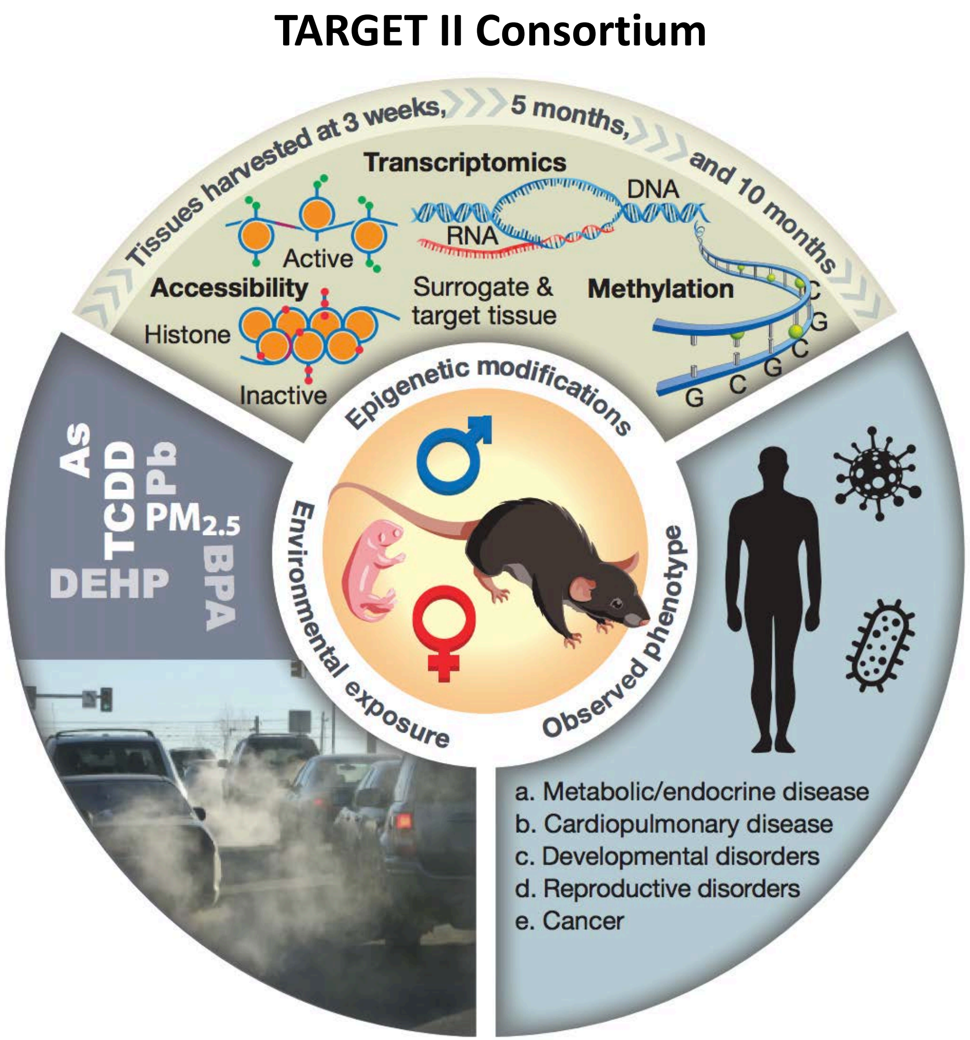
NIEHS U01 TaRGET II CONSORTIUM



- The **NIEHS Toxicant Exposures and Responses by Genomic and Epigenomic Regulators of Transcription (TaRGET)** Program established in 2012 with the goal to increase our understanding of how exposures affect and interact with functional and regulatory processes that cause persistent epigenetic alterations.
- In 2016, **TaRGET II established a mouse consortium** focusing on surrogate (e.g. blood, skin) and target (e.g. liver, brain) tissue analyses of epigenomic signatures across the life course after developmental exposure to environmental toxicants.

DRIVERS FOR TaRGET II INITIATIVE

- **Biomarkers** in human population-based studies are limited to easily obtainable tissues (hair, blood, and saliva).
- It is currently unknown if epigenetic alterations induced in disease-relevant, but often inaccessible target tissues will be reflected in **correlative changes in surrogate tissues.**
- Also unknown is whether toxicant-induced **changes in the epigenome persist** in target or surrogate tissues after exposure cessation and/or change over the life-course.
- Finally, it is increasingly evident that the effects of exposures are highly **sex specific** and influenced by ill-defined **inter-individual differences**, adding another layer of complexity to the interpretation of population-based studies

To fill these knowledge gaps, the NIEHS TaRGET II Consortium is investigating the conservation of toxicant-induced epigenomic changes across tissues and time, in both males and females, in response to a variety of developmental environmental exposures

Organizational Structure TaRGET II Consortium

- Consortium made up of **7 institutions** profiling epigenomic response to **8 toxicants**
- **Data Coordinating Center (DCC)** to which all transcriptomic, epigenomic and meta-data are uploaded into a database for analysis by the DCC, Consortium members, and ultimately non-consortium researchers

NIEHS U01 TaRGET II CONSORTIUM

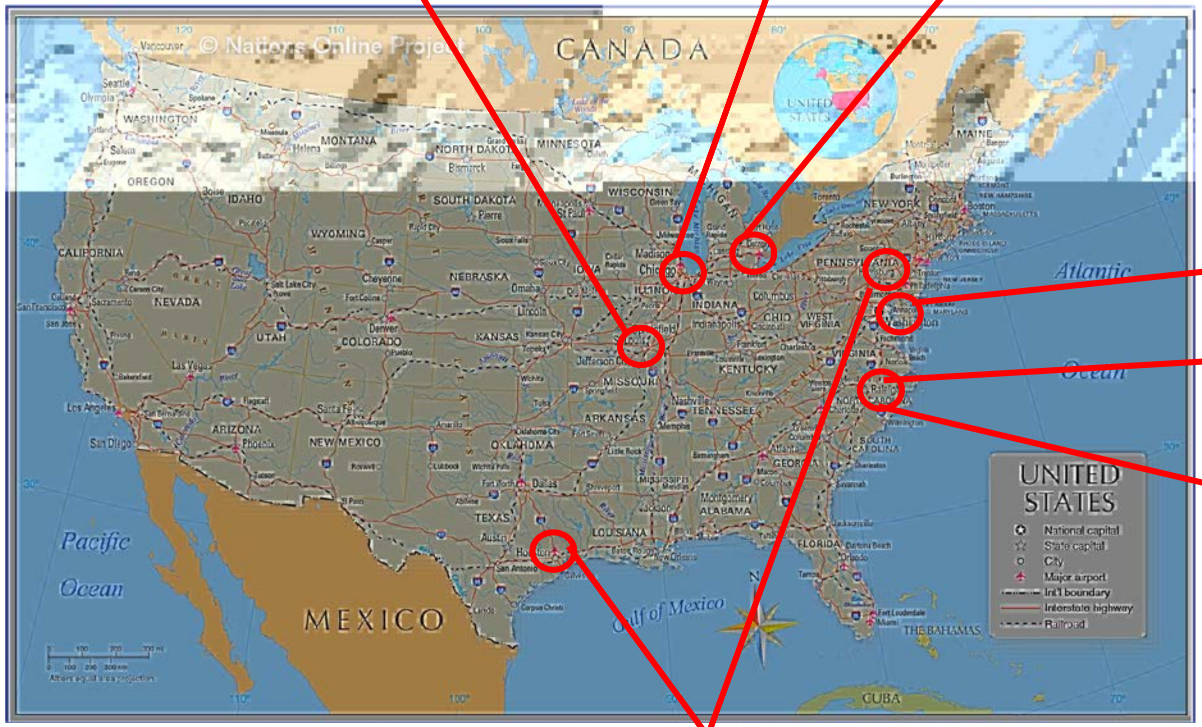
Data Coordinating Center
Wash U
(Wang)

U. Chicago
(Mutlu)

PM2.5

U. Michigan
(Dolinoy)

DEHP, Lead



Johns Hopkins and U. Maryland
(Biswal and Rajagopalan)

PM2.5, As

North Carolina State University
(Aylor)

TCDD

NIEHS RTP North Carolina
(Tyson, Duncan, Chadwick, McAllister)

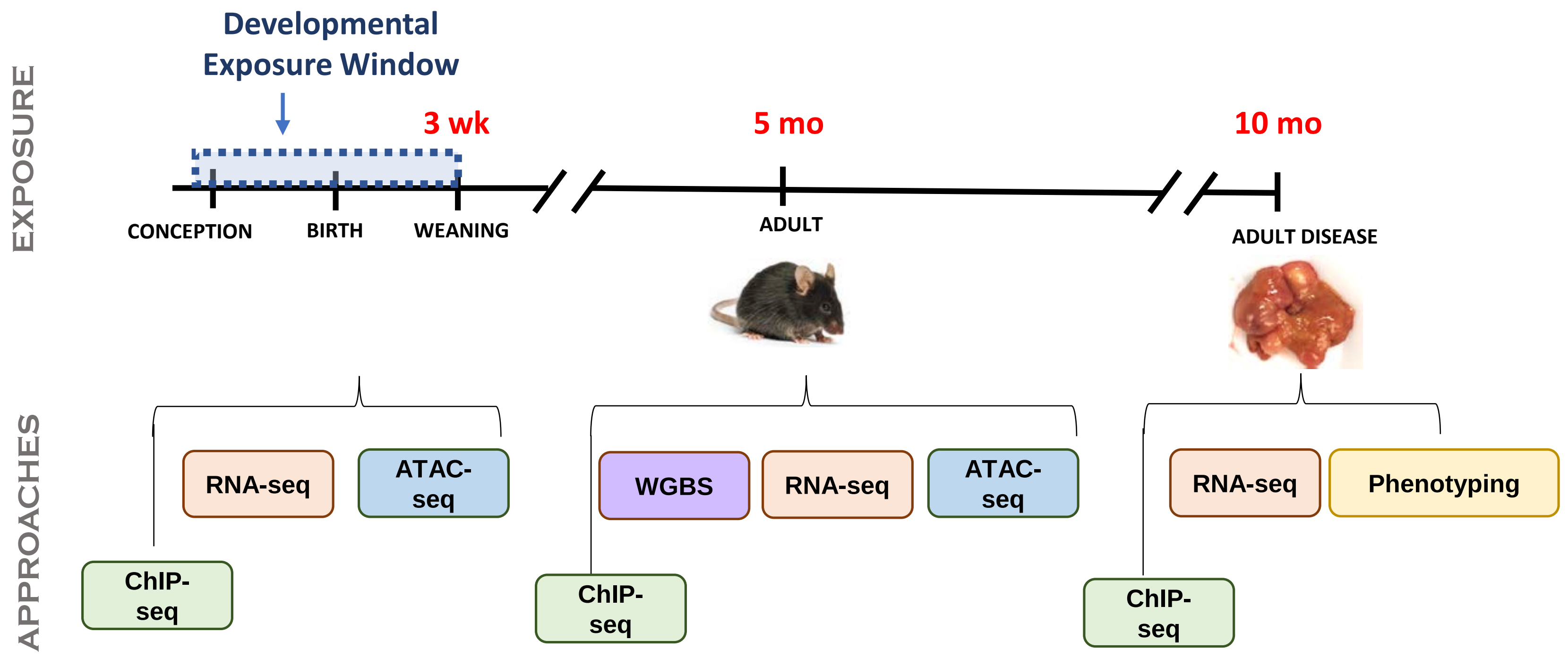
BPA, TBT

Baylor College of Medicine and U. Penn
(Walker, Coarfa & Bartolomei)

Organizational Structure TaRGET II Consortium

- Consortium made up of 7 institutions profiling epigenomic response to 8 toxicants
- Data Coordinating Center to which all transcriptomic, epigenomic and meta-data are uploaded into a database for analysis by the DCC, Consortium members, and ultimately non-consortium researchers
- **T2C Wiki** developed for distribution of consortium information
- Working groups with regularly scheduled meetings established for:
  - **T2C Steering Committee** made up of U01 PIs and NIEHS Program Staff
  - **T2C Bioinformatics** group led by DCC with broad consortium participation
  - **T2C Methods Development** led by U. Mich team (Dolinoy) with broad consortium participation
  - **T2C Manuscript publication guidelines** led by NC State team (Aylor) with broad consortium participation
  - **Scientific Advisory Board** formed and convened
- Annual “all hands” meetings held until COVID pandemic, transitioned to virtual format (not optimal)
- Developing a T2C consortium “package” of manuscripts to be published concurrently in 2022

Adoption of T2C Consortium-wide Exposure Paradigm and Profiling Approaches



QUESTIONS

- What is the direct impact on the developing epigenome and transcriptome?
- How do target (liver) and surrogate (blood) tissues respond to exposure?
- Which epigenetic alterations persist and what is the impact of reprogramming on the transcriptome?
- Can we identify correlative exposure/risk signatures in surrogate tissues?
- What genes and pathways drive pathogenesis downstream of epigenomic reprogramming?

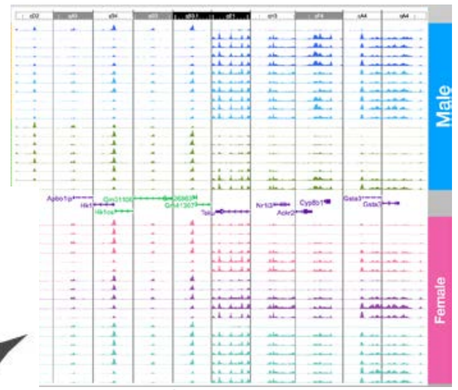
- Use of standardized mouse exposure paradigm across the consortium allowed for **consistency** and the use of **controlled exposures**.
- Use of **same timepoints** (3 weeks, 5 and 10 months) allowed for comparison across exposures and consortium sites
- Developed and deployed **standardized and coordinated methodological approaches** for comparison across exposures and consortium sites
- Developed and shared detailed standardized **metadata schema via the DCC**

Data Coordinating Center Research Activities

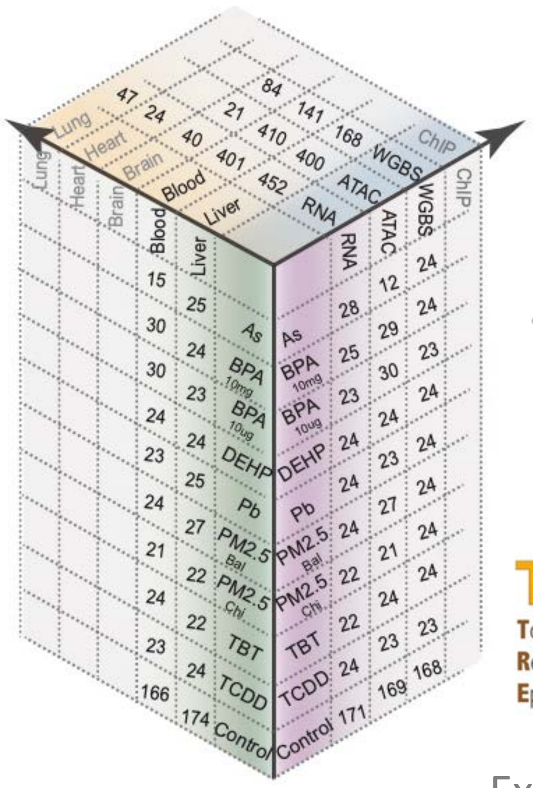
1. **Data curation** of over 2100 TaRGET II omics datasets.
    - Omics data processing, formatting, and management
    - Transparent data sharing through research community
        - TaRGET Data Portal (https://dcc.targetepigenomics.org/)
        - AWS open data (https://registry.opendata.aws/targetepigenomics/)
2. **Development of bioinformatics tools/pipelines** for omics data QC and processing.
    - Construction of version-controlled omics data QC pipeline
    - Development of novel bioinformatics tools to facilitate QC processing
        - AIAP: ATAC-seq Integrative Analysis Package (https://doi.org/10.1016/j.gpb.2020.06.025)
        - BeCorrect: Creates batch corrected visualization file (https://doi.org/10.1038/s41598-020-66998-4)
3. **Development of novel statistical and bioinformatic methods** to analyze omics data.
    - Modified TMM normalization method to enable consortium-wide data normalization
    - Novel statistical framework to analyze WGBS data with large replicates
4. **Identification of signatures** for specific toxicant exposures.
    - Discovering epigenomic signatures corresponding to distinct toxicant exposures
    - Exploring the commonality of toxicant exposures at pathway level
    - Detecting the cross-talk between transcriptome, epigenome landscape and long-term epigenetic memory
    - Understanding the common signatures between target and surrogate tissue as a function of exposure
5. **Creation of database** of epigenomic signatures corresponding to toxicant exposures.

Examples: Data Coordinating Center Activities

WashU Epigenome Browser

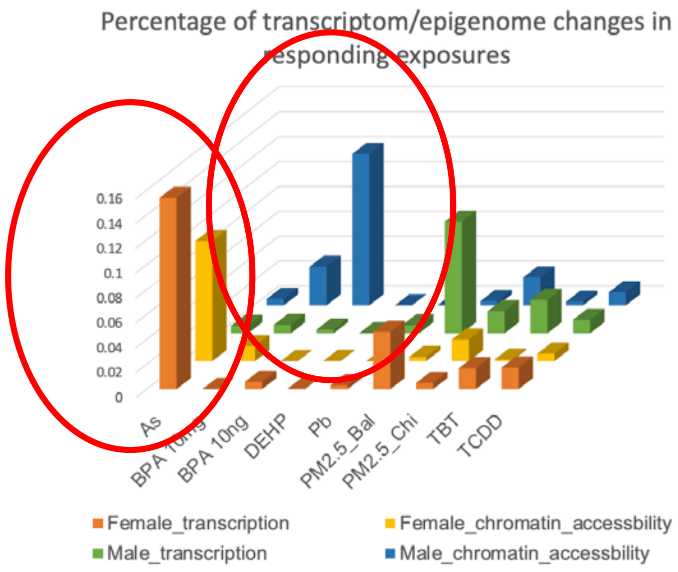


TaRGET Data Portal



Transcriptomic Signature (5 month liver)

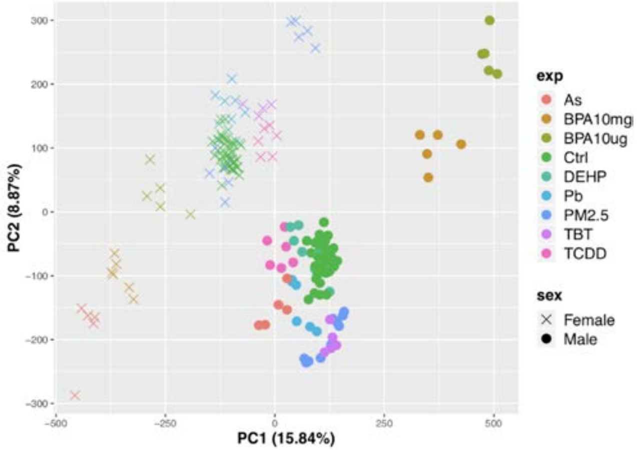
Exposure-responding genes



TaRGET
Toxicant Exposure and Responses by Genomic and Epigenomic Regulators of Transcription

Exposure signature database

Distinct Transcriptomic & Epigenetic Response



Sex-specific genes

Environmental-sensitive genes

TaRGET I and II Consortium Learnings: Consortium Effectiveness

- Harmonizing technologies across multiple sites/labs/investigators is time-consuming and challenging. While we started out using this model (i.e. ATAC-seq), it was replaced by **establishing expertise at a single-sight to generate data consortium-wide** (WGBS, ChIP-seq)

- **Expectations** and **timelines** must be clear, **milestones** agreed upon and investigators held accountable. This includes an understanding that U-type Programs are about data generated and differ from R-type grants.

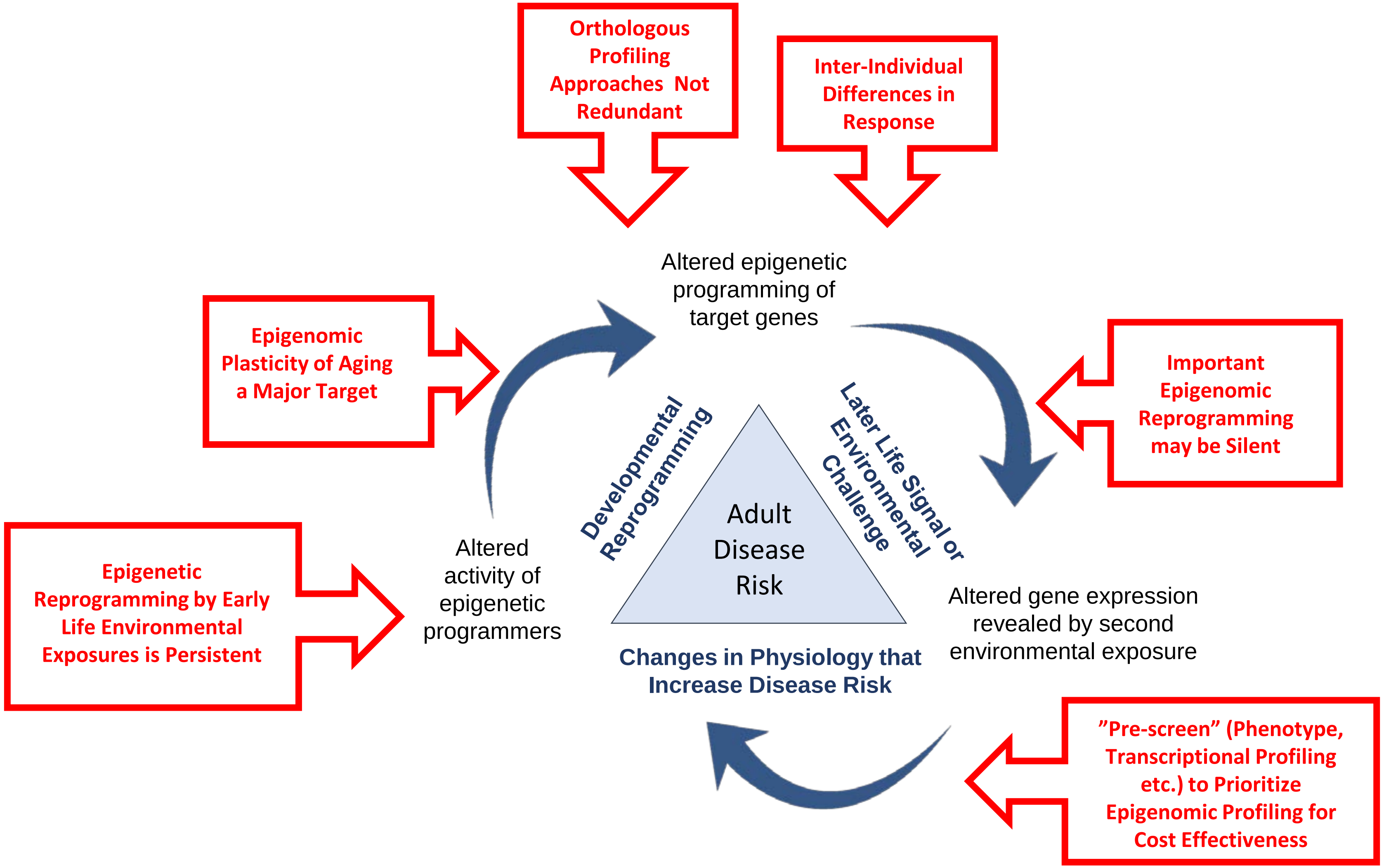
- Would recommend considering an **initial funding period** for participants to help the consortium "gel", provide opportunity to demonstrate responsiveness, and if necessary, reshuffle the deck to enhance productivity **prior to major investment in resources.**

- DCC absolutely essential (and ours was great) but cannot replace **on-site interactive bioinformatics expertise** to generate site/study-specific insights that can benefit the whole consortium. Also, the **DCC needs to be sustainable** beyond the last upload of consortium data for updates, manuscript preparation, community use etc.

- Extensive **tissue banking** made it possible to add other consortium-wide analyses (e.g. ChIP-seq) and respond to other stakeholders (e.g. added cortex at NIA request) after the studies were initiated/concluded

- The concept is always simpler than the actual research- The short **duration and budget** for T2C was unrealistic.

TaRGET I and II Consortium Learnings: Epigenome x Environment Interactions



CHALLENGE OF QUANTITY VS QUALITY

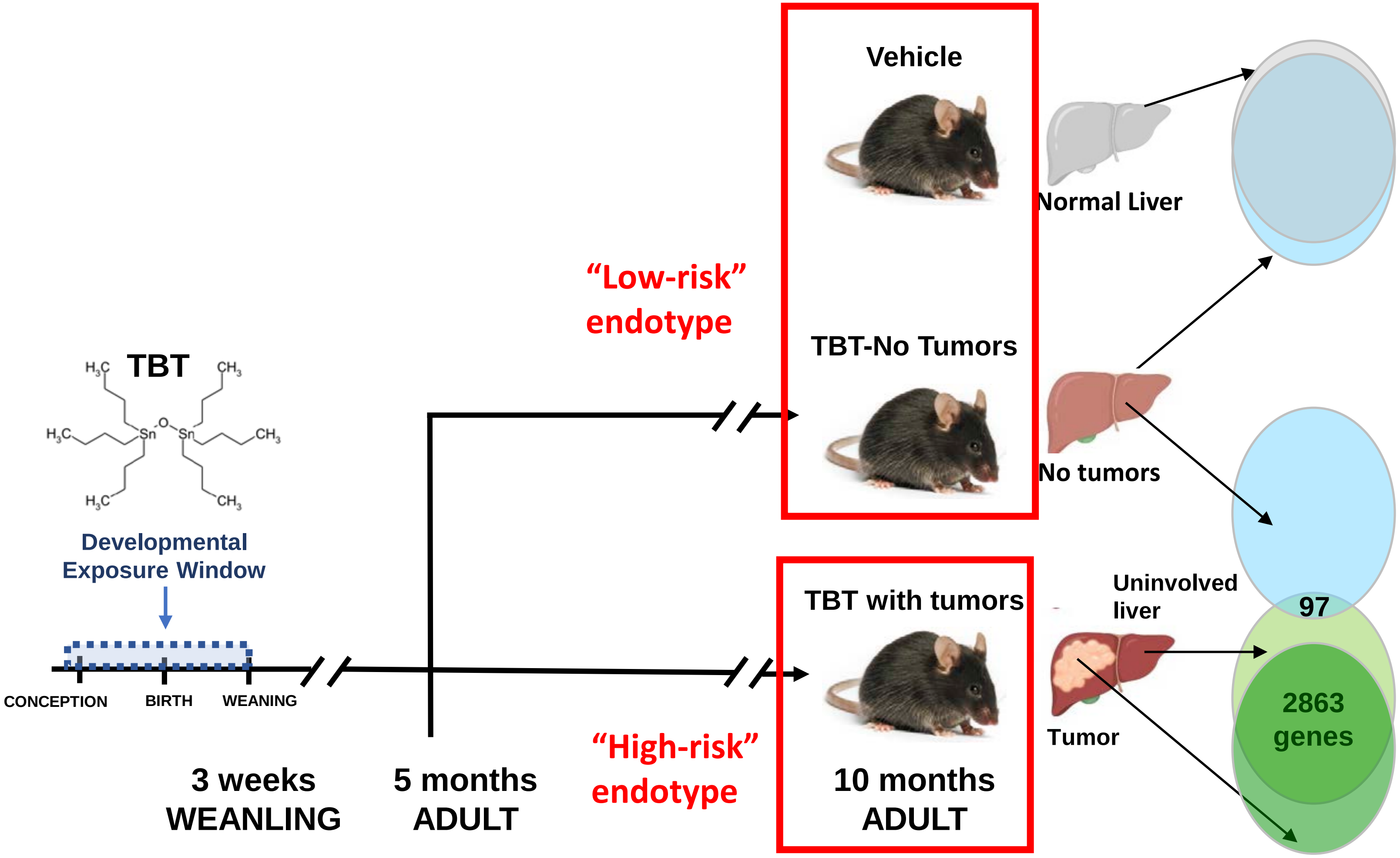
First and most obvious: Sample size vs profiling depth/comprehensiveness

- **Expense of Next-gen profiling vs sample size?** Typical budget for comprehensive epigenomic profiling (RNA-seq, ATAC-seq, ChIP-seq, WGBS) ~$5000/sample (within order of magnitude)
  For 2 sexes x 2 arms (exposure + vehicle) x 2 tissues (target + surrogate)
  N=10 $400,000 N=5 $200,000 etc.

Equally important but less obvious considerations:

- **Number of orthologous approaches vs comprehensiveness of epigenomic profiling?** Using epigenomic profiling, orthologous techniques did not produce redundant information e.g. little overlap between RNA-seq, ATAC-seq, and ChIP-seq signatures. ATAC-seq was least informative.
- **Coverage vs granularity?** Whole tissue RNA-seq vs scRNA-seq, Whole tissue ChIP-seq vs Cut-and-Tag, WGBS vs RRBS etc.
- **Pooling of animals vs individual measurements?** Inter-individual variation in response seen even when using genetically identical inbred C57Bl/6 mice. Sample availability (and in some cases profiling approach) can become limiting when analyses done at the level of the individual animals (e.g. blood from 3 week old mice). However, pooling can mask effects if there are "responders" and "non-responders" in exposure arms

Analysis of Individual Animals Identifies Endotypes Associated with Risk



10 mos ~ 50% mice develop tumors

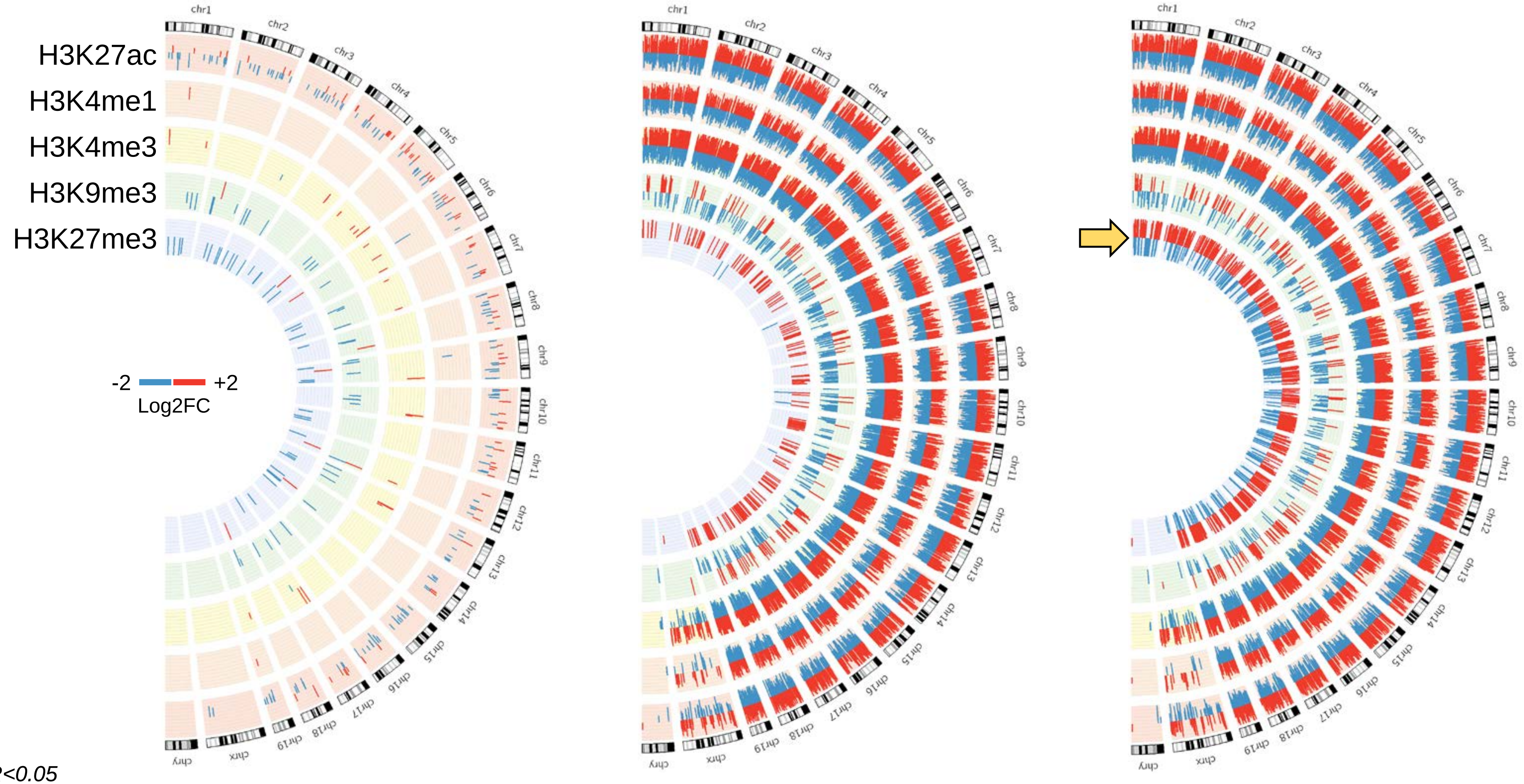
Transcriptional profiles of livers from TBT-exposed mice that do not develop tumors are virtually identical to vehicle-exposed animals by RNA-seq (only 144 DEGs)

Livers from mice without tumors have very few DEGs in common with UL liver or tumors (97 DEGs)

Transcriptional profiles from uninvolved livers (UL) of TBT-exposed mice with tumors share most of their differentially expressed genes (2863 DEGs) with tumors (T)

Distinct Epigenetic Endotypes Linked to Development of TBT-induced Liver Tumors

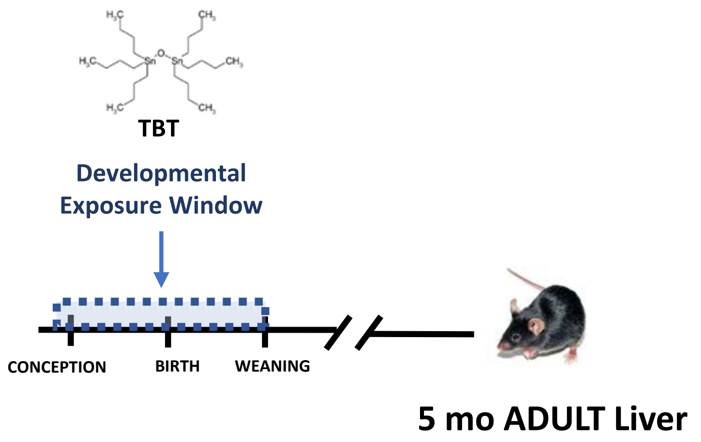
No tumors	Animals with tumors	
10mo “Low-Risk” Endotype TBT over Veh	**10mo “High-risk” Endotype** TBT over Veh	**10mo Tumor** TBT over Veh

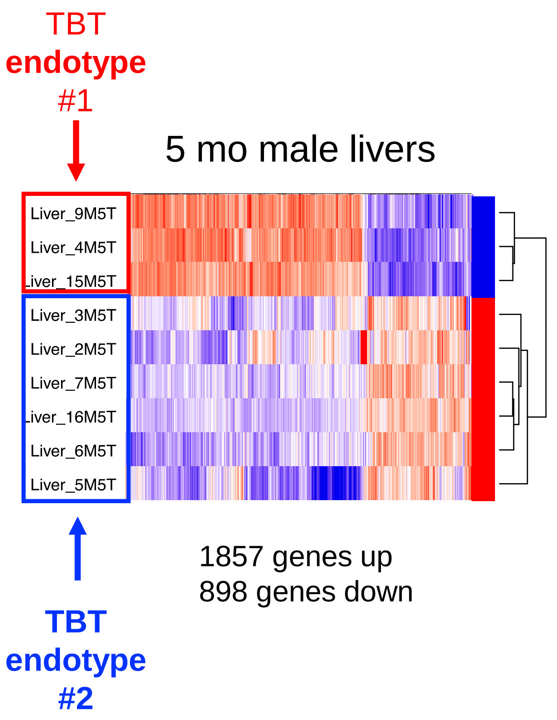


DiffReps FC>1.5, FDR<0.05

TBT-induced Reprogramming and High-risk Endotype Precedes Tumor Development

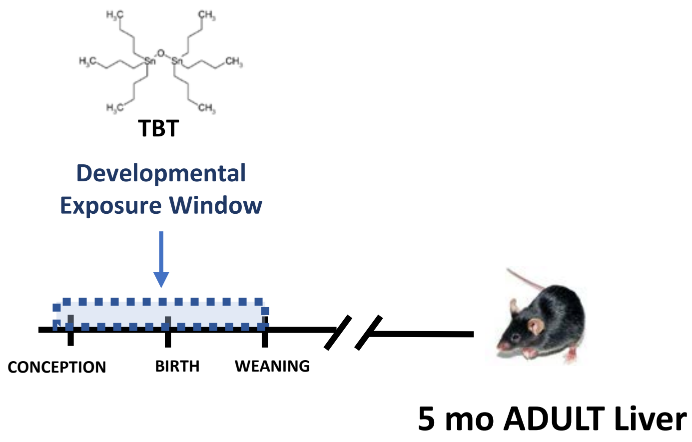
- When combined, 5 month livers of TBT-exposed mice display very few DEGs (227 DEGs) vs vehicle. However hierarchical clustering identified 2 endotypes in 5 month old male livers that differed from each other by > 4000 DEGs and from normal age-matched liver by >1400 DEGs

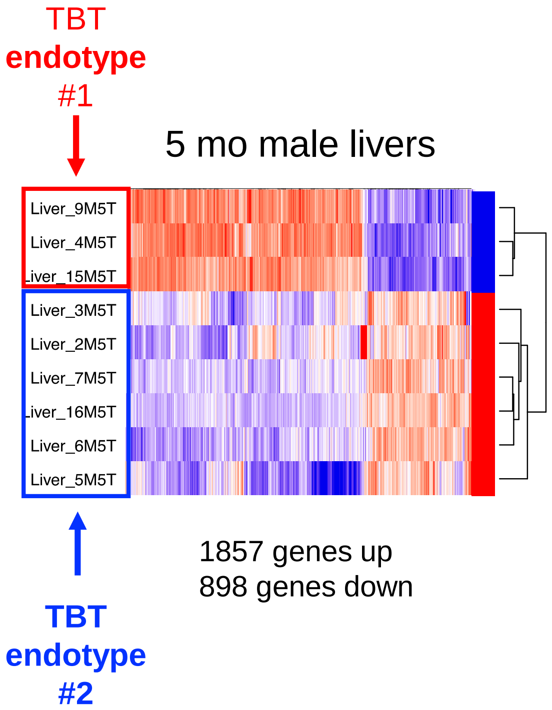


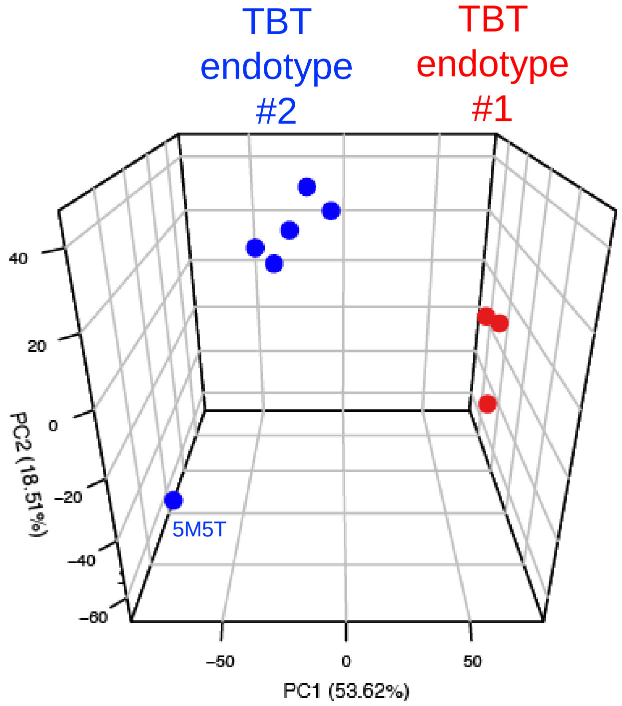


TBT-induced Reprogramming and High-risk Endotype precedes Tumor Development

- When combined, 5 month livers of TBT-exposed mice display very few DEGs (227 DEGs) vs vehicle. However hierarchical clustering identified 2 endotypes in 5 month old male livers that differed from normal age-matched liver by >1000 DEGs
- PCA analysis also separated 5mo TBT liver into 2 endotypes

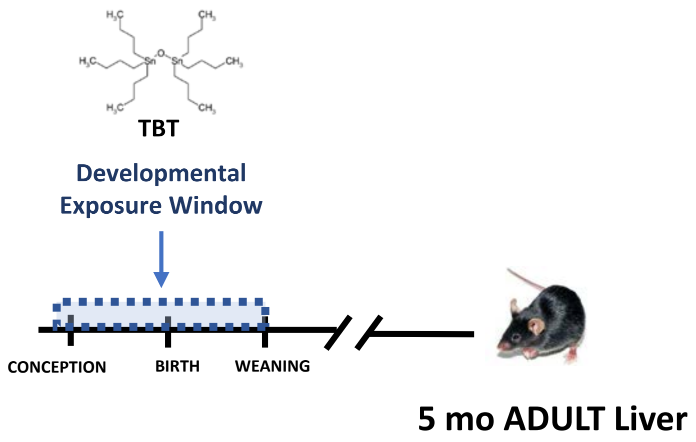


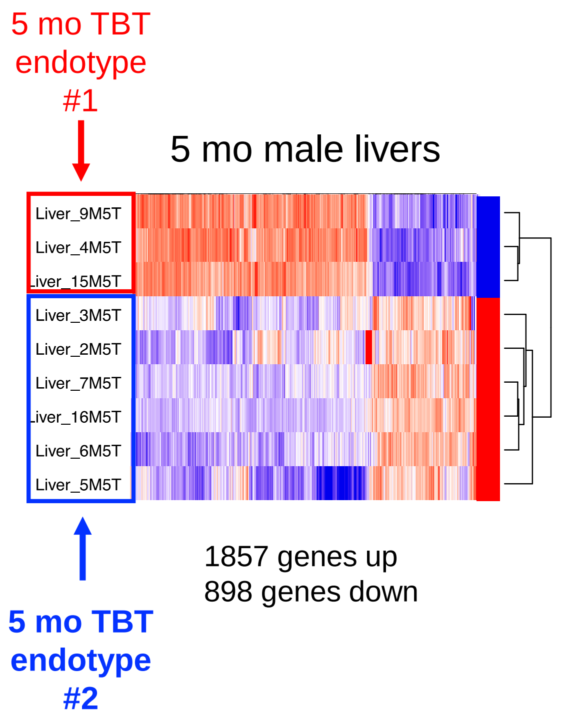


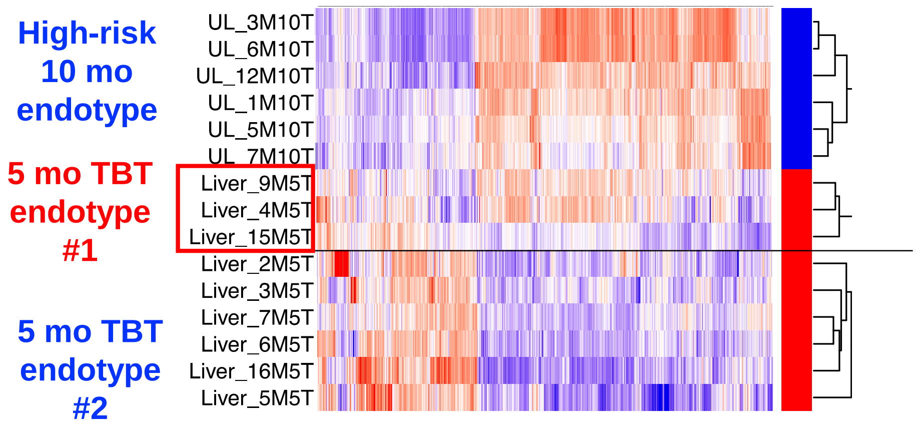


TBT-induced Reprogramming and High-risk Endotype Precedes Tumor Development

- When combined, 5 month livers of TBT-exposed mice display very few DEGs (227 DEGs) vs vehicle. However hierarchical clustering identified 2 endotypes in 5 month old male livers that differed from normal age-matched liver by >1000 DEGs
- PCA analysis also separated 5mo TBT liver into 2 endotypes
- Endotype #1 clusters with the 10 mo “high-risk” TBT endotypes

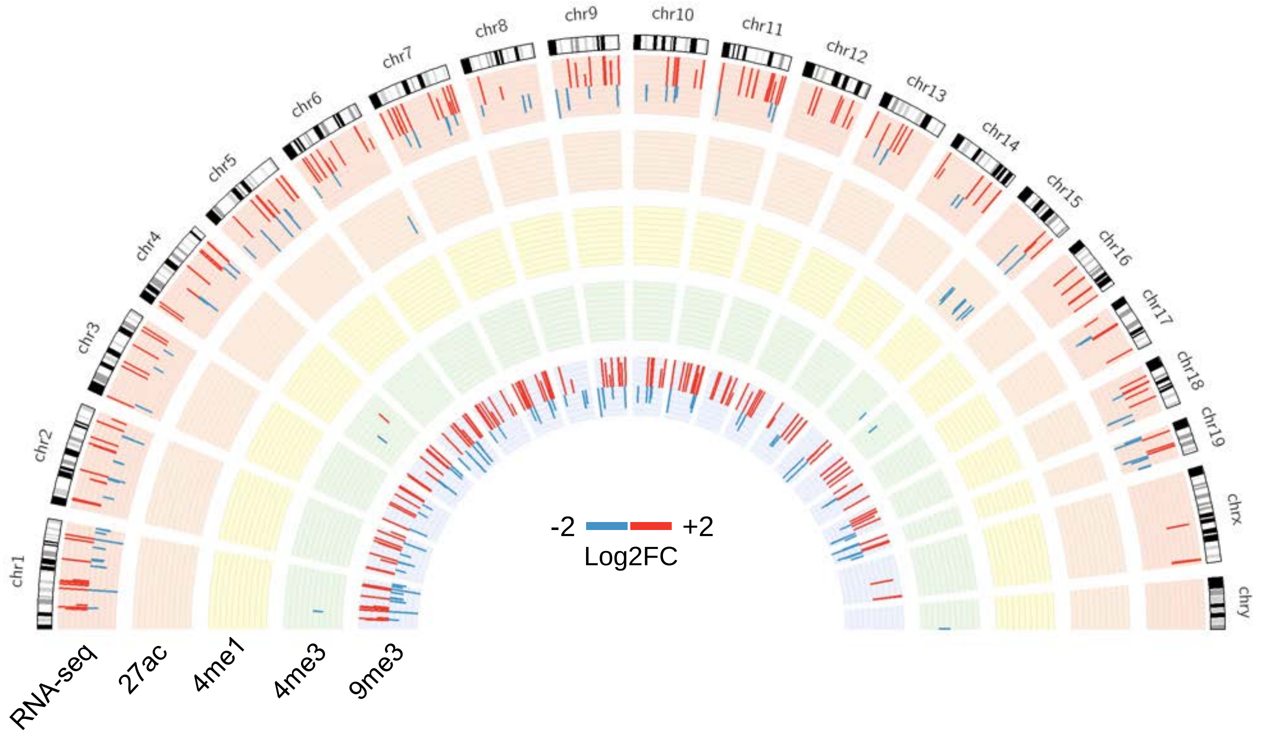






Heterogeneity Can Obscure TBT-induced Reprogramming and High-risk Endotype

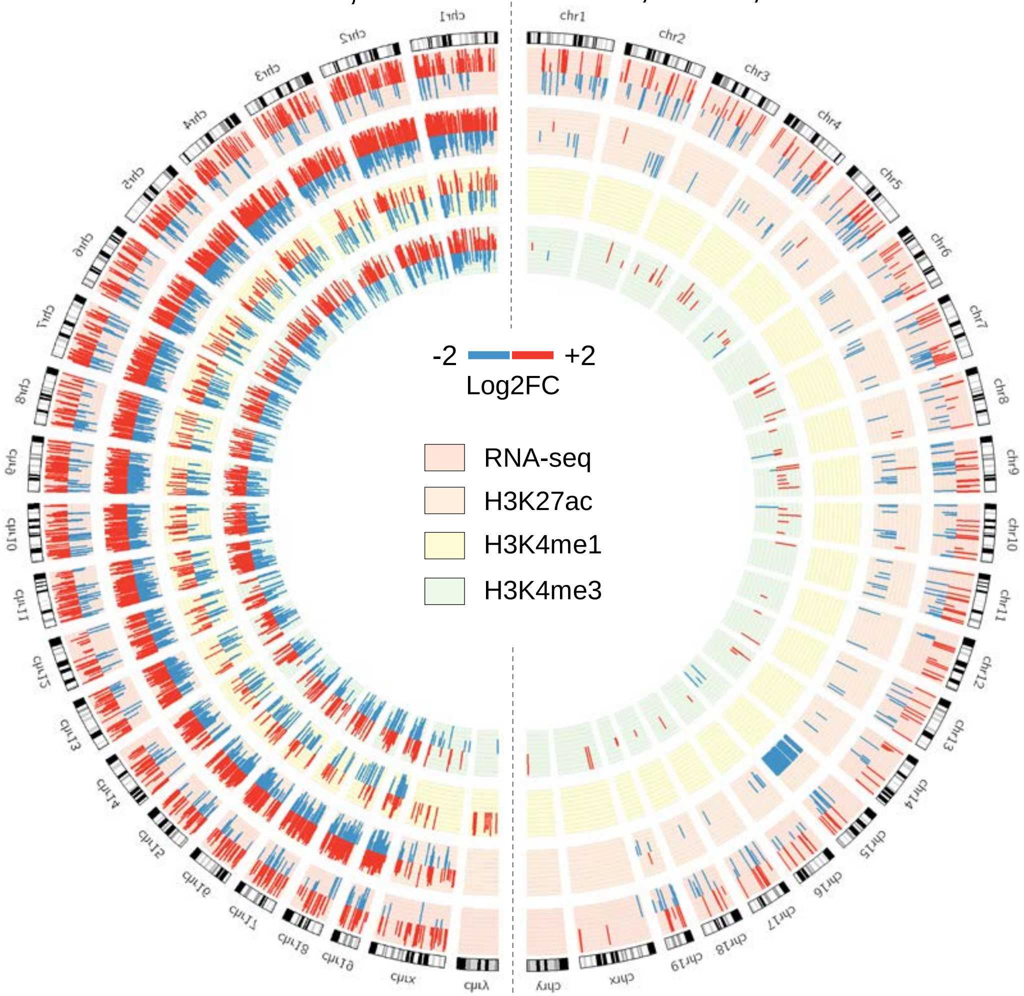
Male liver: 5mo TBT over VEH
All mice combined
(high risk & low risk)



Male liver: 5mo TBT over VEH

High risk:
4M5T, 15M5T

Low risk:
5M5T, 6M5T, 16M5T



Anchoring and Stratifying with Transcriptomics and Phenotyping

First and most obvious: Interpreting epigenomic data with transcriptomic data

- Altered patterns of gene expression as a downstream response to epigenomic reprogramming can aid in interpretation of epigenomic data and provide insights into mechanisms of adverse effects/health outcomes

Equally important but less obvious considerations:

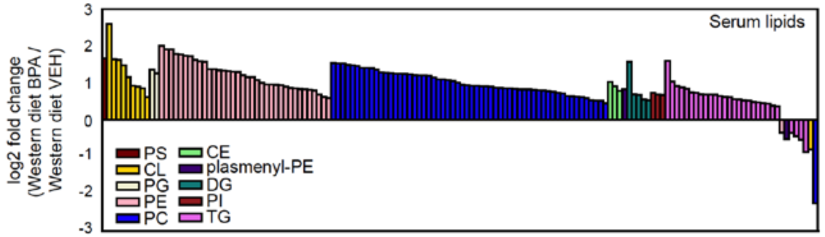
- **Exploratory experiments to determine "value" of in-depth (expensive) epigenomic profiling in specific tissues/cells** Across different exposure paradigms, relevant tissue(cell) targets for an exposure to produce an effect may not be known-how many to profile (see previous cost analysis).

  Similarly, when calibrating across different exposures and consortium sites, choice of a single target (i.e liver for TaRGET II) while providing methodological consistency, may only be informative for a few exposures (i.e. all exposures do not impact all tissues/cells equally-some not al all).

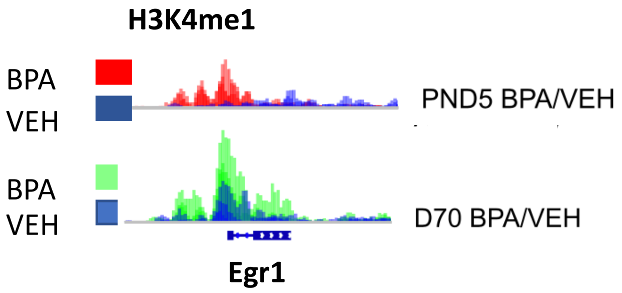
- Over-dependence on Transcriptional profiling can miss important "silent reprogramming" useful as both a biomarker and determinant of later-life effects

Metabolic Dysfunction Caused by Epigenomic Reprogramming Revealed by Later Life Dietary Challenge

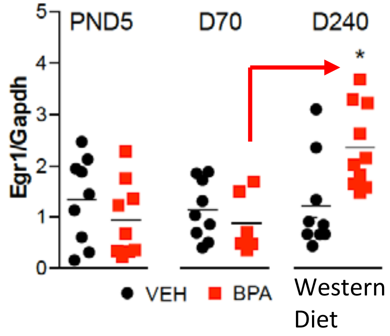
BPA **Developmental Exposure Window PND 3-5**



On Western Diet, rats have larger livers, higher levels of serum cholesterol, and dyslipidemia relative to either BPA-exposed on normal chow or vehicle-exposed on Western diet

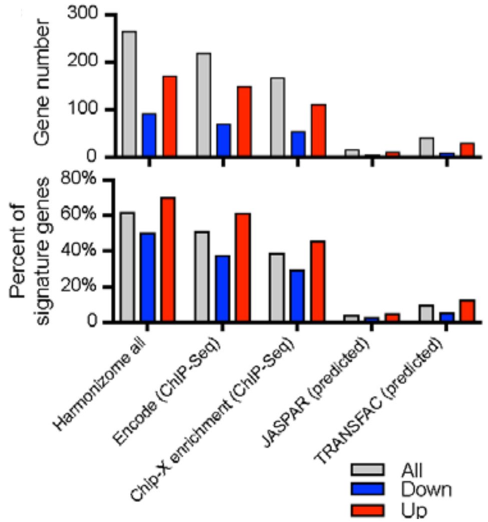


Reprogramming of EGR1 seen acutely after exposure, and is persistent



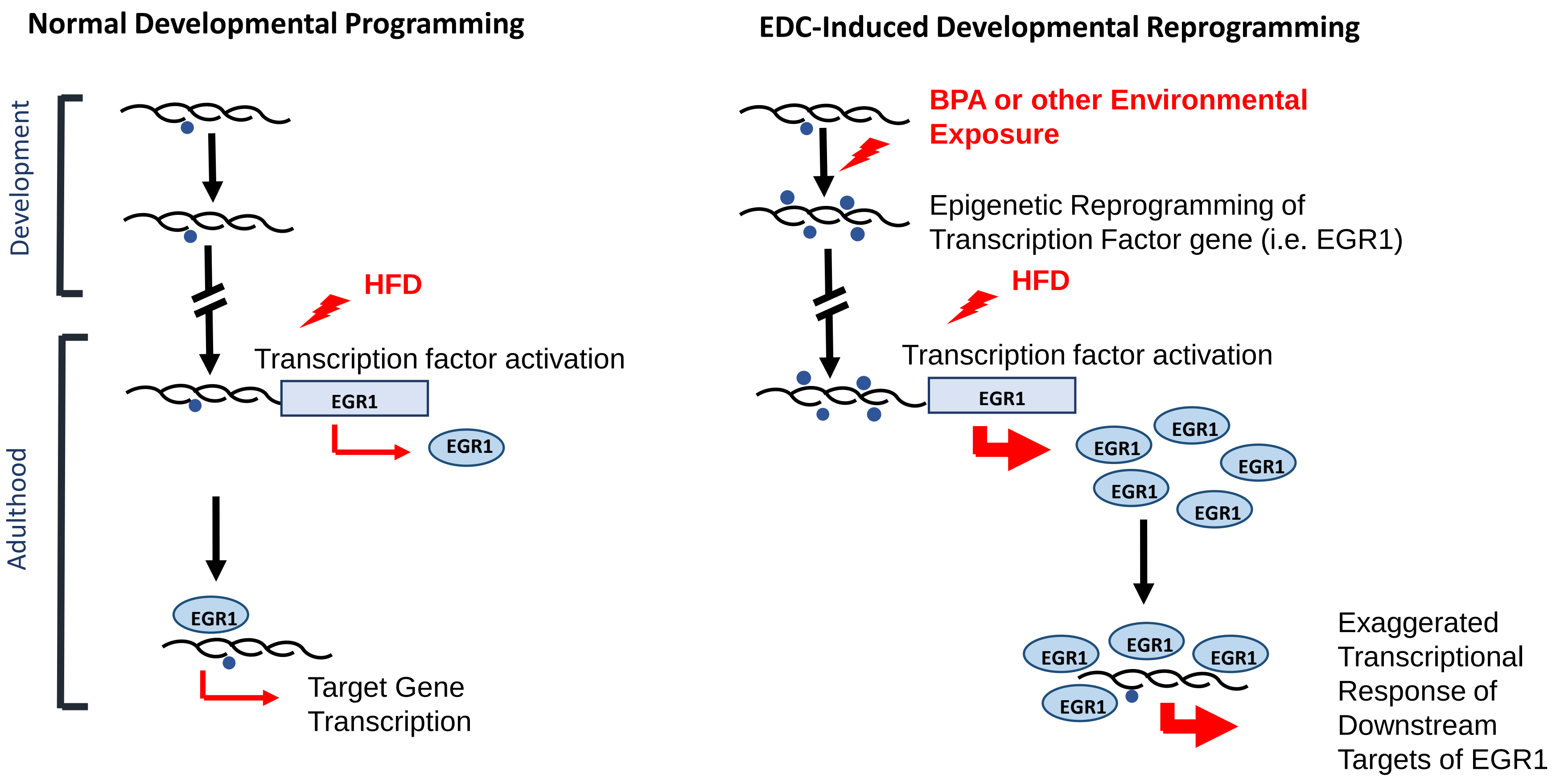
However, no change in EGR1 expression until challenged with Western diet

Reprogramming of the EGR1 transcriptome accounted for > 70% of the aberrant response to Western diet



Trevino et al *Nature Communications* 2020

Metabolic Dysfunction Caused by Epigenomic Reprogramming Revealed by Later Life Dietary Challenge



Trevino et al *Nature Communications* 2020

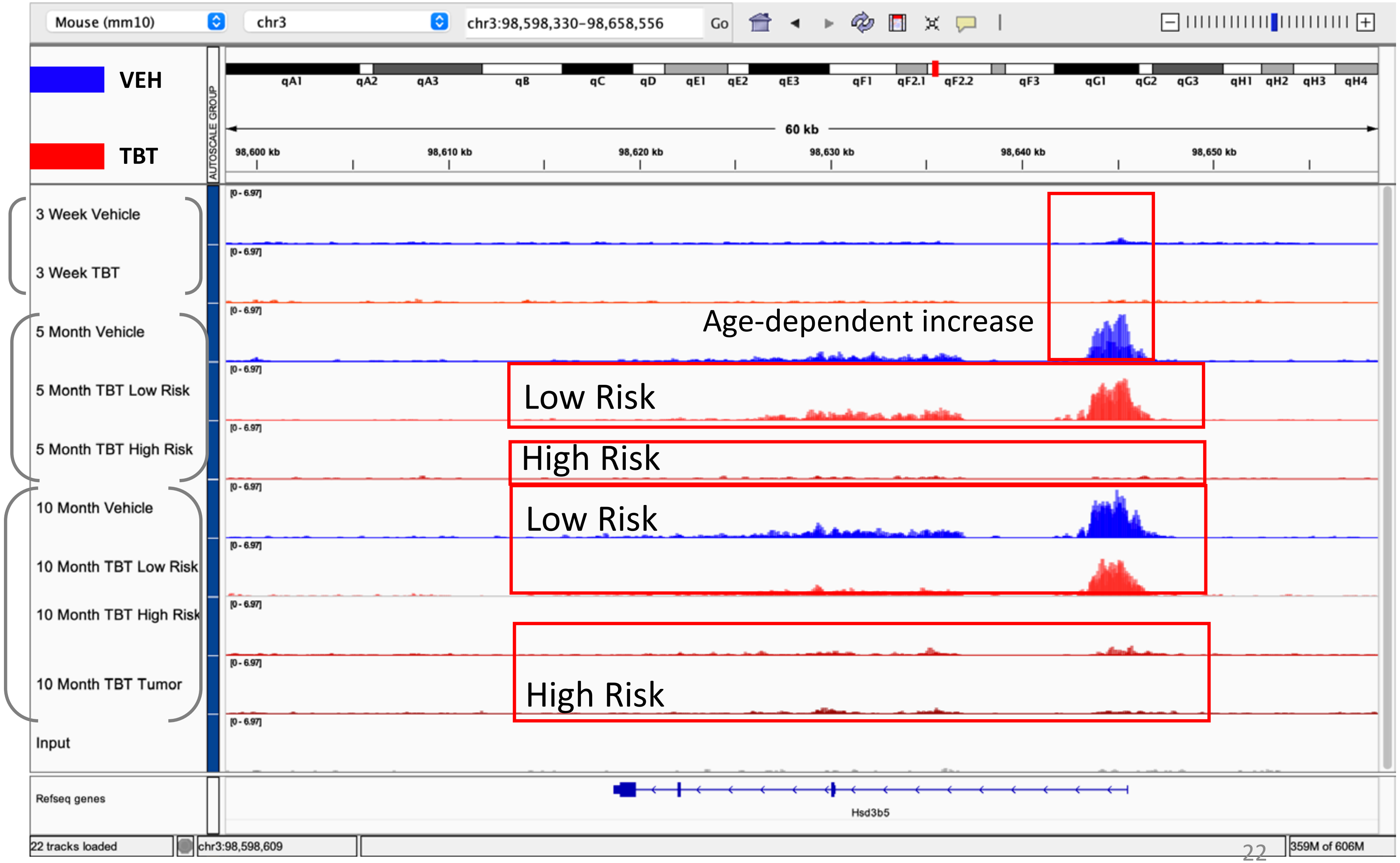
Importance of Longitudinal Analyses: Epigenomic Plasticity of Aging is Vulnerable to Environmental Exposures

- It is now well appreciated **changes in the epigenome occur during the aging process**. What is less well-understood is whether/how these alterations drive age-associated diseases or create vulnerabilities to environmental exposures.
- By capturing both normal and toxicant-induced epigenomic changes across the life-course, we have found the **plasticity inherent to normal epigenomic aging creates a vulnerability** to multiple environmental exposures
- Examples include both **acceleration and attenuation of epigenomic aging** by early-life exposures
- In the case of toxicant-induced attenuation of age-associated changes in the liver epigenome ("anti-aging" signature) linked to development of fatty liver and tumors, these **age-asscociated alterations are able to accurately stratify human patient populations** by disease (HCC) and severity (NAFLD/NASH)

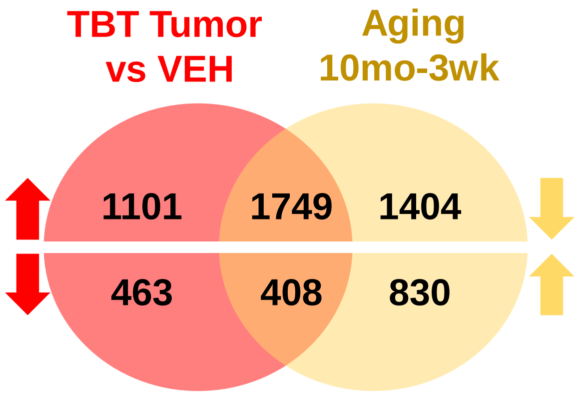
Epigenomic Plasticity of Aging is Vulnerable to Environmental Exposures

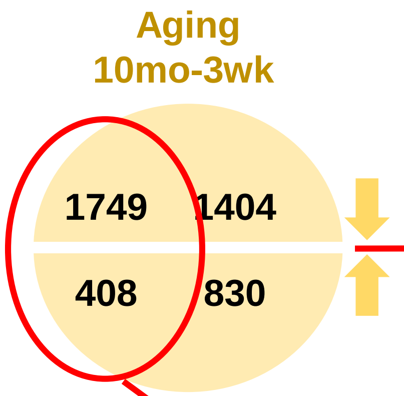
Hsd3b5

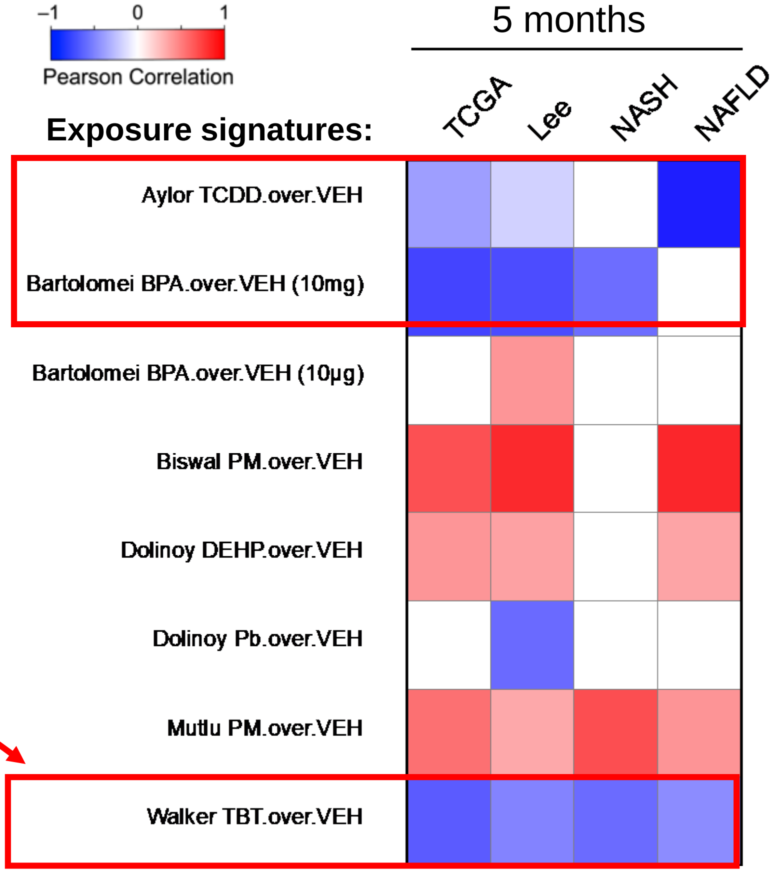
- Hydroxy-delta-5-steroid dehydrogenase
- Normally increases in expression from 3 weeks to 10 months in vehicle animals
- **H3K4me3** at promoter normally increases between neonatal and adult life
- Lack of H3K4me3 in the high-risk endotype at both 5 and 10 mo, indicating epigenetic reprogramming ***preceded*** tumor development



Epigenomic Plasticity of Aging is Vulnerable to Environmental Exposures

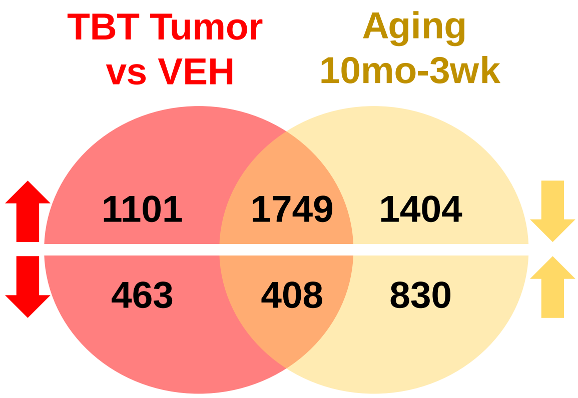


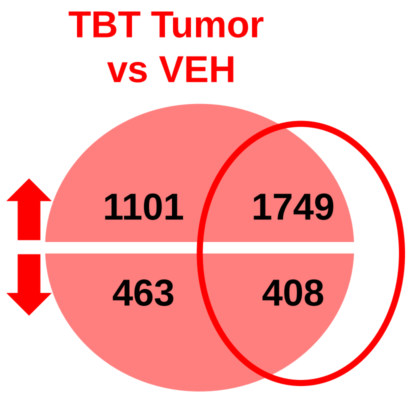


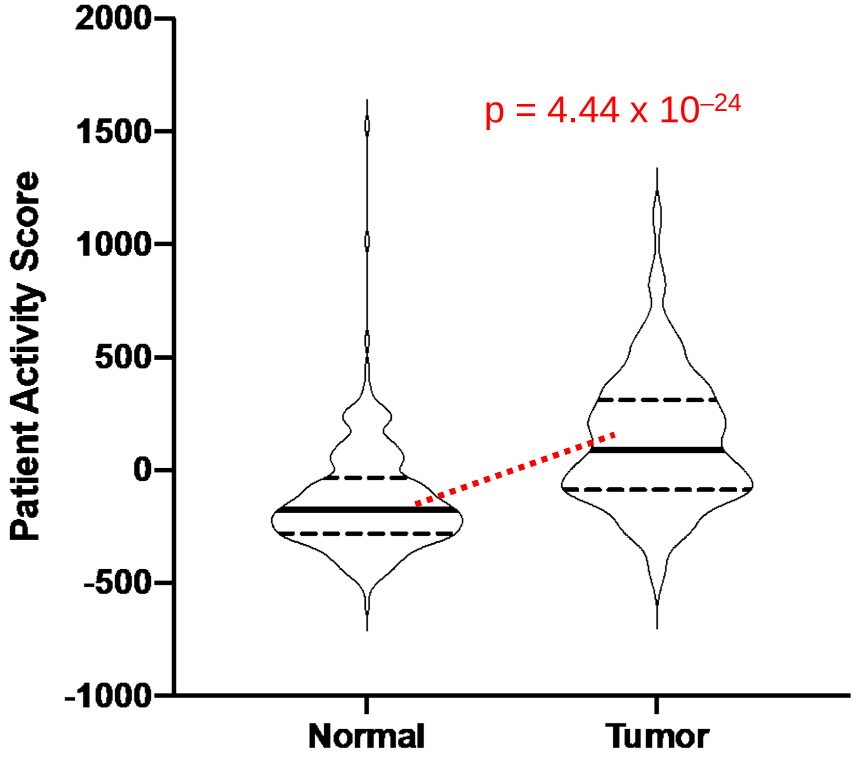


- 2157/3721 **(58%) of the TBT signature** is driven by attenuation of normal age-related changes
- Genes differentially expressed in response to **TBT as well as TCDD and BPA in this "anti-aging" signature** also showed strong negative correlation with human liver disease signatures

Epigenomic Plasticity of Aging is Vulnerable to Environmental Exposures



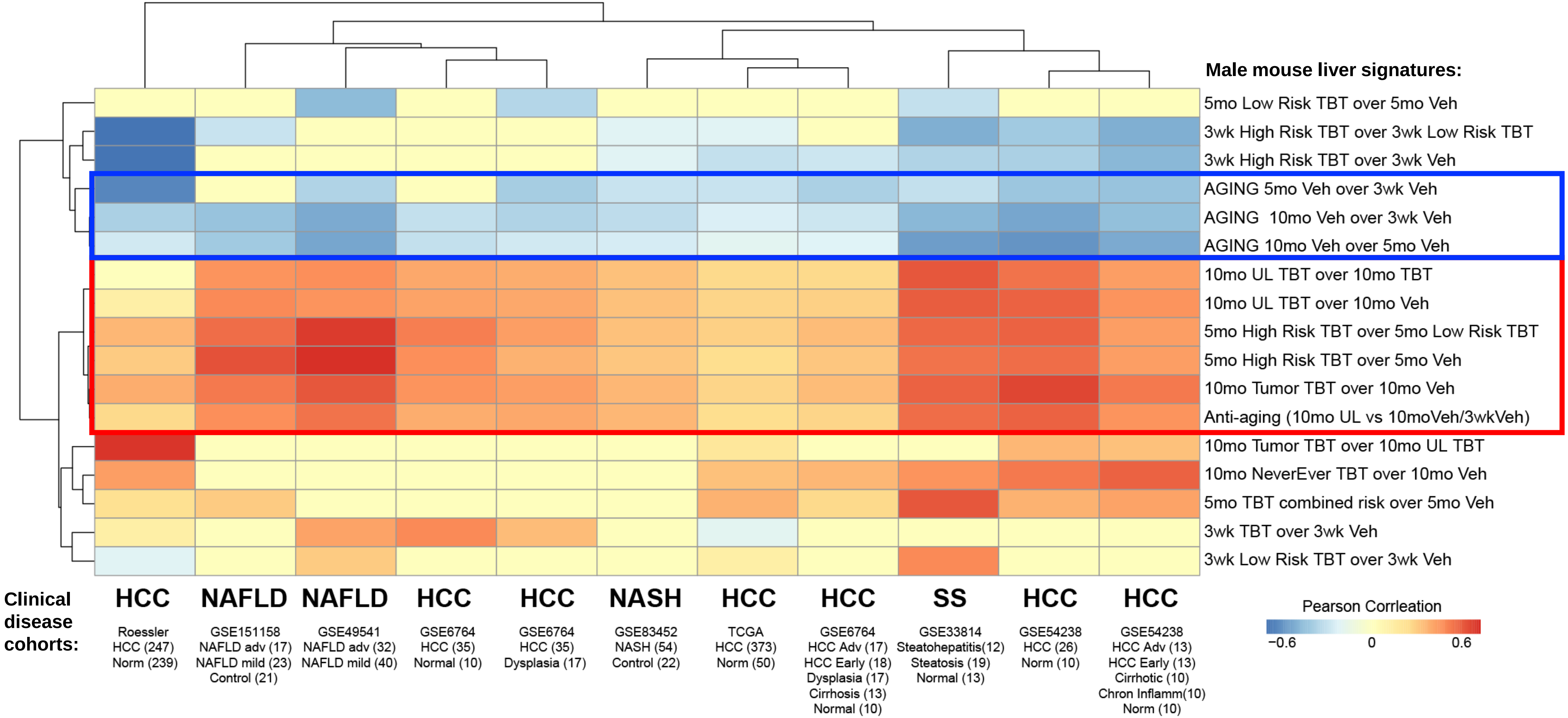




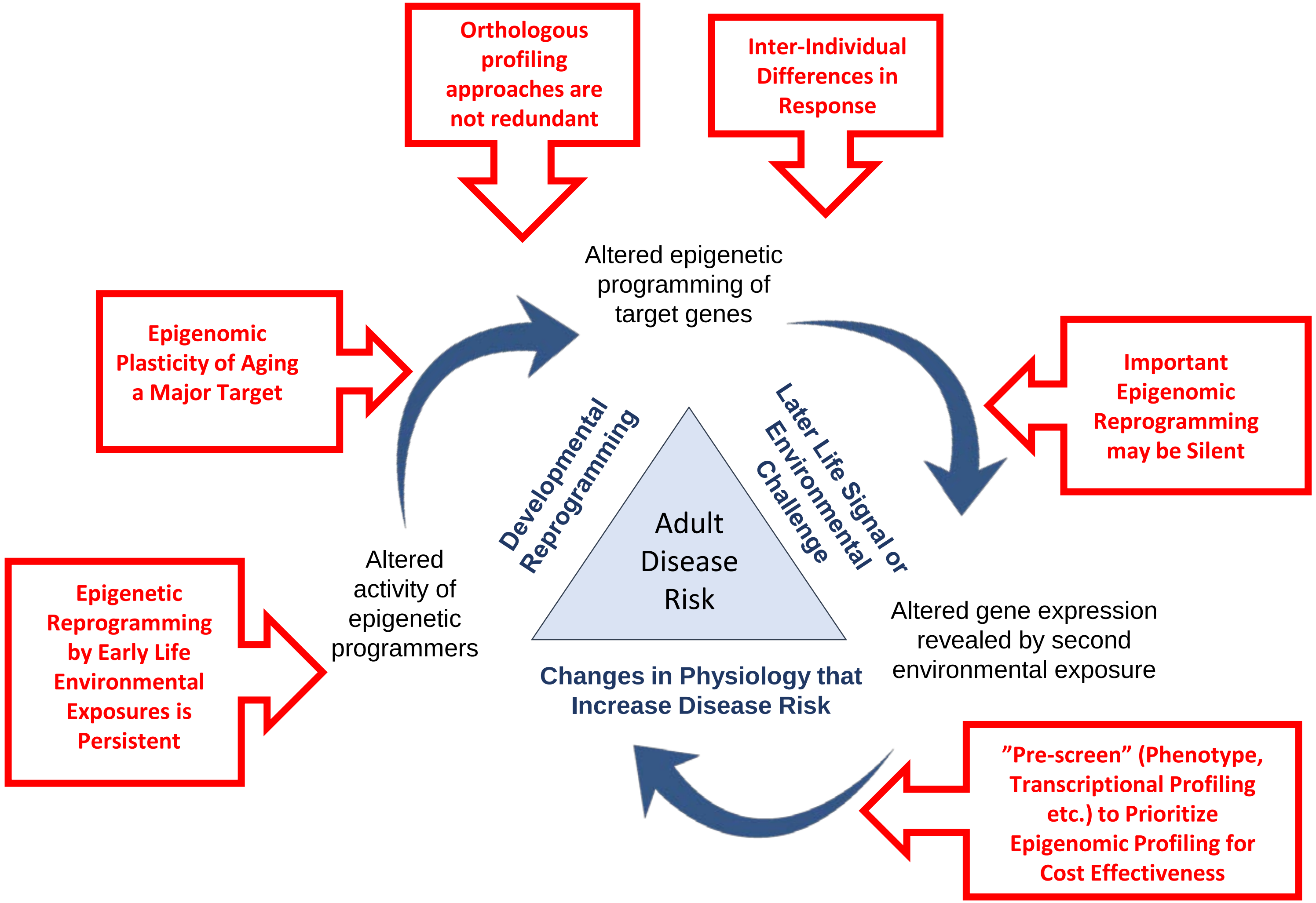
239 patient samples
(Lee et al *Nat Gen* 2014)

- 2157/3721 (58%) **of the TBT signature** is driven by attenuation of normal age-related changes
- Genes differentially expressed in response to **TBT as well as TCDD and BPA in this "anti-aging" signature** also showed strong negative correlation with human liver disease signatures
- The "anti-aging" component of the TBT signature shows significant positive correlation with human HCC signatures

Toxicant-induced Signatures Correlate with Altered Gene Expression in Human Liver Disease



TaRGET I and II Consortium Learnings: Epigenome x Environment Interactions



- Exposures during "first 100 days" (pre-conception through weaning) cause alterations in the epigenome that persist across the life-course
- Epigenomic plasticity associated with normal aging is vulnerable to reprogramming and translates to human disease settings
- Data obtained with different epigenomic profiling approaches provide distinct information
- Heterogeneity in response can be a significant confounder even in genetically homogeneous models
- Epigenomic reprogramming while persistent, may not cause a change in gene expression until triggered by later life environmental stressors