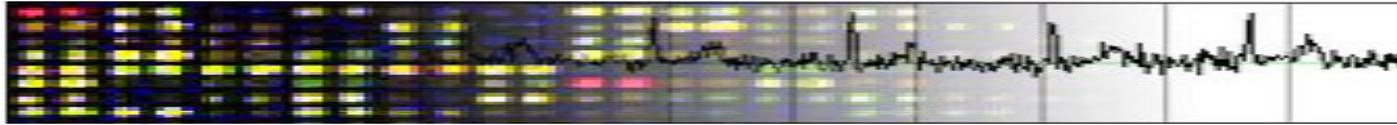


Biomedical Information Technology

2.771J BEH.453J HST.958J Spring 2005

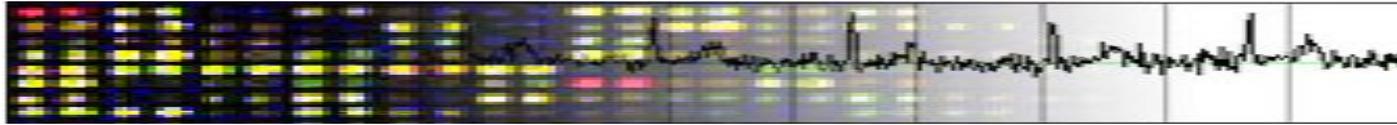
Lecture 5 April 2005

Gel Electrophoresis



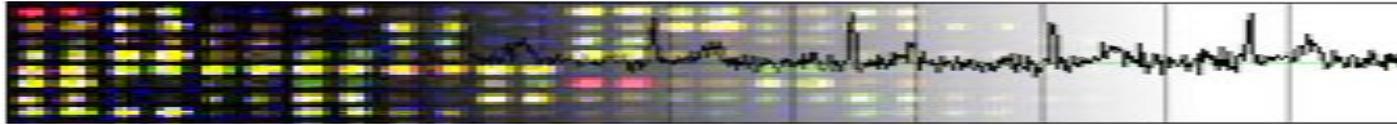
Gel Electrophoresis

- ❖ *Statement of experimental problem*
- ❖ *Defining experimental information objects*
- ❖ *The origins of ExperiBase*
- ❖ *Design of database schema*
- ❖ *Keeping raw image data*
- ❖ *Interaction with external analysis packages*
- ❖ *Creating individual tables and functions*
- ❖ *User query functions for applications*



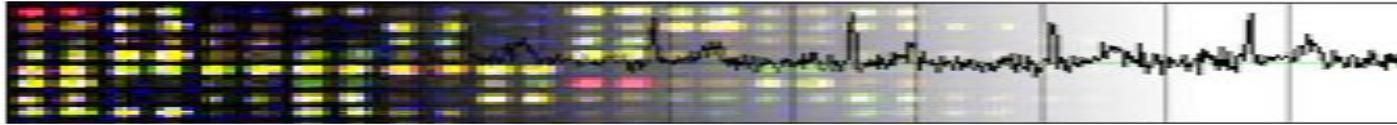
Statement of the experimental problem

- ❖ *Identify the key experimental results and their natural electronic form (text, image, . . .)*
- ❖ *Identify all parameters that define the experimental conditions*
- ❖ *Create a complete set of parameters to define the sample that is being analyzed*
- ❖ *Design the information object to accommodate changes in the experimental protocol*



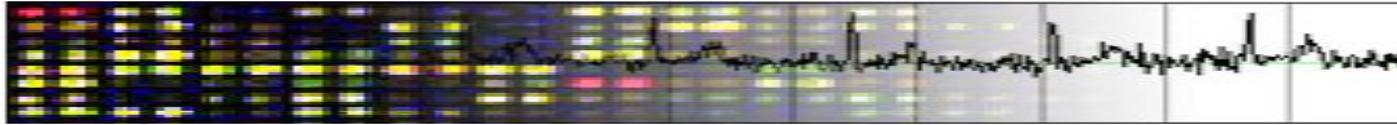
Generalize your customer

- ❖ *What new items will the next researcher have?*
- ❖ *What will the next generation of equipment demand?*
- ❖ *What queries will be run against the data?*
- ❖ *Will you support metadata (computed results) as well?*



Define representative use cases

- ❖ *What procedure will an experimenter perform?*
- ❖ *Who will query the data?*
- ❖ *What queries will be run against the data?*
- ❖ *What user interfaces are available?*
- ❖ *What user interface functionality is required?*



Development process (from NCICB)

- ❖ Initial high level Use Cases were developed for gene components
- ❖ Additional high level Use Cases were added as additional functional areas are mapped (e.g. pathways, therapies, microarray objects, mouse models)
- ❖ Detailed Use Cases are derived from working with domain experts in requirements gathering.

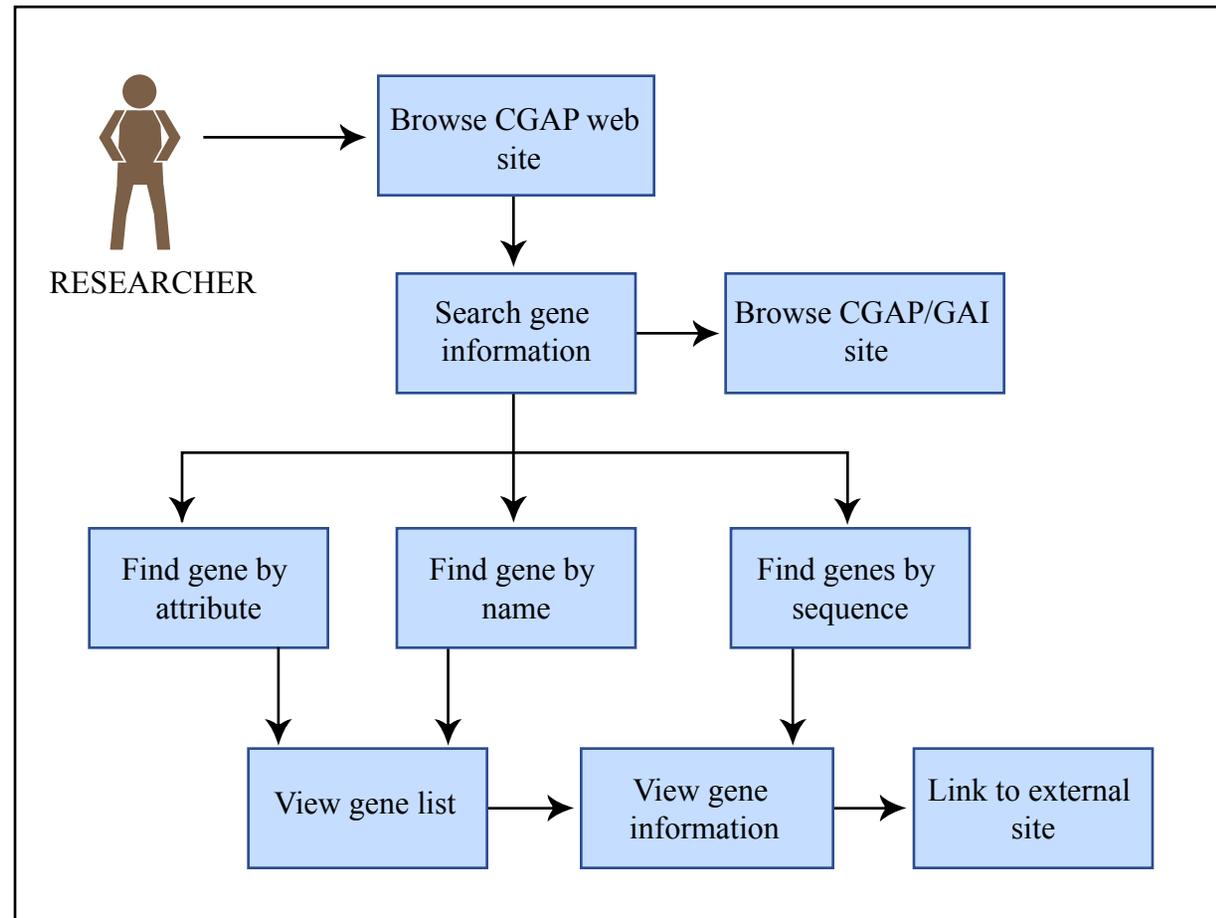
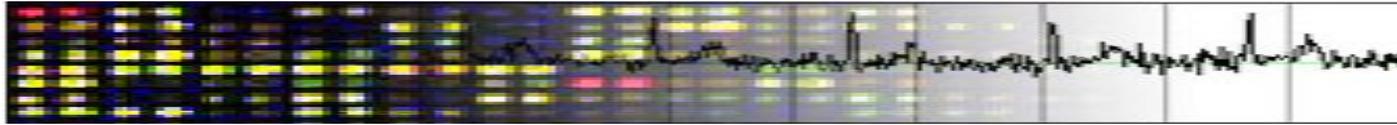


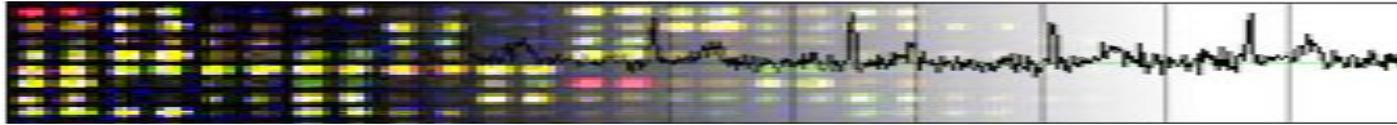
Figure by MIT OCW.



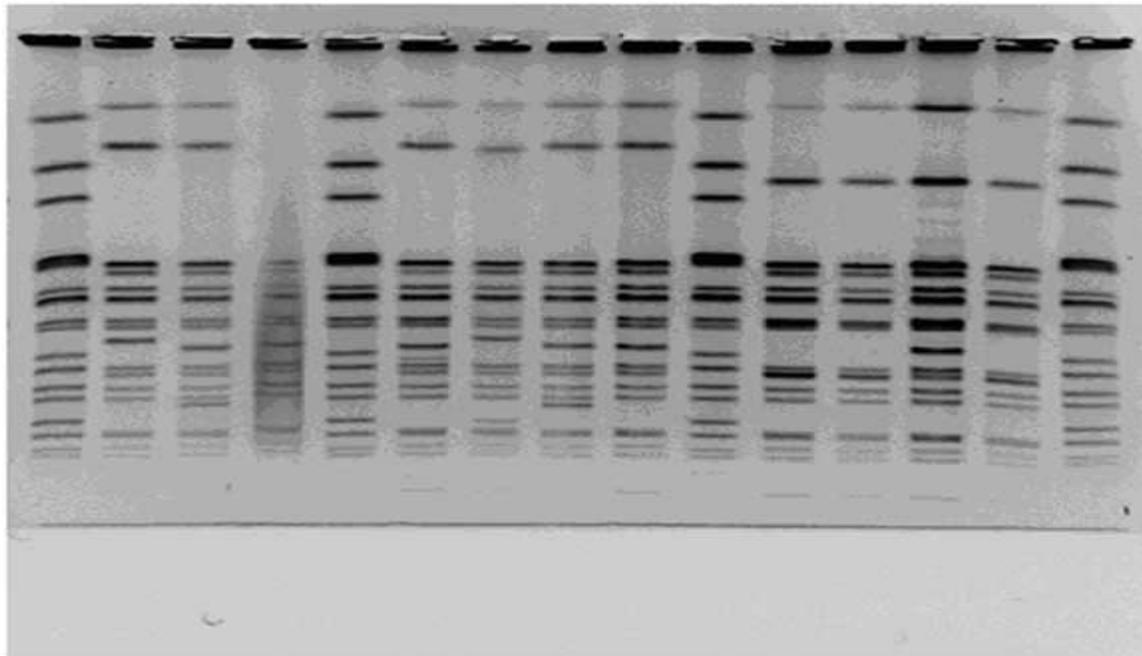
Background on Experibase

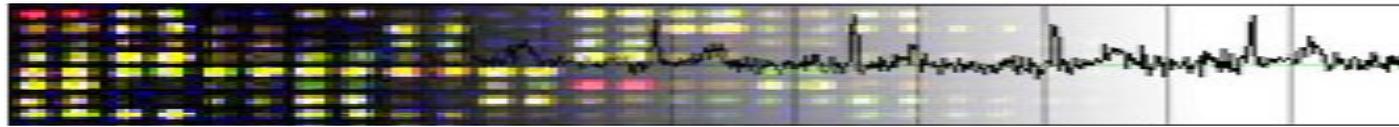
Experibase is conceived as an integrated software platform capable of supporting a variety of experimental methods found in modern biology. It provides: comprehensive database features for searching, classifying, and supporting; web-based client services; data import and export capabilities to accommodate other data repositories; and direct support for metadata produced by analysis programs.

Because Experibase includes many different experimental methods, and the semantic contexts that support them, the creation of this software platform will provide a major contribution to normalizing semantic support for these different methods.

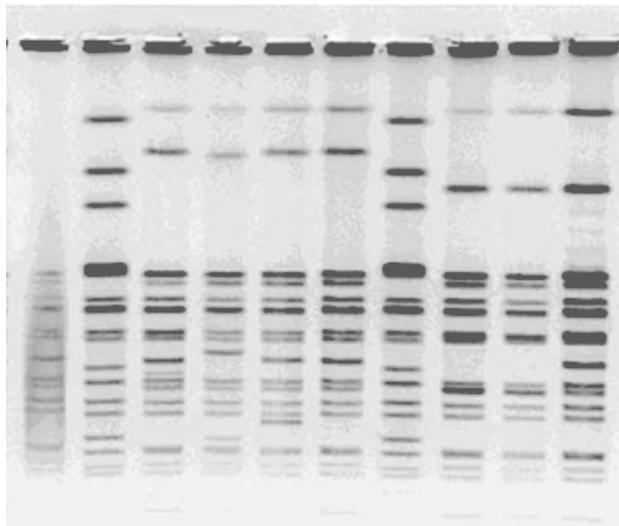


Gel electrophoresis – the first modality of ExperiBase

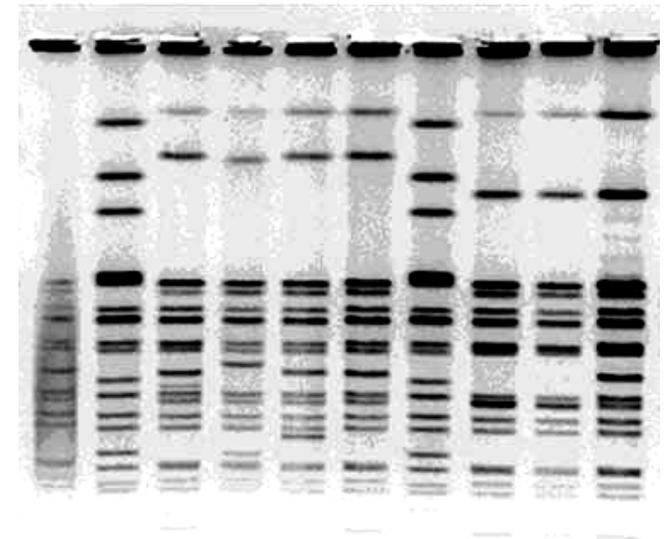




Manipulation of the gel images is required



More contrast
➔
Less intensity





Work flow for gel electrophoresis

Image analyzed
Analysis saved with object

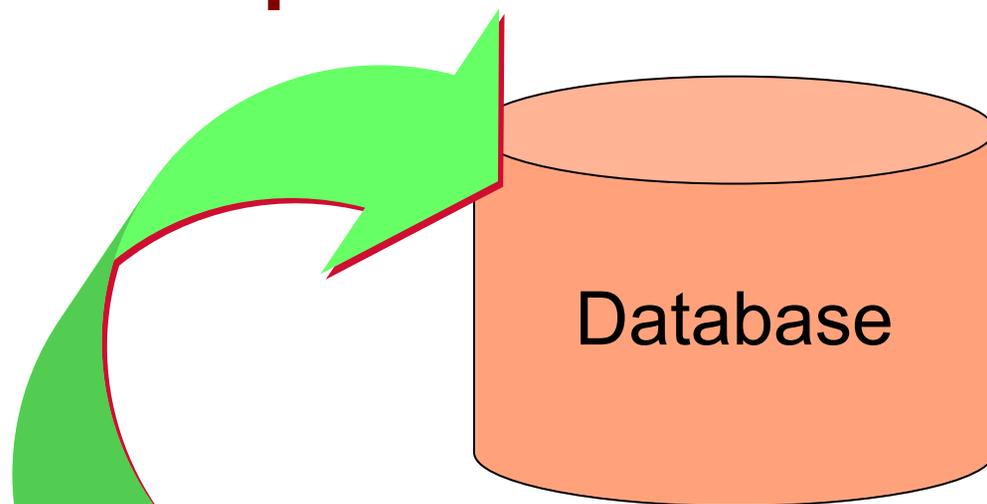
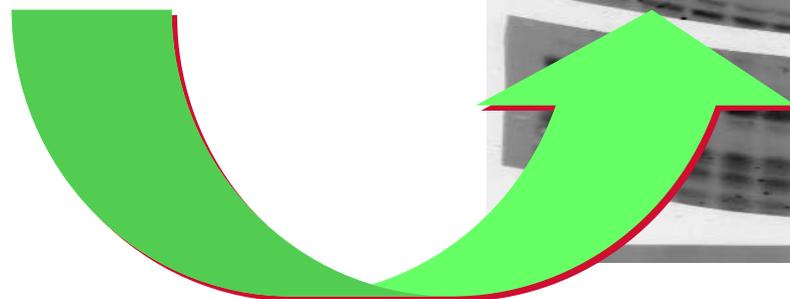
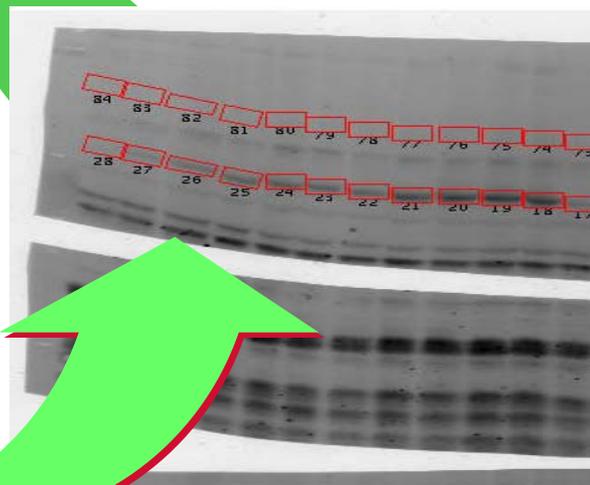


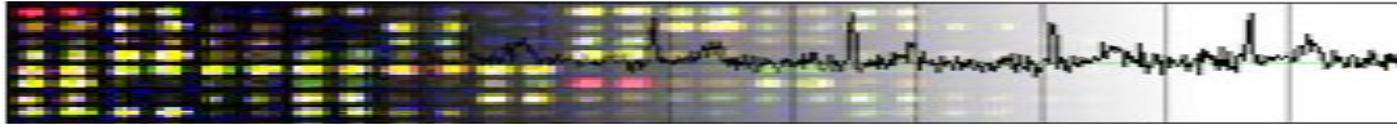
Photo of SDS-Page gel plate
removed for copyright reasons.



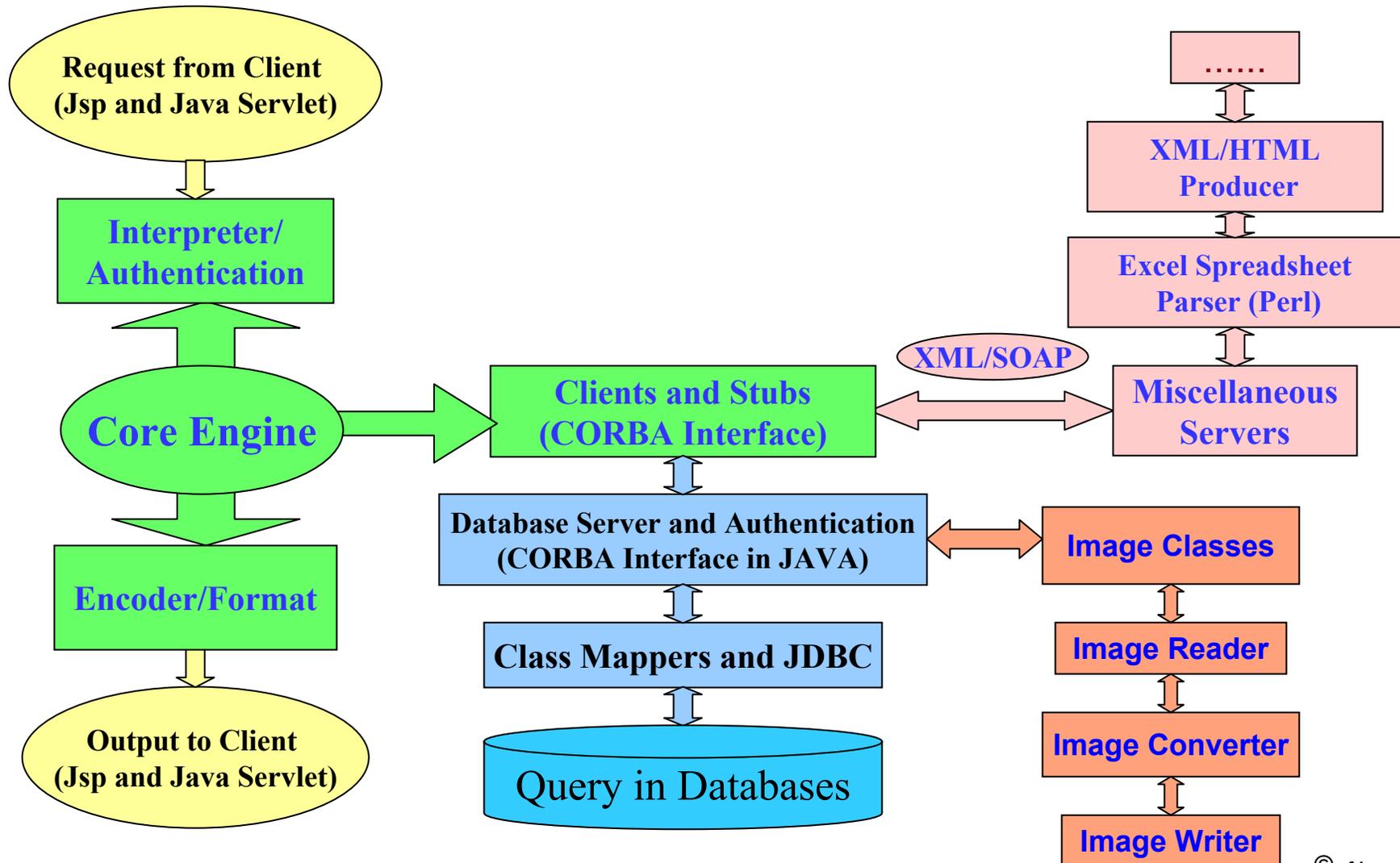


Classes of information for gel electrophoresis

- ❖ *Sample information*
 - *Source*
 - *Description*
- ❖ *Instrument description*
 - *Manufacturer & model*
 - *Settings – time, gel, voltage, etc.*
- ❖ *Preconditioning*
 - *Process method*
 - *Process materials*
 - *Process duration*
- ❖ *Raw image data*
- ❖ *Post-processing protocols*
- ❖ *Derived data from image*

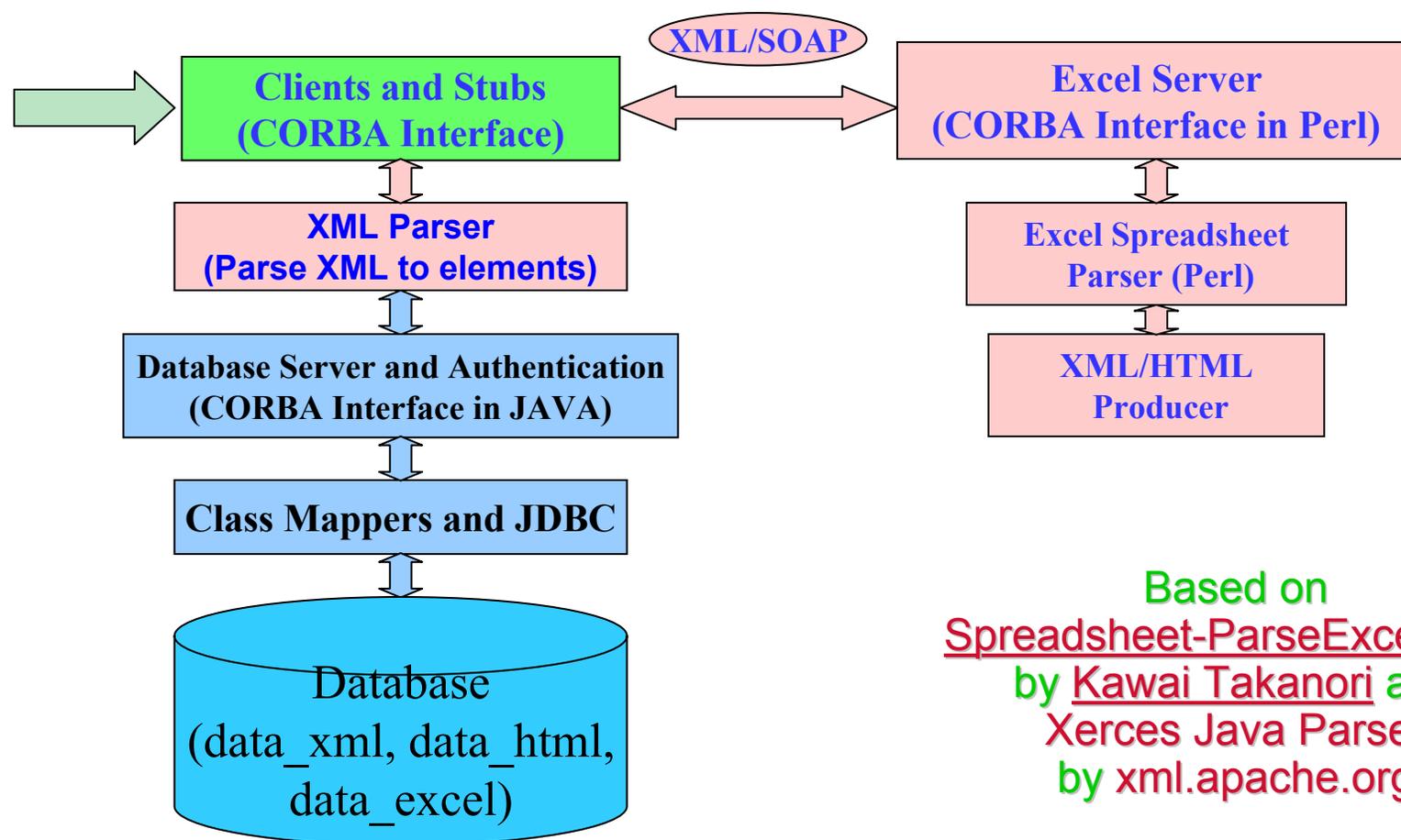


Platform independent Web-enabled user interface

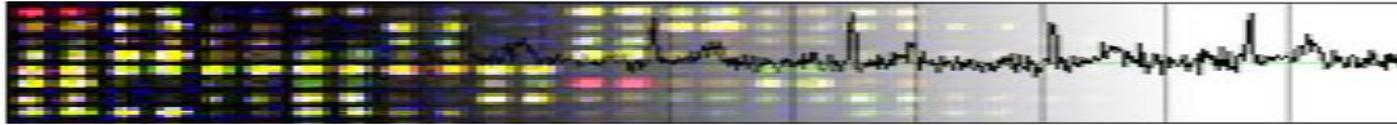




Excel to XML Parser

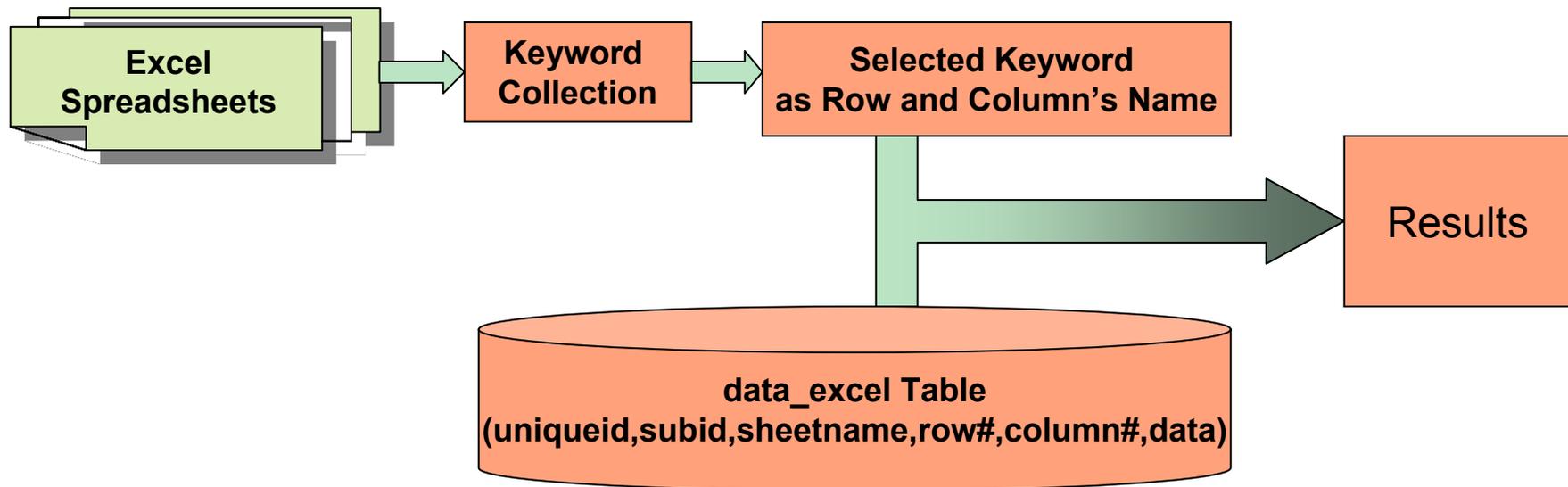


Based on
Spreadsheet-ParseExcel-0.26
by Kawai Takanori and
Xerces Java Parser
by xml.apache.org



Customized Data Selection

--choose interesting columns and rows





Keyword Collections for Excel Spreadsheets

| protocol | word | wordcount |
|-----------------|-------------|------------------|
| kinase_assay | Akt (norm) | 4 |
| kinase_assay | ERK2 (norm) | 4 |
| kinase_assay | JNK1 (norm) | 4 |
| kinase_assay | Replicate # | 4 |
| kinase_assay | Time (min) | 4 |

| | | |
|-------------|------------------|----|
| westernblot | Dec7 correction | 1 |
| westernblot | Height | 14 |
| westernblot | High/MEK | 2 |
| westernblot | IFN 24h | 62 |
| westernblot | IFN 24h, TNF 12h | 62 |

Customized Data Selection

| | | |
|----------------------------|-----------|---|
| The experimental method is | Protocol: | westernblot |
| Choose the experiments | UniqueID: | <ul style="list-style-type: none"> ICMITuniquePrefix_1_HT-29 MEK1 Western Blot Data ICMITuniquePrefix_20_Normalization_actin_data ICMITuniquePrefix_21_P-AKT_AKT_data ICMITuniquePrefix_19_Time course ICMITuniquePrefix_22_P-FKHR_FKHR_P-AFX_AFX_data <li style="background-color: #000080; color: white;">ICMITuniquePrefix_24_P-JNK_data ICMITuniquePrefix_25_P-MEK_MEK_data <li style="background-color: #000080; color: white;">ICMITuniquePrefix_26_TNF-R1_data |
| Choose the columns | Column: | <ul style="list-style-type: none"> UV 60' <li style="background-color: #000080; color: white;">Volume <li style="background-color: #000080; color: white;">Average <li style="background-color: #000080; color: white;">Std. Dev. Min. Int. <li style="background-color: #000080; color: white;">Median Name <li style="background-color: #000080; color: white;">Area Max. Int. Time |
| Choose the rows | Row: | <ul style="list-style-type: none"> IFN 24h, TNF+ins 20h <li style="background-color: #000080; color: white;">IFN 24h, TNF+ins 60' <li style="background-color: #000080; color: white;">IFN 24h, TNF+ins 16h <li style="background-color: #000080; color: white;">IFN 24h, TNF+ins 8h IFN 24h IFN 24h, TNF 12h IFN 24h, TNF+ins 90' IFN 24h, TNF+ins 24h IFN 24h, TNF 4h IFN 24h, TNF+ins 12h |

submit

Results of customized data selection

ExperiBase database - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print

Address http://schiele.mit.edu:8080/research/experibase.htm Go Links >>

ExperiBase Database

----Viewer Page

- [Search in ExperiBase](#)
- [Customized Selection](#)
- [List All Experiments](#)
- [List All Protocols](#)
- [List All Cells](#)
- [List All Chemicals](#)
- [List All Process Methods](#)
- [MAIN PAGE](#)

Request:

```
uniqueid=ICMITuniquePrefix_25_P-MEK_MEK_data
column=Volume
rows=IFN 24h, TNF 60'
rows=IFN 24h, TNF 20h
```

Query Results:

In raw_data_excel, there are:

In proc_data_excel, there are:

```
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 9 1 IFN 24h, TNF 60'
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 59 1 IFN 24h, TNF 60'
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 16 1 IFN 24h, TNF 20h
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 66 1 IFN 24h, TNF 20h
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 1 4 Volume
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 51 4 Volume
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 9 4 5935467.43
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 59 4 5935467.43
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 16 4 4901947.91
ICMITuniquePrefix_25_P-MEK_MEK_data 0 MEK 66 4 4901947.91
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 9 1 IFN 24h, TNF 60'
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 39 1 IFN 24h, TNF 60'
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 16 1 IFN 24h, TNF 20h
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 46 1 IFN 24h, TNF 20h
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 1 4 Volume
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 31 4 Volume
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 9 4 3347762.01
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 39 4 1567023.17
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 16 4 2978229.21
ICMITuniquePrefix_25_P-MEK_MEK_data 0 P-MEK 46 4 1058747.22
ICMITuniquePrefix_25_P-MEK_MEK_data 0 Reanalyzed Oct23 MEK 10 1 IFN 24h, TNF 60'
ICMITuniquePrefix_25_P-MEK_MEK_data 0 Reanalyzed Oct23 MEK 40 1 IFN 24h, TNF 60'
ICMITuniquePrefix_25_P-MEK_MEK_data 0 Reanalyzed Oct23 MEK 17 1 IFN 24h, TNF 20h
ICMITuniquePrefix_25_P-MEK_MEK_data 0 Reanalyzed Oct23 MEK 47 1 IFN 24h, TNF 20h
ICMITuniquePrefix_25_P-MEK_MEK_data 0 Reanalyzed Oct23 MEK 1 4 Volume
```

ICMIT

Done Internet

Start | Inboxes - Outlook Express | Eudora - [In] | ExperiBase database ... | untitled - Paint | Microsoft PowerPoint - [...]

2:03 PM 2005

Users can view HTML and XML format of Excel spreadsheets

ExperiBase database - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://schiele.mit.edu:8080/research/experibase.htm> Go Links >>

ExperiBase Database

----Viewer Page

[Search in ExperiBase](#)

[Customized Selection](#)

[List All Experiments](#)

[List All Protocols](#)

[List All Cells](#)

[List All Chemicals](#)

[List All Process Methods](#)

[MAIN PAGE](#)

Download:

filename: TNFR1_big_Aug16_01.xls

filetype: application/vnd.ms-excel

fileLength: 27648

Please click the "Download" to confirm downloading the file.

[Download](#)

Please click the "View HTML" to view the HTML format of the Excel file.

[View HTML](#)

Please click the "Download XML" to download the XML of the Excel file.

[Download XML](#)

ICMIT

Start | Internet | Inbox - Outlook Express | Eudora - [In] | ExperiBase database ... | untitled - Paint | Microsoft PowerPoint - [...]

2:04 PM

HTML format of Excel spreadsheet

ExperiBase database - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://schiele.mit.edu:8080/research/experibase.htm>

ExperiBase Database

---Viewer Page

[Search in ExperiBase](#)

[Customized Selection](#)

[List All Experiments](#)

[List All Protocols](#)

[List All Cells](#)

[List All Chemicals](#)

[List All Process Methods](#)

[MAIN PAGE](#)

File Name: /disk3/javaServlet/webapps/research/tempfiles/TNFR1_big_Aug16_01.xls

Number of Sheets: 2

Sheet Name: TNFR1

| Sample | Time | Actin | TNFR1 | Background | TNFR1-BG | TNFR1-BG/actin |
|--------------------|------|-------------------|------------|------------|------------|------------------|
| UV 60' | | 1 | 5020157.77 | 1402850.72 | 3617307.05 | 3617307.05 |
| Insulin 2h | | 1.04365451818392 | 6377401.53 | 1432943.87 | 4944457.66 | 4737638.34089842 |
| IFN 24h | 0' | 0.899334856449538 | 3467356.4 | 1311942.47 | 2155413.93 | 2396675.62592793 |
| IFN 24h, untreated | | 0.890245541082783 | 3368563.85 | 1272649.69 | 2095914.16 | 2354310.20238618 |
| IFN 24h, TNF 5' | 5' | 0.835358599751162 | 3826144.45 | 1368897.56 | 2457246.89 | 2941547.36748023 |
| IFN 24h, TNF 15' | 15' | 0.989184708659882 | 4497633.93 | 1400106.41 | 3097527.52 | 3131394.46342275 |
| IFN 24h, TNF 30' | 30' | 1.03012590715752 | 4088877.35 | 1338636.02 | 2750241.33 | 2669810.8560233 |
| IFN 24h, TNF 60' | 60' | 0.977973019090633 | 3167623.18 | 1301482.17 | 1866141.01 | 1908172.28448207 |
| IFN 24h, TNF 90' | 90' | 0.934683964988254 | 3476619.17 | 1338998.85 | 2137620.32 | 2286997.95874519 |
| IFN 24h, TNF 2h | 2h | 0.916268906331079 | 4199541.8 | 1410079.72 | 2789462.08 | 3044370.55620446 |
| IFN 24h, TNF 4h | 4h | 0.867008945778204 | 4165832.19 | 1417914.48 | 2747917.71 | 3169422.55714967 |
| IFN 24h, TNF 8h | 8h | 0.777416212042982 | 3439693.26 | 1357684.66 | 2082008.6 | 2678113.17508888 |
| IFN 24h, TNF 12h | 12h | 0.71768148759466 | 3276979.47 | 1324928.66 | 1952050.81 | 2719940.31299648 |
| IFN 24h, TNF 16h | 16h | 0.723969784287343 | 3763710.13 | 1403387.55 | 2360322.58 | 3260250.12538809 |
| IFN 24h, TNF 20h | 20h | 0.76119686642846 | 3856582.86 | 1363990.31 | 2492592.55 | 3274570.16697304 |

ICMIT

http://schiele.mit.edu/temp/tempfiles/TNFR1_big_Aug16_01.xls.html

Start | In-box - Outlook Express | Eudora - [In] | ExperiBase database ... | untitled - Paint | Microsoft PowerPoint - [...]

Internet 2:05 PM

XML format of Excel spreadsheet

ExperiBase Database

----Viewer Page

[Search in ExperiBase](#)

[Customized Selection](#)

[List All Experiments](#)

[List All Protocols](#)

[List All Cells](#)

[List All Chemicals](#)

[List All Process Methods](#)

[MAIN PAGE](#)

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
- <DOCUMENT>
  <FILE>/disk3/javaServlet/webapps/research/tempfiles/TNFR1_big_Aug16_01.xls</FILE>
  <COUNT>2</COUNT>
- <SHEETS>
  - <SHEET>
    <SHEETNAME>TNFR1</SHEETNAME>
  - <RECORDS>
    - <RECORD>
      <ROW>1</ROW>
      <COLUMN>1</COLUMN>
      <DATA>Sample</DATA>
    </RECORD>
    - <RECORD>
      <ROW>1</ROW>
      <COLUMN>2</COLUMN>
      <DATA>Time</DATA>
    </RECORD>
    - <RECORD>
      <ROW>1</ROW>
      <COLUMN>3</COLUMN>
      <DATA>Actin</DATA>
    </RECORD>
    - <RECORD>
      <ROW>1</ROW>
      <COLUMN>5</COLUMN>
      <DATA>TNFR1</DATA>
    </RECORD>
    - <RECORD>
      <ROW>1</ROW>
      <COLUMN>6</COLUMN>
      <DATA>Background</DATA>
    </RECORD>
    - <RECORD>
      <ROW>1</ROW>
      <COLUMN>7</COLUMN>
      <DATA>TNFR1-BG</DATA>
    </RECORD>
    - <RECORD>
      <ROW>1</ROW>
      <COLUMN>8</COLUMN>
      <DATA>TNFR1-BG/actin</DATA>
    </RECORD>
  </RECORDS>
</SHEET>
</SHEETS>
</DOCUMENT>
```

ICMIT

Done

Start | InBox - Outlook Express | Eudora - [In] | ExperiBase database ... | untitled - Paint | Microsoft PowerPoint - [...] | 2:05 PM

List All Experiments

ExperiBase database - Microsoft Internet Explorer

Address: http://schiele.mit.edu:8080/research/experibase.htm

ExperiBase Database

---Viewer Page

- [Search in ExperiBase](#)
- [Customized Selection](#)
- [List All Experiments](#)
- [List All Protocols](#)
- [List All Cells](#)
- [List All Chemicals](#)
- [List All Process Methods](#)
- [MAIN PAGE](#)

| | | | | |
|---|---|---|------------|--------------------------|
| ICMITuniquePrefix_20_Normalization_actin_data | Normalization_actin_data | raw data: (actin_May2201_900.gel,sg_dec04_01_750_530_488.gel); Processed data: (Actin_May22.xls, Actin_750_Dec4_01.xls, Actin_Dec07_01.xls) | suzanne | 2001-05-22 12:00:00.0 |
| ICMITuniquePrefix_4_Kevin's Kinase Assay experiment | Kevin's Kinase Assay experiment | There are 2 different treatments (EGF, insulin) for kinase assay experiment. Here is the EGF treated. Another is in Kevin's kinase assay (EGF treated); 7 timepoints (0, 5, 10, 20, 30, 60, 90 min.); 3 kinases (ERK2, Akt, JNK1); 3 replicates | kevinjanes | 2002-04-05 12:00:00.0 |
| ICMITuniquePrefix_6_Kevin's Kinase Assay experiment (insulin treated) | Kevin's Kinase Assay experiment (insulin treated) | There are 2 different treatments (EGF, insulin) for kinase assay experiment. Here is the insulin treated. Another is in Kevin's kinase assay (EGF treated); 7 timepoints (0, 5, 10, 20, 30, 60, 90 min.); 3 kinases (ERK2, Akt, JNK1); 3 replicates | kevinjanes | 2002-04-05 12:00:00.0 |
| ICMITuniquePrefix_21_P-AKT_AKT_data | P-AKT_AKT_data | raw_data (AKT_PAKT_Casp3_Casp7_May2101_900.gel, AKT_PAKT_IKB_PIKB_Jun0801_900.gel); process_data (AKT-06-08-01_forDB.xls) | suzanne | 2001-06-08 12:00:00.0 |
| ICMITuniquePrefix_19_Time course | Time course | Probing for 3 different targets | suzanne | 2002-04-22 12:00:00.0 |
| ICMITuniquePrefix_22_P-FKHR_FKHR_P-AFX_AFX_data | P-FKHR_FKHR_P-AFX_AFX_data | raw_data (P-AKT_P-FKHR_May1001.gel); process_data (P-AKT_P-FKHR_May10.xls, P-AKT_P-FKHR_May22.xls, P-AKT_P-FKHR_May22b.xls) | suzanne | 2001-05-10 12:00:00.0 |
| ICMITuniquePrefix_24_P-JNK_data | P-JNK_data | raw_data: (Casp3_Casp7_PJNK_JNK_May2901_900.gel, suz_aug9_01_800_590_pjnk.gel); process_data: (P-JNK_Aug9_01.xls, PJNK_May29_01.xls) | suzanne | 2001-05-29 12:00:00.0 |
| ICMITuniquePrefix_25_P-MEK_MEK_data | P-MEK_MEK_data | raw_data: (MEK_PMEK_May1001_750.gel, MEK_PMEK_May1001_900.gel, sg_dec07_01_110um_488_530_750.gel); process_data: (MEK_10_5_01_750.xls, MEK_10_5_01_900.xls, P-MEK_Dec7_01.xls, P-Mek-Dec4_750.xls) | suzanne | 2001-05-10 12:00:00.0 |
| ICMITuniquePrefix_26_TNF-R1_data | TNF-R1_data | raw_data: (sg_16_08_01_700_tnfr1.gel, sg_16_08_01_700_tnfr1_small.gel); process_data: (TNF-R1_16_08_01_700_tnfr1_small.xls, TNF-R1_16_08_01_700_tnfr1.xls) | suzanne | 2001-08-16 12:00:00.0 |

ICMIT

Done

Start | Internet Explorer | Outlook Express | Eudora - [In] | ExperiBase database ... | 1:58 PM

© cfedewey 2005