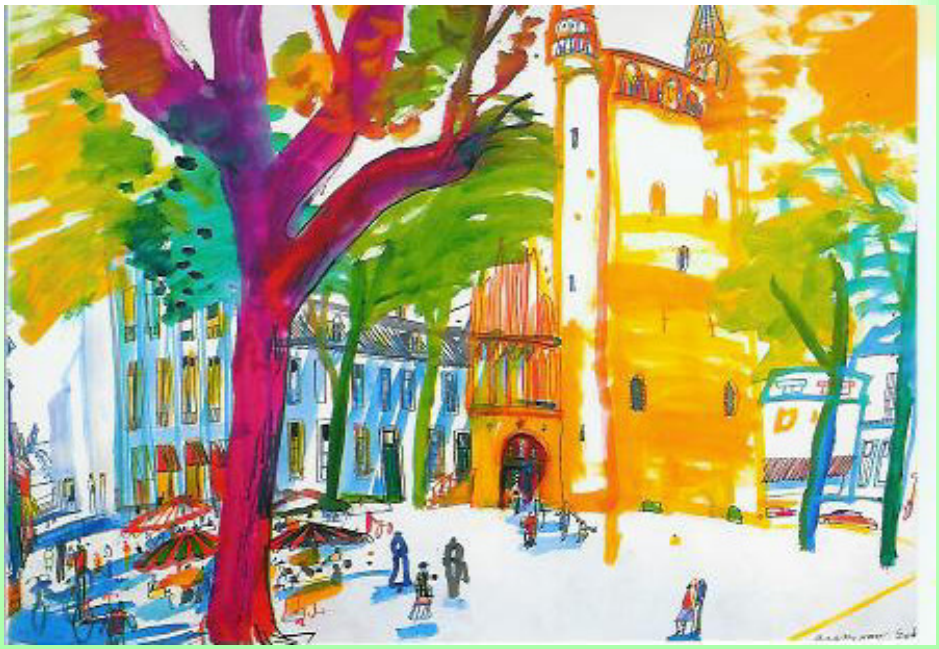
# The nutritional phenotype database
## A real data structure for systems biology

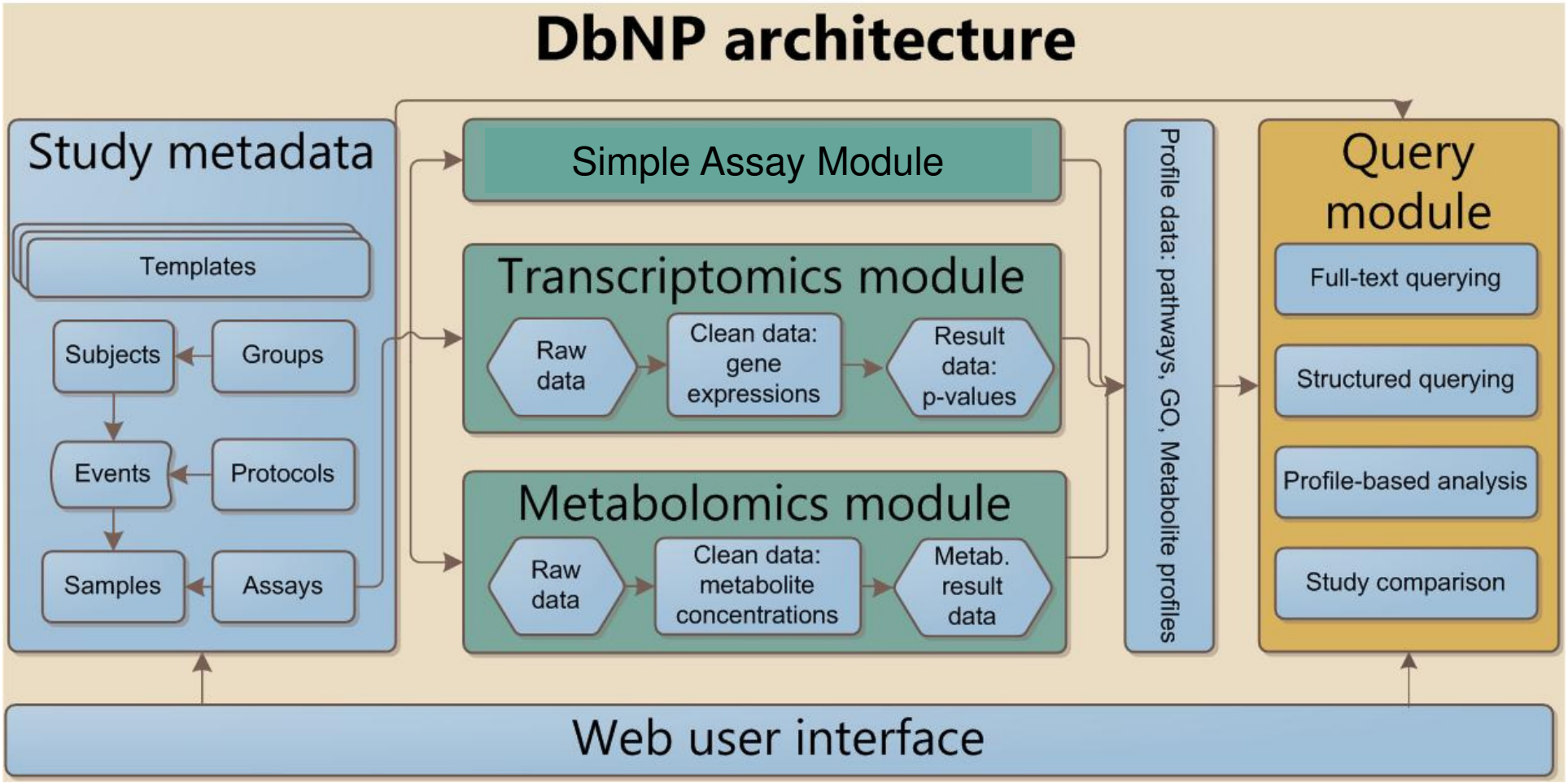
Chris Evelo

Department of Bioinformatics - BiGCaT

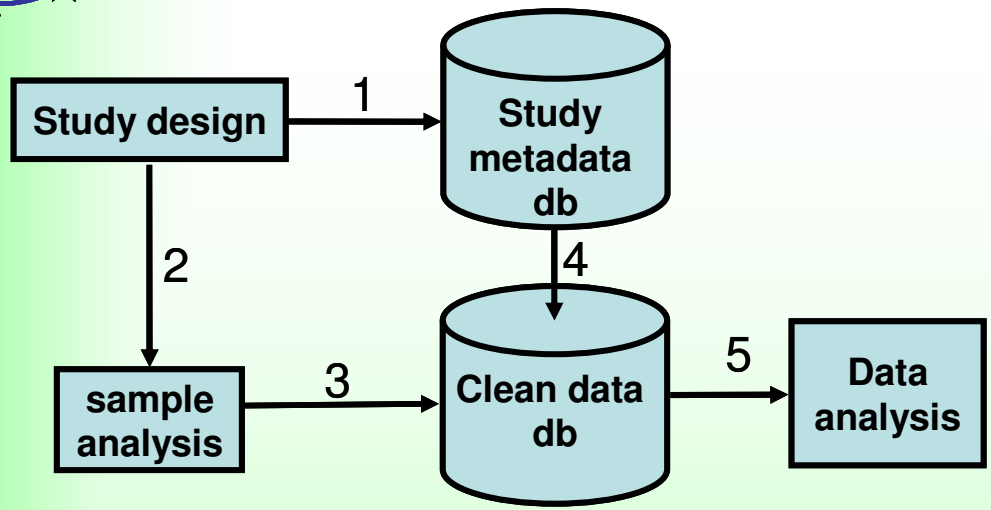
Maastricht University



A paper based on this presentation will soon be published in Genes and Nutrition: http://dx.doi.org/10.1007/s12263-010-0190-x

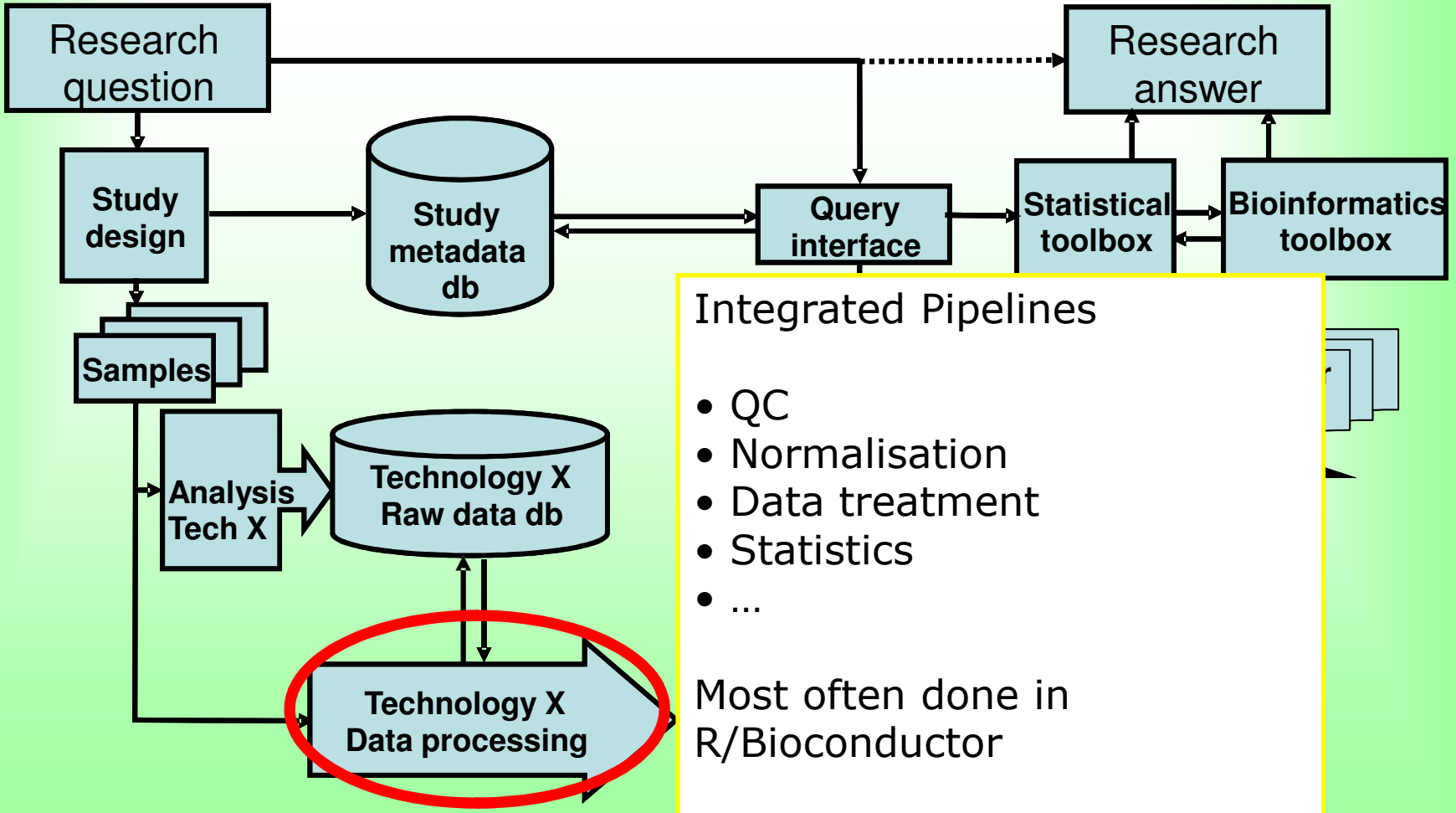


# The basic structure




1) *Protocols are stored in the study metadata database.*
2) *Analytical procedures on study samples*
3) *Processing to 'clean data'*
4) *By interrogating the study metadata database, data subsets of multiple studies can be selected*
5) *and analyzed by statistical and bioinformatics tools.*

# The overall structure



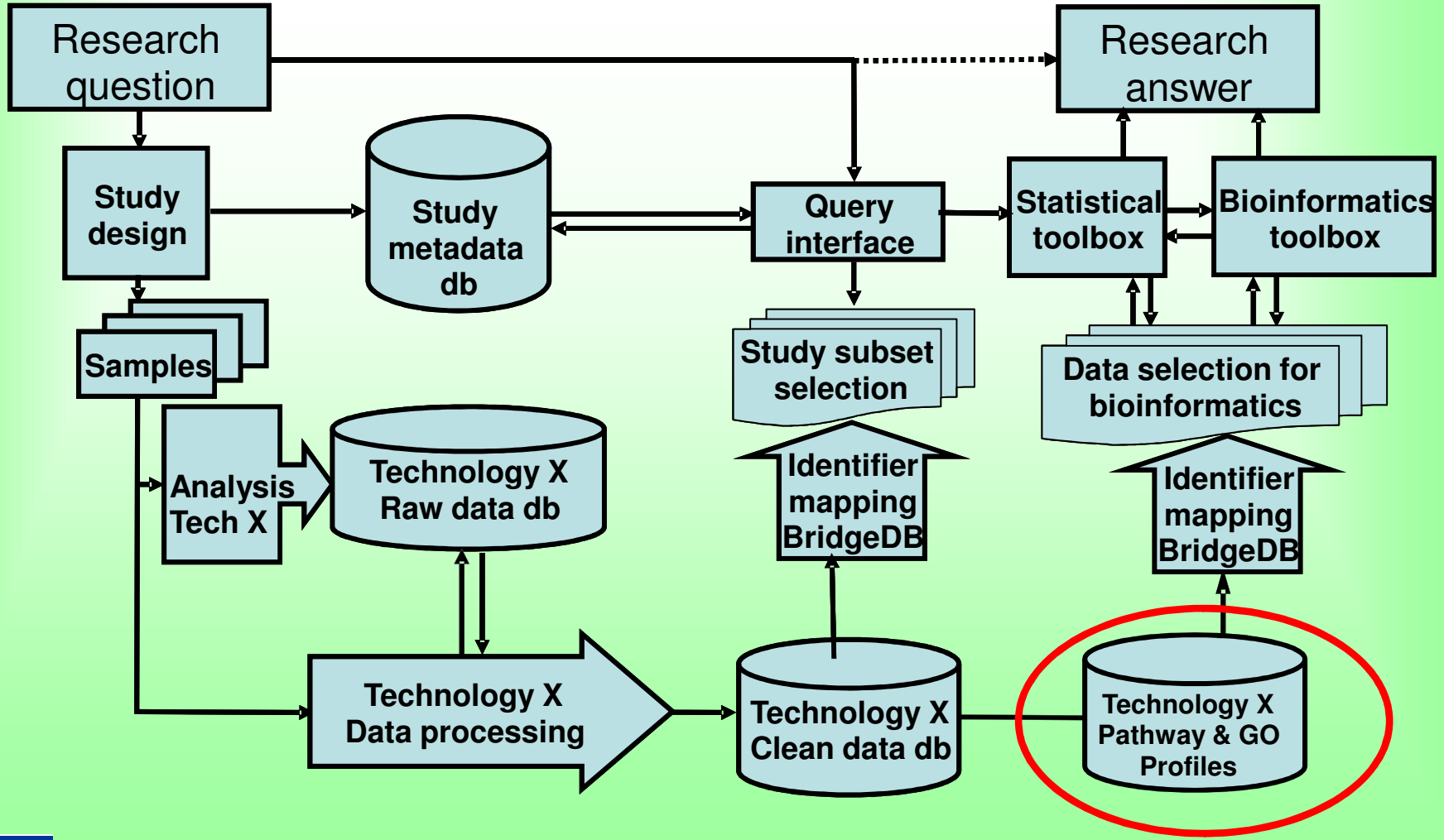

Integrated Pipelines

- QC
- Normalisation
- Data treatment
- Statistics
- …

Most often done in R/Bioconductor

Using NuGO R-servers or Grid

# The overall structure



# Biological processes in duodenal mucosa affected by glutamine administration

| Pathway | number of genes | | | | | Z Score |
|---|---|---|---|---|---|---|
| | Changed | Up | Down | Measured | Total | |
| Hs_Mitochondrial_fatty_acid_betaoxidation | 6 | 6 | 0 | 16 | 16 | 4.456 |
| Hs_Electron_Transport_Chain | 17 | 17 | 0 | 85 | 105 | 4.278 |
| Hs_Fatty_Acid_Synthesis | 5 | 5 | 0 | 21 | 22 | 2.757 |
| Hs_Fatty_Acid_Beta-Oxidation | 6 | 6 | 0 | 31 | 32 | 2.424 |
| Hs_mRNA_processing_Reactome | 16 | 6 | 10 | 118 | 127 | 2.402 |
| Hs_Unsaturated_Fatty_Acid_Beta_Oxidation | 2 | 2 | 0 | 6 | 6 | 2.342 |
| Hs_HSP70_and_Apoptosis | 4 | 4 | 0 | 18 | 18 | 2.299 |
| Hs_Oxidative_Stress | 5 | 5 | 0 | 27 | 28 | 2.097 |
| Hs_Fatty_Acid_Omega_Oxidation | 3 | 3 | 0 | 14 | 15 | 1.915 |
| Hs_Proteasome_Degradation | 8 | 8 | 0 | 60 | 61 | 1.629 |
| Hs_RNA_transcription_Reactome | 5 | 5 | 0 | 38 | 40 | 1.25 |
| Hs_Irinotecan_pathway_PharmGKB | 2 | 1 | 1 | 12 | 12 | 1.154 |
| Hs_Synthesis_and_Degradation_of_Ketone_Bodies_KEGG | 1 | 1 | 0 | 5 | 5 | 1.023 |

# Find the relevant pathways

- And ask what other studies contain high z-scores in these pathways

- Essentially asking questions like:

"What studies showed an effect on e.g. Apoptosis"

Querying pathway by pathway instead of gene by gene

# Biological processes in duodenal mucosa affected by glutamine administration

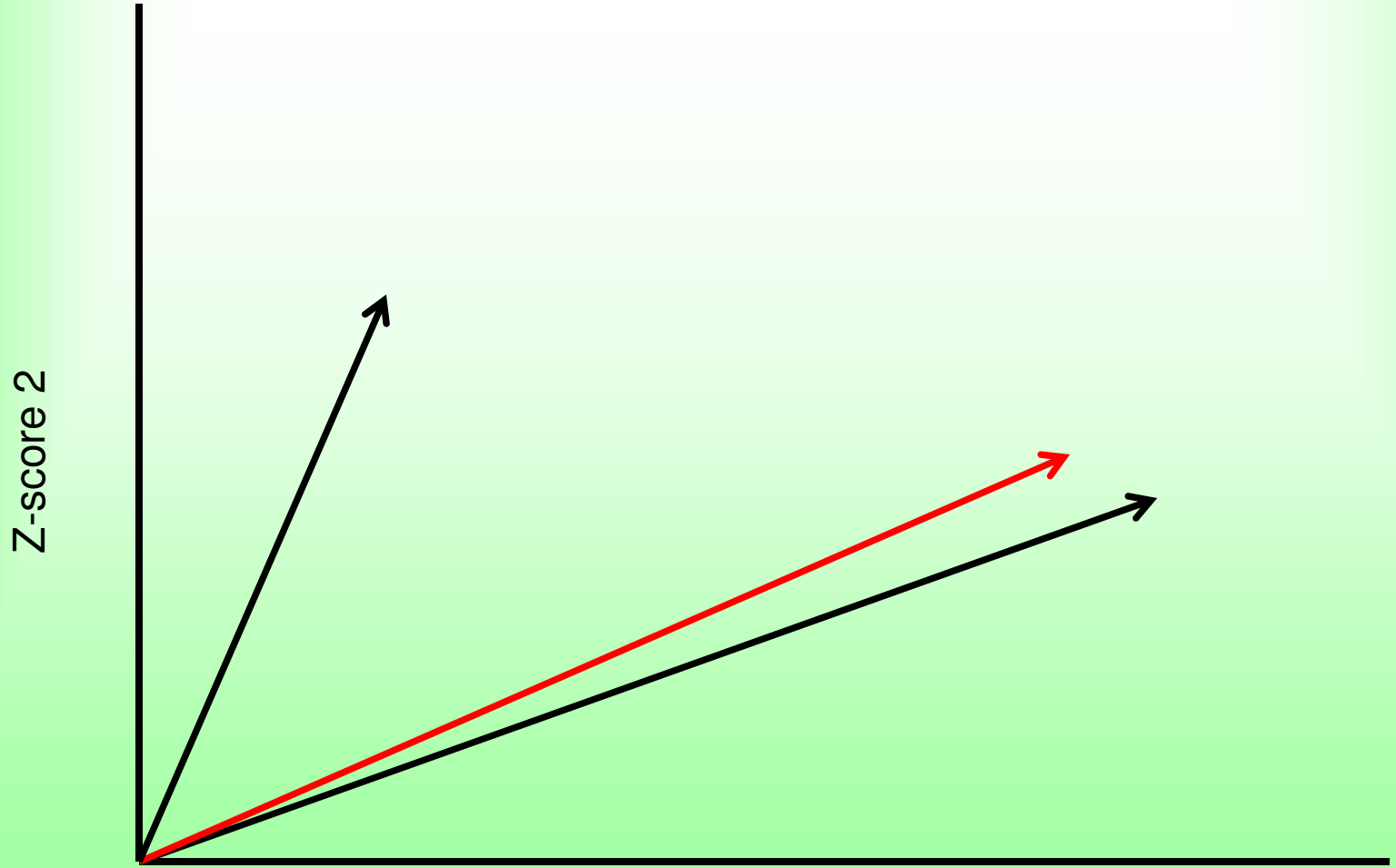
| Pathway | number of genes | | | | | Z Score |
|---|---|---|---|---|---|---|
| | Changed | Up | Down | Measured | Total | |
| Hs_Mitochondrial_fatty_acid_betaoxidation | 6 | 6 | 0 | 16 | 16 | 4.456 |
| Hs_Electron_Transport_Chain | 17 | 17 | 0 | 85 | 105 | 4.278 |
| Hs_Fatty_Acid_Synthesis | 5 | 5 | 0 | 21 | 22 | 2.757 |
| Hs_Fatty_Acid_Beta-Oxidation | 6 | 6 | 0 | 31 | 32 | 2.424 |
| Hs_mRNA_processing_Reactome | 16 | 6 | 10 | 118 | 127 | 2.402 |
| Hs_Unsaturated_Fatty_Acid_Beta_Oxidation | 2 | 2 | 0 | 6 | 6 | 2.342 |
| Hs_HSP70_and_Apoptosis | 4 | 4 | 0 | 18 | 18 | 2.299 |
| Hs_Oxidative_Stress | 5 | 5 | 0 | 27 | 28 | 2.097 |
| Hs_Fatty_Acid_Omega_Oxidation | 3 | 3 | 0 | 14 | 15 | 1.915 |
| Hs_Proteasome_Degradation | 8 | 8 | 0 | 60 | 61 | 1.629 |
| Hs_RNA_transcription_Reactome | 5 | 5 | 0 | 38 | 40 | 1.25 |
| Hs_Irinotecan_pathway_PharmGKB | 2 | 1 | 1 | 12 | 12 | 1.154 |
| Hs_Synthesis_and_Degradation_of_Ketone_Bodies_KEGG | 1 | 1 | 0 | 5 | 5 | 1.023 |

# Biological processes in duodenal mucosa affected by glutamine administration

| Pathway | Pathway ID | Z Score |
|---|---|---|
| Hs_Mitochondrial_fatty_acid_betaoxidation | xxxx | 4.456 |
| Hs_Electron_Transport_Chain | …. | 4.278 |
| Hs_Fatty_Acid_Synthesis | | 2.757 |
| Hs_Fatty_Acid_Beta-Oxidation | | 2.424 |
| Hs_mRNA_processing_Reactome | | 2.402 |
| Hs_Unsaturated_Fatty_Acid_Beta_Oxidation | | 2.342 |
| Hs_HSP70_and_Apoptosis | | 2.299 |
| Hs_Oxidative_Stress | | 2.097 |
| Hs_Fatty_Acid_Omega_Oxidation | | 1.915 |
| Hs_Proteasome_Degradation | | 1.629 |
| Hs_RNA_transcription_Reactome | | 1.25 |
| Hs_Irinotecan_pathway_PharmGKB | | 1.154 |
| Hs_Synthesis_and_Degradation_of_Ketone_Bodies_KEGG | | 1.023 |

That's a vector!

# Find a nearby vector

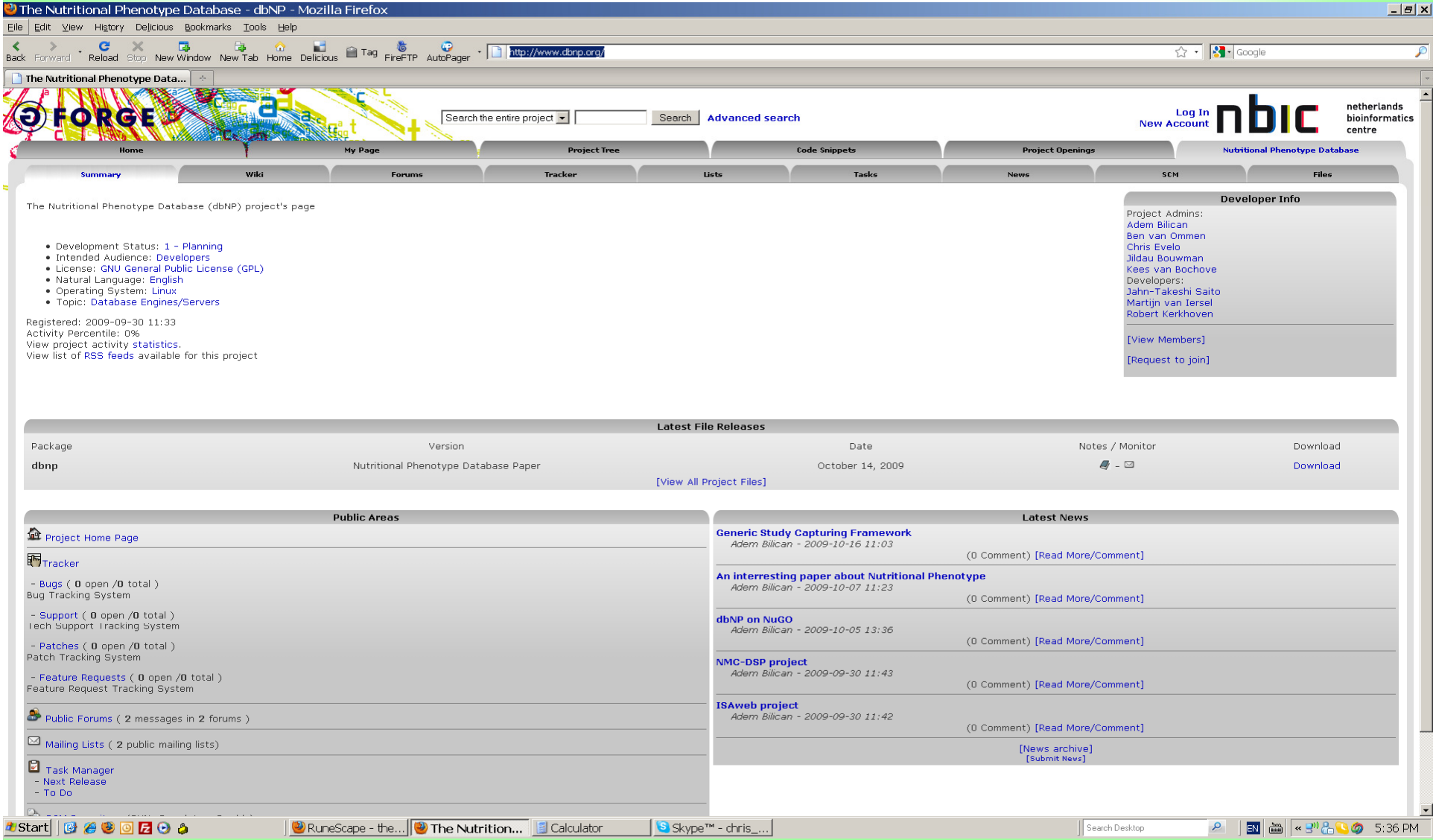


# Vectors look alike?

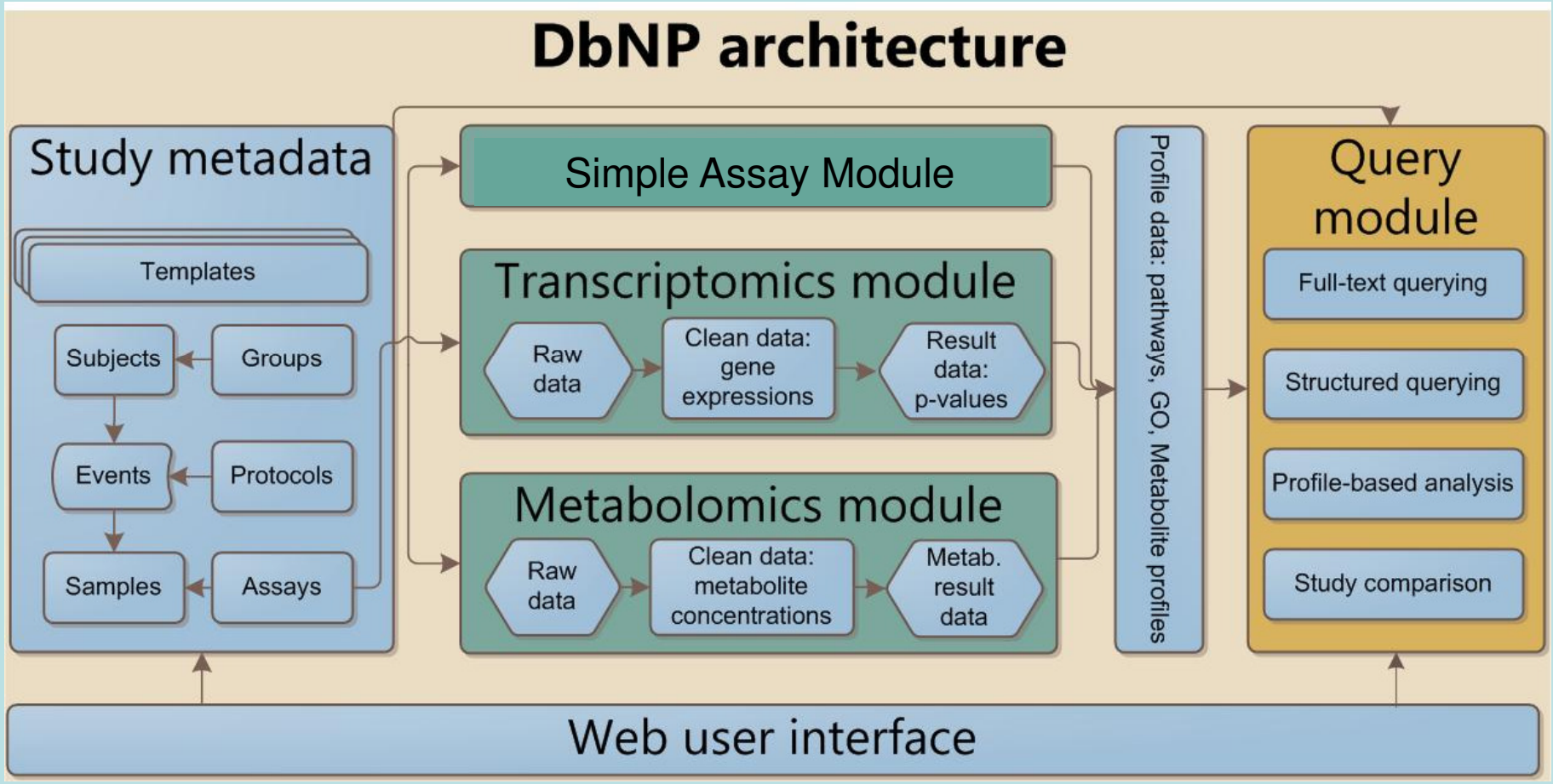
- 1 study is a vector
- So 2 studies is 2 vectors
- And we can calculate the distance
- And find the other vector closest
- So we can ask questions like:

"What other studies showed an overall biological effect like this one"

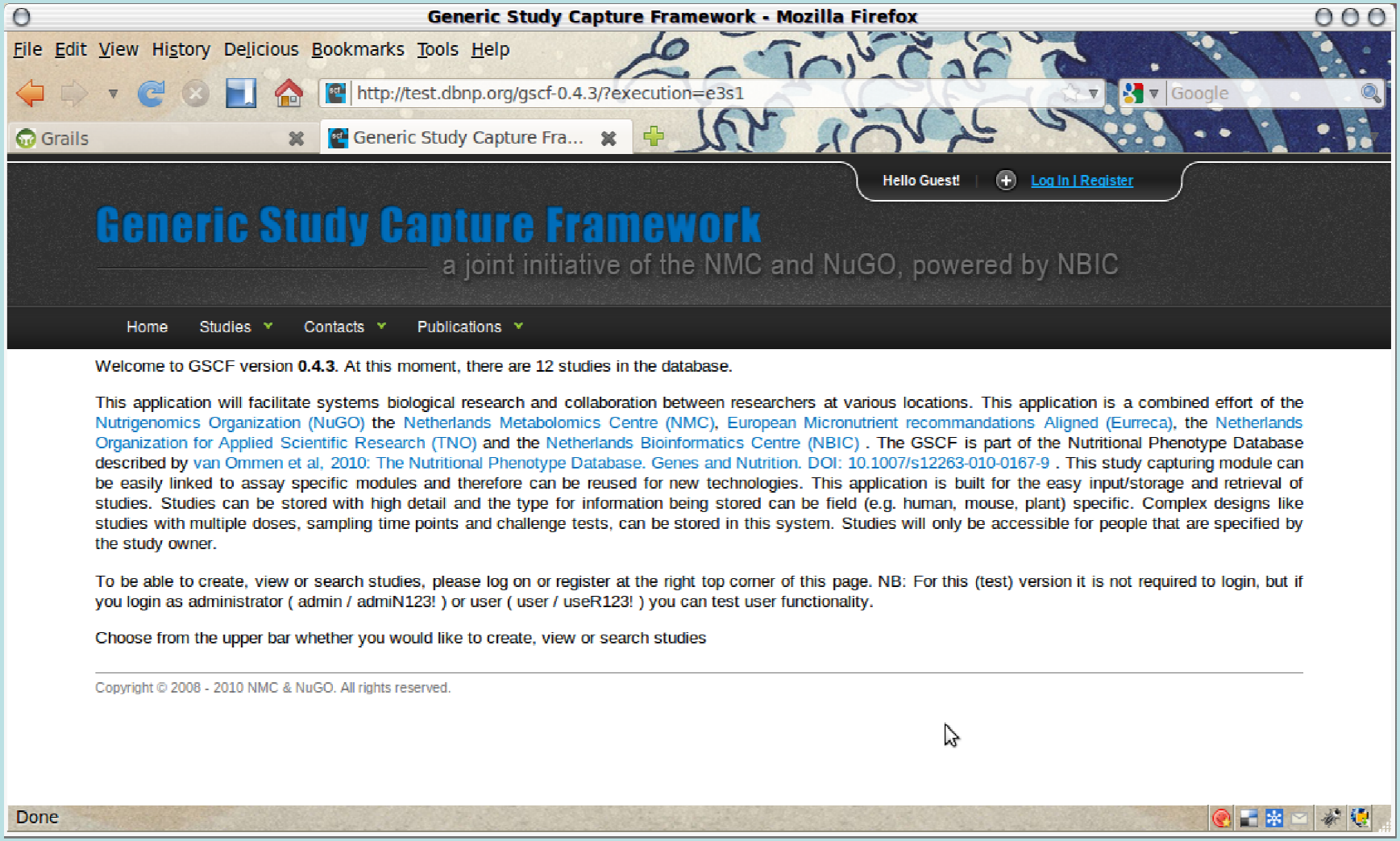
# www.dbNP.org



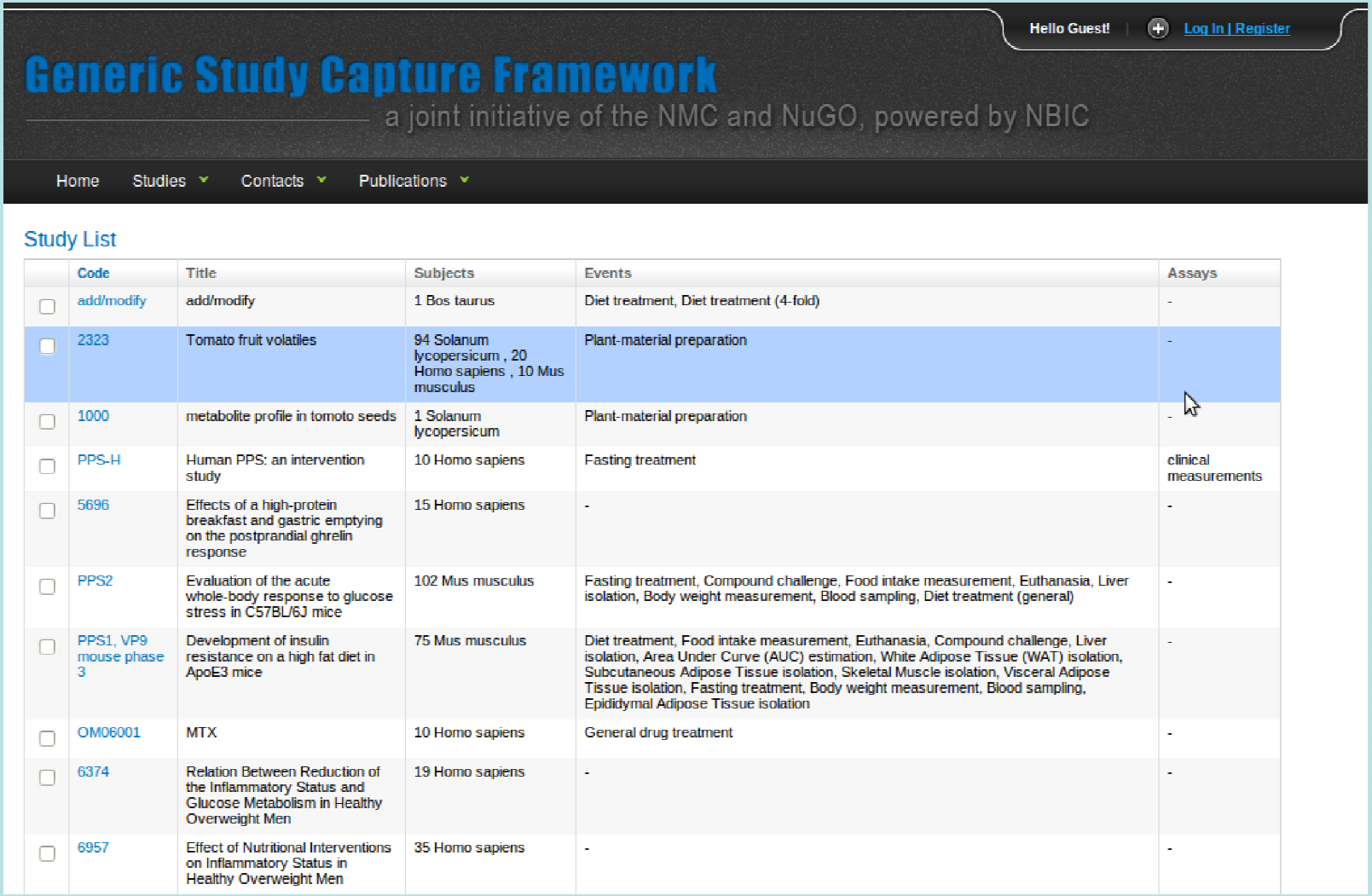
# The real thing…



# Browsing the GSCF

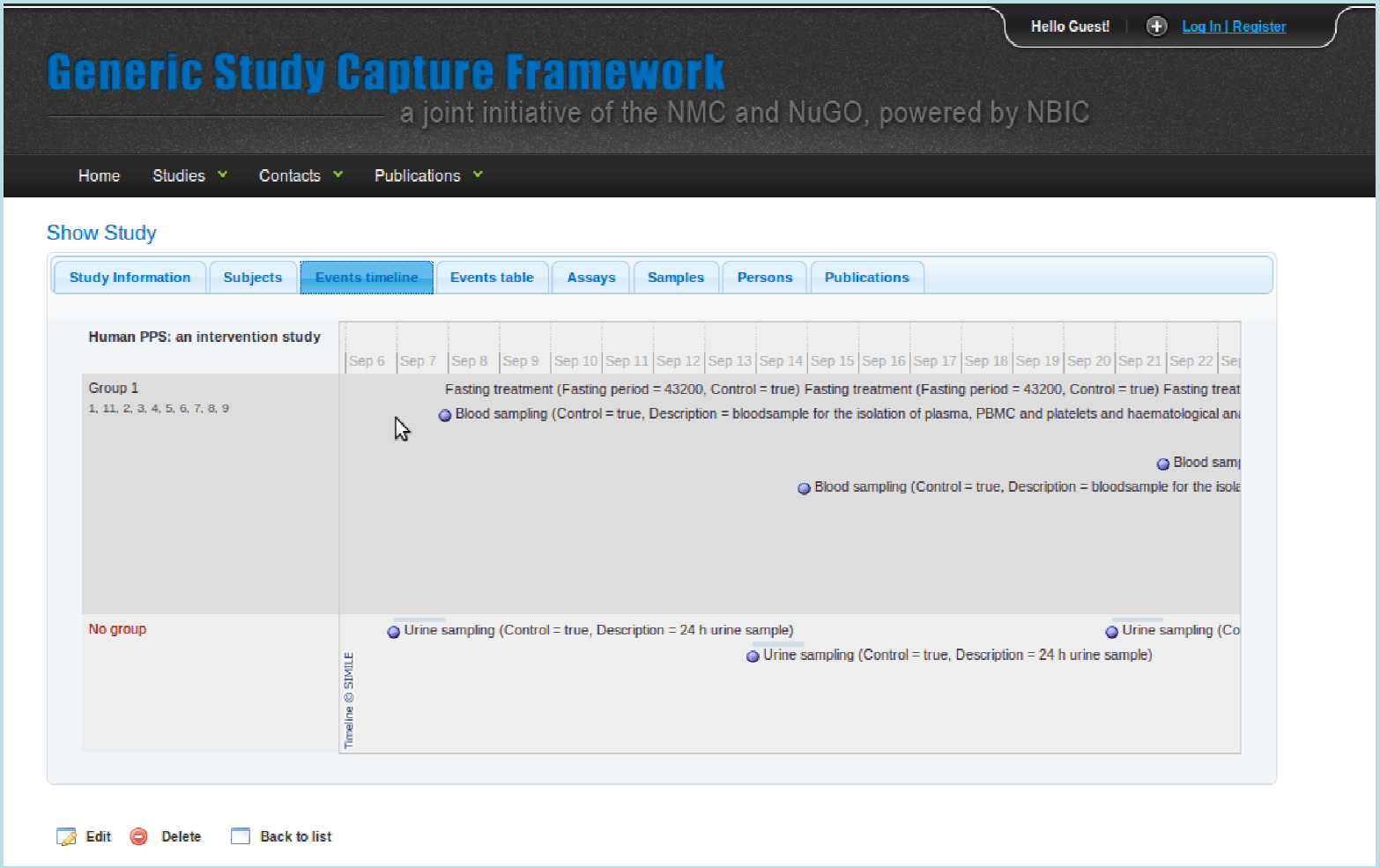


# Study overview

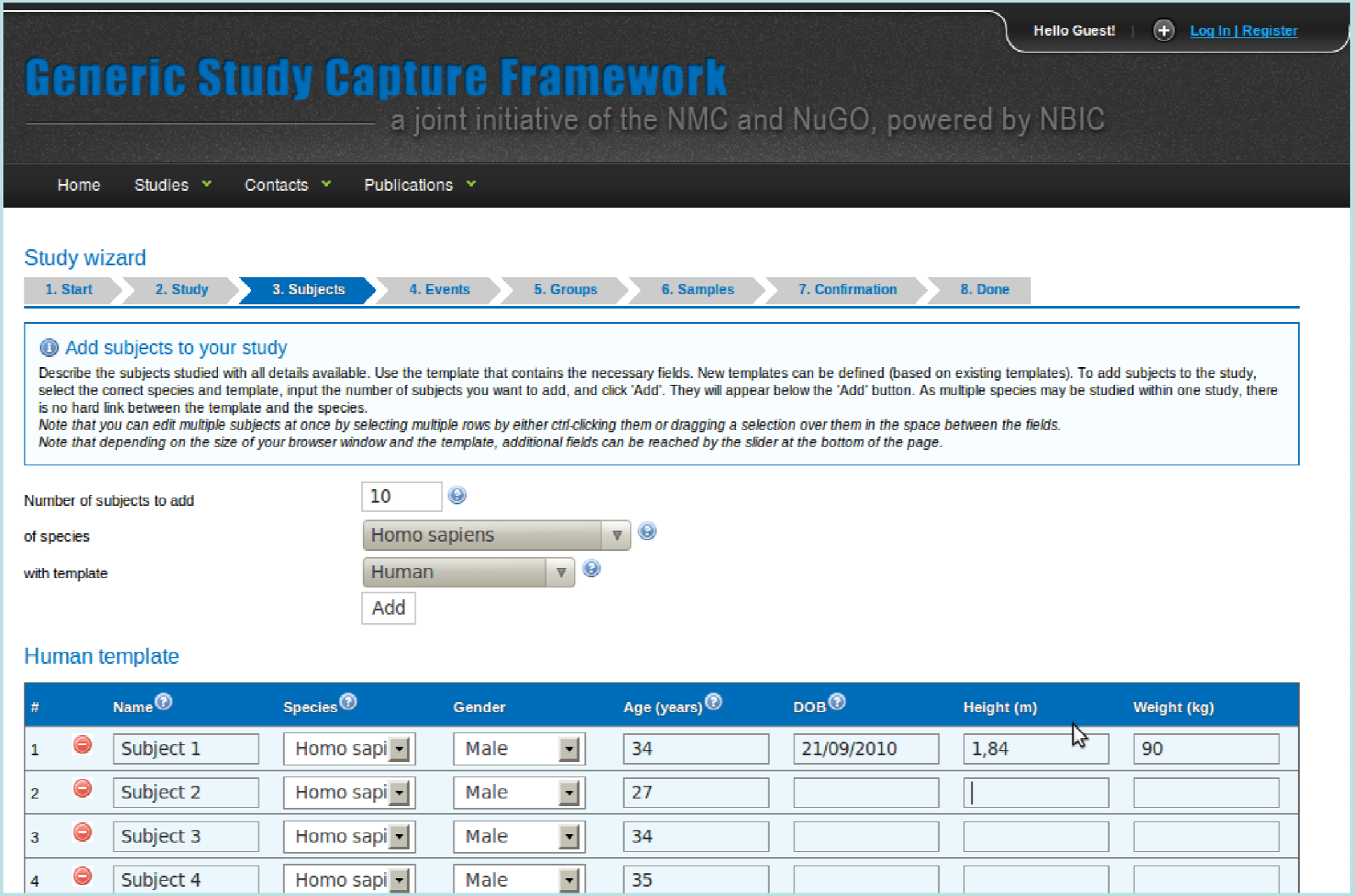



| | Code | Title | Subjects | Events | Assays |
|---|---|---|---|---|---|
| ☐ | add/modify | add/modify | 1 Bos taurus | Diet treatment, Diet treatment (4-fold) | - |
| ☐ | 2323 | Tomato fruit volatiles | 94 Solanum lycopersicum , 20 Homo sapiens , 10 Mus musculus | Plant-material preparation | - |
| ☐ | 1000 | metabolite profile in tomoto seeds | 1 Solanum lycopersicum | Plant-material preparation | - |
| ☐ | PPS-H | Human PPS: an intervention study | 10 Homo sapiens | Fasting treatment | clinical measurements |
| ☐ | 5696 | Effects of a high-protein breakfast and gastric emptying on the postprandial ghrelin response | 15 Homo sapiens | - | - |
| ☐ | PPS2 | Evaluation of the acute whole-body response to glucose stress in C57BL/6J mice | 102 Mus musculus | Fasting treatment, Compound challenge, Food intake measurement, Euthanasia, Liver isolation, Body weight measurement, Blood sampling, Diet treatment (general) | - |
| ☐ | PPS1, VP9 mouse phase 3 | Development of insulin resistance on a high fat diet in ApoE3 mice | 75 Mus musculus | Diet treatment, Food intake measurement, Euthanasia, Compound challenge, Liver isolation, Area Under Curve (AUC) estimation, White Adipose Tissue (WAT) isolation, Subcutaneous Adipose Tissue isolation, Skeletal Muscle isolation, Visceral Adipose Tissue isolation, Fasting treatment, Body weight measurement, Blood sampling, Epididymal Adipose Tissue isolation | - |
| ☐ | OM06001 | MTX | 10 Homo sapiens | General drug treatment | - |
| ☐ | 6374 | Relation Between Reduction of the Inflammatory Status and Glucose Metabolism in Healthy Overweight Men | 19 Homo sapiens | - | - |
| ☐ | 6957 | Effect of Nutritional Interventions on Inflammatory Status in Healthy Overweight Men | 35 Homo sapiens | - | - |

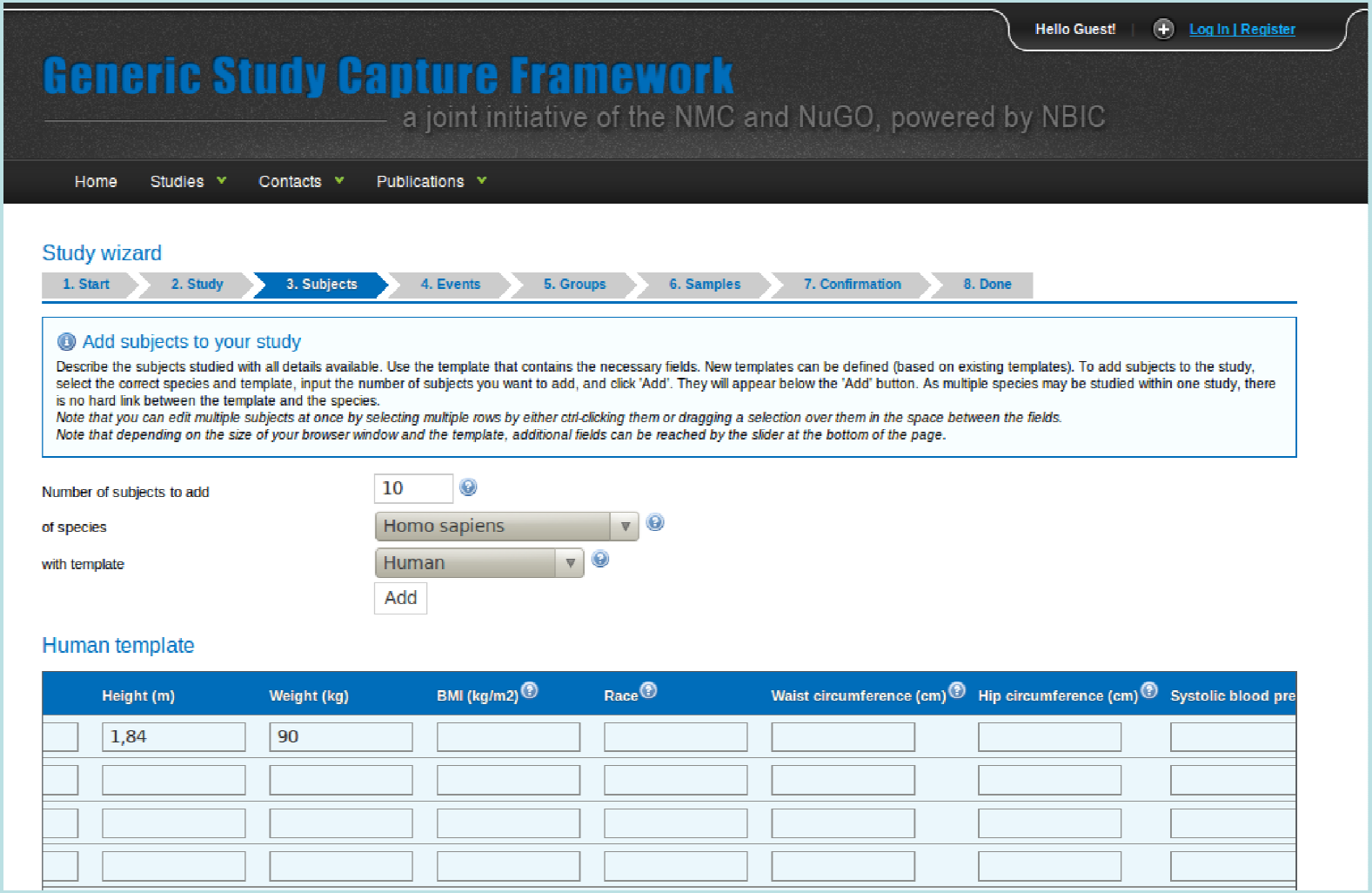
# Example view: event time line



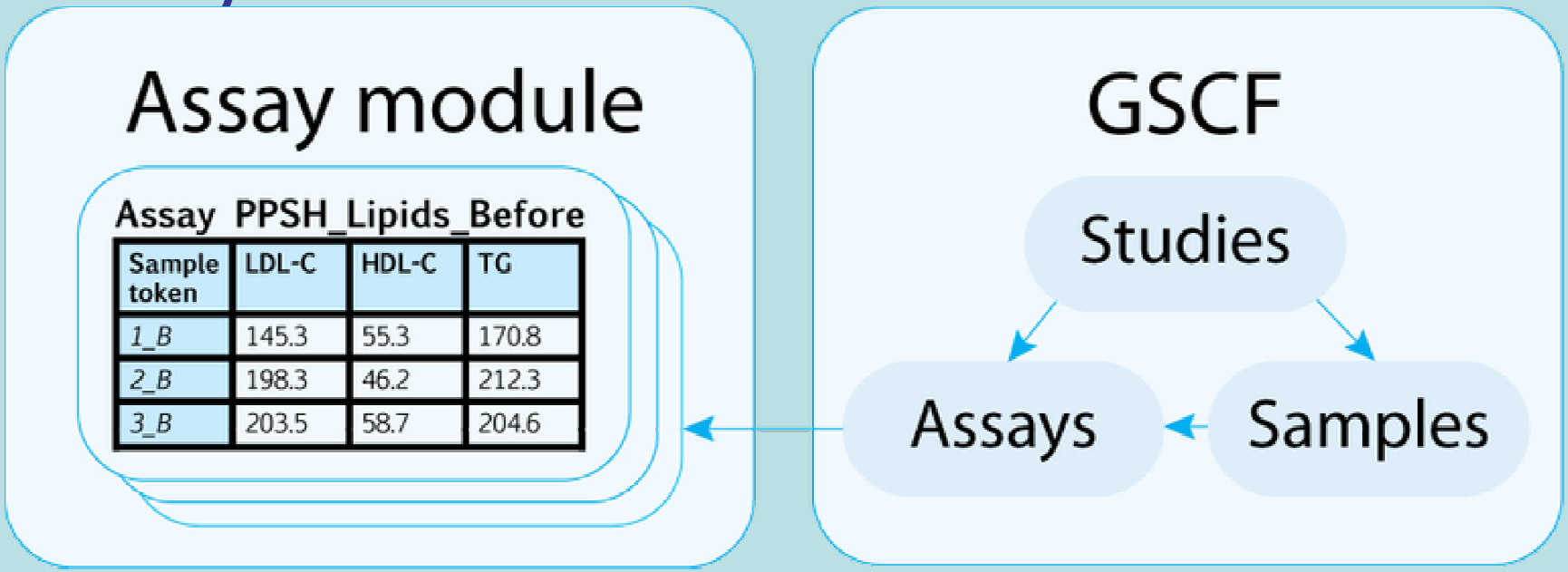
# Web flow for creating studies



# Templates for a human nutrition study

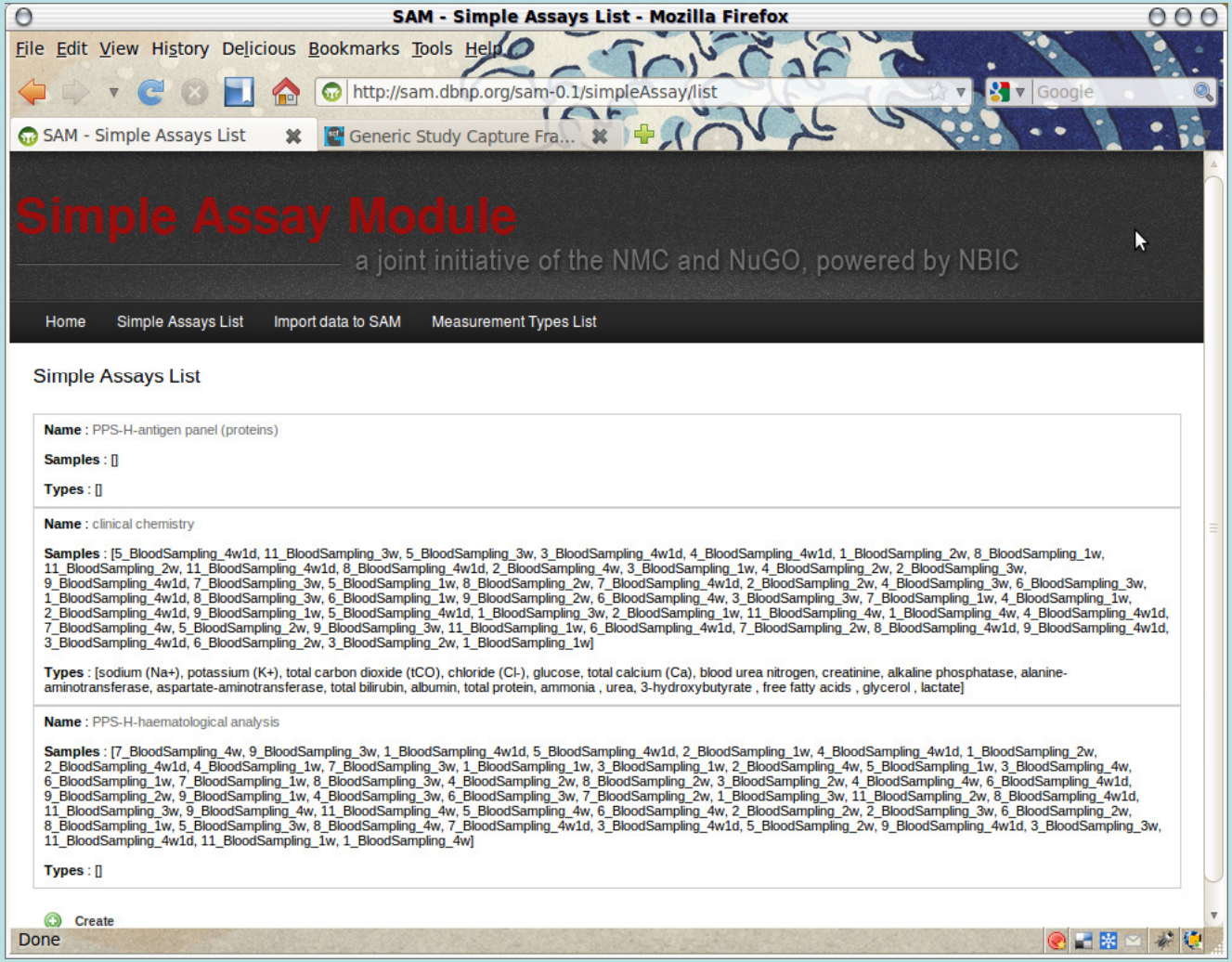


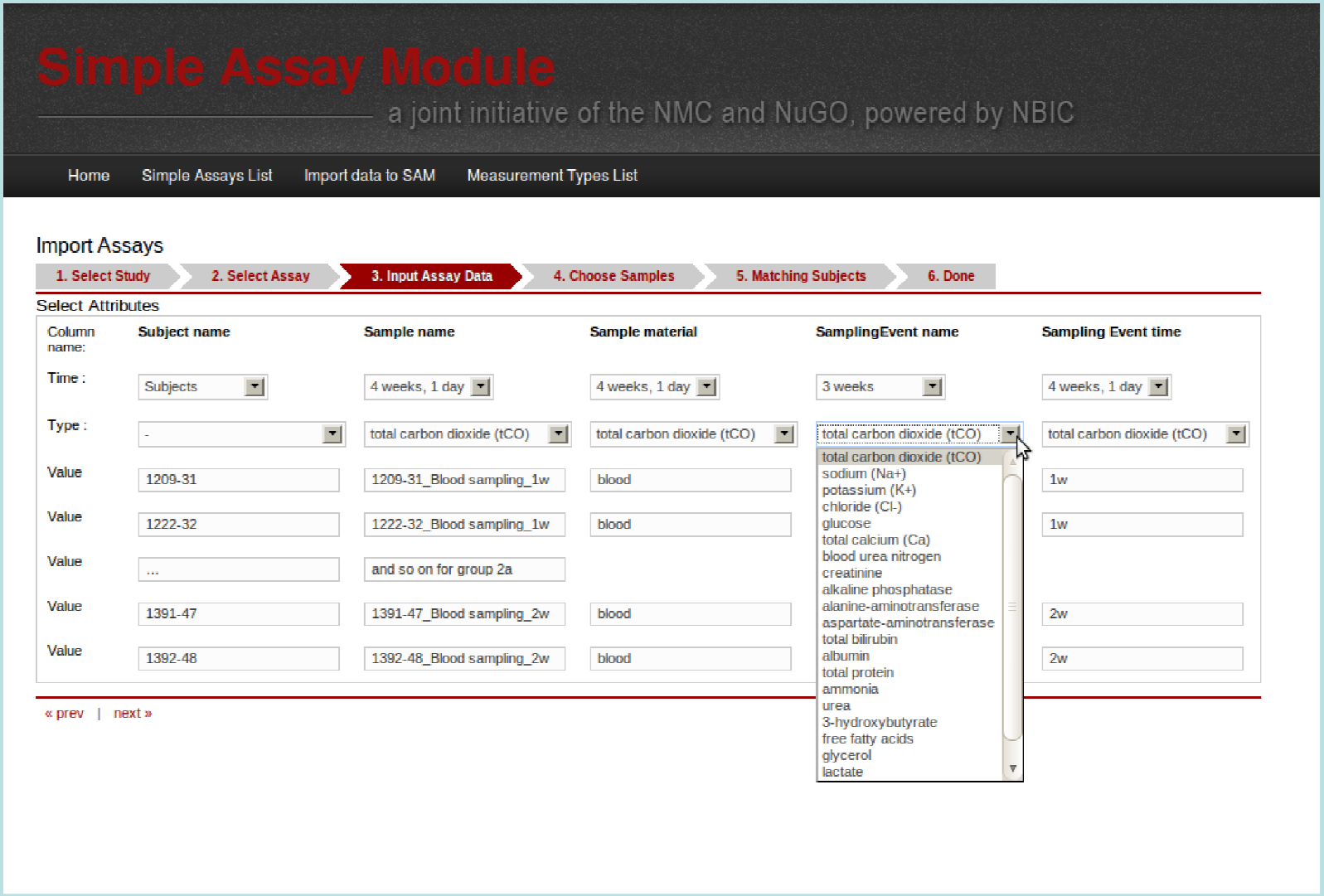
# Assays and modules in dbNP



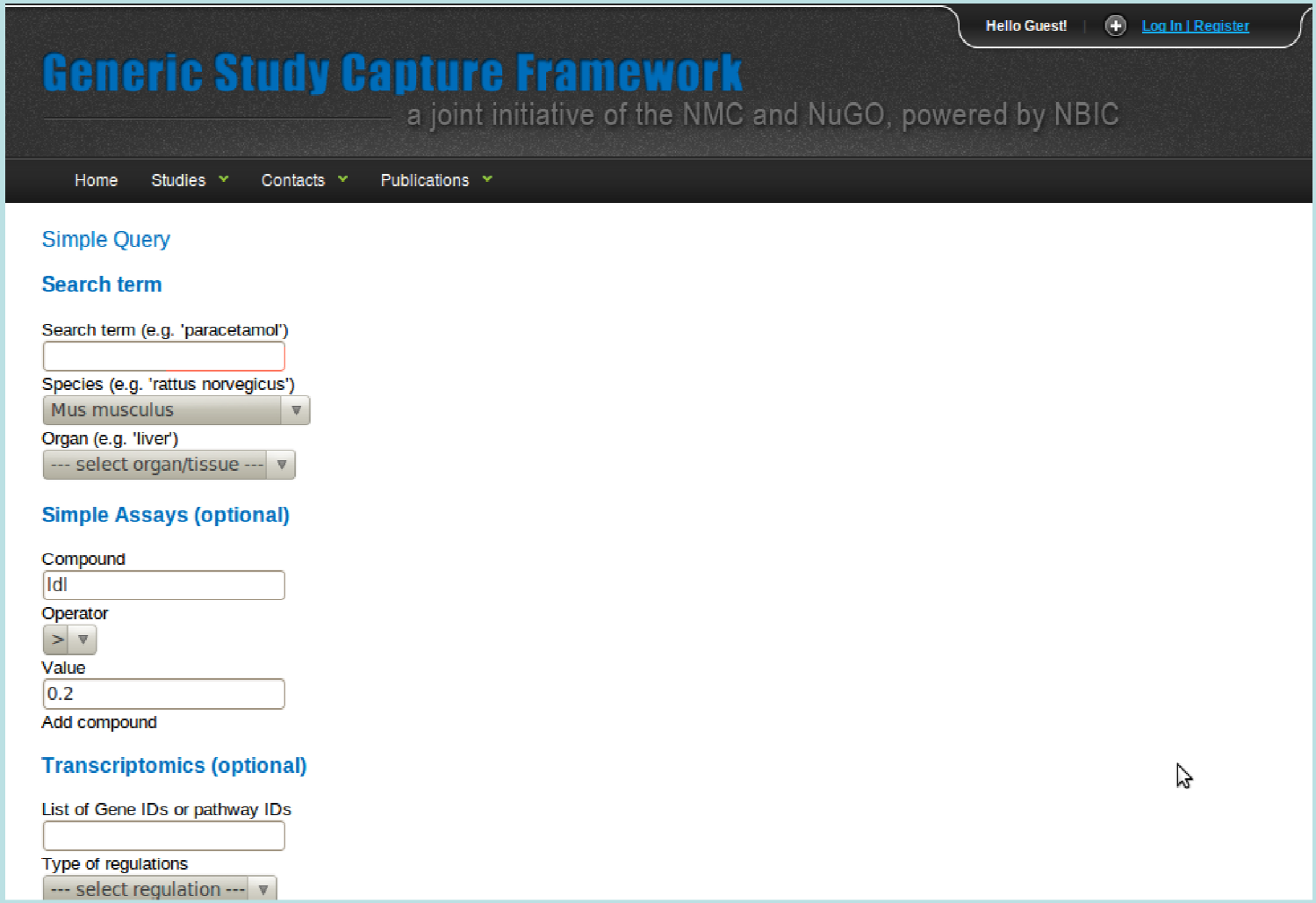

- Modules store and process types of assays
- Simple Assay (SAM)
- Transciptomics assay (Transcriptomics module)
- Metabolomics assay (NMC-DSP)

# Example Module: SAM





# Querying assays from SAM in GSCF



# Many people involved

**Challenges of molecular nutrition research 6: The Nutritional Phenotype database to store, share and evaluate nutritional systems biology studies.**

Ben van Ommen, Jildau Bouwman, Lars Dragsted, Christian A. Drevon, Ruan Elliott, Philip de Groot, Jim Kaput, John C. Mathers, Michael Müller, Fre Pepping, Jahn Saito, Augustin Scalbert, Marijana Radonjic, Philippe Rocca-Serra, Tony Travis, Suzan Wopereis and Chris Evelo. Genes and Nutrition, 2010

Developers on dnNP.org lead by Kees van Bochove

Pathway profiling: Martijn van Iersel, Martina Kutmon, Jahn Saito

Investments (time) by: NuGO, NMC, NBIC, FDA

Many more to come…(including innovative medicine initiative FP7)

A paper based on this presentation will soon be published in Genes and Nutrition: http://dx.doi.org/10.1007/s12263-010-0190-x