# ILSI/HESI Genomics Committee: Overcoming Challenges to Share Data on Non-Proprietary Compounds

**Current Committee leadership:**

Syril Petit, HESI

Cindy Afshari, Amgen

George Orphanides, Syngenta

**Database Steering Committee:**

Jennifer Fostel, LMIT, NCT

Mike Lawton, Pfizer

# ILSI / HESI Genomics Committee

- ILSI = International Life Sciences Institute
- HESI = Health and Environmental Sciences Institute

- Genomics Committee = Technical Committee on the Application of Genomics to Mechanism-Based Risk Assessment

# ILSI / HESI Genomics Committee

- The mission of HESI's Technical Committee on Application of Genomics to Mechanism-Based Risk Assessment is to advance the scientific basis for the development and application of genomic and proteomic methodologies to risk assessment.

# Long-term objectives:

- Provide a scientific forum to facilitate further development and advances in genomics and proteomics methodologies to increase the utility of gene expression data for mechanism-based risk assessment
- Facilitate a broad-based, multinational collaboration involving academic, governmental and industrial institutions
- Develop public information of interest to the specialist user and the scientific community at large – to include details of experimental design, materials and methods, and study results
- Contribute to the greater understanding of toxic mechanisms and identification of appropriate biomarkers so that the toxicity observed in animals and the extrapolation of these observations to humans results in meaningful risk assessment
- Contribute to the development of publicly available databases, to include details of experimental design, methods, and study results

# Technical Challenges

- Genomics Committee formed in 1999
- Technical challenges at that time:
  - No public databases available
  - Technology not validated; multiple array platforms
  - Lack of standard tools, methods, study designs
  - Lack of robust analysis tools
  - Lack of consensus on interpretation; relevance of single gene change
  - Uncertain regulatory applications, environment

# Scientific Questions:

- Can known pathways and mechanisms of toxicity can be associated with gene expression profiles?
- Can this information be used for risk assessment?
  - develop data using well-characterized toxicants with well-known mechanism of toxicity under controlled conditions
  - evaluate gene expression changes in the context of other toxicity measures
  - use the information obtained to identify potential biomarkers and interspecies extrapolation issues
  - help to build consensus on optimal approaches for interpretation of gene expression data and its application to risk assessment

# Four Working Groups

- Hepatotoxicity:
  - 2 agents, 2 doses, 3 times, 2 in-life studies
  - 14 institutions, 7 array platforms
- Nephrotoxicity:
  - cisplatin, gentamicin, puramicin, 2 doses, 3 times
  - 8 institutions, 6 array platforms
- Genetic Toxicity:
  - 14 genotoxicants, 5 array platforms
- Database Steering Committee
  - oversee contract with EBI to extend ArrayExpress

# Data “ownership” not an issue

- Committee members agreed that the data and tissue / RNA samples from the in-life studies were to be shared with all participating members of the Committee.
- Individual labs “owned” the data they generated but committed to sharing outcomes with the Committee, and to publishing data jointly.
- All data were to be made publicly available per HESI Committee requirements
- Some microarray sequence probes were considered to be proprietary.
  - these reporters were not identified to the public

# Data transfer became an issue

- Development of database infrastructure followed generation of data, often with significant time lag
- Transfer of data (originally) was a very manual process
- Committee members were new to large-scale data exchange / management; some inefficiencies surfaced as a result

# Current status

- Tox/MIAMExpress, web interface to EBI curation database, used to load data from several thousand hybs
- Tox/MIAMExpress captures treatment information, clinical chemistry and histopathology observations about the sample.

# Data sharing, current projects

- Gathering data from current members for retrospective analysis
- How to manage donor identity
  - what if “my” data are the outlier?
- How to facilitate data transfer
  - if it’s not easy as pie I don’t have time
- How to maximize benefit
  - What does sharing data do for me?

# Managing donor identity

- Anonymize the donors to the public.
  - only identify the institutions within the Committee
- How to provide useful information that will not spill the beans?
  - If company X is the only one using array type Y, then array type will disclose the ID

# Facilitating data transfer

- Make it rapid
  - close in time to the data generation / sharing (have the DB ready)
  - make it electronic as much as possible
- Make it conform to standards
  - controlled vocabularies commonly used within the donor community is a bonus
  - use a standard form for data submission

# What is the benefit to me?

- Support the public good
- Understand “how I stack up”
- Address questions I couldn’t answer alone
- Provide a standard analysis platform, comparator data
- Provide useful, validated analysis tools

# Acknowledgements.....


- HESI
  - Syril Petit
  - Candace Hamilton
- Genomics Committee
  - Cindy Afshari, George Orphanides (current co-chairs)
  - Bill Pennie, Pete Lord (previous co-chairs)
  - Mike Lawton (co-chair of Database Working Group)
- EBI (European Bioinformatics Institute)
  - Susanna-Asunta Sansone
  - Philippe Rocca-Serra