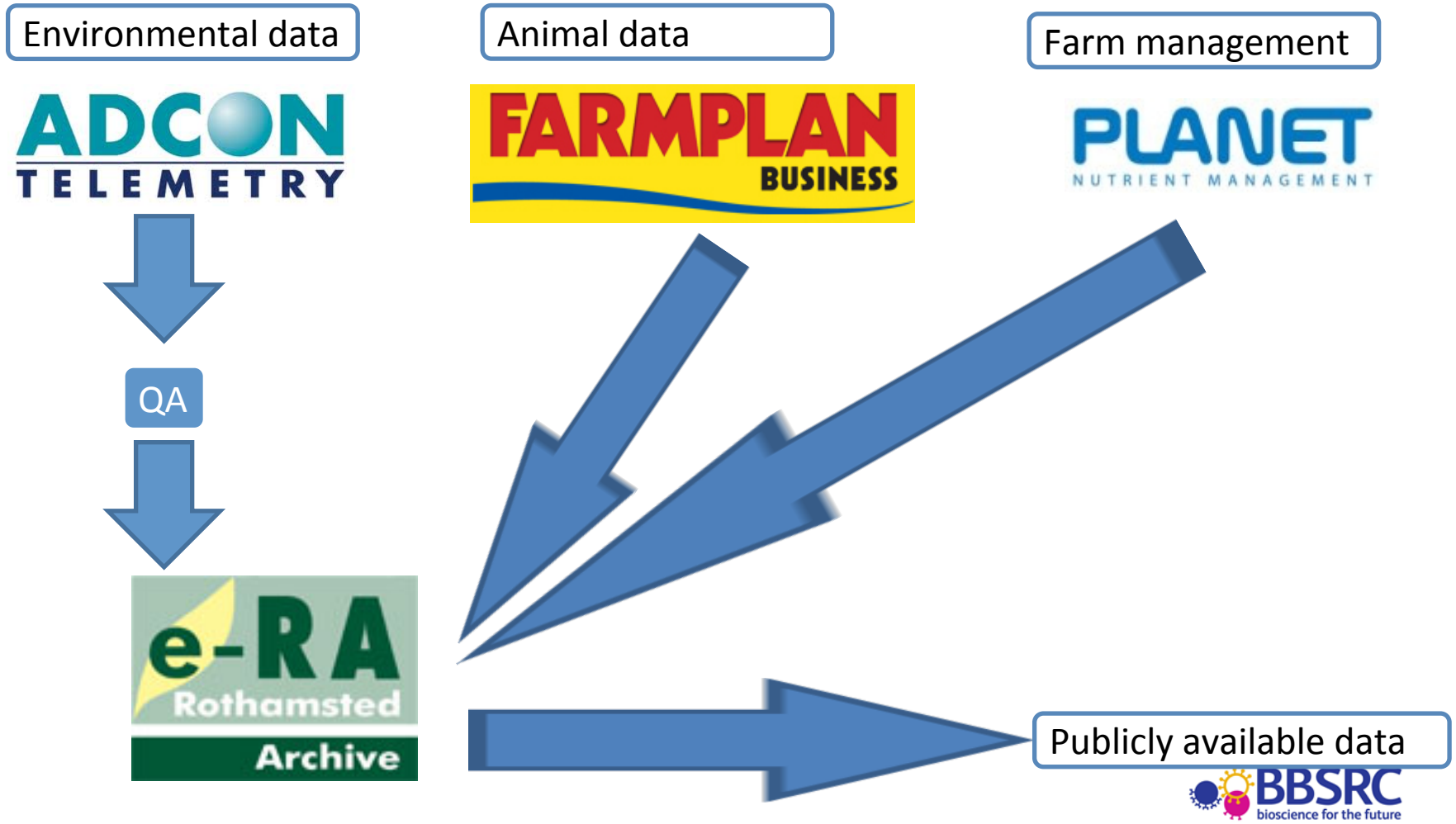
ROTHAMSTED RESEARCH

Main features



- Three “farmlets”
- Highly instrumented
  - Most instrumented farm in Europe(?)
  - Realtime data capture from 15 monitoring stations
- Known topology and hydrology
- Long term experiments – just starting
  - Baseline data 2 years
- Integration with remote sensing data
  - Satellite, hyperspectral imaging

Databases


Environmental data
Animal data
Farm management
ADCON TELEMETRY
FARMPLAN BUSINESS
PLANET NUTRIENT MANAGEMENT
QA
e-RA Rothamsted Archive
Publicly available data
BBSRC bioscience for the future