Overview: Accession

Data Accessioning Program

http://open2dprot.sourceforge.net/Accession

Revised: 03-01-2004, P. Lemkin

Basic Open n-D Analysis Pipeline

1. Accession sample images or n-D data and experiment information — *interactive*
   XML
2. Segment n-D data to quantify or extract "spots" for all samples (2D-gels, 2D-LC-MS, etc.) — *batch*
   XML
3. Create a landmark database between reference sample and remaining samples by spot pairing algorithm *(if spot pairing is not automatic)* — *interactive*
   XML
4. Pair spots between a reference sample and the rest of samples — *batch*
   XML
5. Construct Composite Database, CSD, by merging paired spot lists — *batch*
   RDBMS and caches
6. Explore the CSD data using exploratory data analysis techniques: statistics, clustering, direct-manipulation graphics and reports, etc. — *interactive* / *batch*

Introduction

- Data must be accessioned or entered into the program database. This consists of several activities:
- 1. Entering the name of the sample and sample information
- 2. Entering the Region Of Interest of sample
- 3. Entering sample calibration info if any

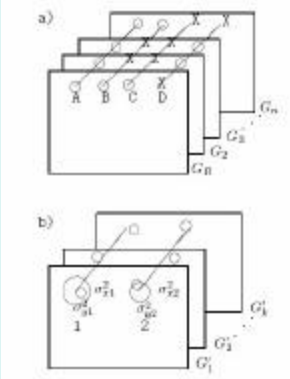
2

Initial Open n-D Data-Mining Tools

- Accession n-D sample images or n-D data and experiment data
- Quantify 'spots' from sample images or peptide clusters
- Pair spots between samples and a reference sample
- Construct composite sample database for exploratory data analysis
- Manage subsets of proteins in the database
- Manage replicate samples and condition sets of samples
- Analyze expression profiles for multiple conditions
- Data-filter protein sets by statistics, clustering, set membership
- Direct-manipulation of data in graphics, spreadsheets
- Integrate R language statistical, clustering, classifiers, class prediction, and other methods
- Integrate access to Internet proteomic/genomic/function data servers for user-specified protein sets

5

Composite Samples Database (CSD) Paradigm



Proteomic composite samples database (CSD) consisting of a set of n samples $G_1, G_2, \ldots, G_n$ with representative sample $G_r = G_1$

Expression profiles A,B,C, ...

A canonical sample database is a statistical representation of the CSD spot geometry and quantification that could be used for data mining

in Lemkin *et al.*, *Computers Biomedical Research*, 1981

3

Open2Dprot Pipeline Subprojects

6

01-09-2005

Associated or Related Projects

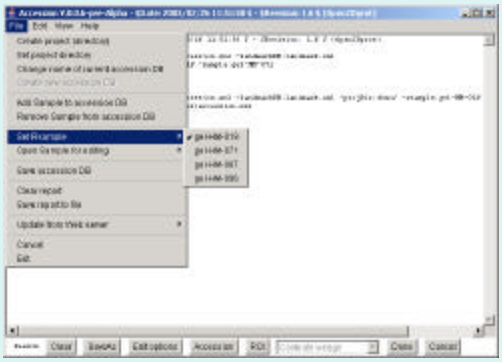
We had added some additional non-pipeline open source projects that may use similar data or common software modules. They may be useful for performing other types of analysis on data used by Open2Dprot or alternate types of analysis.

Contributed Project Name	Download	Documentation	Overview (PDF)	PDF documents	Version	Revision history	Status
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

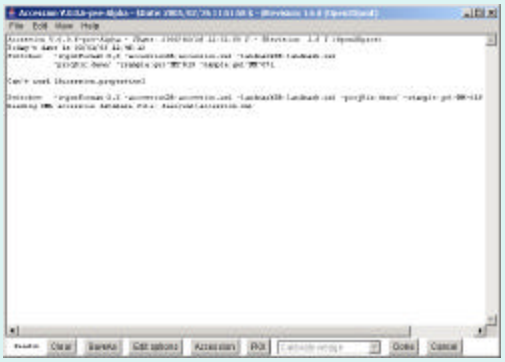
09-12-2004

7

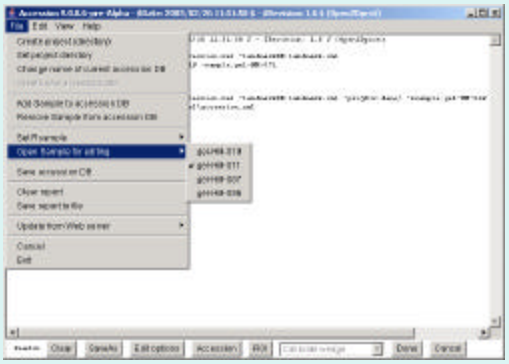
File Menu - Select Reference Sample



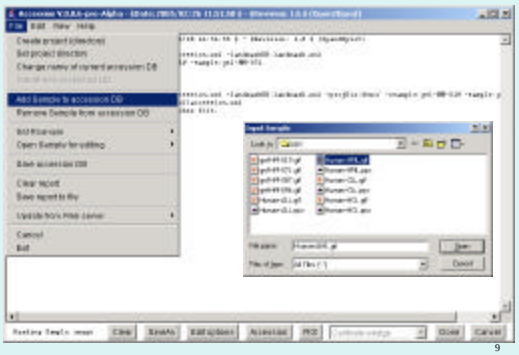
Main Window



File Menu - Select Current Sample

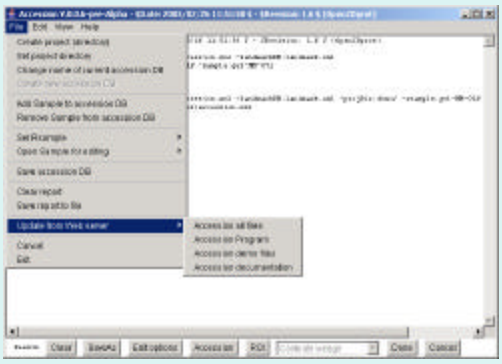


File Menu – Add New Sample

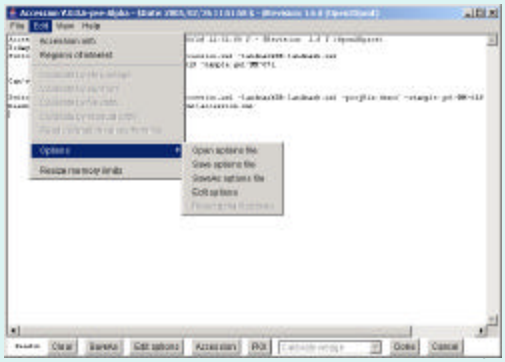


9

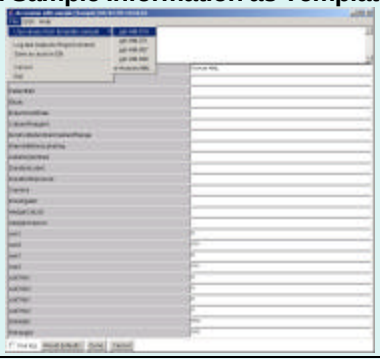
File Menu - Update From Web Server



Edit Menu



Edit Menu – Accession New Sample Using Old Sample Information as Template

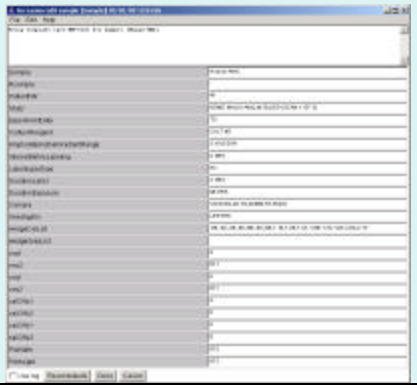


16

Edit Menu - Accession

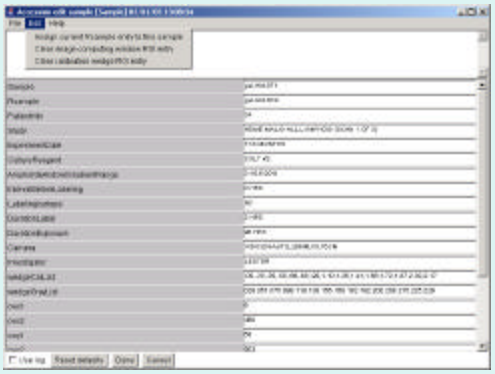
14

Edit Menu – Accession Sample after Template

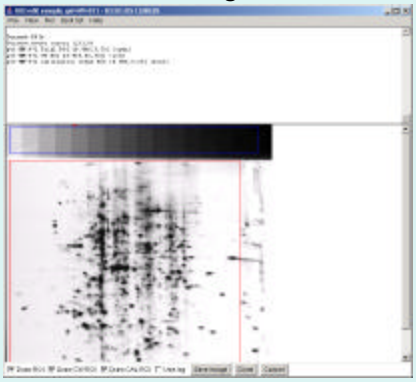


17

Edit Menu – Accession editing options

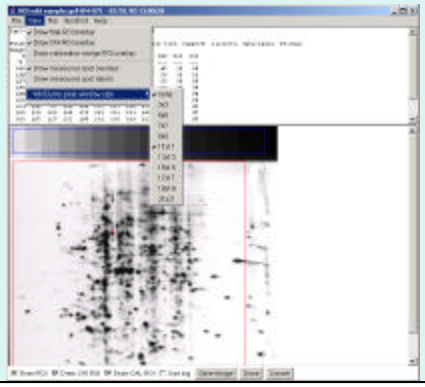


Edit Menu – Region Of Interest



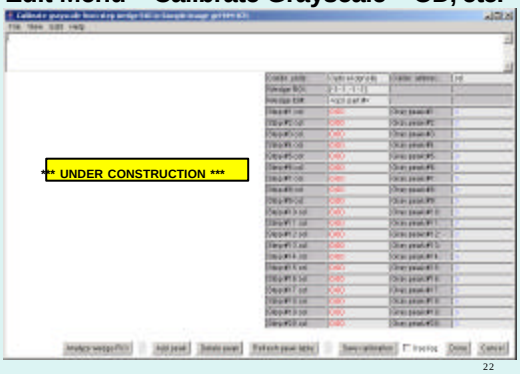
18

Edit Menu – Region Of Interest overlay views



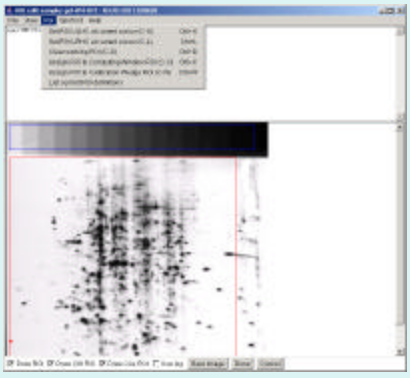
19

Edit Menu – Calibrate Grayscale – OD, etc.



22

Edit Menu – Region Of Interest commands

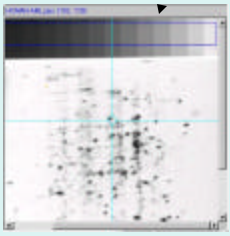


20

Flicker - Calibrating Grayscale with a ND Step-Wedge

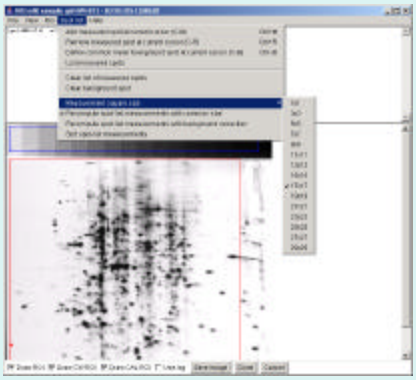
1. The ND step wedge must be scanned with the image and the corresponding OD values known for each step
2. A region of interest (ROI) is overlaid on the step step-wedge
3. The ND wedge calibration wizard is invoked to analyze the data and estimate the calibration

ROI over step-wedge



23

Edit Menu – Region Of Interest Spot List

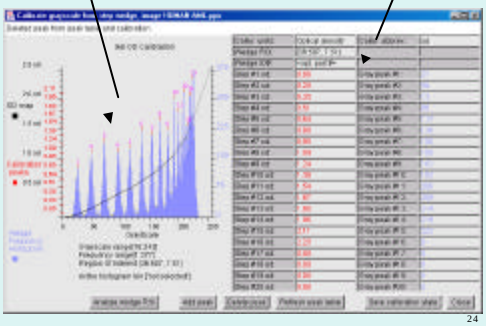


21

Flicker - Calibrating Grayscale from ND Wedge Data

ROI histogram, peaks found and extrapolated calibration curve

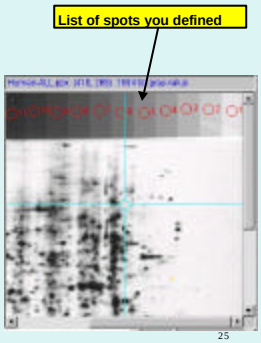
OD vs gray-peaks table



24

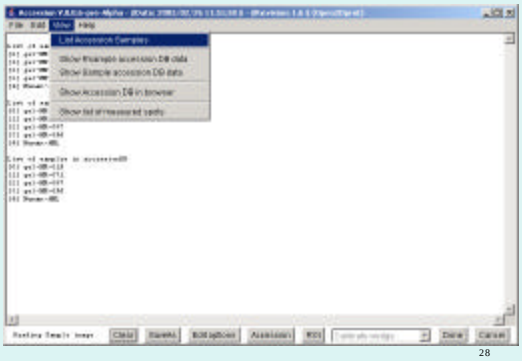
Flicker- Calibrating Grayscale with a Spot List of Calibrated Data

1. The image must contain calibrated regions with known concentrations or corresponding OD values known for each spot
2. You define a set of spots using (C-M) or (ALT-click)
3. The Spot List Calibration wizard is invoked to analyze the data and estimate the calibration



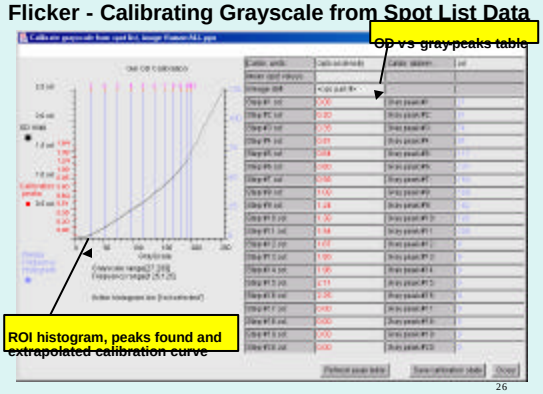
25

View Menu



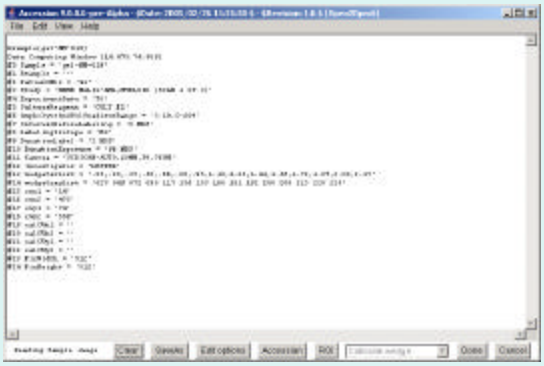
28

Flicker - Calibrating Grayscale from Spot List Data

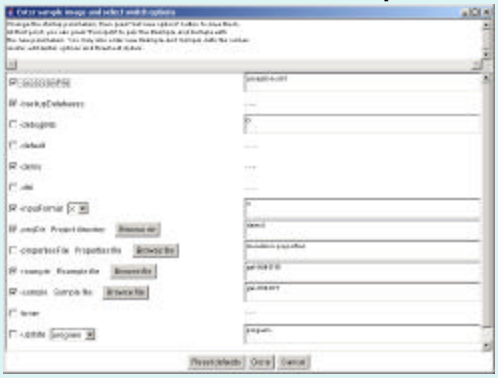


26

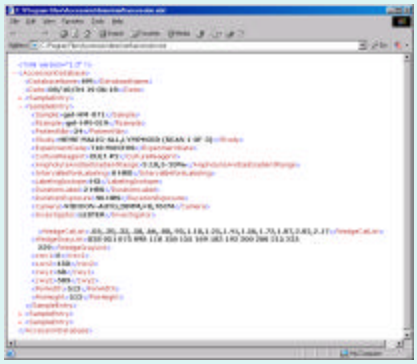
View Menu - Show Current Sample Acc DB Data



Edit Menu - command line options

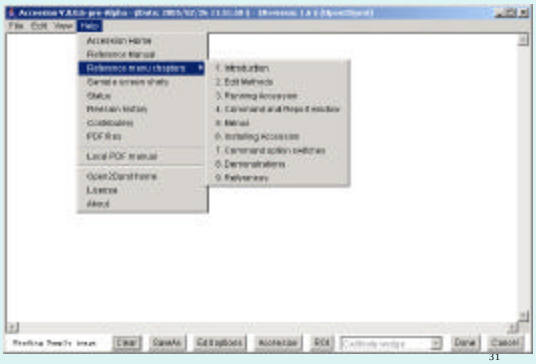


View Menu - XML Accession DB in Browser



30

Help Menu



31

Summary

- Accession is an open-source sample accessioning Java program freely available at **http://open2dprot.sourceforge.net/Accession**
- Useful for adding sample experiment information, regions of interest and grayscale calibration (if available).
- It will be used as one of the step [1] alternative modules in the analysis pipeline in the Open2Dprot project at **http://open2dprot.sourceforge.net**

32