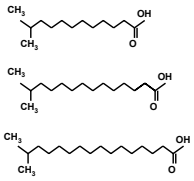
Finding functional gene clusters responsive to changes in the Monomethyl Branched-Chain Fatty Acid levels.

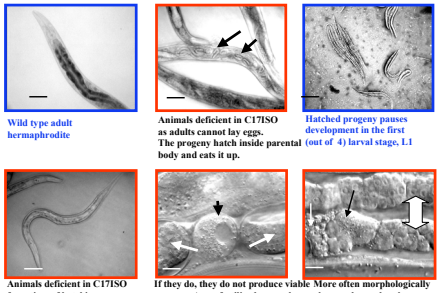
Andrew Knyazev (MATH, CU Denver) and Min Han (MCDB, CU Boulder)

Monomethyl Branched-Chain Fatty Acids



… are essential for reproductive growth
are needed at all developmental stages
are not physiologically identical
are sufficient in minute quantities
in *C. elegans*

Monomethyl Branched-Chain Fatty Acids Deficiency



Wild type adult hermaphrodite

Animals deficient in C17ISO as adults cannot lay eggs. The progeny hatch inside parental body and eats it up.

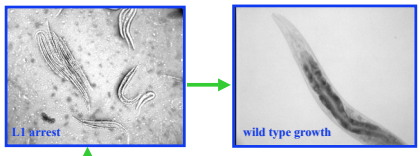
Hatched progeny pauses development in the first (out of 4) larval stage, L1

Animals deficient in C17ISO from time of hatching rare reach adulthood

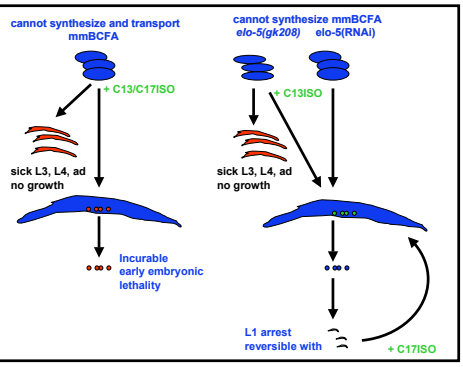
If they do, they do not produce viable progeny. An unfertilized egg and dead embryos shown

More often morphologically abnormal gonad make no oocytes

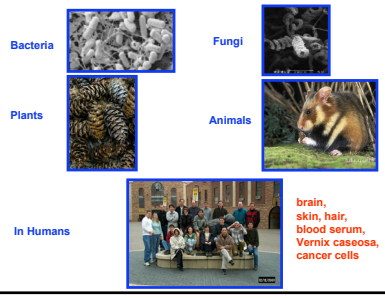
C15ISO and C17ISO supplements fed to the L1-arrested worms rescue them to full wild type growth.



C15ISO and C17ISO
But not C16:0, C16:1, C18:0, C18:1, C18:2 n6, C18:3 n3, C20:4 n6, C20:5 n3, or phytanic



Monomethyl Branched-Chain Fatty Acids have been detected in many living beings among



brain, skin, hair, blood serum, Vernix caseosa, cancer cells

Monomethyl Branched-Chain Fatty Acids were sporadically studied:

Chevreul (France, 1823): discovery of BCFA
Hradec (Czechoslovakia, 1971): C17anteISO functions in protein biosynthesis by binding eukaryotic peptide-initiation and -elongation factors
Jones (Australia, 1997): C20anteISO is major covalently bound FA in hair fiber
Yang (USA, 2000): C15ISO induces apoptotic death and inhibits growth of human cancer cells in vivo
Oku (Japan, 2002): C15ISO suppresses FA biosynthesis
Anti-tumor activity in human breast cancer

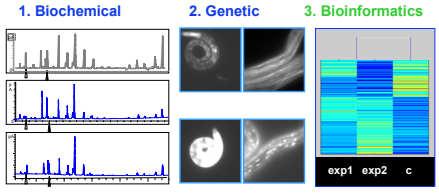
We have shown:
mmBCFA are physiologically important molecules that regulate essential function in eukaryotes (Kniazeva et al., PLoS, 2004)

Monomethyl Branched-Chain Fatty Acids

The **Main Goals** of the study are:
to understand the exceptional physiological importance of mmBCFA in *C. elegans* and to discern their emerging importance for the human health.

Approaches:

1. Biochemical
2. Genetic
3. Bioinformatics



3. Bioinformatics Aim:

identify clusters of genes that serve particular functions (functional clusters) and that are sensitive to mmBCFA

Preliminary data:

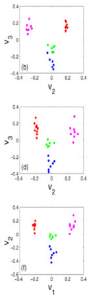
I. We identified genes that are
   1. essential for mmBCFA biosynthesis and transport
   2. involved in regulation of mmBCFA biosynthesis
   3. regulated by genetically induced deficiency of mmBCFA

II. We characterized mutants and/or RNAi strains that experience suppression of these genes

III. We collected microarray hybridization results for the mutants/RNAi strains and developmentally corresponding controls

IV. We identified publicly available the *C. elegans* microarray data for the future cluster analysis

V. We probed our already available microarray data with the newly developed bi-clustering techniques



Strategy:

Create a microarray data set that combines our experimental and publicly available microarray hybridization results

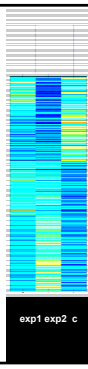
Perform cluster analysis using the bi-clustering weighted method with the gene annotation

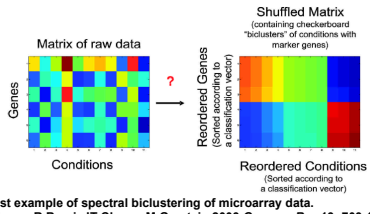
Compare the obtained results with those generated using common clustering methods such as K-means

Test stability of the clusters by introducing noise to the original microarray data and compare with other methods

Test relevance of the clusters experimentally using biochemical and genetic approaches

Once meaningful clusters are identified in *C. elegans*, the cluster analysis on publicly available data bases of different organisms, e.g., humans, will be performed in order to verify if a similar functionality exists there





A test example of spectral biclustering of microarray data.
Y Kluger, R Basri, JT Chang, M Gerstein 2003 *Genome Res* 13: 703-16.

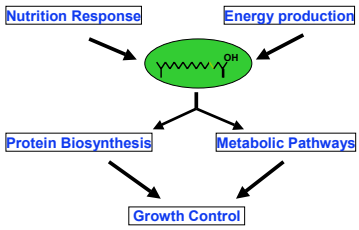
Spectral biclustering uses the Singular Value Decomposition to co-cluster the genes and the experiments simultaneously. Spectral biclustering is a special case of spectral clustering that relies on the Principal Components Analysis (PCA) to cluster the data. For large data sets, the PCA needs modern eigenvalue solvers, such as developed by Knyazev, and advanced computing resources.

The spectral clustering has been tested in image segmentation:

It is fully automatic, allows to introduce weights in the analysis, produces high quality clusters and can be used for very large mega-size datasets if matched with a proper software.

Multiplatform PCA computer codes in MATLAB, C and FORTRAN, including implementations for modern highly parallel computers are being written. They will be used for biclustering the microarray data in available databases against the data obtained for this project.

Expected Outcome:



Proposal © 2004 Knyazev and Han