MANCHESTER 1824

The University of Manchester

A Systematic approach to the Large-Scale Analysis of Genotype-Phenotype correlations

Paul Fisher

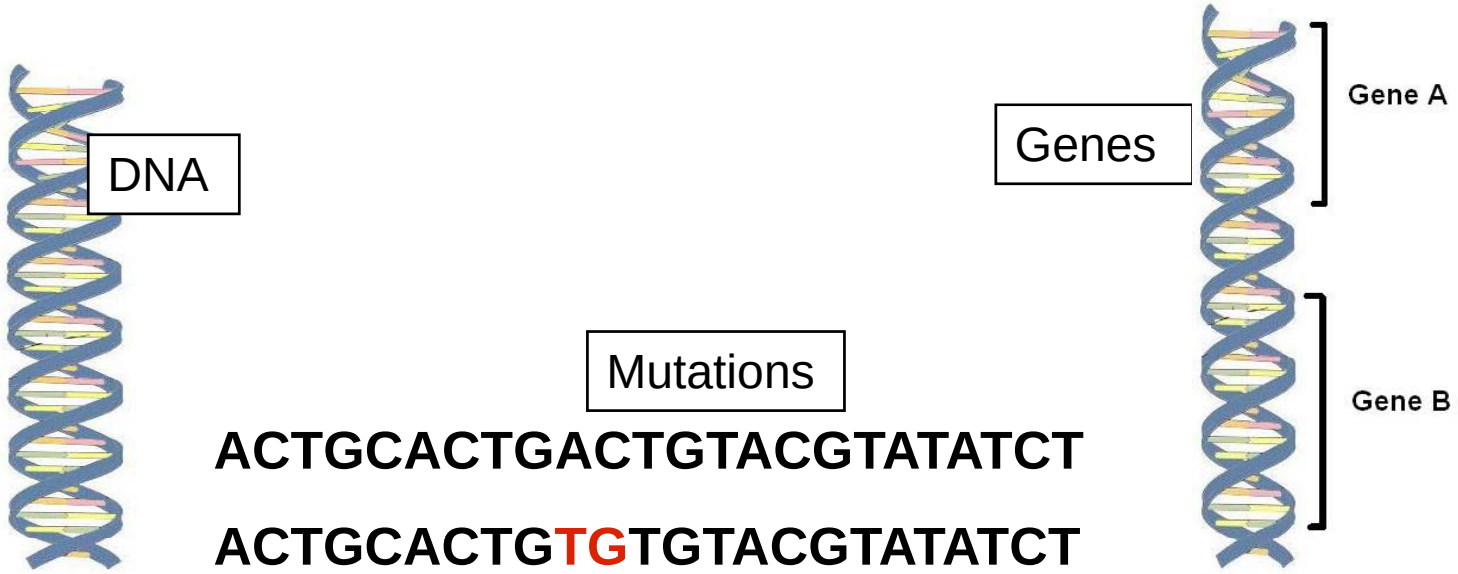
Dr. Robert Stevens

Prof. Andrew Brass

BHIG

Genotype

The entire genetic identity of an individual that **does not show** any outward characteristics, *e.g.* Genes, mutations



Phenotype

(harder to characterise)

The observable expression of gene's producing **notable characteristics** in an individual, *e.g.* Hair or eye colour, body mass, resistance to disease

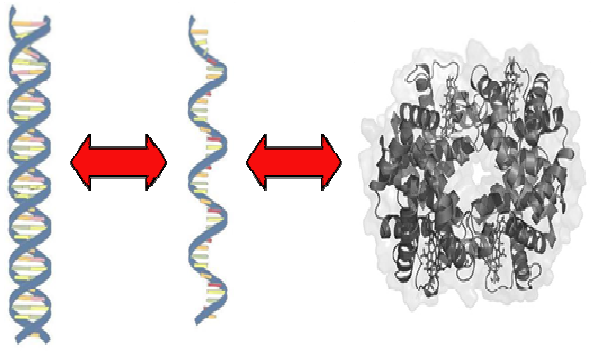
Brown

vs.

White and Brown

Genotype to Phenotype

Genotype



Phenotype

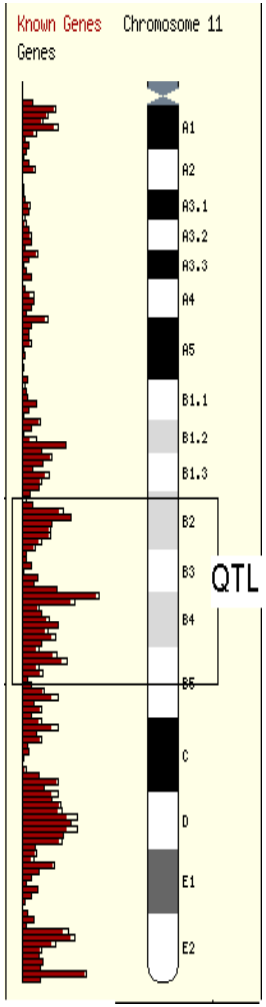
DNA RNA Protein Protein-Protein interaction Pathway Trait

MANCHESTER 1824

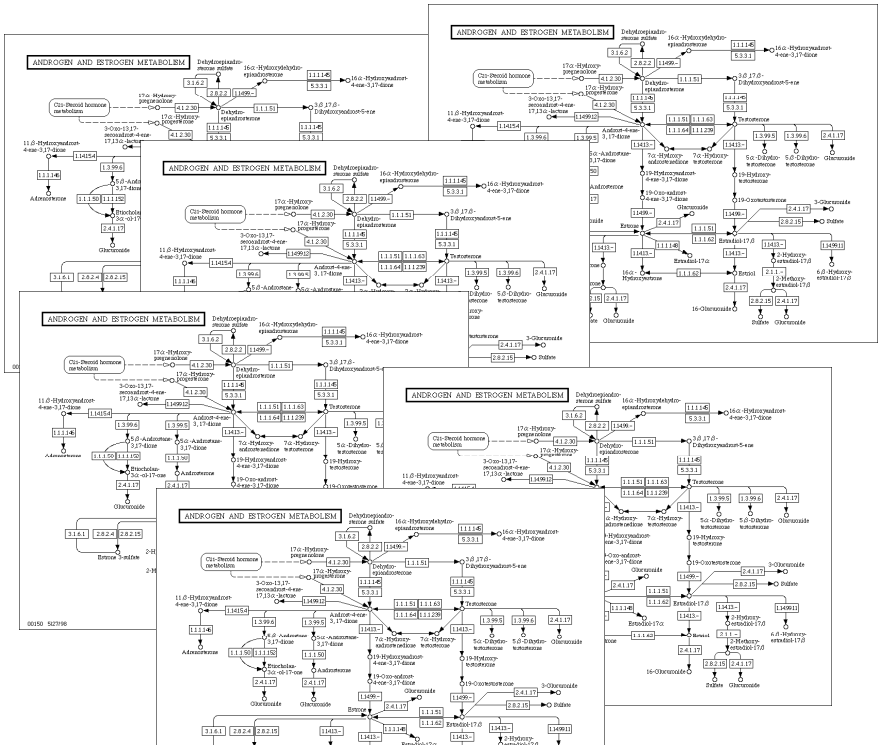
The University of Manchester

Current Methods

Genotype



200



Phenotype

?

What processes to investigate?

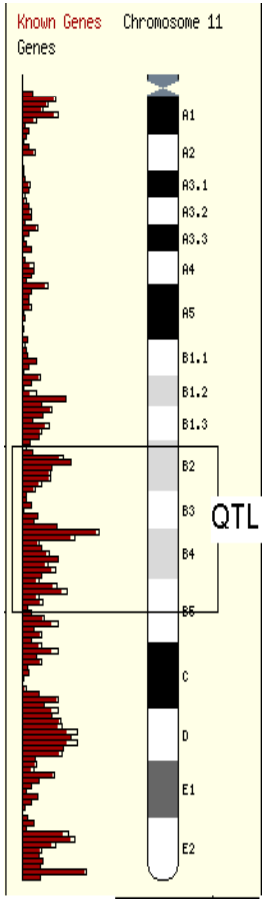
BHIG

MANCHESTER 1824

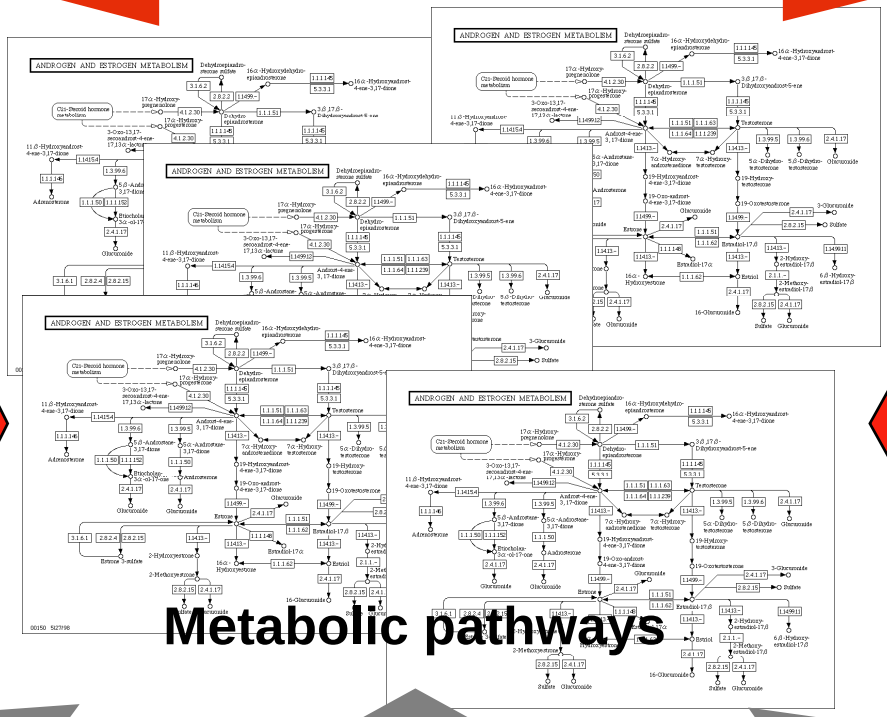
The University of Manchester

Genotype

Phenotype

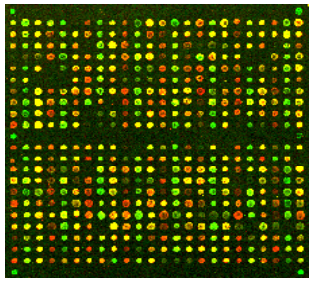


200



Metabolic pathways

?



Microarray + QTL

Genes captured in microarray experiment and present in QTL (**Quantitative Trait Loci**) region

Phenotypic response investigated using microarray in form of expressed genes or evidence provided through QTL mapping

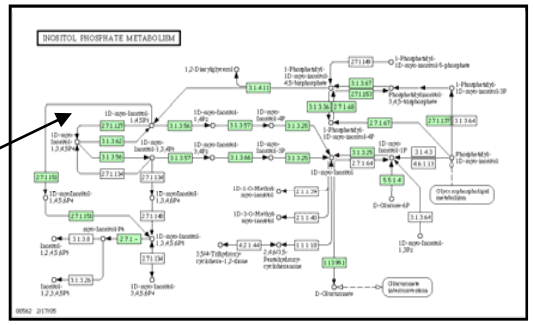
BHIG

MANCHESTER 1824

The University of Manchester

Phenotype

Pathway A



literature

Pathway linked to phenotype – high priority

CHR

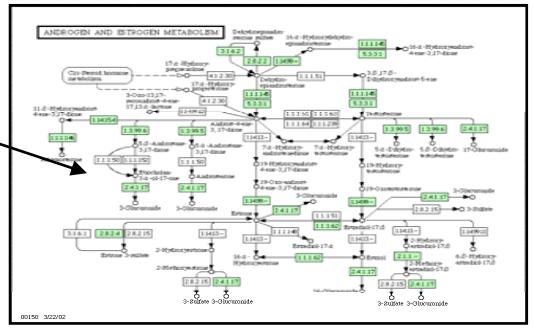
QTL

Gene A

Gene B

Gene C

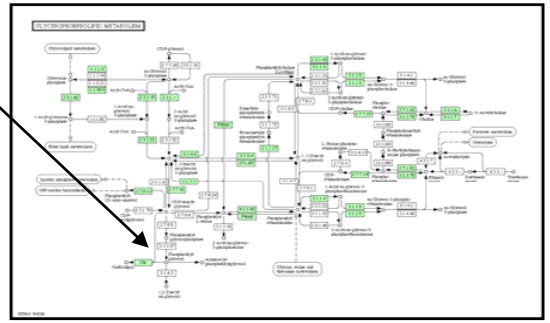
Pathway B



literature

Pathway not linked to phenotype – medium priority

Pathway C



literature

Pathway not linked to OTL – low priority

Genotype

BHIG

Issues with current approaches

- Scale of analysis task
- User bias and premature filtering
- Hypothesis-Driven approach to data analysis
- Constant flux of data - problems with re-analysis of data
- Implicit methodologies (hyper-linking through web pages)
- Error proliferation from any of the listed issues

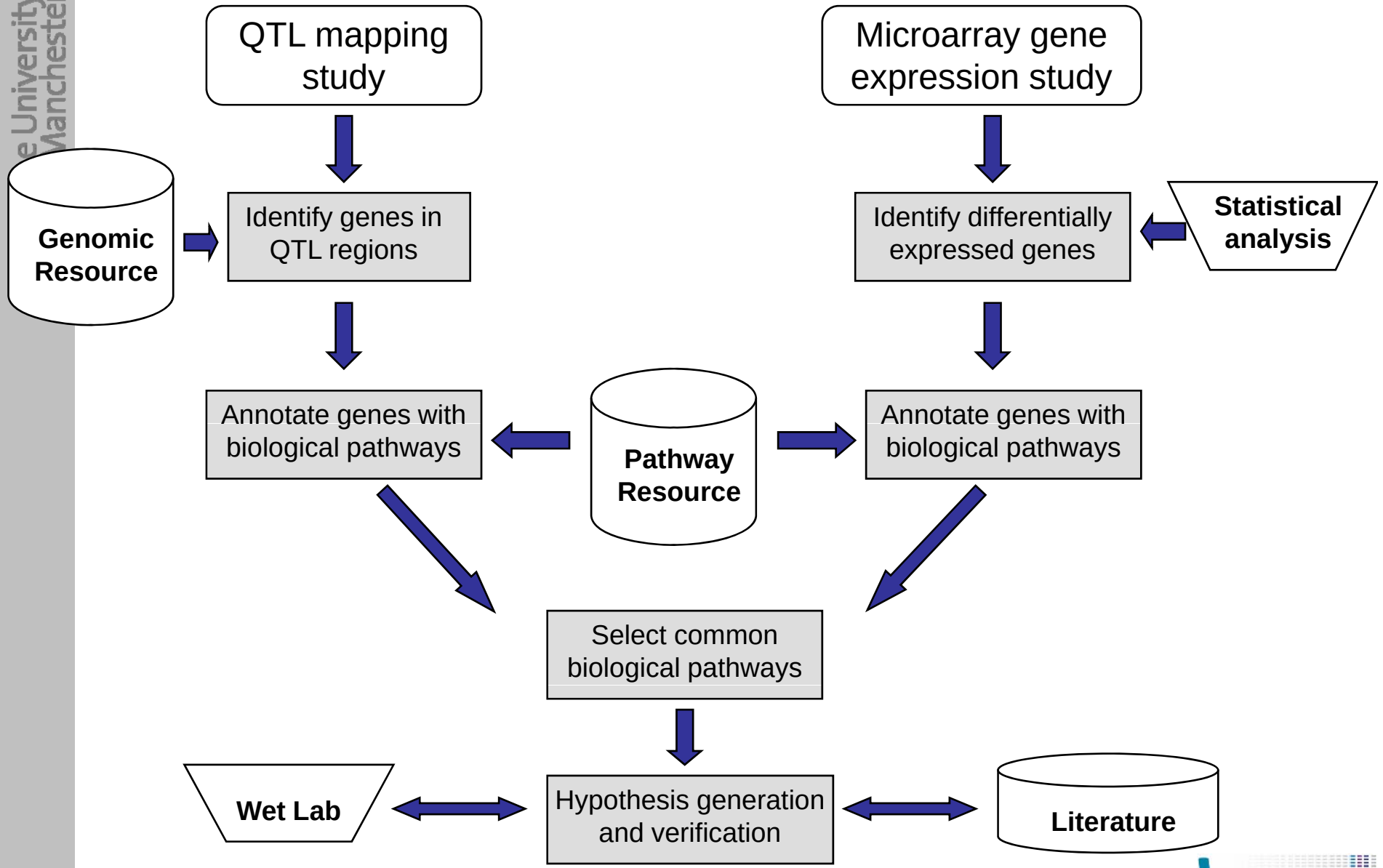
Solution – Automate through workflows

Hypothesis

Utilising the capabilities of workflows and the pathway-driven approach, we are able to provide a more:

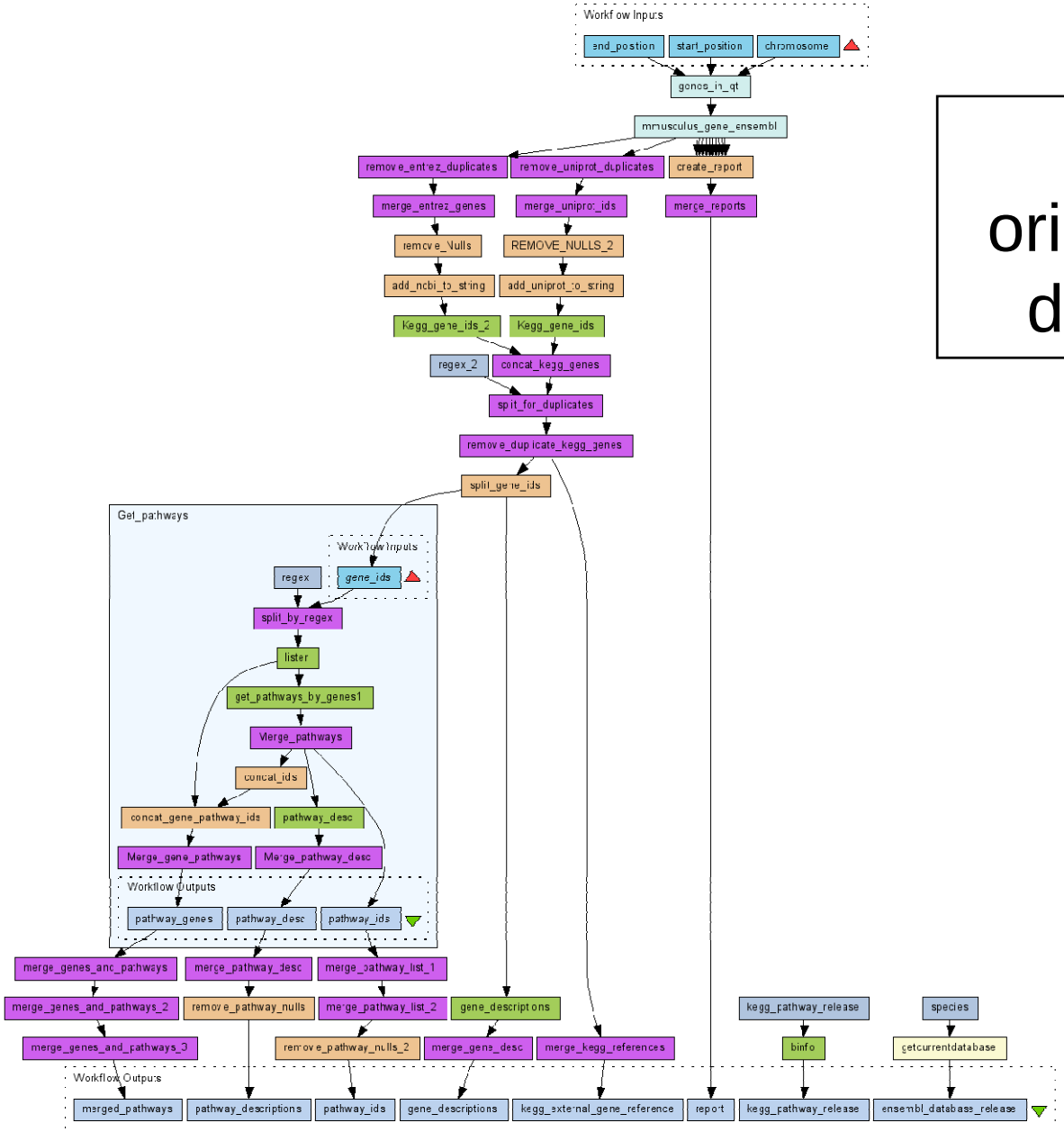
- systematic
- explicit
- scalable
- un-biased

the benefit will be that **new biology** results will be derived, increasing community knowledge of genotype and phenotype interactions.


QTL mapping study
Microarray gene expression study
Genomic Resource
Identify genes in QTL regions
Identify differentially expressed genes
Statistical analysis
Annotate genes with biological pathways
Pathway Resource
Annotate genes with biological pathways
Select common biological pathways
Wet Lab
Hypothesis generation and verification
Literature

MANCHESTER 1824

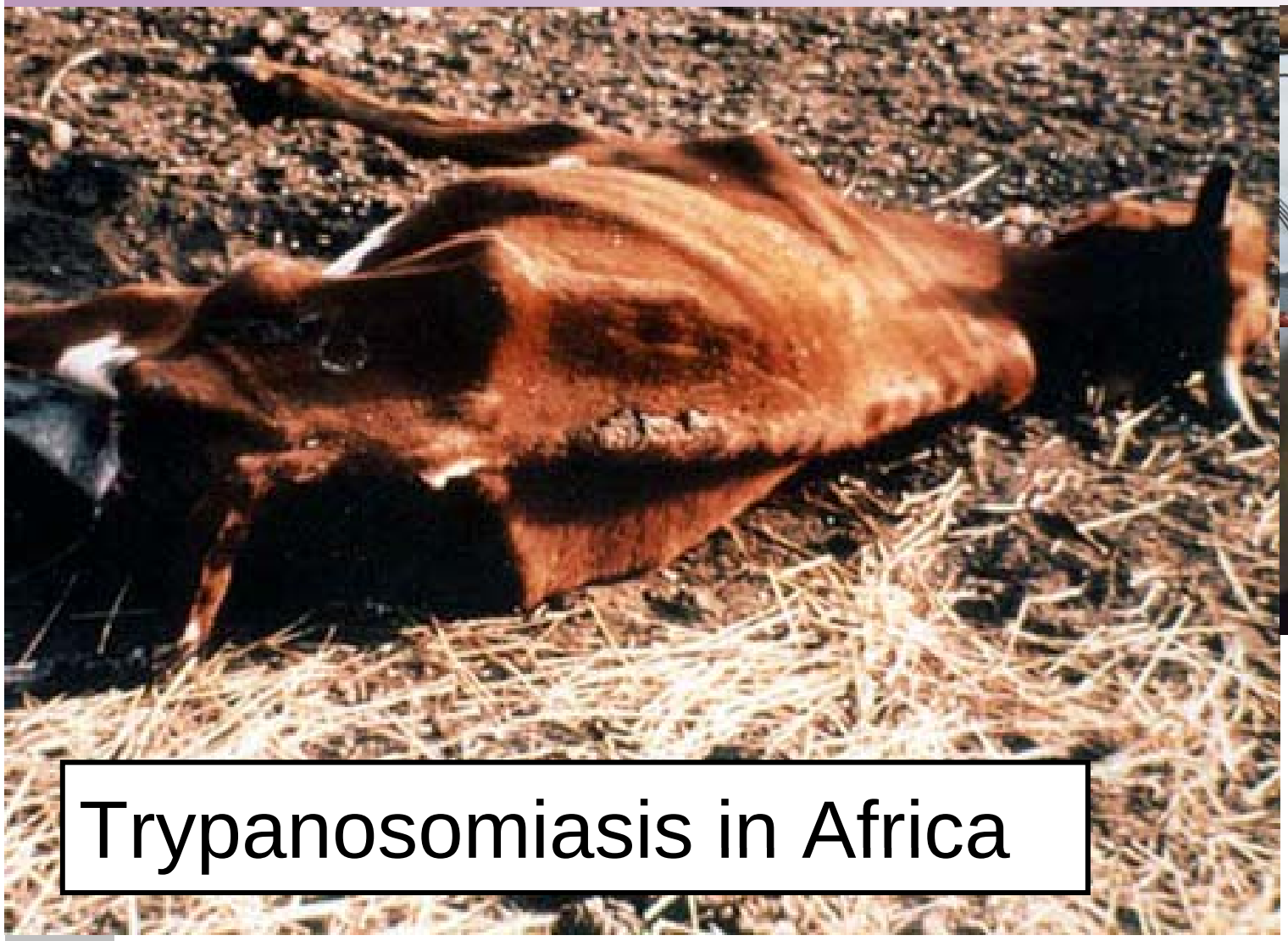
The University of Manchester



Replicated original chain of data analysis

BHIG

The Wellcome Trust Funded Host-Pathogen Project



Steve Kemp

Andy Brass

+ many Others

Trypanosomiasis in Africa

http://www.genomics.liv.ac.uk/tryps/trypsindex.html

MANCHESTER 1824

The University of Manchester

Preliminary Results

Trypanosomiasis resistance

A strong candidate gene was found

- **Daxx** gene not found using manual investigation methods
- The gene was identified from analysis of biological pathway information
- Possible candidate identified by Yan et al (2004): Daxx SNP info
- Sequencing of the Daxx gene in **Wet Lab** showed mutations that is thought to change the structure of the protein
- Mutation was published in scientific literature, noting its effect on the binding of Daxx protein to p53 protein – **p53 plays direct role in cell death and apoptosis, one of the Trypanosomiasis phenotypes**
- More genes to follow (hopefully) in publications being written

BHIG

MANCHESTER 1824

The University of Manchester

Shameless Plug!

A Systematic Strategy for Large-Scale Analysis of Genotype-Phenotype Correlations: Identification of candidate genes involved in African Trypanosomiasis

Fisher *et al*., (2007) Nucleic Acids Research
doi:10.1093/nar/**gkm623**

- Explicitly discusses the methods we used for the Trypanosomiasis use case
- Discussion of the results for Daxx and shows mutation
- Sharing of workflows for re-use, re-purposing

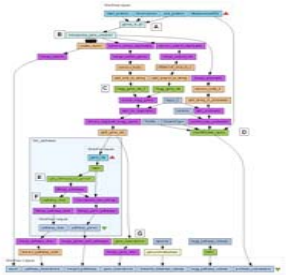
BHIG

Recycling, Reuse, Repurposing

The University of Manchester

Here's the **Science**!

- Identified a candidate gene (Daxx) for Trypanosomiasis resistance.
- Manual analysis on the microarray and QTL data failed to identify this gene as a candidate.
- Unbiased analysis. Confirmed by the wet lab.



Here's the **e-Science**!

- Trypanosomiasis mouse workflow reused without change in *Trichuris muris* infection in mice
- Identified biological pathways involved in sex dependence
- Previous manual two year study of candidate genes had failed to do this.

Workflows now being run over Colitis/ Inflammatory Bowel Disease in Mice (without change)

BHIG

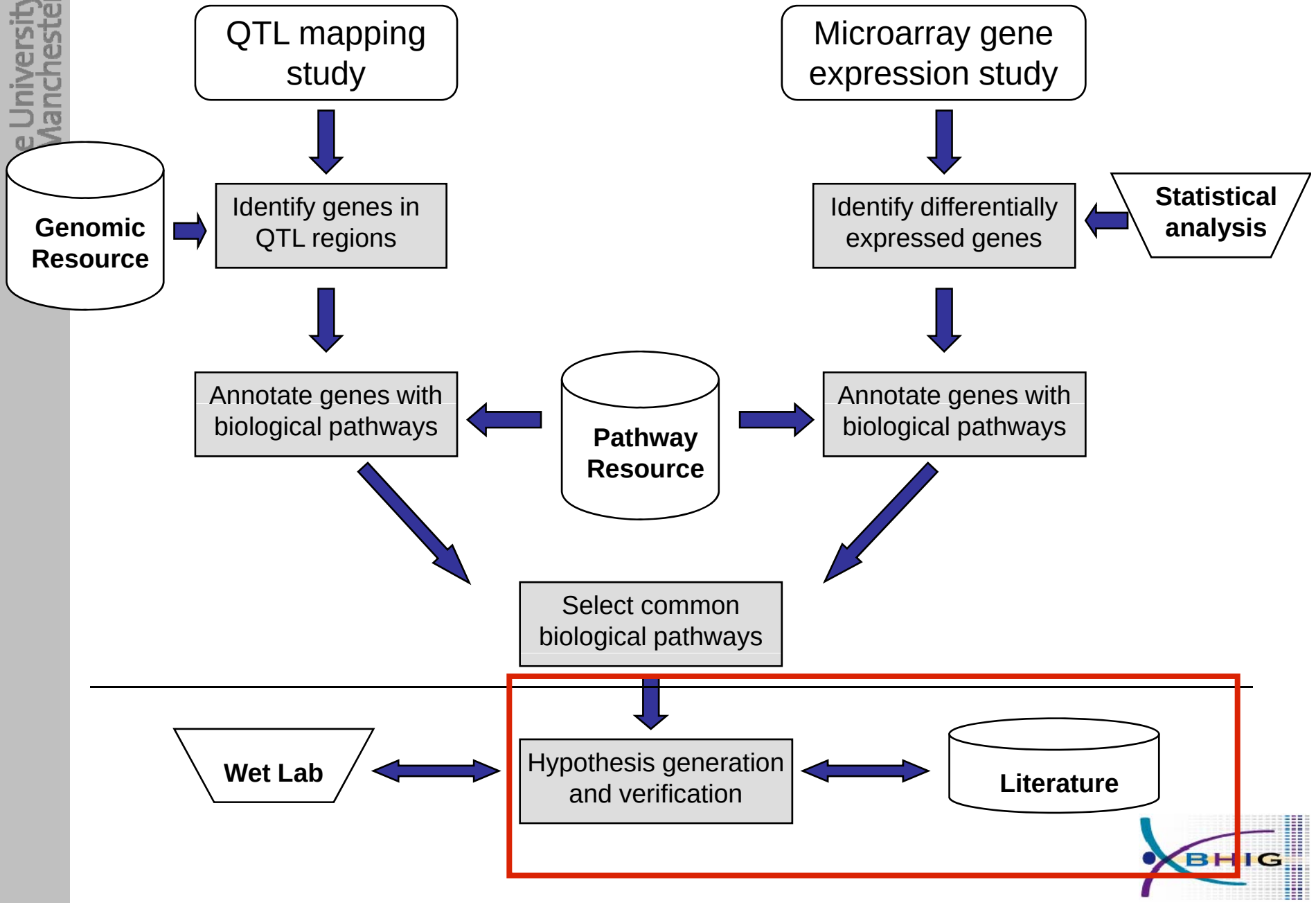
Recycling, Reuse, Repurposing

- Share
- Search
- Re-use
- Re-purpose
- Execute
- Communicate
- Record

http://www.myexperiment.org/

What next?

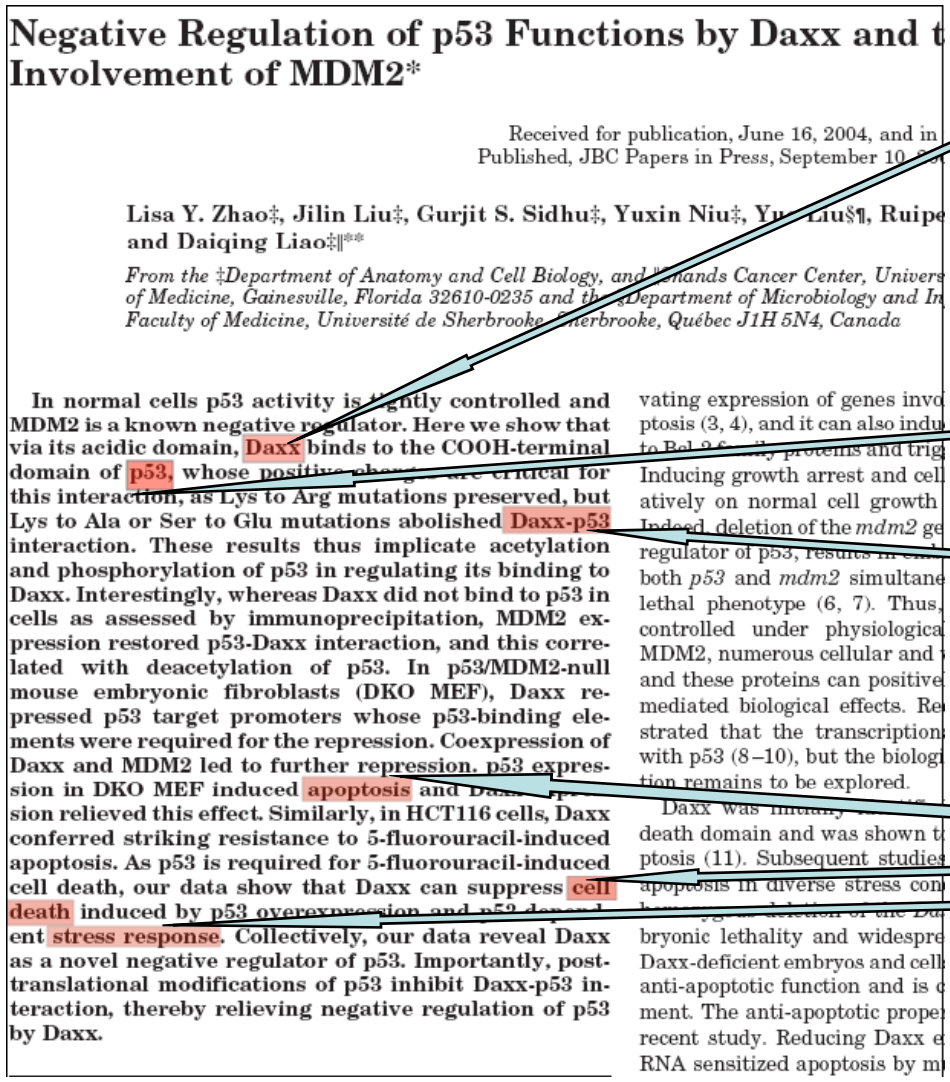
- More use cases??
  - Can be done, but not for my project
- Text Mining !!!
  - Aid biologists in identifying novel links between pathways
  - Link pathways to phenotype through literature


QTL mapping study
Microarray gene expression study
Genomic Resource
Identify genes in QTL regions
Identify differentially expressed genes
Statistical analysis
Annotate genes with biological pathways
Pathway Resource
Annotate genes with biological pathways
Select common biological pathways
Wet Lab
Hypothesis generation and verification
Literature
BHIG

MANCHESTER 1824

The University of Manchester

What Does the Text Hold?



Negative Regulation of p53 Functions by Daxx and t
Involvement of MDM2*

Received for publication, June 16, 2004, and in
Published, JBC Papers in Press, September 10,

Lisa Y. Zhao‡, Jilin Liu‡, Gurjit S. Sidhu‡, Yuxin Niu‡, Yu Liu§¶, Ruipe
and Daiqing Liao‡‖**

*From the ‡Department of Anatomy and Cell Biology, and Shands Cancer Center, Univers
of Medicine, Gainesville, Florida 32610-0235 and the Department of Microbiology and In
Faculty of Medicine, Université de Sherbrooke, Sherbrooke, Québec J1H 5N4, Canada*

In normal cells p53 activity is tightly controlled and MDM2 is a known negative regulator. Here we show that via its acidic domain, Daxx binds to the COOH-terminal domain of p53, whose positive charges are critical for this interaction, as Lys to Arg mutations preserved, but Lys to Ala or Ser to Glu mutations abolished Daxx-p53 interaction. These results thus implicate acetylation and phosphorylation of p53 in regulating its binding to Daxx. Interestingly, whereas Daxx did not bind to p53 in cells as assessed by immunoprecipitation, MDM2 expression restored p53-Daxx interaction, and this correlated with deacetylation of p53. In p53/MDM2-null mouse embryonic fibroblasts (DKO MEF), Daxx repressed p53 target promoters whose p53-binding elements were required for the repression. Coexpression of Daxx and MDM2 led to further repression. p53 expression in DKO MEF induced apoptosis and Daxx expression relieved this effect. Similarly, in HCT116 cells, Daxx conferred striking resistance to 5-fluorouracil-induced apoptosis. As p53 is required for 5-fluorouracil-induced cell death, our data show that Daxx can suppress cell death induced by p53 overexpression and p53-dependent stress response. Collectively, our data reveal Daxx as a novel negative regulator of p53. Importantly, post-translational modifications of p53 inhibit Daxx-p53 interaction, thereby relieving negative regulation of p53 by Daxx.

vating expression of genes invo
ptosis (3, 4), and it can also indu
to Bcl-2 family proteins and trig
Inducing growth arrest and cell
atively on normal cell growth
Indeed, deletion of the *mdm2* ge
regulator of p53, results in e
both *p53* and *mdm2* simultane
lethal phenotype (6, 7). Thus,
controlled under physiologica
MDM2, numerous cellular and
and these proteins can positive
mediated biological effects. Re
strated that the transcription
with p53 (8–10), but the biologi
tion remains to be explored.

Daxx was initially
death domain and was shown t
ptosis (11). Subsequent studies
apoptosis in diverse stress con
homozygous deletion of the Da
bryonic lethality and widespre
Daxx-deficient embryos and cell
anti-apoptotic function and is c
ment. The anti-apoptotic prope
recent study. Reducing Daxx e
RNA sensitized apoptosis by m

Protein Info

Related Proteins

Protein-Protein Interactions

Pathways

Biological processes

BHIG

What Next ?

Negative Regulation of p53 Functions by Daxx and t
Involvement of MDM2*

Received for publication, June 16, 2004, and in
Published, JBC Papers in Press, September 10, 20

Lisa Y. Zhao‡, Jilin Liu‡, Gurjit S. Sidhu‡, Yuxin Niu‡, Yue Liu§¶, Ruipe
and Daiqing Liao‡‖**

*From the ‡Department of Anatomy and Cell Biology, and ‖Shands Cancer Center, Univers
of Medicine, Gainesville, Florida 32610-0235 and the §Department of Microbiology and I
Faculty of Medicine, Université de Sherbrooke, Sherbrooke, Québec J1H 5N4, Canada*

In normal cells p53 activity is tightly controlled and MDM2 is a known negative regulator. Here we show that via its acidic domain, Daxx binds to the COOH-terminal domain of p53, whose positive charges are critical for this interaction, as Lys to Arg mutations preserved, but Lys to Ala or Ser to Glu mutations abolished Daxx-p53 interaction. These results thus implicate acetylation and phosphorylation of p53 in regulating its binding t Daxx. Interestingly, whereas Daxx did not bind to p5 in cells as assessed by immunoprecipitation, MDM2 ex pression restored p53-Daxx interaction, and this cor lated with deacetylation of p53. In p5 MDM2 null mouse embryonic fibroblasts (DKO M F), Daxx re pressed p53 target promoters whose p53-b ding ele ments were required for the repress on. C expression of Daxx and MDM2 led to further epress on. p53 expres sion in DKO MEF induced apoptos and Daxx expres sion relieved this effect. Similarly n HCT116 cells, Daxx conferred striking resistance 5-fluorouracil-induced apoptosis. As p53 is require for 5-fluorouracil-induced cell death, our data sho that Daxx can suppress cell death induced by p53 overexpression and p53-depend ent stress response. Collectively, our data reveal Daxx as a novel negative regulator of p53. Importantly, post translational modifications of p53 inhibit Daxx-p53 in teraction, thereby relieving negative regulation of p53 by Daxx.

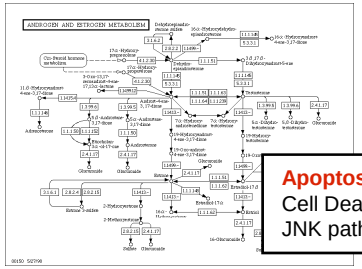
vating expression of gene involved in
ptosis (3, 4), and it can also indu
to Bcl-2 family proteins and tri
Inducing growth arrest and cell
atively on normal cell growth
Indeed, deletion of the *mdm2* ge
regulator of p53, results in emb
both *p53* and *mdm2* simultane
lethal phenotype (6, 7). Thus,
controlled under physiological
MDM2, numerous cellular and
and these proteins can positive
mediated biological effects. Re
strated that the transcription
with p53 (8–10), but the biologi
tion remains to be explored.
Daxx was initially identified
death domain and was shown t
ptosis (11). Subsequent studies
apoptosis in diverse stress con
homozygous deletion of the *Da*
bryonic lethality and widespre
Daxx-deficient embryos and cell
anti-apoptotic function and is c
ment. The anti-apoptotic prope
recent study. Reducing Daxx e
RNA sensitized apoptosis by m

Biological processes

Generate a Profile for Pathway / Phenotype

Apoptosis
Cell Death
Stress response
……..

Score and Rank Terms



Apoptosis
Cell Death
JNK pathway

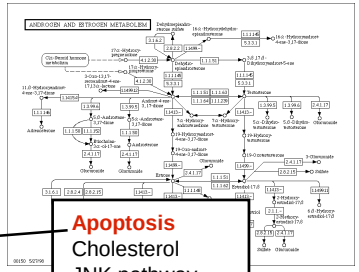
13.27

Common terms

Phenotype Terms

Apoptosis

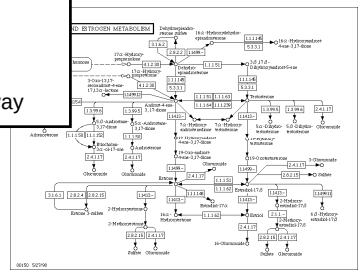
0.15



Apoptosis
Cholesterol
JNK pathway

28.35

Apoptosis
Cholesterol
Diabetes
Apoptosis
JNK pathway
Another pathway

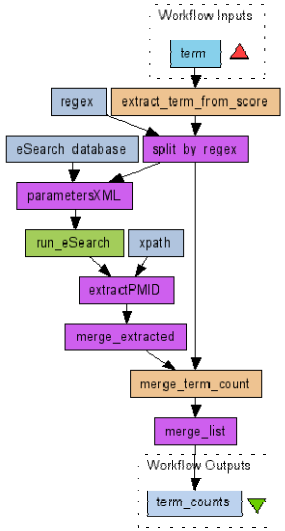


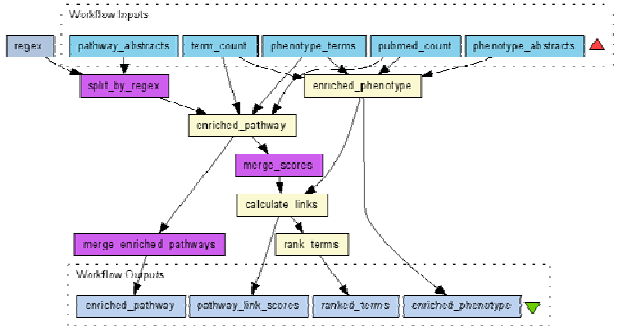
Score pathway links based on occurrence of phenotype term in pathway abstracts

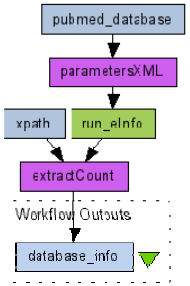
MANCHESTER 1824

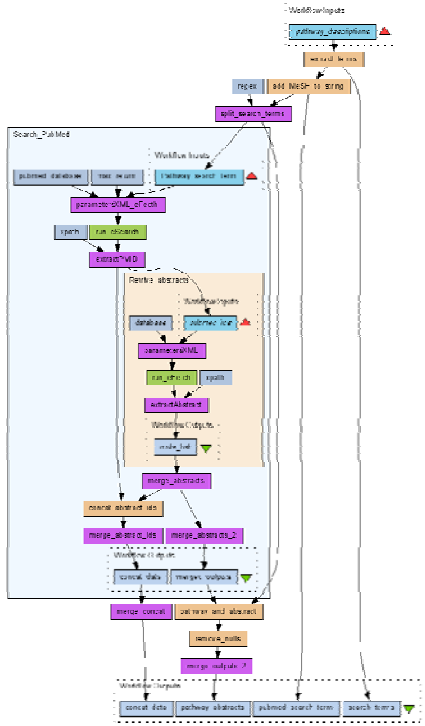
The University of Manchester

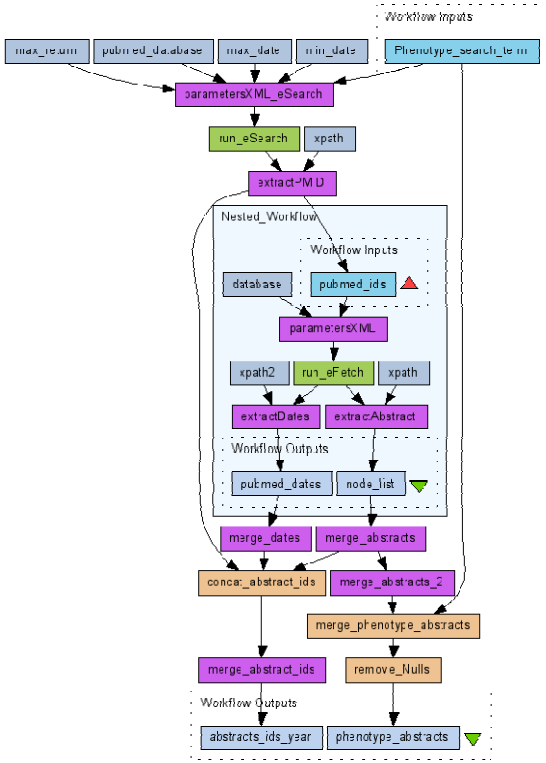
The Workflows

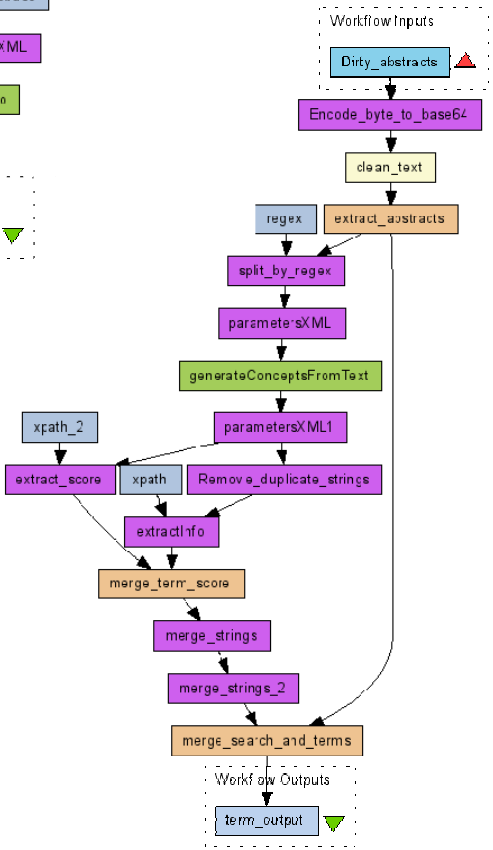












MANCHESTER 1824

The University of Manchester

To Sum Up ….

- Need for Genotype-Phenotype correlations with respect to disease control
- High-throughput data can provide links between Genotype and Phenotype
- Highlighted issues with manually conducted *in silico* experiments
- Improved the methods of current microarray and QTL based investigations through systematic nature
- Increased reproducibility of our methods
  - workflows stored in XML based schema
  - explicit declaration of services, parameters, and methods of data analysis
- Shown workflows are capable of deriving new biologically significant results
  - African Trypanosomiasis in the mouse
  - Infection of mice with *Trichuris muris*
- The workflows require expansion to accommodate new analysis techniques – text mining

BHIG

Many thanks to:

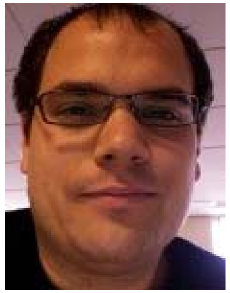


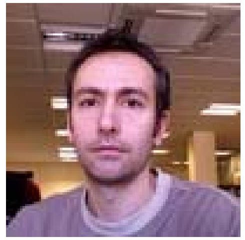
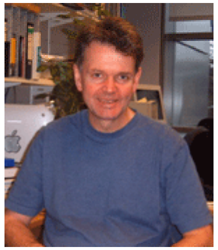















including: Joanne Pennock, EPSRC, OMII, myGrid, and lots more people