A MIAME for Toxicogenomics - Towards Harmonization of a New Field

Susanna-Assunta Sansone[1], Weida Tong[2] and Michael D. Waters[3]

[1]European Bioinformatics Institute,
EMBL-EBI, Cambridge, UK

[2]Center for Toxicoinformatics,
NCTR, FDA, Jefferson, AR, USA

[3]National Center for Toxicogenomics,
NIEHS, NIH, Research Triangle Park, NC, USA

NETTAB, Bologna, Italy, November 2003

Talk Structure

- Toxicogenomics
  - Definition and objectives
  - Potentials, challenges and priority
- Towards harmonization of this new field
  - Required, ongoing and accomplished efforts
- Focus on array-based toxicogenomics experiments
  - Tox Working Group within MGED
  - MIAME/Tox and MIAME/Tox-compliant databases

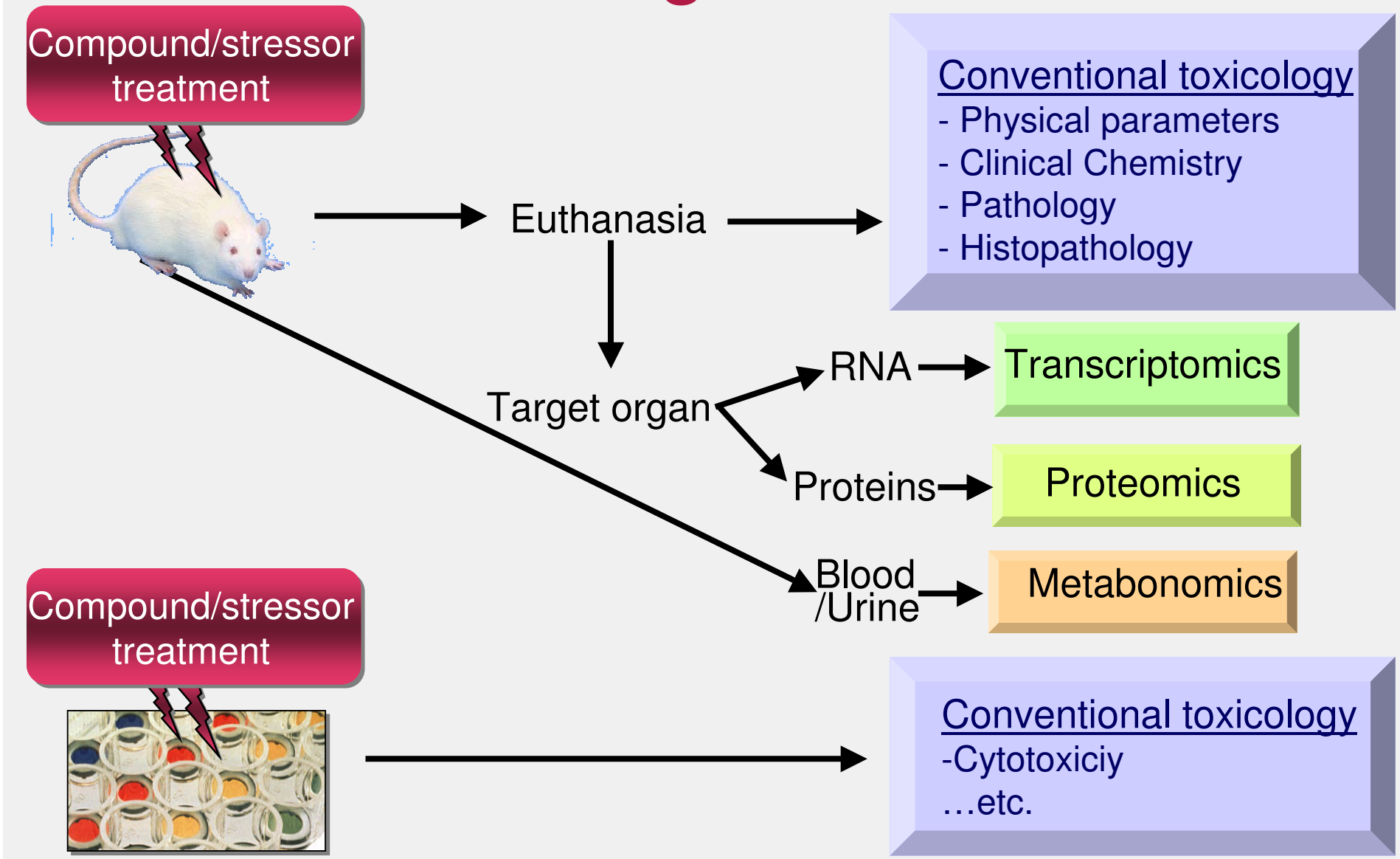
Talk Structure

- Toxicogenomics
  - Definition and objectives
  - Potentials, challenges and priority
- Towards harmonization of this new field
  - Required, ongoing and accomplished efforts
- Focus on array-based toxicogenomics experiments
  - Tox Working Group within MGED
  - MIAME/Tox and MIAME/Tox-compliant databases

The Definition

Toxicogenomics
is the study of the
response of a genome to
environmental stressors and toxicants

Toxicogenomics



The Objectives

- Compare toxicogenomic effects of chemicals/ stressors across species
  - Yielding signatures of altered gene/protein expression
- 'Phenotypically anchor' these changes with conventional toxicology data
  - Classifying effects as well as disease phenotypes
- Delineate global changes as pharmacologic, adaptive or toxic outcomes
  - Defining biomarkers, sequence of key events, mechanisms of action

The Potentials

- Holding promises for
  - Drug/biologics discovery and development
    - Target selection, risk assessment and quality control
  - Chemical/drug-induced disease processes
    - Evaluation and prediction
  - Medical practice
    - Diagnostic, therapeutic decisions and monitoring
  - Regulatory science *(health and environmental)*
    - Support and facilitate the decision-making process
- Improving methods to assess toxicity
  - Gain insight into the molecular mechanisms
  - Reduce the length of long-term toxicology study
  - Limit the numbers of species used

Toxicogenomics in Europe

PUBLIC HEALTH

E.U. Starts a Chemical Reaction

BRUSSELS—If the European Unio
gets its way, toxicology will
booming in Europe. And, to hear
manufacturers tell it, their industr
in decline.

Last week, officials from the
Commission, Europe's executive b
with industry and environmenta
here to discuss proposed legisla
would require chemical manufac
run extensive safety tests over 11
the 30,000 most common chemica
market, many of which have been
decades. The proposal would also
restrict the use of an estimated 150
cals considered the most hazardo
mans and the environment. "The n
introduces a radical paradigm

Paradigm shift. The E.U.'s Margot Wallström wants the chemical industry to shoulder the burden of safety.

TOXICOLOGY

Europe Whittles Down Plans for Massive Chemical Testing Program

MADRID—The European Commission (EC) has scaled back a major piece of legislation on safety testing of commercial chemicals. Yet even in its revised form, the proposed law would represent one of the most ambitious toxicological programs ever undertaken.

An earlier version of the legislation, which has been in the works for more than 2 years, would have required chemical makers to perform extensive toxicological and environmental tests on the 30,000 chemicals most commonly used in commerce (*Science*, 18 April, p. 405). Under the latest draft, released by the EC last week, the testing requirements would apply only to chemicals produced in amounts greater than 10 tons, covering about one-third of
inally envisioned. Some 15
European regulators deem
ardous to human health—

nated flame retardants, phthalates used as plastic softeners, and perfluorinated compounds—are likely to be severely restricted or banned, the EC says.

The testing program, to be called REACH (Registration, Evaluation, and Authorization of Chemicals), would require some safety tests of chemicals produced in amounts of between 1 and 10 tons. But such substances would be exempt from tests of reproductive effects and environmental persistence. The changes mean that "we will have no idea how far the chemicals get into the environment," contends Stefan Scheuer of the European Environmental Bureau, a coalition of 140 nongovernmental organizations.

that the tests would cost as much as $12 billion. In reworking the legislation, the EC stated that it wants a program that would not unduly crimp European competitiveness. Industry and environmental groups concur that the legislation, although watered down, still amounts to a radical change. "At the end of the day, this will still be the biggest such program in the world," says Véronique Scailteur, director of external relations at Procter & Gamble's headquarters in Brussels. The testing is now expected to cost about $2.3 billion.

Scheuer says, however, that he and other activists are planning a lobbying counterattack to try to persuade the European Parliament and the Council of Ministers to re-

R&D. U
ities of
than 1 n
istration
proposed
striction
years, w
extensio
Patrick
stewards
The exi
tight th
testing,
Some
tive sid

Alternative (Non-animal) Methods for Chemicals Testing: Current Status and Future Prospects

A Report prepared by ECVAM and the ECVAM Working Group on Chemicals

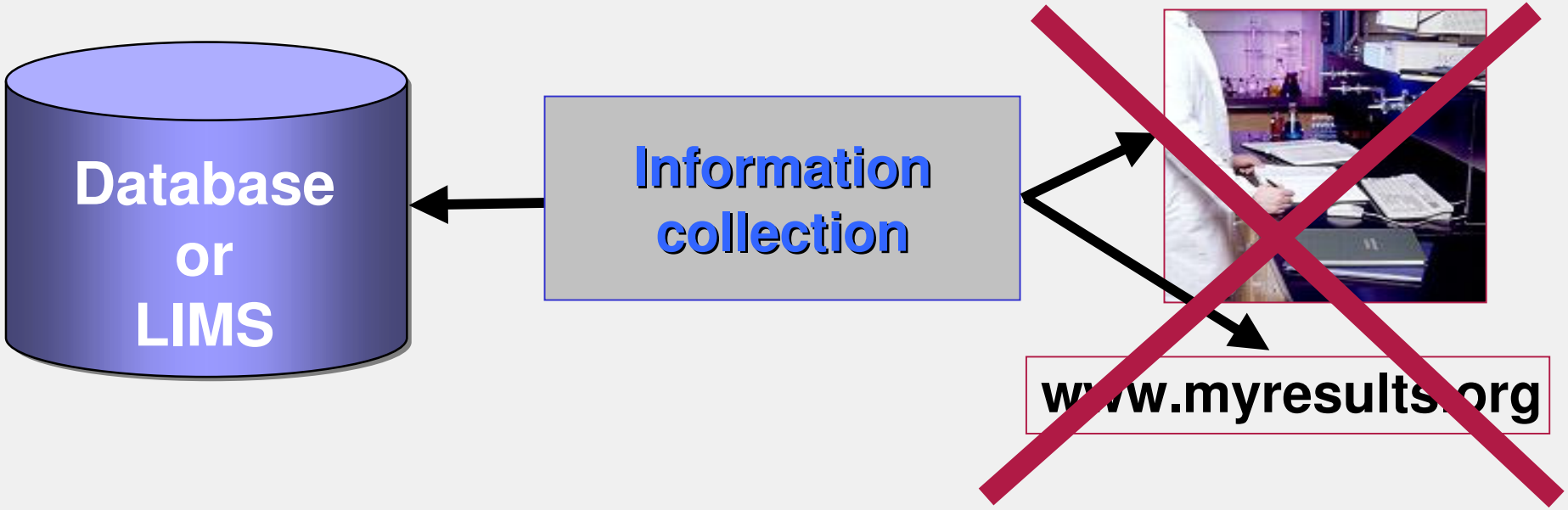
The Challenges

- Technical issues
  - Reproducibility, specificity, sensitivity and accuracy
  - International 'gold standards' for
    - Reference RNA
    - Genes, platform level annotation
    - Proteins, biomarker validation
    - Body fluids, target-surrogate relationships
    - Histopathology, images and pattern recognition
    - Reference algorithms, computational toxicogenomics
- Regulatory issues
  - Consensus on the application/interpretation of these data
    - Technical validation
    - Establishment of QC and QA systems
  - Electronic submission format
- Infrastructures required!

Priority – Establish Databases

"…Encouraging and empowering scientists to provide results in a structured and computable format alongside publication"

(Mark Boguski)



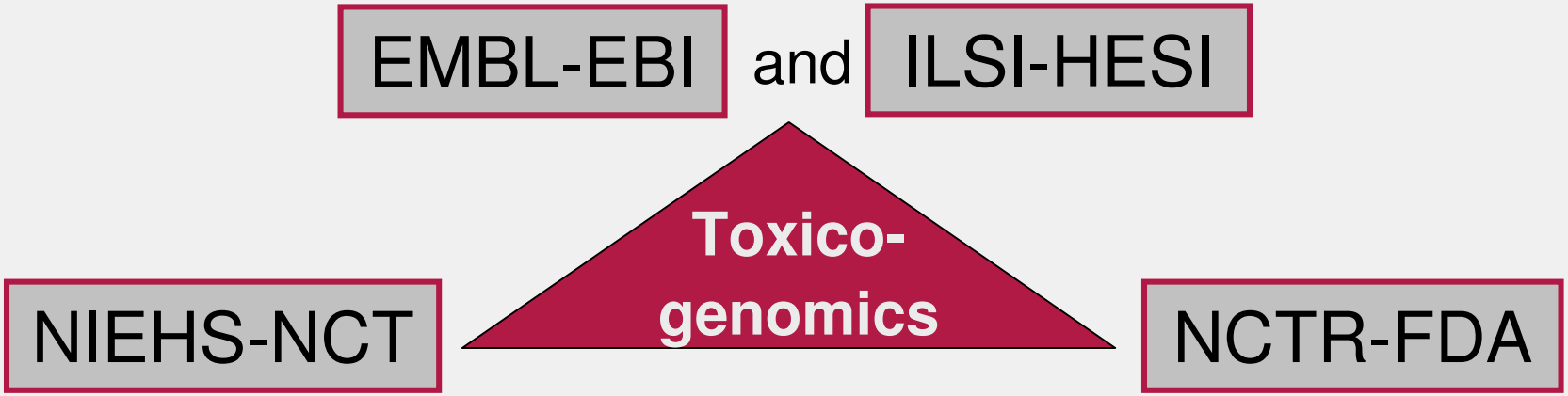
Database – The Potentials

- Centralization of the information
  - Providing a long term storage
  - Easy data access and data sharing
- Data comparisons
  - Giving critical mass to the datasets
  - Allowing validation of technology
- Integration the different domains
  - Genomics, conventional toxicology and experimental
- Annotation harmonization
  - Allowing curation
  - Improving annotation based on new computational predictions or experimental evidences

Database - The Challanges

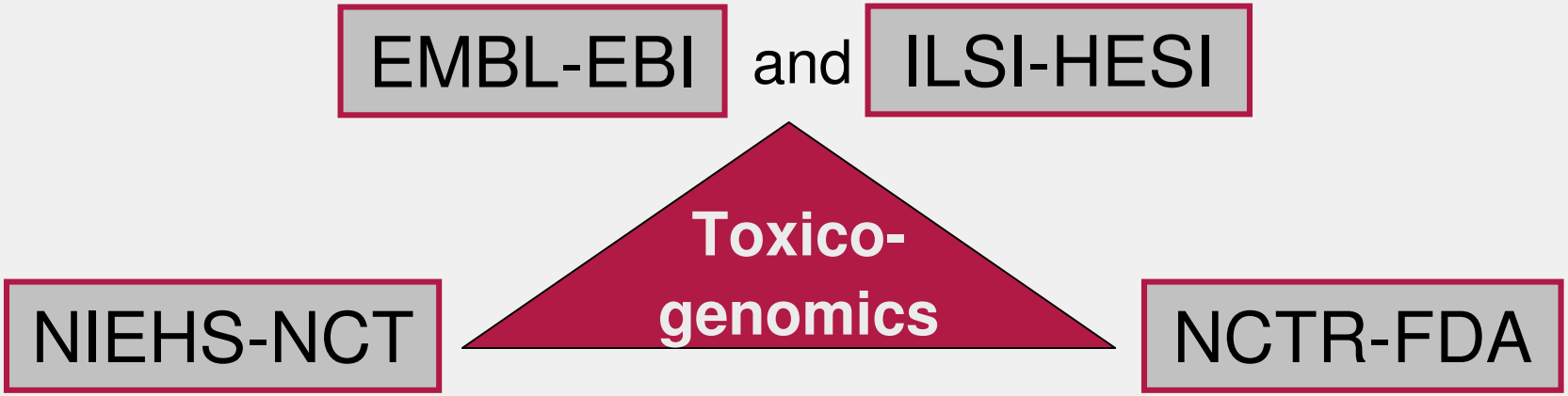
- Breadth, depth, and uniformity of the information
  - Information intensive field => define minimal descriptors
  - Lots of free text descriptions => use CVs
  - No lack of terminologies => promote harmonization
  - Semantic heterogeneity => develop ontologies
  - Heterogeneous formats => use existing standards
- Interoperability on an international scale to allow comparability
- Standards are essential !

An International Public Forum



- EMBL- EBI and International Life Sciences Institute (ILSI) Health and Environmental Sciences Institute (HESI)
- NIH NIEHS National Center for Toxicogenomics (NCT) and National Toxicology Program
- FDA National Center for Toxicological Research (NCTR) and Center for Toxicoinformatics

An International Public Forum



- Establishing a common, public infrastructure for toxicogenomics data on an international scale
- Towards harmonization
- But…. where start from?….

Talk Structure

- Toxicogenomics
  - Definition and objectives
  - Potentials, challenges and priority
- Towards harmonization of this new field
  - Required, ongoing and accomplished efforts
- Focus on array-based toxicogenomics experiments
  - Tox Working Group within MGED
  - MIAME/Tox and MIAME/Tox-compliant databases

Talk Structure

- Toxicogenomics
  - Definition and objectives
  - Potentials, challenges and priority
- Towards harmonization of this new field
  - Required, ongoing and accomplished efforts
- Focus on array-based toxicogenomics experiments
  - Tox Working Group within MGED
  - MIAME/Tox and MIAME/Tox-compliant databases

Toxicogenomics

Compound treatment

Euthanasia

Conventional toxicology
- Physical parameters
- Clinical Chemistry
- Pathology
- Histopathology

Target organ

RNA → Transcriptomics

Proteins → Proteomics

Blood /Urine → Metabonomics

Compound treatment

Conventional toxicology
-Cytotoxiciy
…etc.

Toxicogenomics

Conventional toxicology
- Physical parameters
- Clinical Chemistry
- Pathology
- Histopathology

Transcriptomics

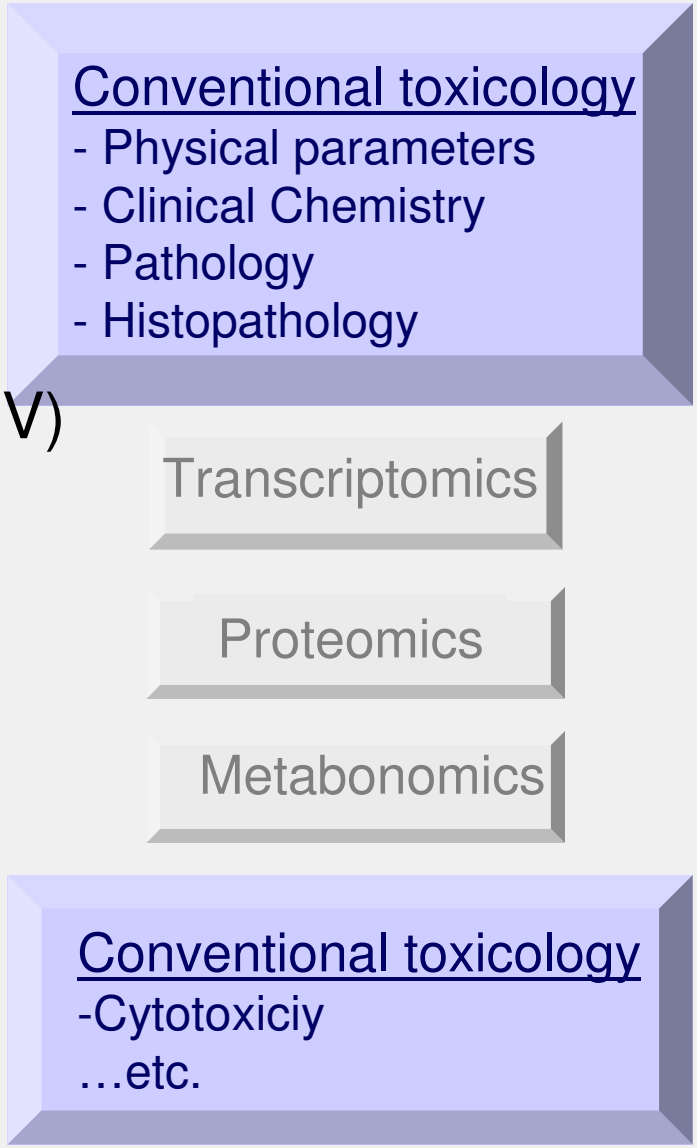
Proteomics

Metabonomics

Conventional toxicology
-Cytotoxiciy
…etc.

Conventional Toxicology

- Harmonization is required!
- No initiative is known at the present time
- Lots of controlled vocabularies (CV)
  - Public, e.g.:
    - NIEHS-NTP, LOINC, RENI
  - Proprietary, e.g.:
    - UMLS, SNOMED
- Problems:
  - Specie-specific CVs hamper cross species comparisons
  - Just CVs not ontologies!

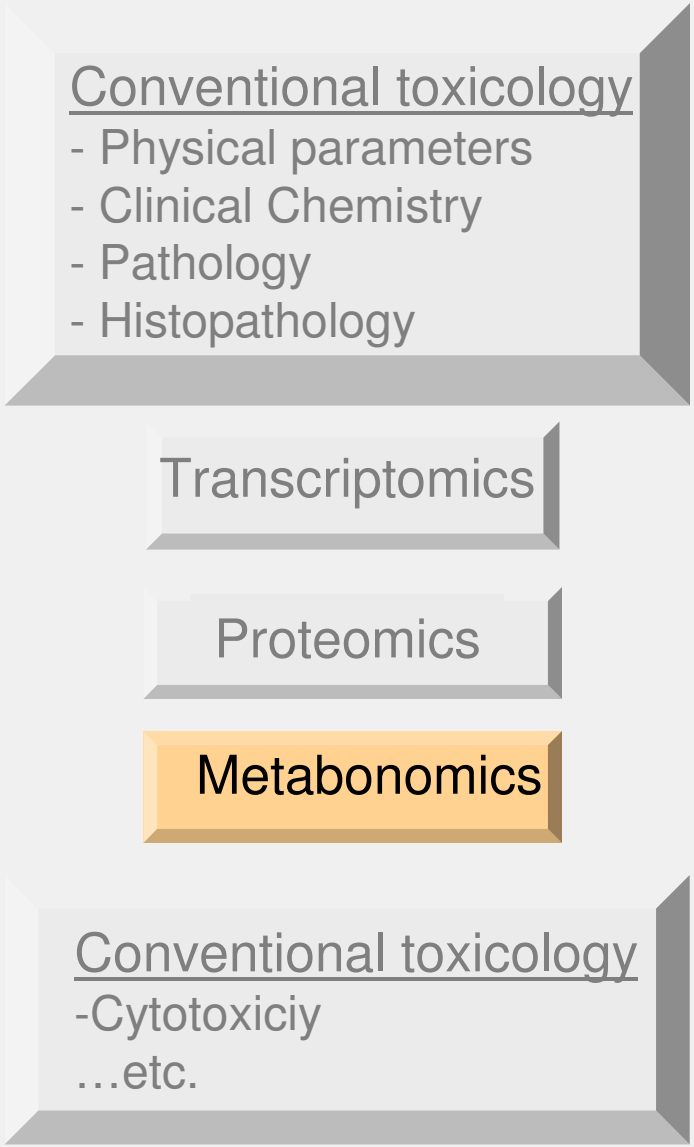


Metabonomics

- Starting…

- November 2003, London

first meeting to establish a harmonization steering committee for Metabonomic/Tox standards



Proteomics

- Ongoing…
- HUPO-PSI (**P**roteomics **S**tandards **I**nitiative)
Data exchange format in XML
  - PSI-PI Working Group
    - Protein interaction
    - IntAct, MINT, BIND, DIP…
  - PSI-MS Working Group
    - Mass Spectroscopy
  - Proteomics integration Working Group
    - PSI-MI+ PSI-MS+ 2D-Gel + Others
    - PEDRo Schema

Conventional toxicology
- Physical parameters
- Clinical Chemistry
- Pathology
- Histopathology

Transcriptomics

Proteomics

Metabonomics

Conventional toxicology
-Cytotoxiciy
…etc.

Transcriptomics

- Nearly accomplished!
- MGED Society
  - Standard for contextual information
    - MIAME Working Group
  - Standard for experiment annotation
    - MGED Ontology Working Group
  - Standard for recording controls and normalisation methods
    - Normalization Working Group
  - Standard for data model and exchange
    - MAGE Working Group
  - Standard for toxicogenomics
    - Tox Working Group

Conventional toxicology
- Physical parameters
- Clinical Chemistry
- Pathology
- Histopathology

Transcriptomics

Proteomics

Metabonomic

Conventional toxicology
-Cytotoxiciy
…etc.

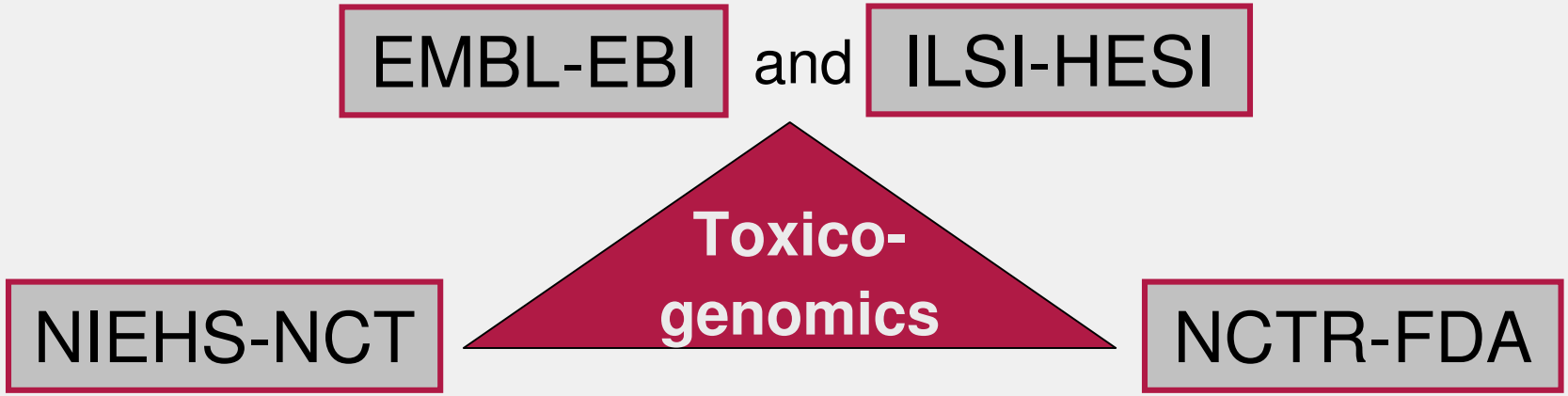
Talk Structure

- Toxicogenomics
  - Definition and objectives
  - Potentials, challenges and priority
- Towards harmonization of this new field
  - Required, ongoing and accomplished efforts
- Focus on array-based toxicogenomics experiments
  - Tox Working Group within MGED
  - MIAME/Tox and MIAME/Tox-compliant databases

Talk Structure

- Toxicogenomics
  - Definition and objectives
  - Potentials, challenges and priority
- Towards harmonization of this new field
  - Required, ongoing and accomplished efforts
- Focus on array-based toxicogenomics experiments
  - Tox Working Group within MGED
  - MIAME/Tox and MIAME/Tox-compliant databases

MGED Tox Working Group



Starting from array-based toxicogenomics experiments

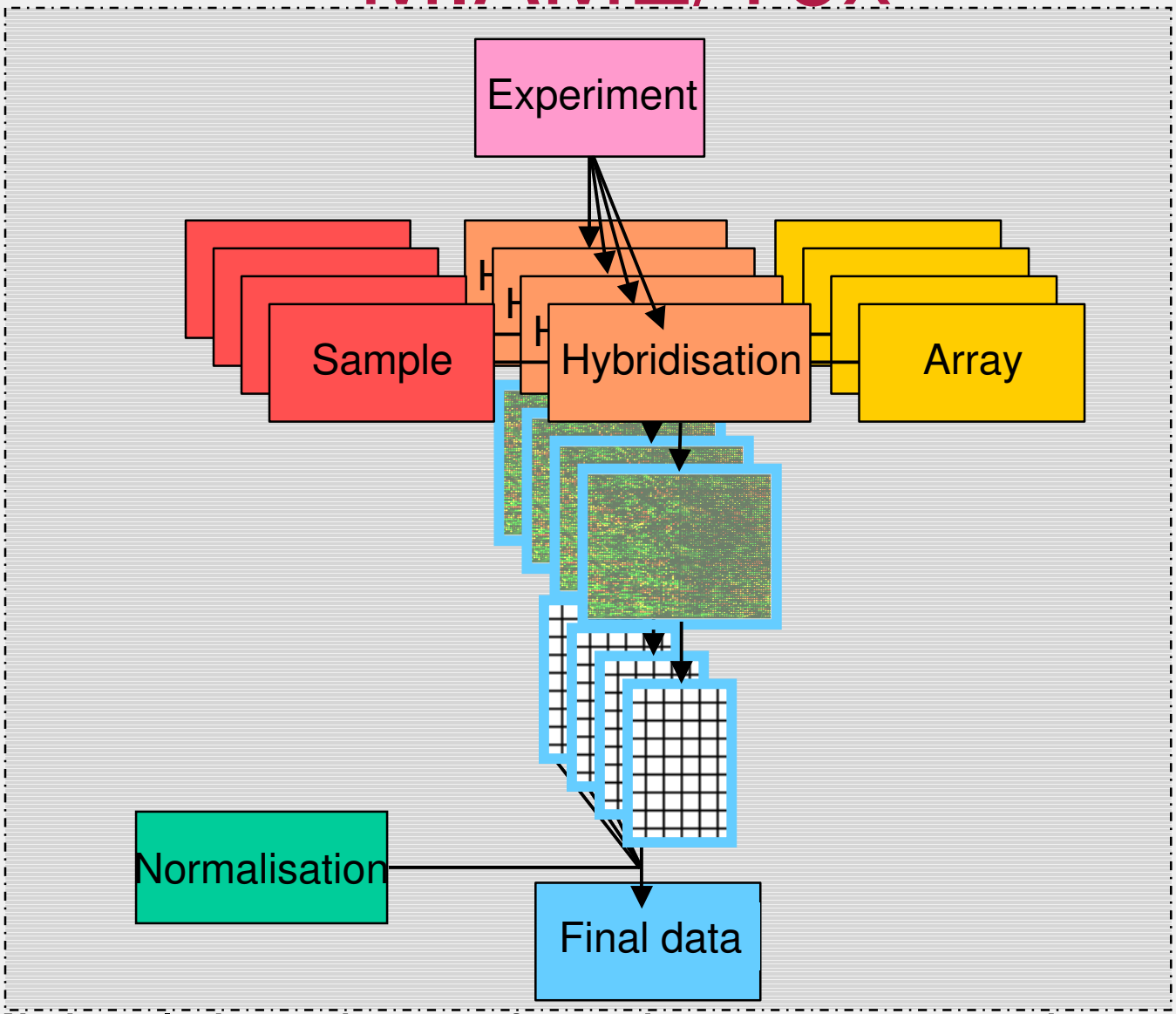
A MIAME for Toxicogenomics !

MIAME/Tox

MIAME/Tox objectives

- Standard data representation
  - Worldwide scientific consensus on the **minimal information descriptors** for toxicogenomics experiments
- Data harmonization for conventional toxicology
  - **Controlled vocabularies and ontologies**
- Data storage and data sharing
  - Use of **MAGE-OM** and **MAGE-ML** standards by MGED
    - Data management softwares
    - LIMSs (Laboratory Information Management Systems)
    - Public databases
- Common, public infrastructure for toxicogenomics data on an international scale

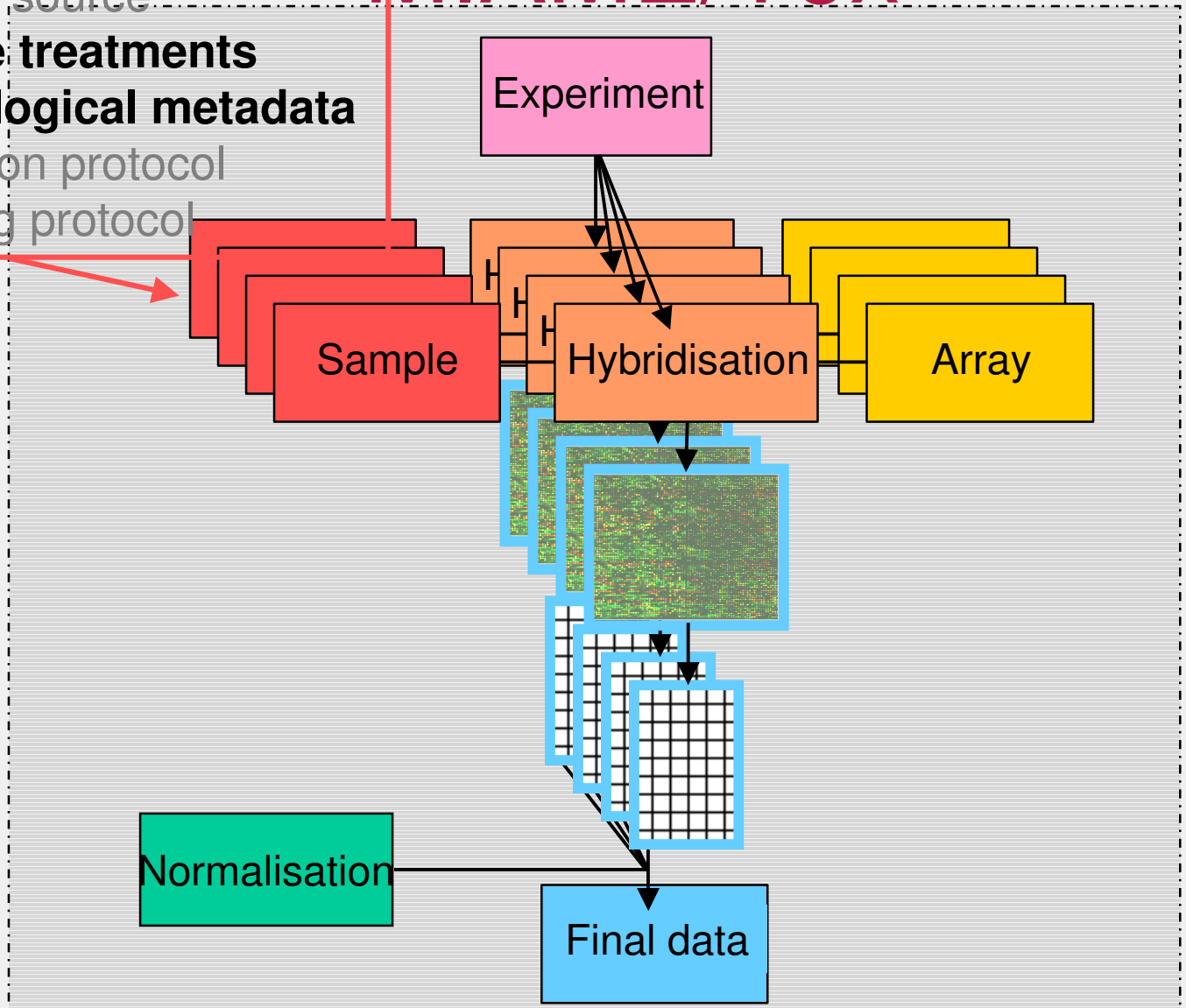
MIAME/Tox



Minimal descriptors for microarray experiments

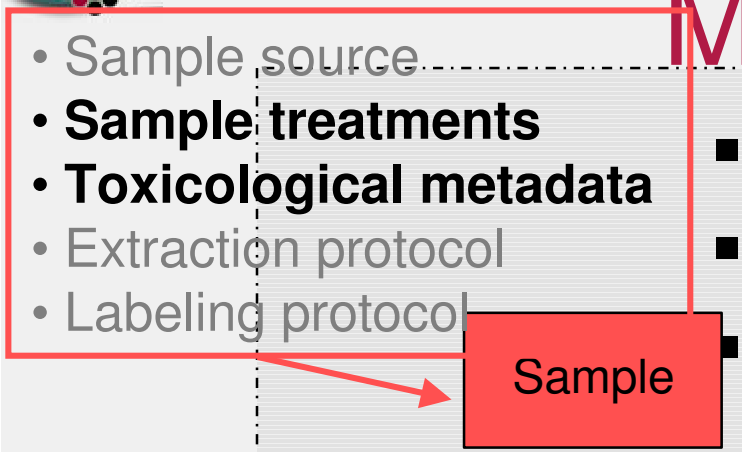
MIAME/Tox

- Sample source
- **Sample treatments**
- **Toxicological metadata**
- Extraction protocol
- Labeling protocol



Minimal descriptors for microarray experiments

MIAME/Tox



- MIAME/Tox document/checklist
- Minimal descriptors
- Source:

"Specifications for the conduct of the studies" (NIEHS-NTP)

- A 210 page document covering all aspects of NIH/NIEHS NTP toxicology studies
- NOT *minimal* descriptors BUT *maximal* descriptors !
  From facility requirements and animal care to data collection and storage……

MIAME/Tox

- Sample source
- **Sample treatments**
- **Toxicological metadata**
- Extraction protocol
- Labeling protocol

Sample

Sample treatments

- Facilities details
- Cell culture and growth conditions
- Animal husbandry/housing details
- Treatment type
  - Compound
  - CAS#, chemical structure/molecular formula
  - Vehicle for exposure
  - Exposure method
  - Duration
  - Dose (and unit)
- Date/time at death or at sacrifice
- Sacrifice method

MIAME/Tox

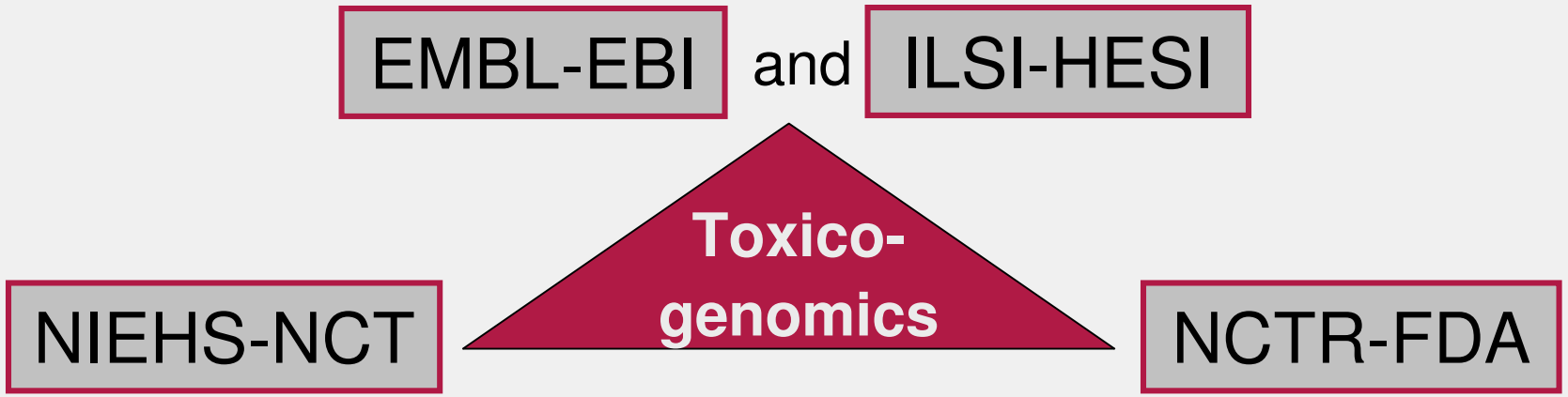
- Sample source
- **Sample treatments**
- **Toxicological metadata**
- Extraction protocol
- Labeling protocol

Sample

Toxicological assessments

- Clinical observations, e.g.
  - Weight, survival
  - Signs (e.g., general, behavior)
  - Lesions…etc.
- Gross necropsy examination, e.g.
  - Organs and tissues examination, weight lists…etc.
- Clinical pathology, e.g.
  - Hematology
  - Clinical chemistry
  - Other parameters Sperm morphology and vaginal Cytology evaluation (SMVCE), etc.
- Histopathology evaluation, e.g.
  - System
  - Organ
  - Sites
  - Morphology (s) and qualifier (s)

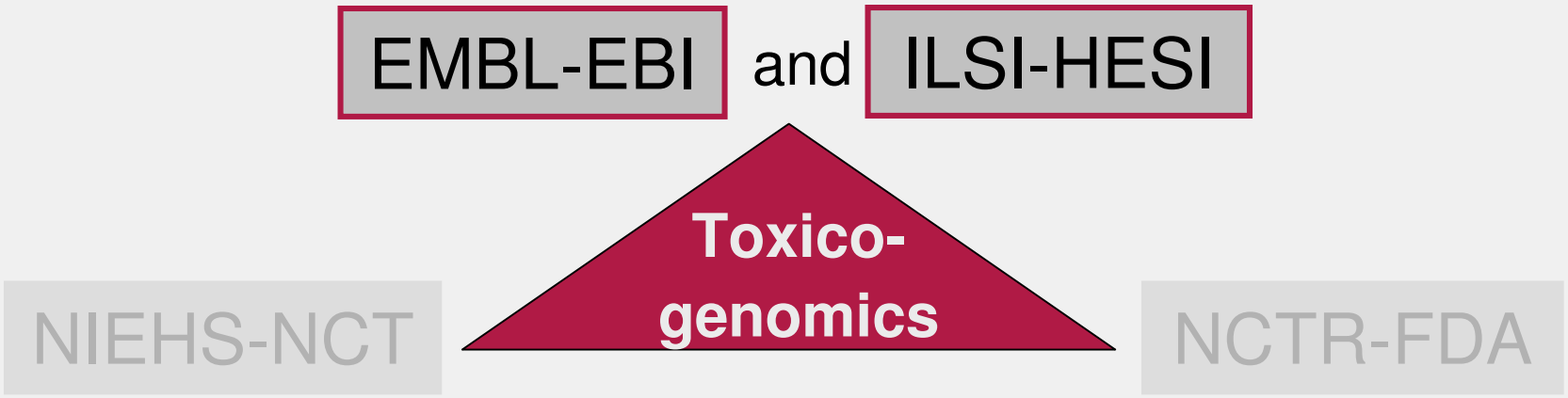
MIAME/Tox-compliant databases



Common, public infrastructure for toxicogenomics data on an international scale

…..Current status…..

MIAME/Tox-compliant databases



- ArrayExpress public repository for array-based data
  - MGED standards-supportive, and MIAME/Tox-compliant
  - Curation Team
  - Supporting data in publications (acc. num., reviewers access)
- Not an archive but a dynamic environment !
  - Mining agent to update genome annotation
    - DNA, protein and functional level

EMBL-EBI Toxicogenomics MIAME/Tox

Pathology Observations data sheet

	A	B	C	D	E	F	G	H	I	J
1										
2	Sample identifier	Body weight (g)	Brain weight (mg)	Target Organ	Organ weight (mg)	Site	Morphology	qualifier	value	qualifier
3	S10-[GSK]	150	2500	Liver	7850	Left Lat Lobe	Necrosis	Distribution	_	Duration
4	S10-[GSK]	150	2500	Kidney	1950	_	_	_	_	_
5				Liver		_	_	_	_	_
6				Kidney		_	_	_	_	_
7				Liver		_	_	_	_	_
8				Kidney		_	_	_	_	_
9				Liver		_	_	_	_	_
10				Kidney		_	_	_	_	_
11				Liver		_	_	_	_	_
12				Kidney		_	_	_	_	_
13				Liver		_	_	_	_	_
14				Kidney						

Accessory Spleen
Accumulation
Adenocarcinoma
Adenolipoma
Adenoma
Adhesion
Alveolar/Bronchiolar Carcinoma

HELP Toxicological endpoints data

Clinical Observation | Clinical Path

Pathology Observations: | Download

Extracts

Extracts (create/view/edit)

Leukocytes 10E9/l

Macroc +,clotted

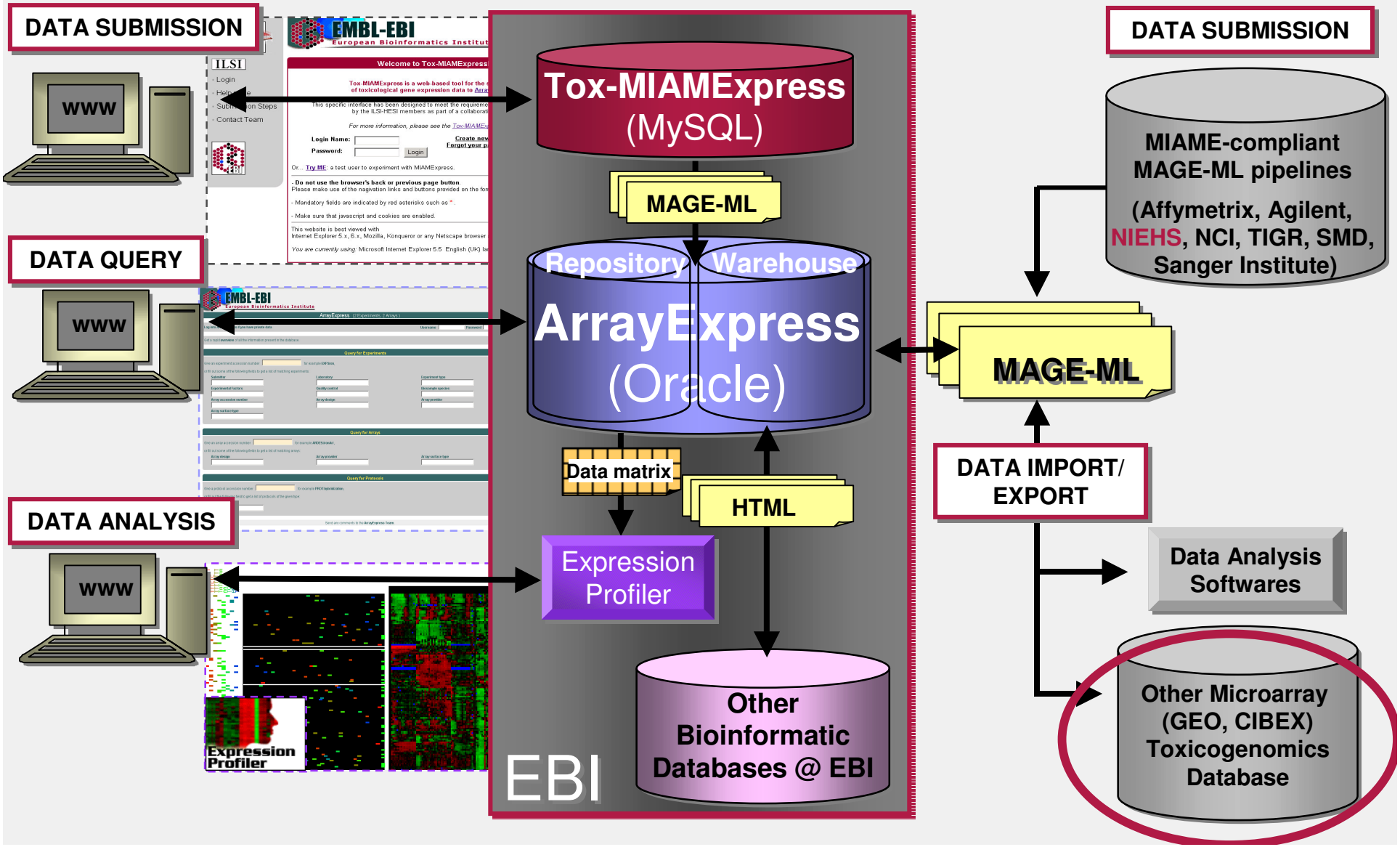
Mean ce pmol/l

Mean cell haem g/dl

Controlled vocabulary from NIH/NIEHS-NTP Pathology Code Tables (PCT)

EMBL-EBI Toxicogenomics infrastructures @ EBI

MIAME/Tox



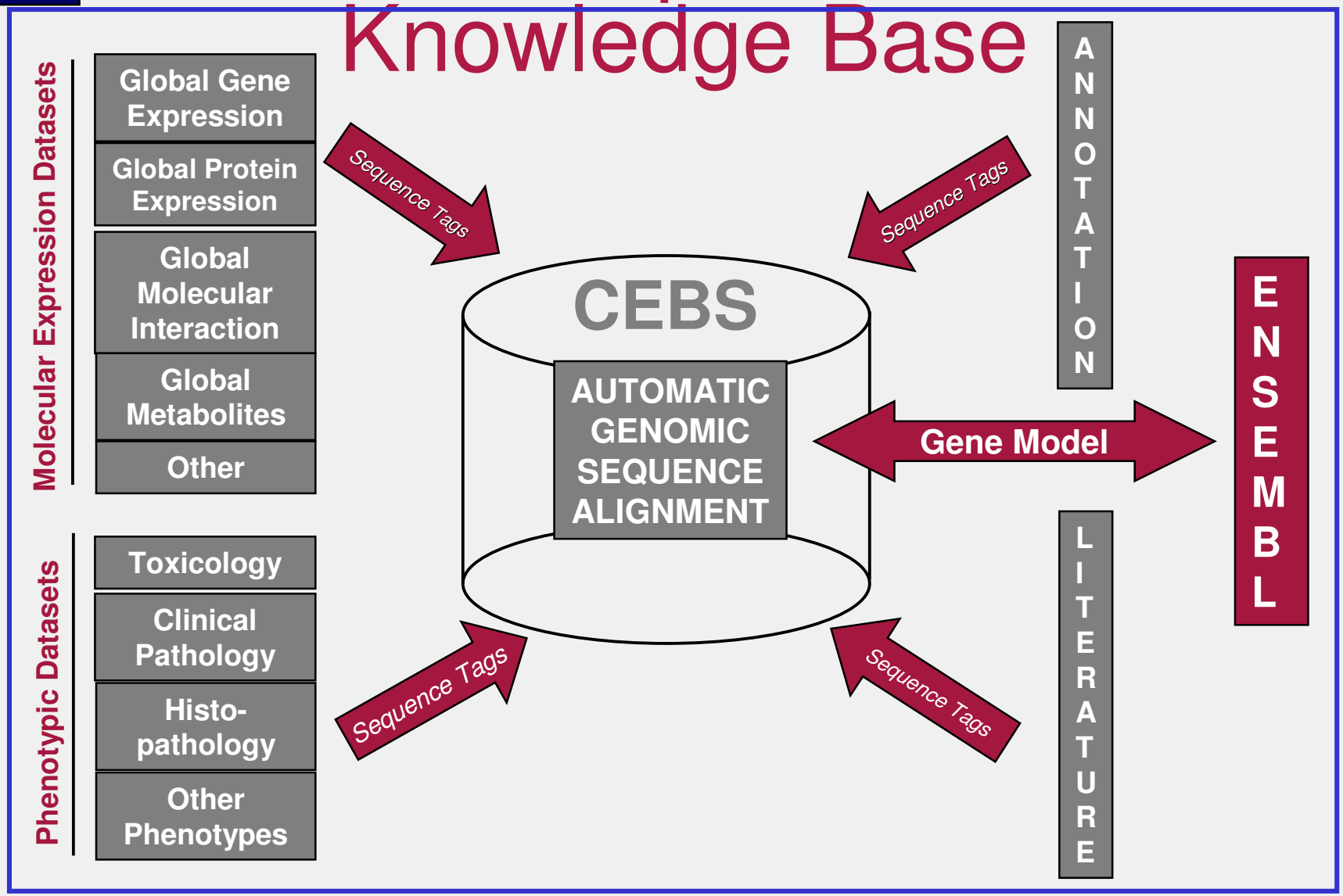
MIAME/Tox-compliant databases

EMBL-EBI and ILSI-HESI

- **C**hemical **E**ffects in **B**iological **S**ystems (**CEBS**) Knowledge base
  - MGED standards-supportive, and MIAME/Tox-compliant
- Reference toxicogenomic information system
  - Studies on environmental chemicals/ stressors and their effects
- Relational and descriptive data compendia
  - Toxicologically important genes, SNPs, mutants, and biological phenotypes
- Hypothesis-driven and discovery research
  - Environmental toxicology and risk assessment

MIAME/Tox

CEBS A Sequence-Driven Knowledge Base



PATENT PENDING: LVM 402604, DHHS E-026-2003/0-US-01

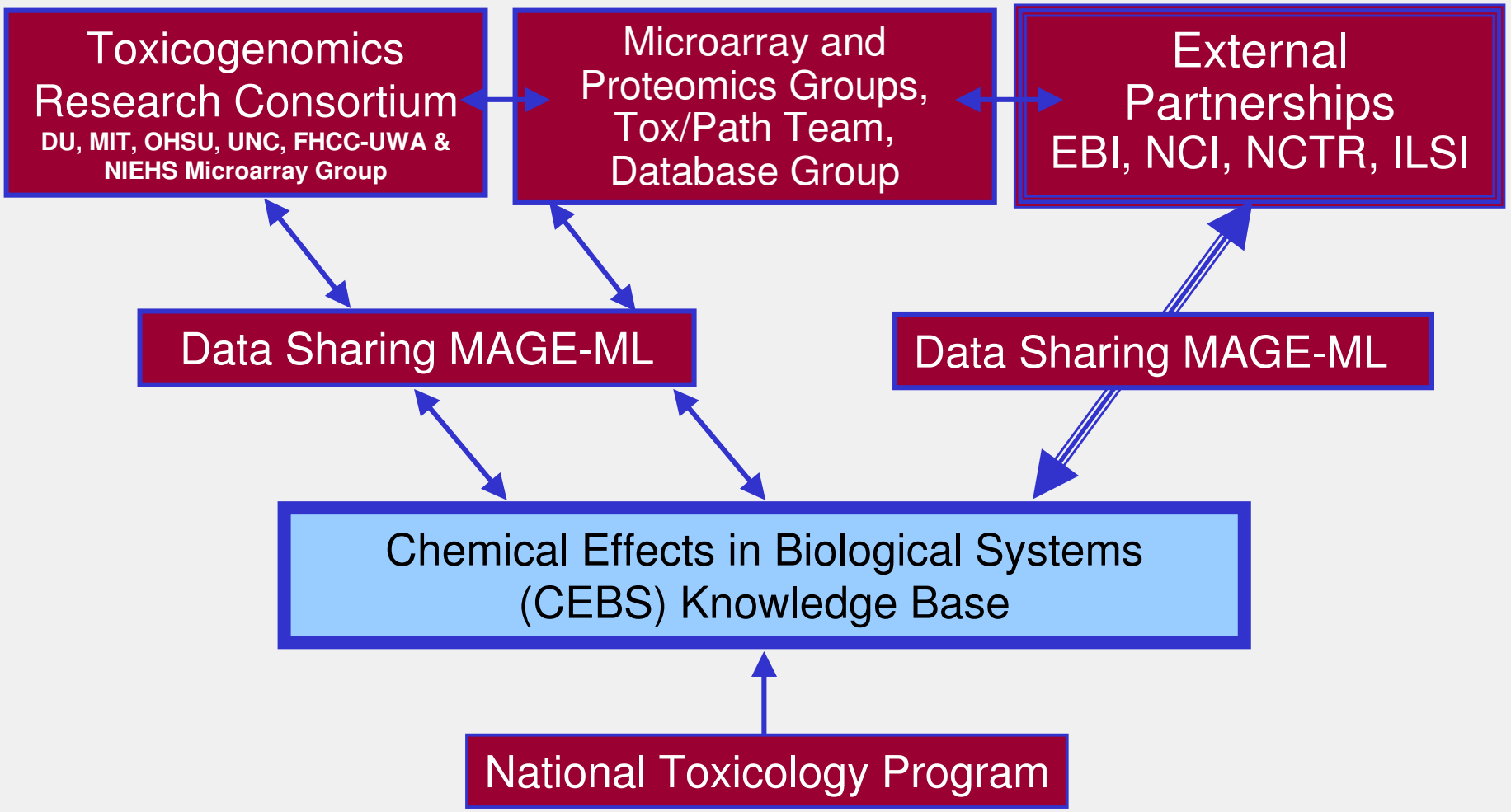
METHOD AND SYSTEM FOR DEVELOPING AND QUERYING A SEQUENCE DRIVEN CONTEXTUAL KNOWLEDGE BASE

INVENTORS: M. WATERS, J. SELKIRK, R. TENNANT

MIAME/Tox

CEBS-Content

Intramural and Extramural Partnerships



MIAME/Tox-compliant databases

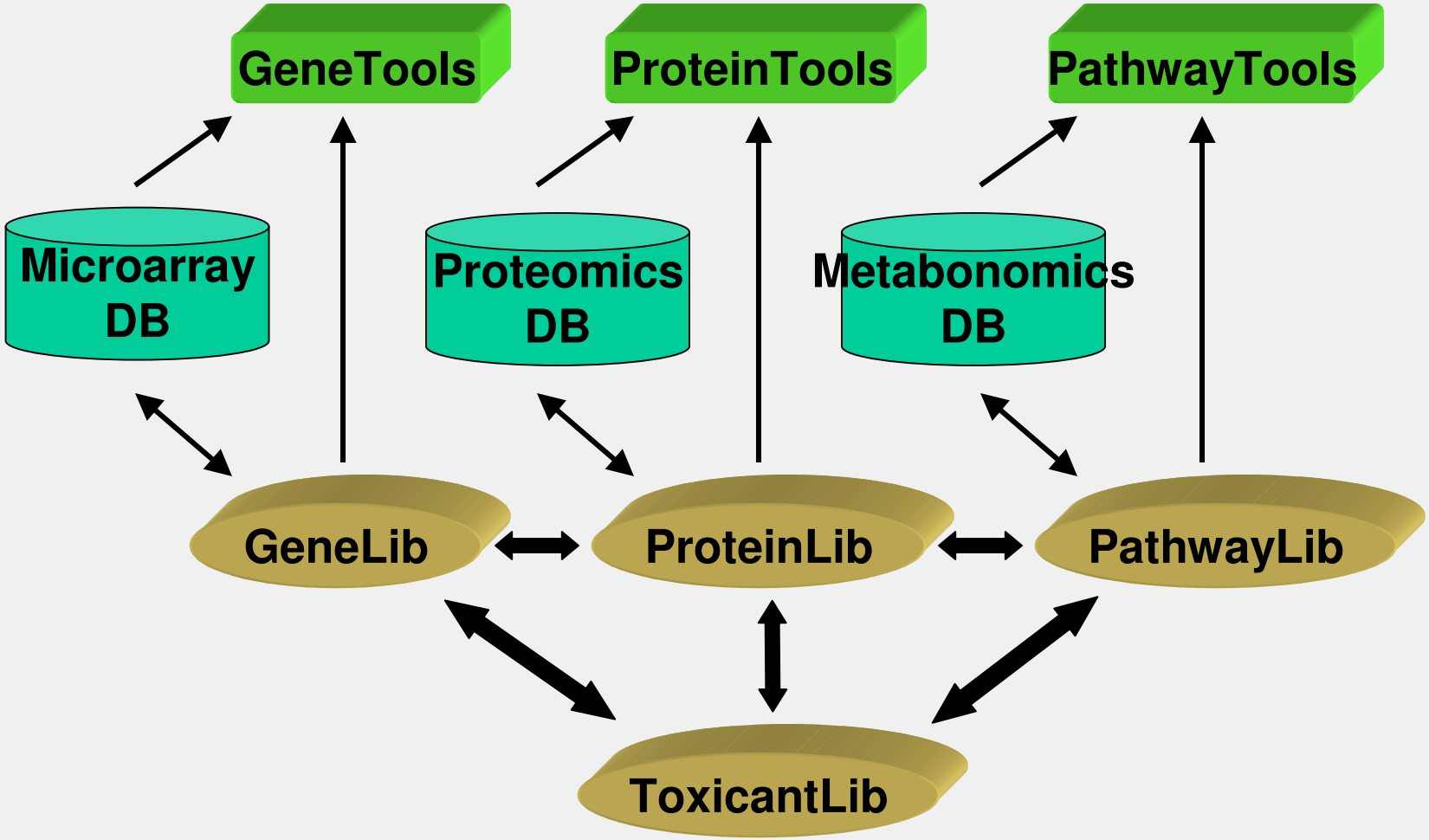
EMBL-EBI and ILSI-HESI

NIEHS-NCT

NCTR-FDA

- Toxicoinformatics Integrated System (**TIS**)
  - MIAME/Tox-compliant, will be MGED standards supportive
  - NCTR and FDA researches
  - Local installation
- Phenotypically anchor the –omics data and chemical structure information
- ArrayTrack first module for microarray data

Toxicoinformatics Integrated System (TIS)



MIAME/Tox progressing

- Presented and circulate for consensus
  - **MGED Tox Working Group**
    - Includes Toxicogenomics, Pharmacogenomics, Ecotoxicogenomics communities (*MIAME/Env*) and Nutrigenomics (*MIAME/Nut*)
  - **Journals**
    - *Nature*, the *Nature* group of journals, *Cell*, *The Lancet*, *EMBO*, *Toxicology Pathology and EHP* require MIAME-compliant information
  - **ECVAM-ICCVAM-NICEATM**
    - Committee on 'Validation Principles And Approaches For Toxicogenomics-Based Test Systems', December 2003
  - **Society of Toxicology**
    - Database and Standards Symposia, March 2004
  - **FDA Pharmacology/Toxicology Advisory Committee**
    - Microarray Database Projects as first step in guidance process, June 2003

Guidance for Industry

Pharmacogenomic Data Submissions

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact (CDER) Lawrence Lesko 301-594-5690, (CBER) Raj Puri 301-827-0471, or (CDRH) Steve Gutman 301-594-3084.

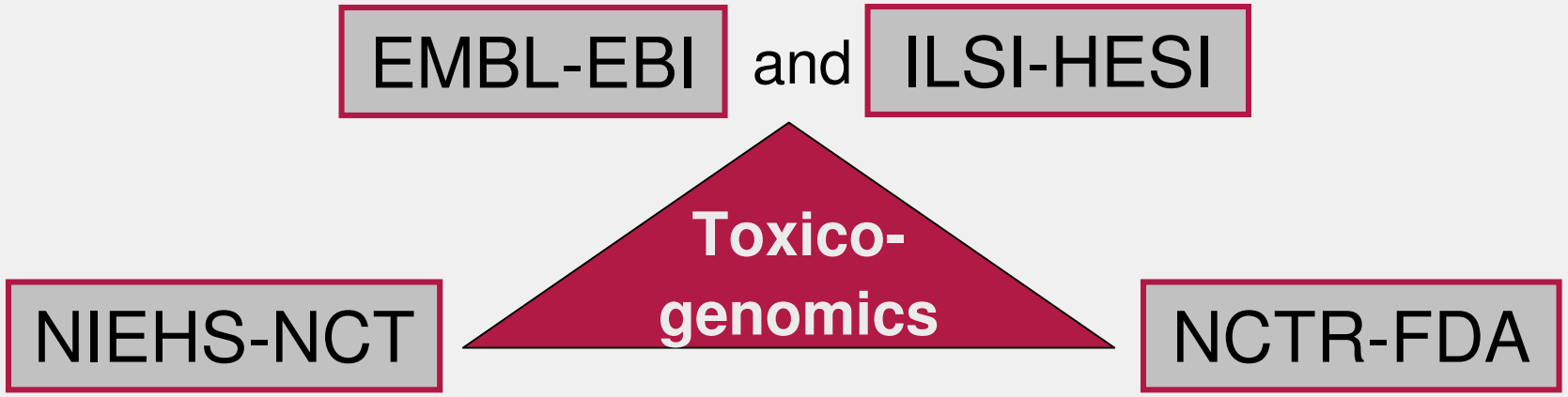
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Center for Devices and Radiological Health (CDRH)

November 2003
Procedural

FDA Pharmacology/Toxicology Advisory Committee

http://www.fda.gov/cder/guidance/5900dft.doc

Acknowledgments



- EMBL-EBI
  - Alvis Brazma
  - Microarray Informatics Team
- NCTR-FDA
  - Dan Casciano (Director)
  - **Weida Tong**
- MGED Society
  - Working Groups
- NIH-NIEHS NCT and NTP
  - Ray Tennant (Director)
  - **Mike Waters**
  - Pierre Bushel
  - Jennifer Fostel
- ILSI-HESI Genomics Committee
  - Syril Pettit
  - Bill Mattes

Resources

www.mged.org

- MGED Society and Working Groups and mailing lists

mged-toxico@lists.sourceforge.net

- Toxicogenomic Working Group mailing list

www.niehs.nih.gov/nct

- CEBS Knowledge base

www.ebi.ac.uk/microarray

- ArrayExpress and Tox-MIAMExpress schema-access to code

www.nctr.fda.gov

- ArrayTrack and TIS