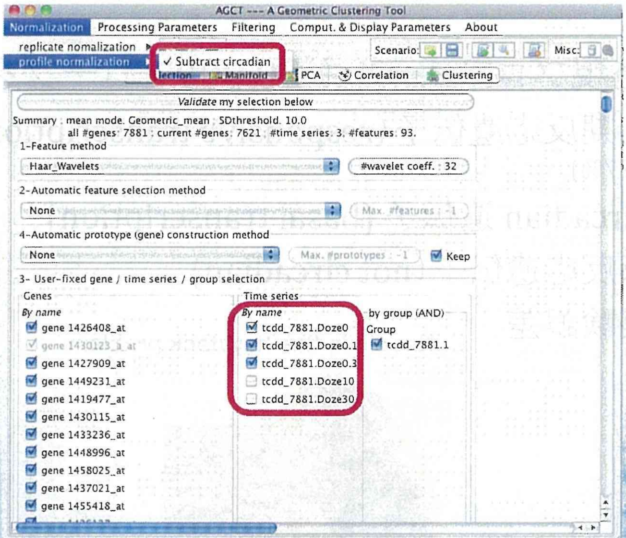
AGCT 2.0 Interface



Summary : mean mode. Geometric_mean ; SDthreshold. 10.0
all #genes: 7881 ; current #genes: 7621 ; #time series: 3, #features: 93.

1-Feature method
Haar_Wavelets
#wavelet coeff. : 32

2-Automatic feature selection method
None
Max. #features : -1

4-Automatic prototype (gene) construction method
None
Max. #prototypes : -1
☑ Keep

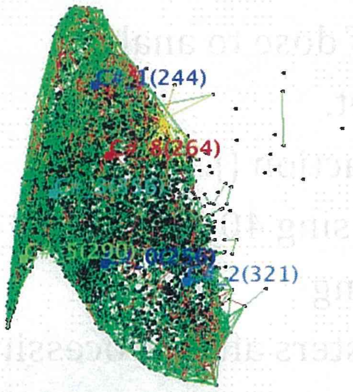
3- User-fixed gene / time series / group selection

Genes
By name
☑ gene 1426408_at
☑ gene 1430123_a_at
☑ gene 1427909_at
☑ gene 1449231_at
☑ gene 1419477_at
☑ gene 1430115_at
☑ gene 1433236_at
☑ gene 1448996_at
☑ gene 1458025_at
☑ gene 1437021_at
☑ gene 1455418_at

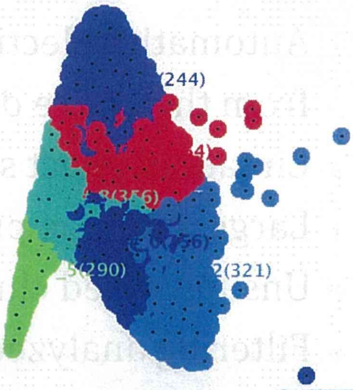
Time series
By name
☑ tcdd_7881.Doze0
☑ tcdd_7881.Doze0.1
☑ tcdd_7881.Doze0.3
☐ tcdd_7881.Doze10
☐ tcdd_7881.Doze30

by group (AND)
Group
☑ tcdd_7881.1

Spectral manifold representation



Delaunay triangulation

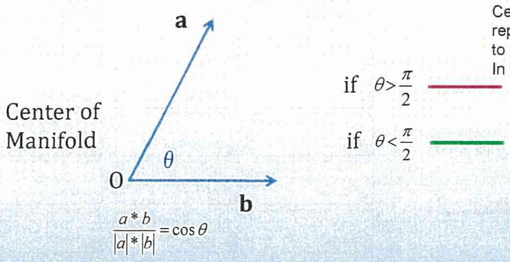


Clustering in 7 clusters

Delaunay triangulation

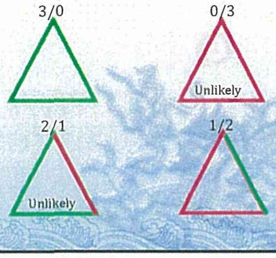
To evaluate if the obtained graph is reasonable or not, we defined a measure used to filter out any Delaunay edges between two genes g and g' for which:

$$\cos^{-1}(\boldsymbol{x}_g, \boldsymbol{x}_{g'}) \in [\pi \times p/2, \pi \times (1-p/2)] \pmod{\pi} . \tag{2}$$

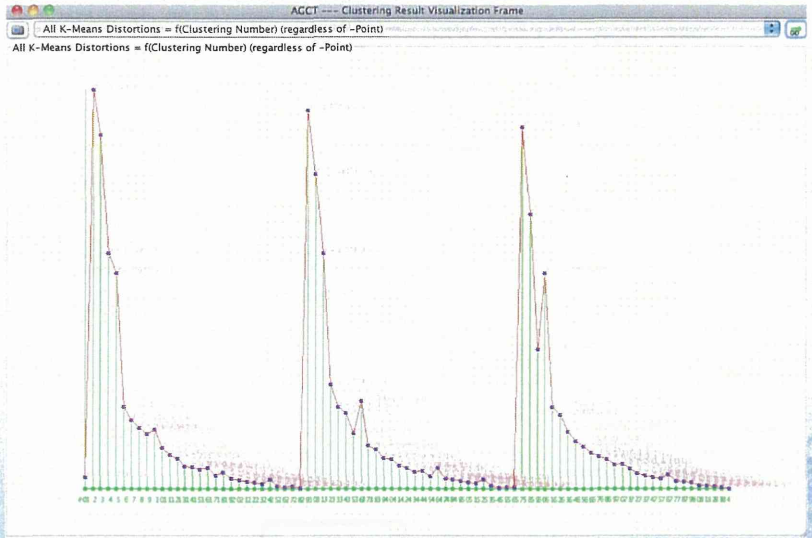


$$\frac{a*b}{|a|*|b|} = \cos\theta$$

Cells (triangles) and genes are in bijection, each cell representing the volume composed of all points closer to the gene than to any other gene.
In the computed structure local consistency is tested.

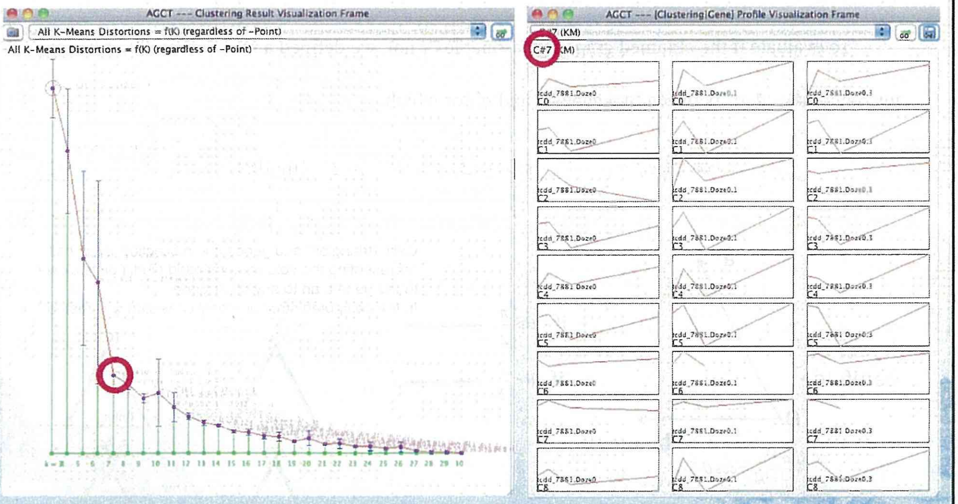


Run k-means with ***k*** {3..30}

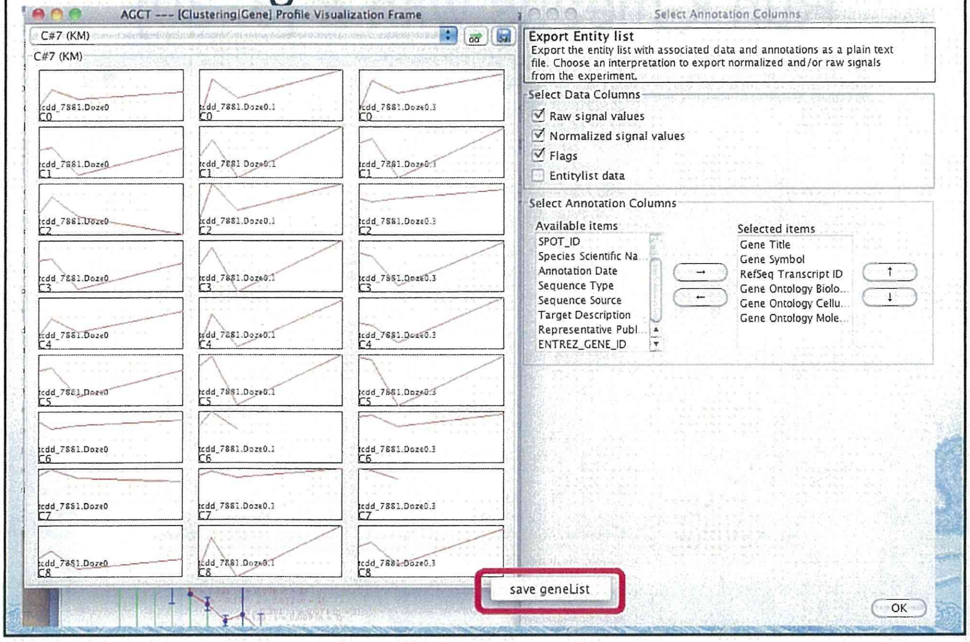


All K-Means Distortions = f(Clustering Number) (regardless of -Point)

Select the best ***k*** which gives the smallest standard deviation in different runs



Save gene cluster information



Export Entity list

Export the entity list with associated data and annotations as a plain text file. Choose an interpretation to export normalized and/or raw signals from the experiment.

Select Data Columns

- ☑ Raw signal values
- ☑ Normalized signal values
- ☑ Flags
- ☐ Entitylist data

Select Annotation Columns

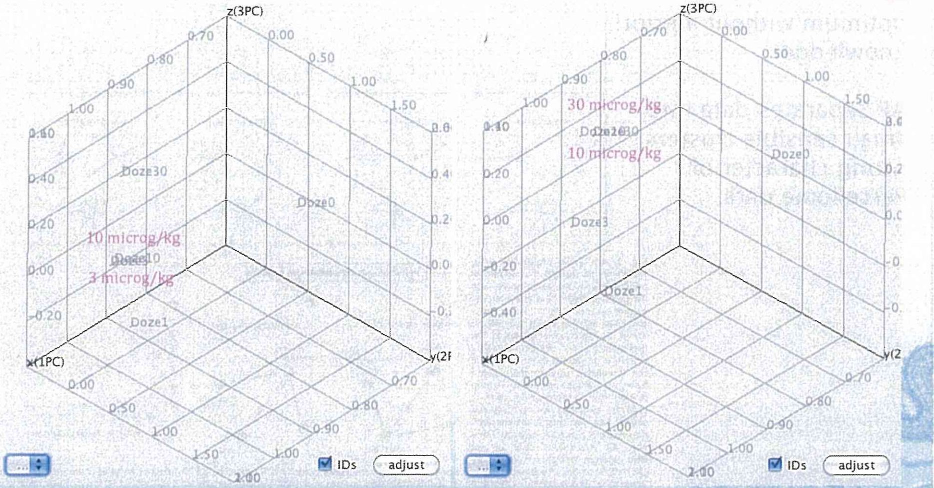
Available items
SPOT_ID
Species Scientific Na...
Annotation Date
Sequence Type
Sequence Source
Target Description
Representative Publ...
ENTREZ_GENE_ID

Selected items
Gene Title
Gene Symbol
RefSeq Transcript ID
Gene Ontology Biolo...
Gene Ontology Cellu...
Gene Ontology Mole...

TCDD and TCDF: PCA on 5 doses

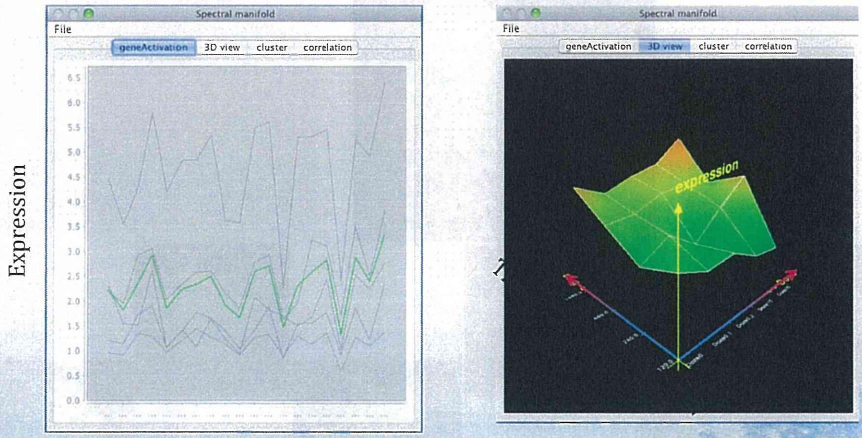
7,881 probes

8,674 probes

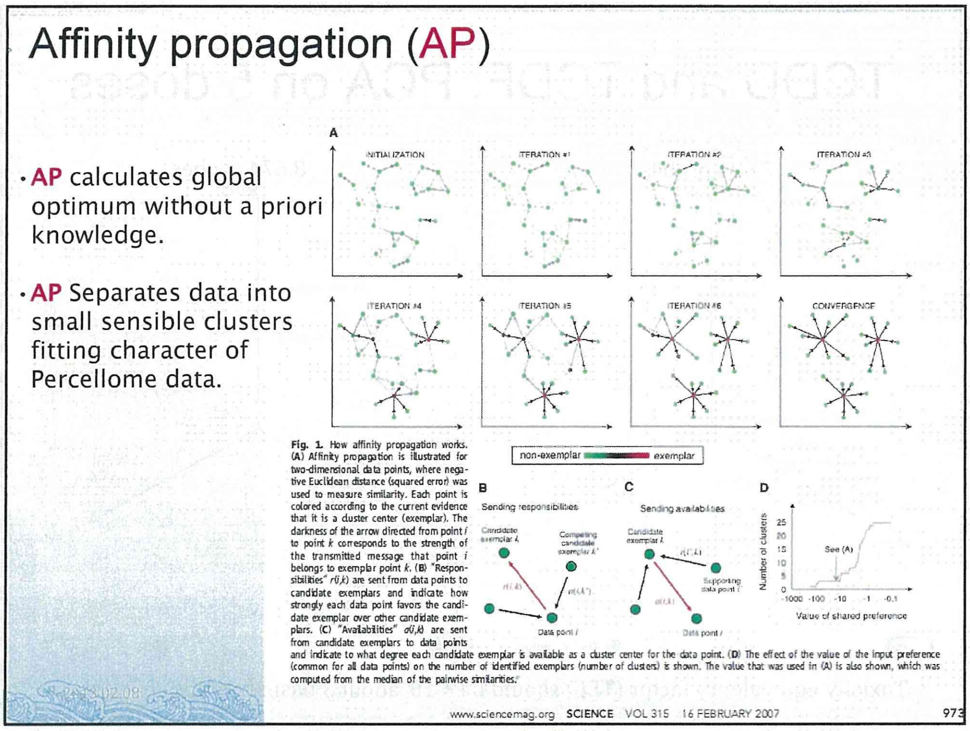


Toxicity equivalency factor (TEF) should be < 10, about 3 basing on this analysis.

2D and 3D visualization of clusters



Dozes 0, 1, 3, 10, 30

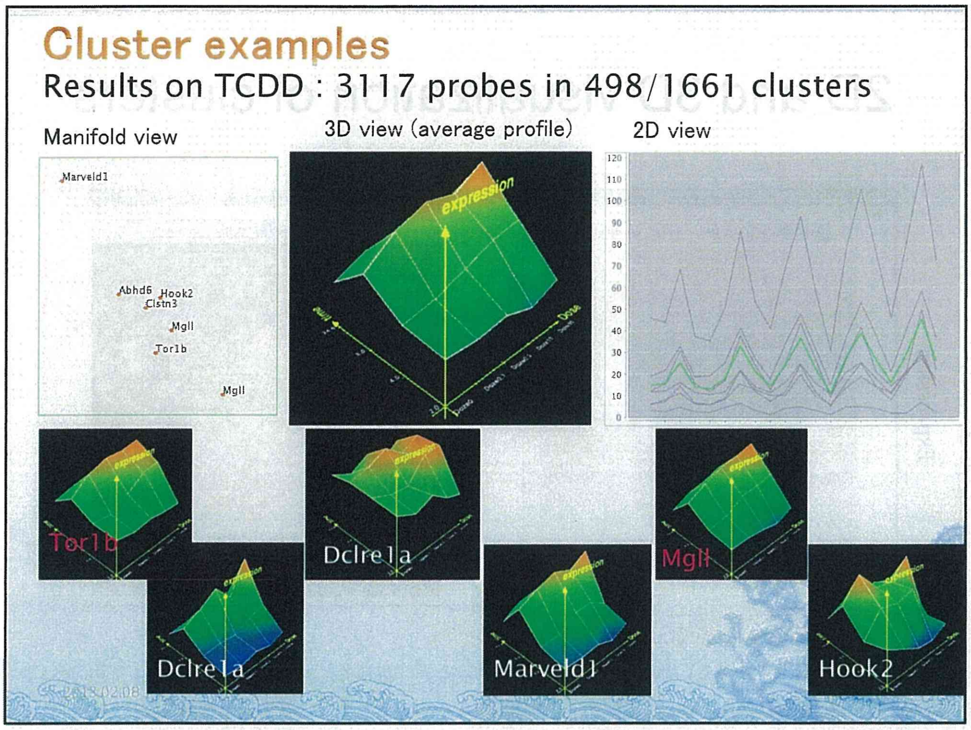


Affinity propagation (AP)

- **AP** calculates global optimum without a priori knowledge.
- **AP** Separates data into small sensible clusters fitting character of Percellome data.

Fig. 1. How affinity propagation works. **(A)** Affinity propagation is illustrated for two-dimensional data points, where negative Euclidean distance (squared error) was used to measure similarity. Each point is colored according to the current evidence that it is a cluster center (exemplar). The darkness of the arrow directed from point *i* to point *k* corresponds to the strength of the transmitted message that point *i* belongs to exemplar point *k*. **(B)** "Responsibilities" $r(i,k)$ are sent from data points to candidate exemplars and indicate how strongly each data point favors the candidate exemplar over other candidate exemplars. **(C)** "Availabilities" $a(i,k)$ are sent from candidate exemplars to data points and indicate to what degree each candidate exemplar is available as a cluster center for the data point. **(D)** The effect of the value of the input preference (common for all data points) on the number of identified exemplars (number of clusters) is shown. The value that was used in (A) is also shown, which was computed from the median of the pairwise similarities.

www.sciencemag.org SCIENCE VOL 315 16 FEBRUARY 2007 973



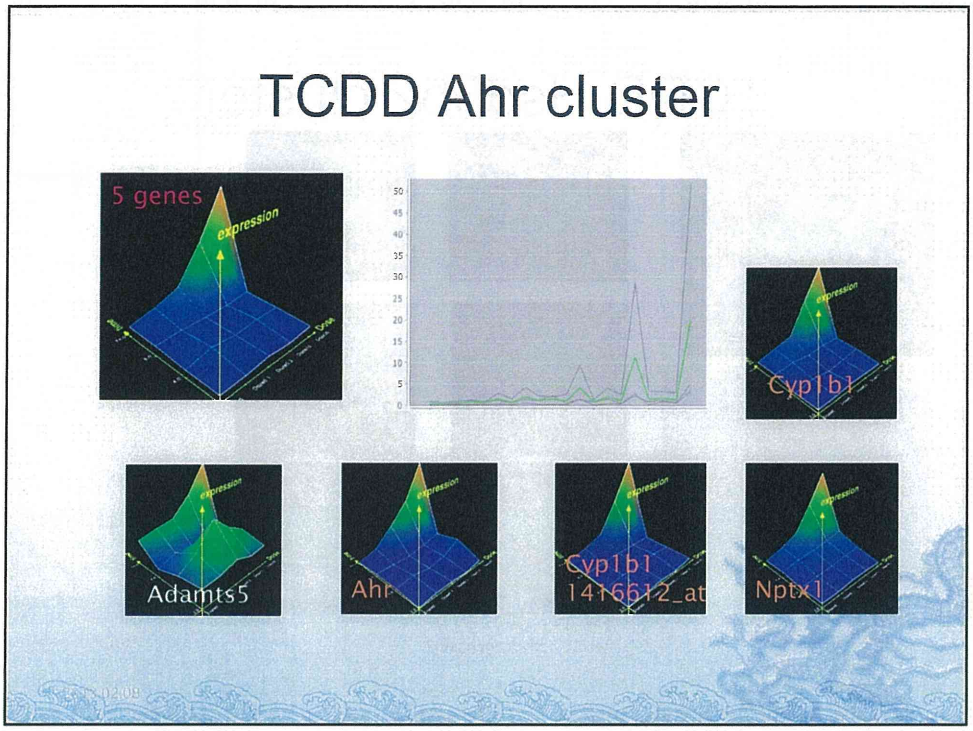
Cluster examples

Results on TCDD : 3117 probes in 498/1661 clusters

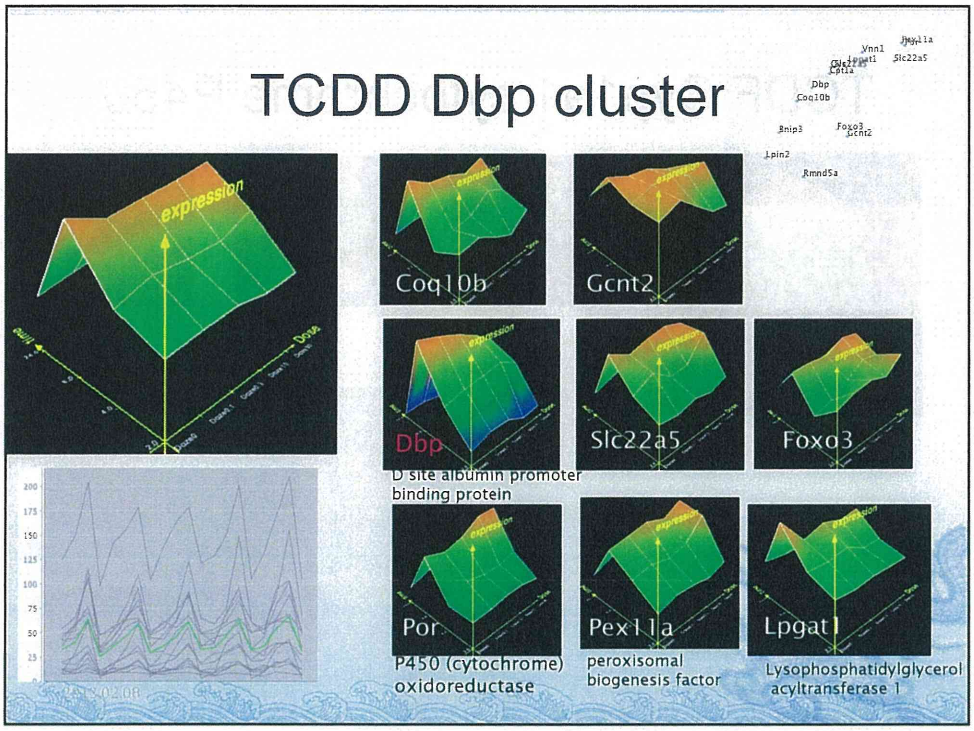
Manifold view

3D view (average profile)

2D view



TCDD Ahr cluster



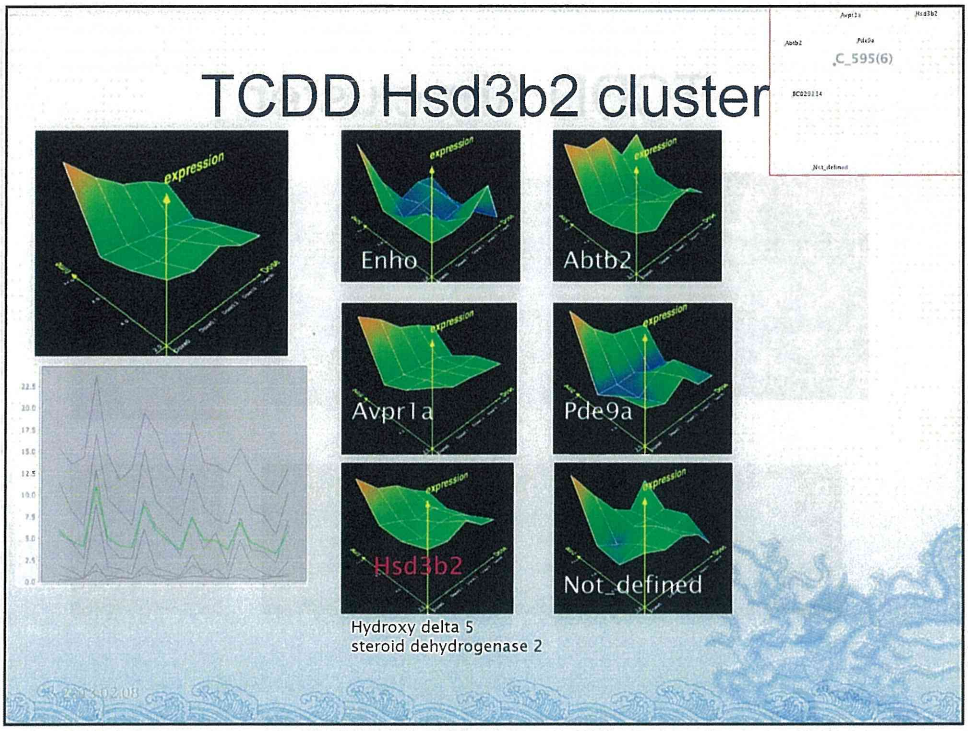
TCDD Dbp cluster

D site albumin promoter binding protein

P450 (cytochrome) oxidoreductase

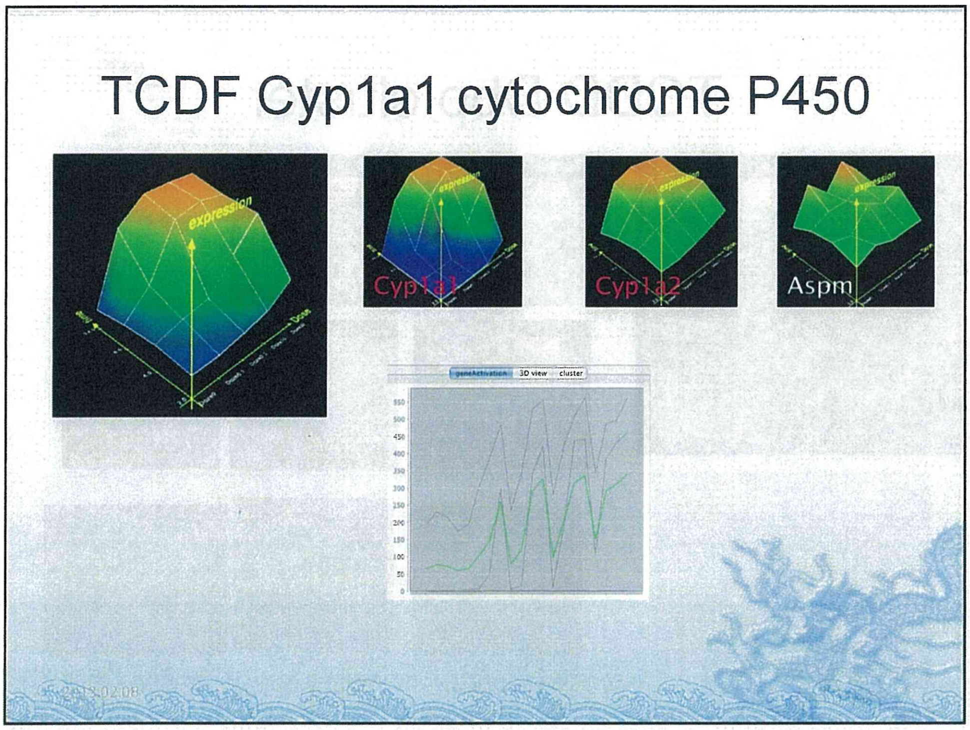
peroxisomal biogenesis factor

Lysophosphatidylglycerol acyltransferase 1

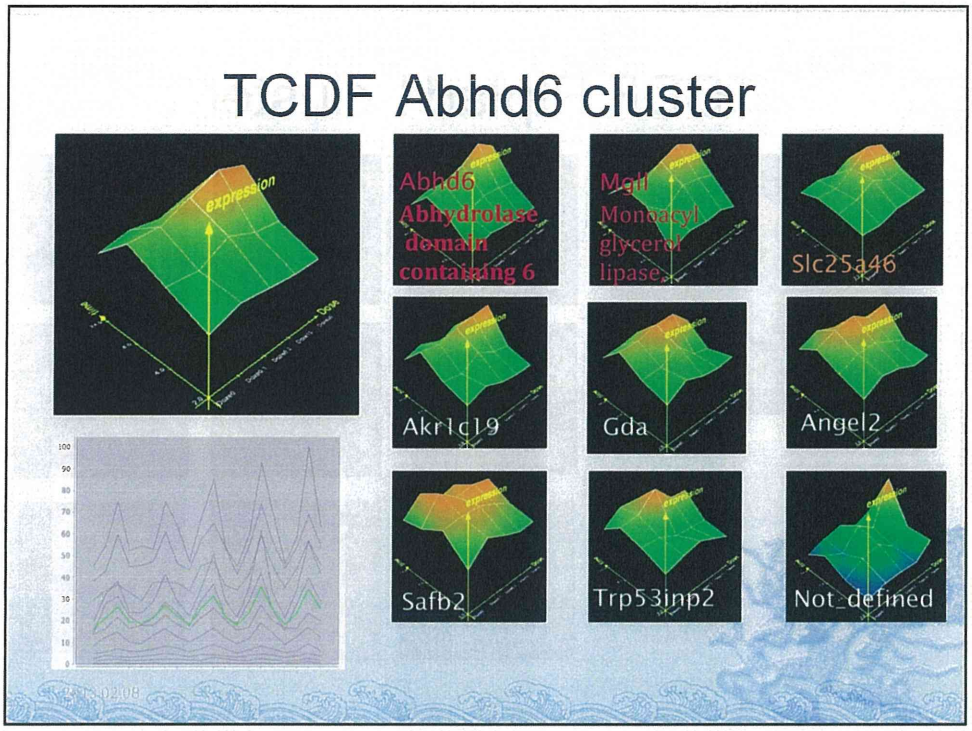


TCDD Hsd3b2 cluster

Hydroxy delta 5
steroid dehydrogenase 2

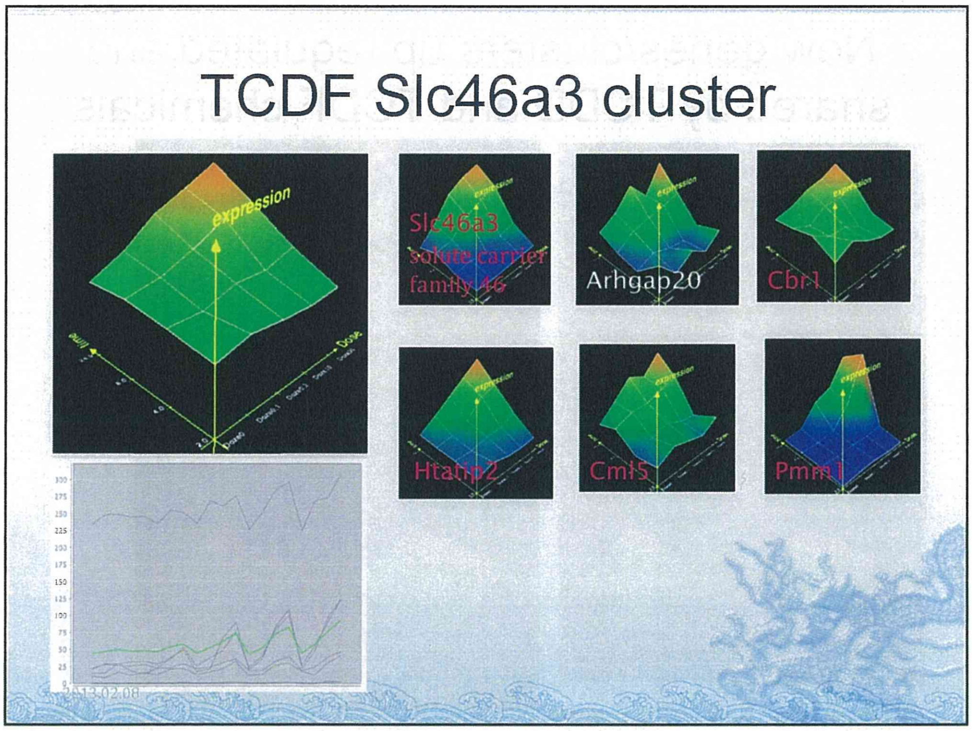


TCDF Cyp1a1 cytochrome P450



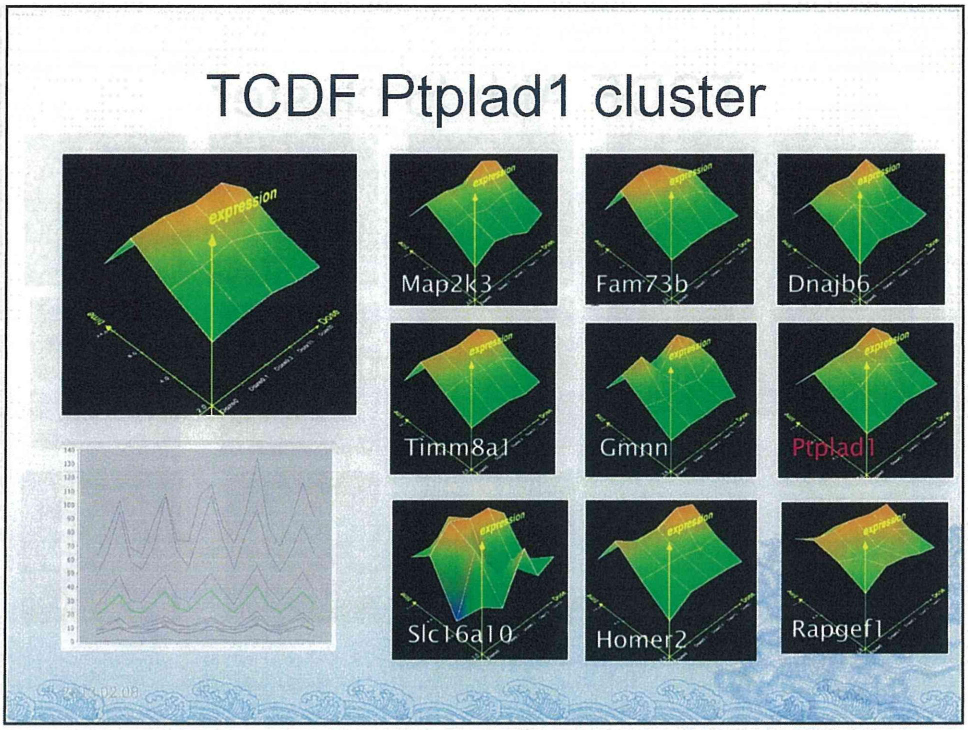
TCDF Abhd6 cluster

2013.02.08

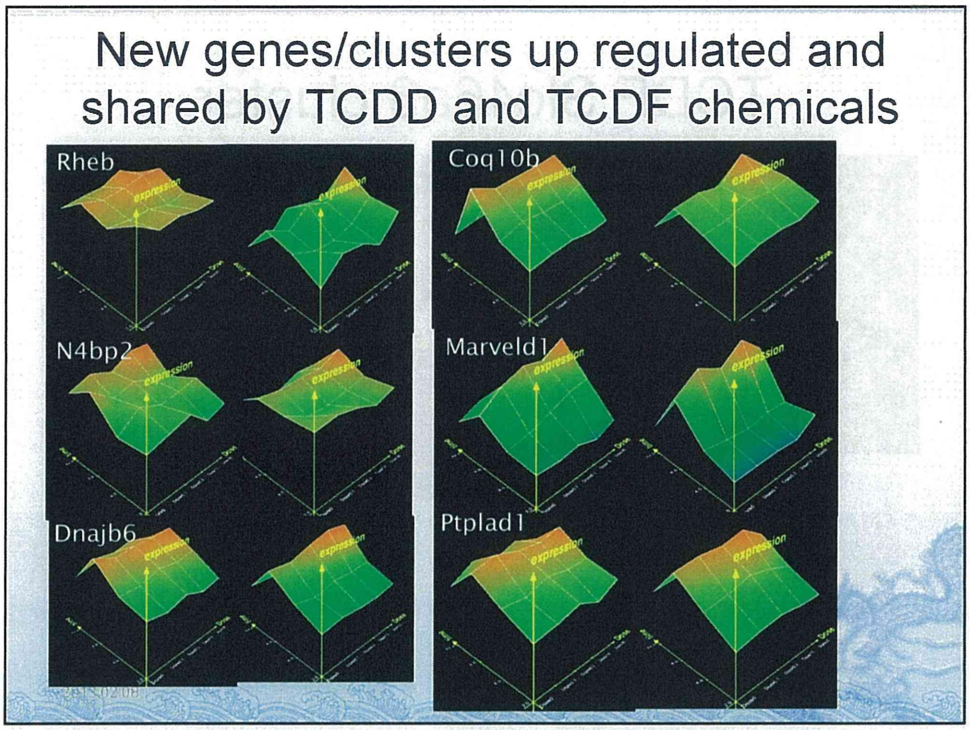


TCDF Slc46a3 cluster

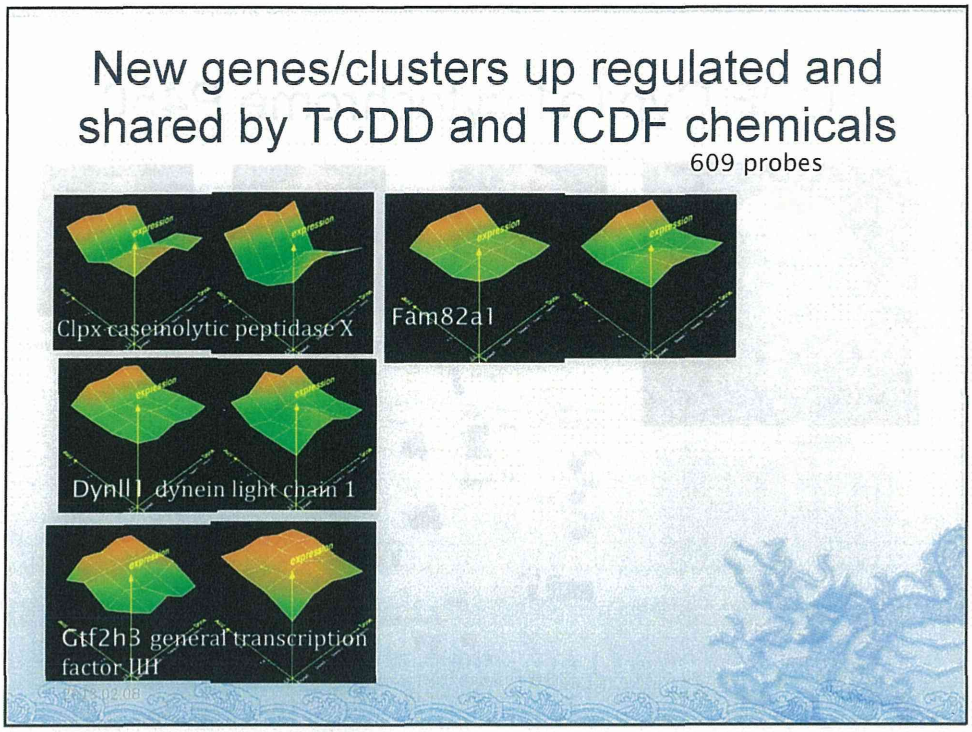
2013.02.08



TCDF Ptplad1 cluster



New genes/clusters up regulated and shared by TCDD and TCDF chemicals



New genes/clusters up regulated and shared by TCDD and TCDF chemicals

609 probes

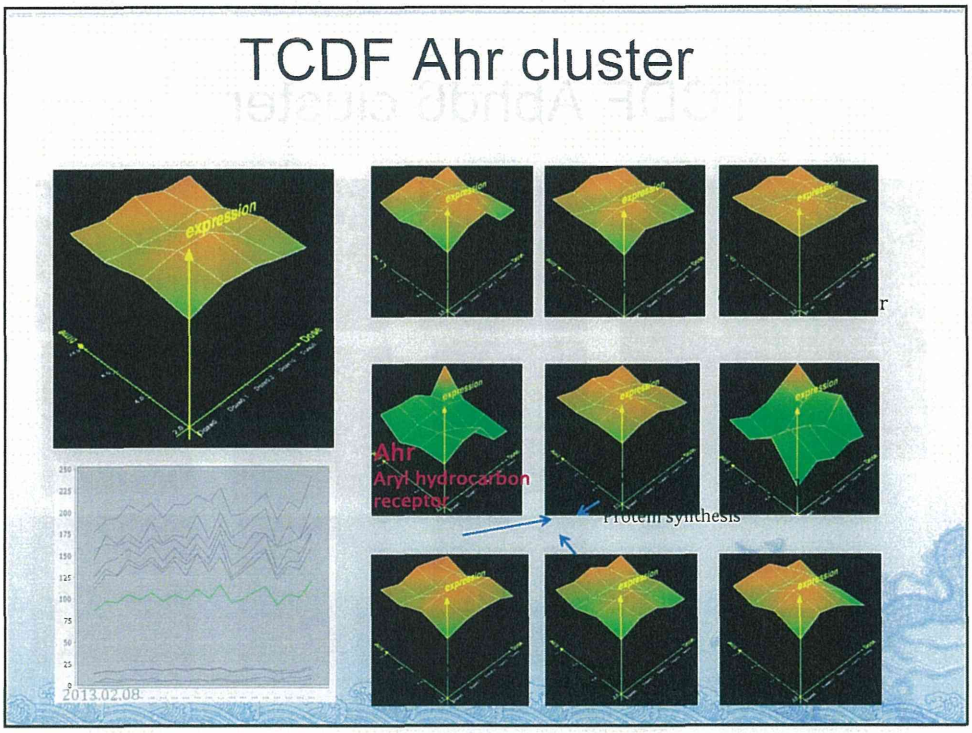
Clpx caseinolytic peptidase X

Fam82a1

Dynll1 dynein light chain 1

Gtf2h3 general transcription factor IIH

2013.02.08



TCDF Ahr cluster

Ahr
Aryl hydrocarbon receptor

Protein synthesis

r

2013.02.08