Micro-array data analysis
(Practical)
Value added Course
Lecture 9

Er. Brijendra Singh
Assistant Professor, Department of Bioinformatics
UIET, CSJMU University, Kanpur

Microarray Data analysis

•Microarray technology is a laboratory approach that involves binding an array of thousands to millions of known nucleic acid fragments to a solid surface, referred to as a “chip.” The chip is then bathed with DNA or RNA isolated from a study sample (such as cells or tissue)

•The development of DNA microarray technology in the mid-1990s allowed for the first time to simultaneously profile and study the transcriptome, in other words, to study cells’ real-time “chatter” in more detail. That is, microarrays are a tool for gene expression analysis.

•The first DNA arrays were constructed by immobilizing cDNAs onto filter paper.

Conti…

- A basic protocol for using a DNA microarray for gene expression profiling.

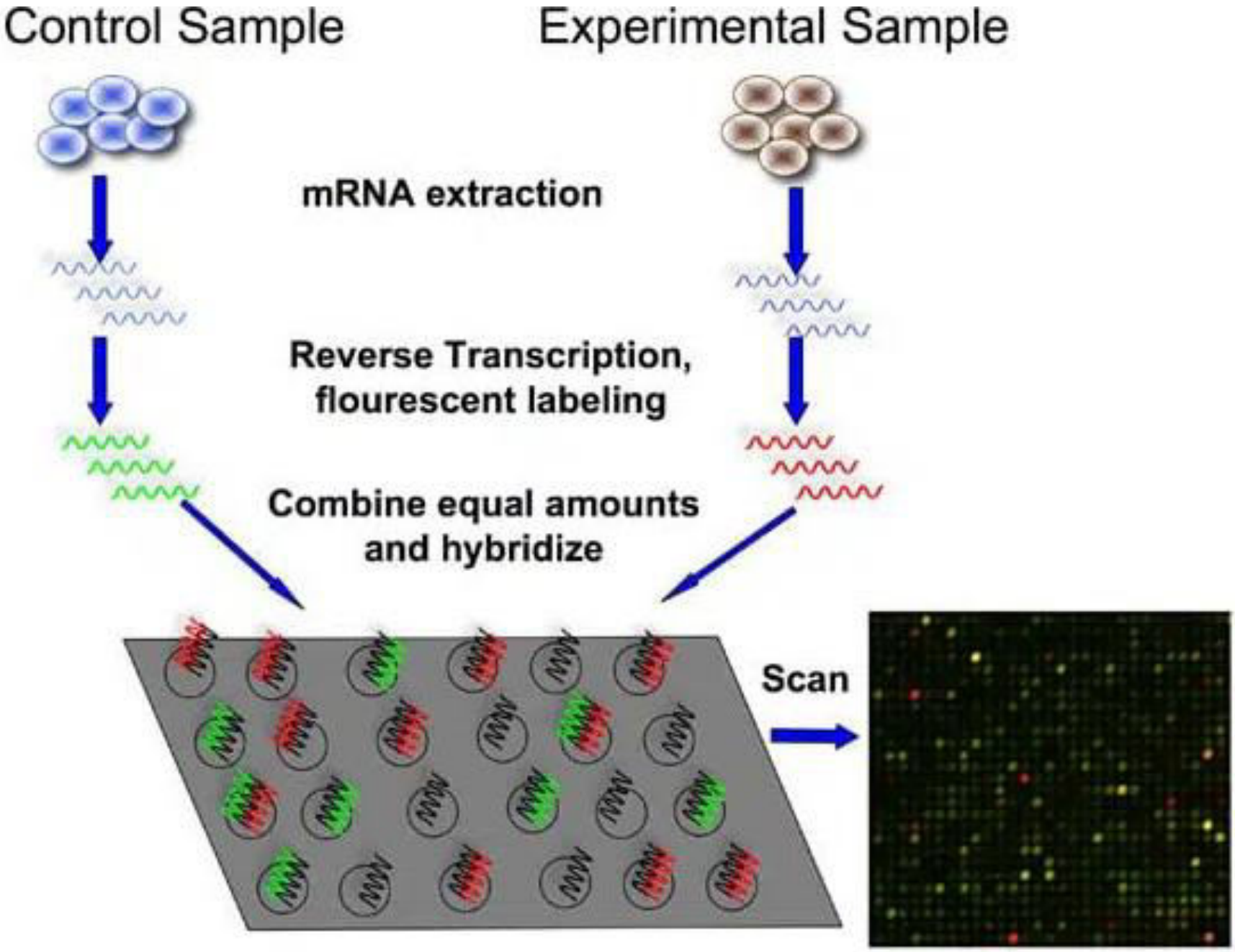
i) Isolate and Purify - The first step is to isolate and purify mRNA from samples of interest. Since we are interested in comparing gene expression, one sample usually serves as a control, and another sample would be the experiment (e.g., healthy vs. disease)

ii) Reverse Transcription and Labeling - The next step is to reverse transcribe and label the mRNA. In order to detect the transcripts by hybridization, they need to be labeled, and because starting material may be limited, an amplification step is also used.

iii) Hybridization - Now it's time to hybridize the labeled target to the microarray. This step involves placing labeled cDNAs onto a DNA microarray where they will hybridize to their synthetic complementary DNA probes attached to the microarray.

iv) Scanning and Quantitation - The fluorescent tags on the bound cDNA are excited by a laser and the fluorescently labeled target sequences that bind to a probe generate a signal.

- The total fluorescent intensity of the signal depends upon the amount of target sample binding to the probes present on that spot.
- Thus, the amount of target sequence bound to each probe correlates to the expression level of various genes expressed in the sample. The signals are detected, the signal intensity is quantified, and used to create a digital image of the array.


Control Sample
Experimental Sample
mRNA extraction
Reverse Transcription,
flourescent labeling
Combine equal amounts
and hybridize
Scan

Tools and Software's

- Array Mining
- Genesis
- GEPAS (gene expression pattern analysis)
- Genomic-Space
- SAGE (serial analysis of gene expression)
- Web Array

Note: Data can be collected GEO databases and Stanford Databases.

schdule - brijendra.bioinfo@gn | Online Microarray Analysis | Arraymining.net - Workflow | DNA Microarrays: An Easy Expl

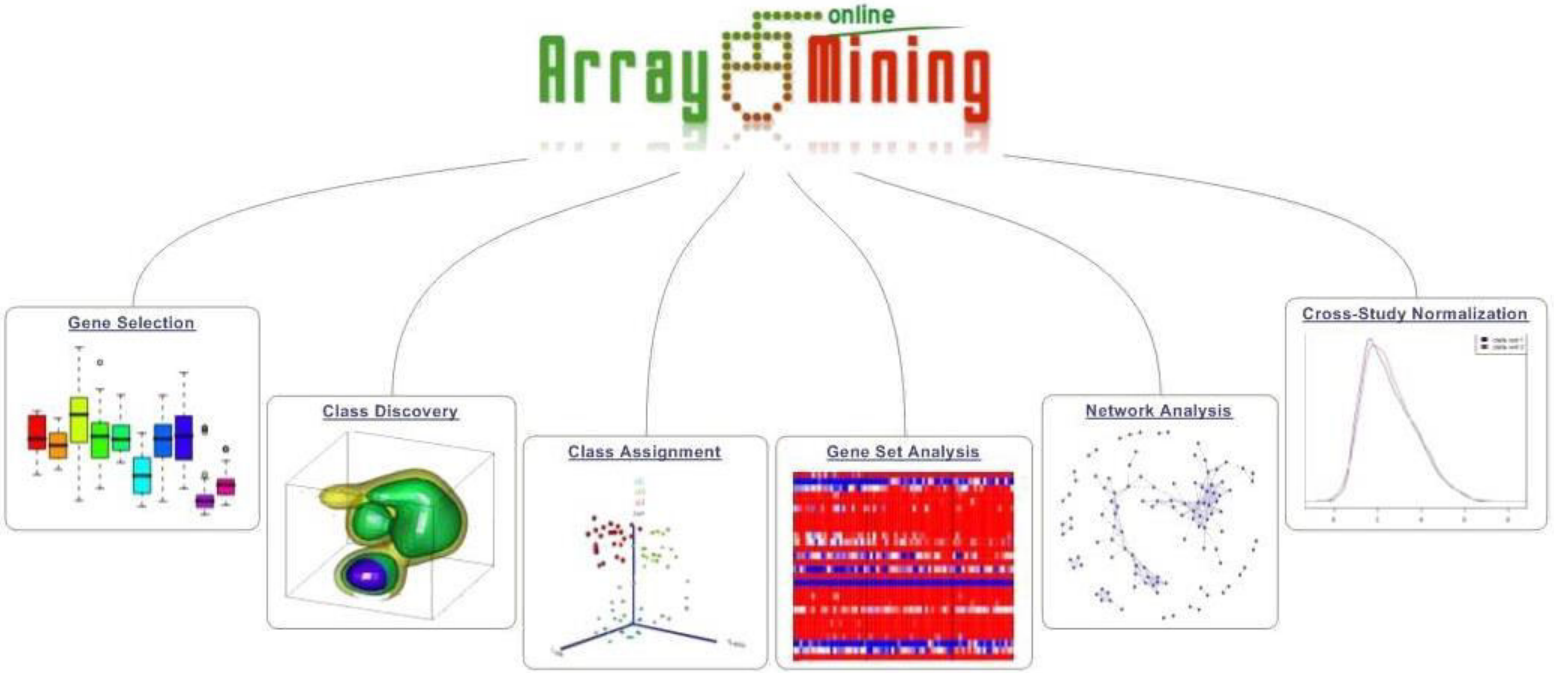
Not secure | https://arraymining.net

English MP3 | FileHippo | HP OFFICEJET 4355... | Hindi MP3 | Hotmail | Lotus Notes Tutorials | Mininova | Web Slice Gallery | Other bookmarks

ArrayMining - Online Microarray Data Mining

Ensemble and Consensus Analysis Methods for Gene Expression Data

Home | Features | Tutorials | Feedback | Contact | Newsletter | Disclaimer | How to cite | Related Software | Help | FAQ



This project is funded by Marie Curie grant MEST-CT-2004-007597, EPSRC grant EP/E017215/1 and BBSRC grant BB/F01855X/1.

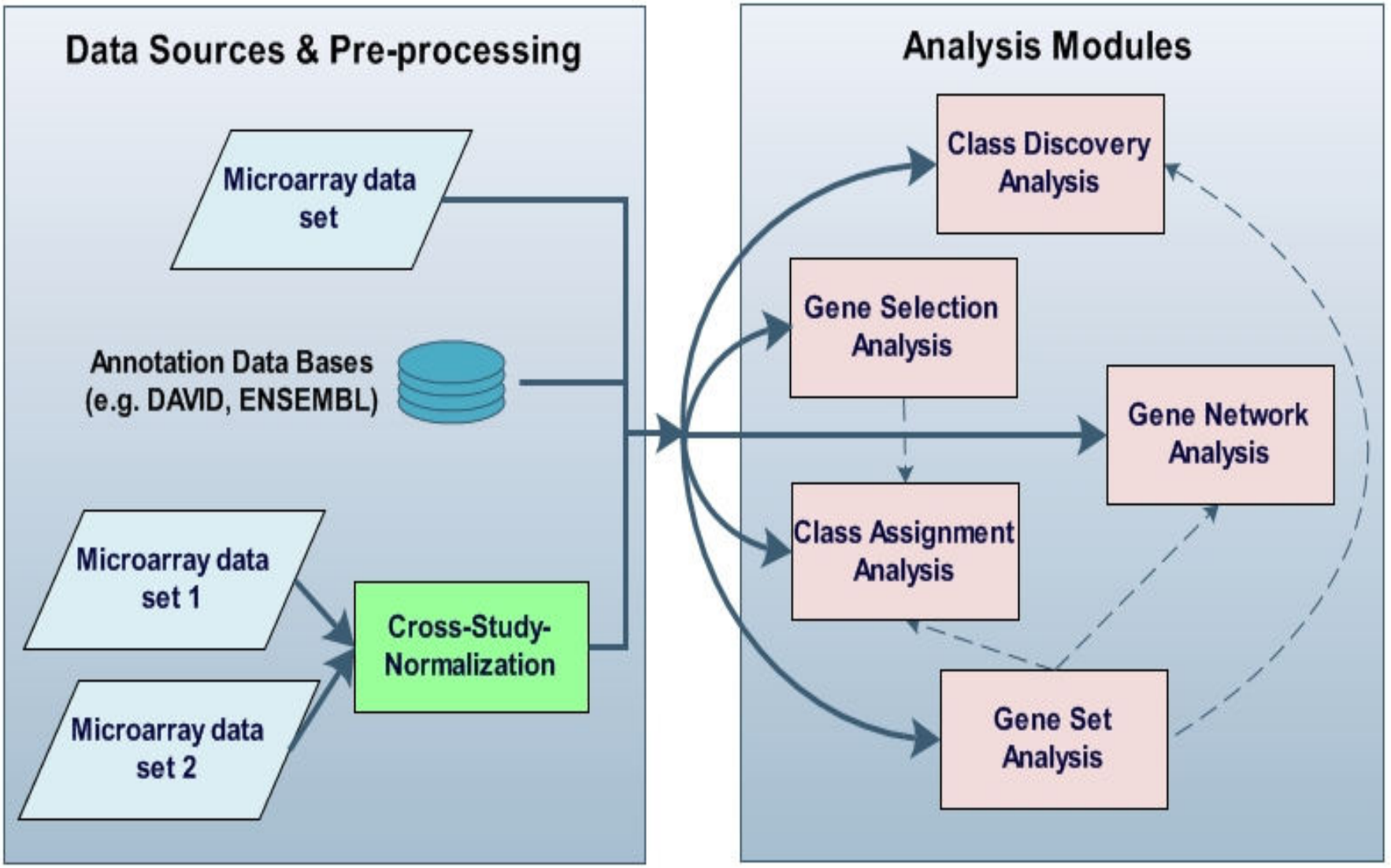
Published in BMC Bioinformatics:

BBSRC EPSRC

full text free on....

9:34 AM
6/19/2022

Workflow and Features



ArrayMining - Online Microarray Data Mining

Ensemble and Consensus Analysis Methods for Gene Expression Data

Gene Selection Analysis (Supervised Feature Selection)

This module allows you to select differentially expressed genes for microarray data with labelled samples.

To obtain instructions click on **help**.

1) Data Set

UPLOAD your own data:

Get help | See example input

Please upload a tab-delimited matrix file or a zip-archive with CEL-files and label-file (max. size: 100 MB):

Choose File | No file chosen

Upload

(After submission, please wait until the upload has been confirmed)

OR

use an EXAMPLE data set:

- (•) Golub
- ( ) Alon
- ( ) van't Veer
- ( ) Yeoh
- ( ) Singh
- ( ) Shipp
- ( ) Shin
- ( ) Armstrong

(After submission, please wait until the upload has been confirmed)

Shin

Armstrong

2) Feature Selection Method

- (•) eBayes
- ( ) SAM
- ( ) PLS-CV
- ( ) CFS
- ( ) RF-MDA
- ( ) ENSEMBLE

3) Parameters

Maximum feature subset size: 100

4) E-Mail Notification (optional)

Your e-mail address:

Submit

(After submission, please wait until the upload has been confirmed)

2) Clustering method

- (•) k-Means
- ( ) PAM
- ( ) SOM
- ( ) SOTA
- ( ) HCL
- ( ) DIANA
- ( ) HYBRID
- ( ) ALL

3) Parameters

Standardization method: No standardization

Filtering of genes:

- (•) Variance filter - size of filtered gene set: 2000
- ( ) SUMCOV (automatic filter size detection; Tritchler et al., 2009)
- ( ) sparse PCA filtering (quadratic penalty lambda = Infinity, Zou et al., 2004)

4) E-Mail Notification (optional)

3) Prediction method

- (•) SVM
- ( ) PAM
- ( ) RF
- ( ) kNN
- ( ) BioHEL
- ( ) ENSEMBLE

4) Parameters

Choose evaluation method:

- (•) cross-validation:
  - Number of cross-validation folds: 10
  - balanced [ ]
- ( ) user-specified training/test set partition:

Maximum feature subset size (must be >= 2): 30

4) E-Mail Notification (optional)

Gene Set Analysis (GSA)

Using the web-form below users can identify whether sets of functionally related genes are significantly differentially expressed in different microarray sample classes.
To obtain instructions click **help**.

1) Data Set

UPLOAD your own data: **OR** **use an EXAMPLE data set:**

Get help | See example input

Please upload a tab-delimited matrix file or a zip-archive with CEL-files and label-file, max. size: 100 MB):

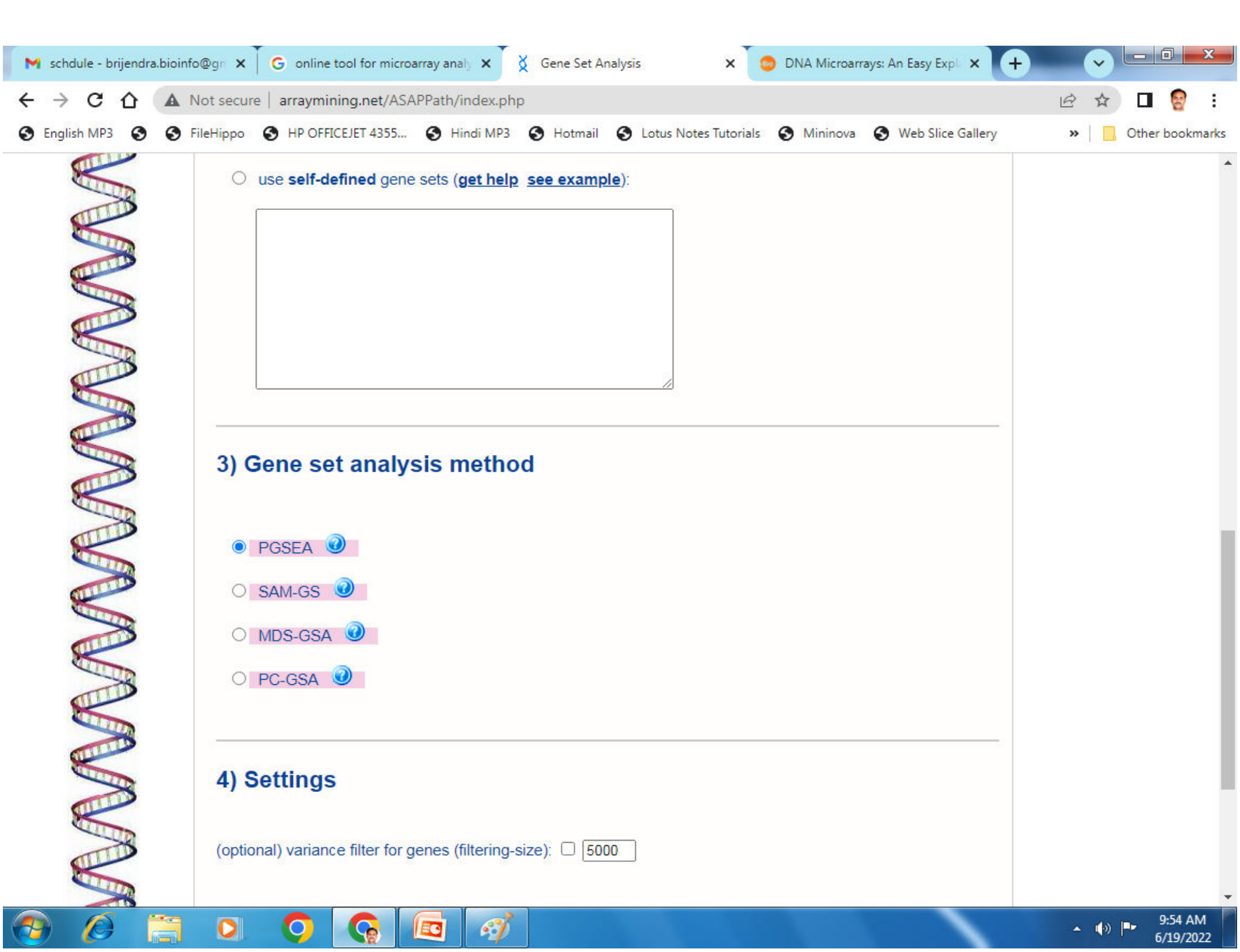
Choose File | No file chosen

Upload

(After submission, please wait until the upload has been confirmed)

- (•) Golub
- ( ) Alon
- ( ) van't Veer
- ( ) Yeoh
- ( ) Singh
- ( ) Shipp
- ( ) Shin
- ( ) Armstrong

2) Functional gene annotation data source


schdule - brijendra.bioinfo@gn
online tool for microarray anal
Gene Set Analysis
DNA Microarrays: An Easy Expl
Not secure | arraymining.net/ASAPPath/index.php
English MP3
FileHippo
HP OFFICEJET 4355...
Hindi MP3
Hotmail
Lotus Notes Tutorials
Mininova
Web Slice Gallery
Other bookmarks
use self-defined gene sets (get help see example):
3) Gene set analysis method
PGSEA
SAM-GS
MDS-GSA
PC-GSA
4) Settings
(optional) variance filter for genes (filtering-size):
5000
9:54 AM
6/19/2022

Gene Co-Expression Network Analysis

This module constructs a weighted gene co-expression network and applies a simple topological network analysis for input data from microarray study. To obtain instructions click **help**.

1) Data Set

UPLOAD your own data: **OR** **use an EXAMPLE data set:**

Get help See example input

Please upload a tab-delimited matrix file or a zip-archive with CEL-files and label-file, max. size: 100 MB):

Choose File | No file chosen

- (•) Armstrong
- ( ) Golub
- ( ) van't Veer
- ( ) Yeoh
- ( ) Singh
- ( ) Shipp

2) Network visualization method

- (•) Fruchterman-Reingold
- ( ) Graphopt

2) Network visualization method

- (●) Fruchterman-Reingold
- ( ) Graphopt
- ( ) DrL
- ( ) Singular Value Decomposition
- ( ) Kamada Kawai
- ( ) Circle

3) Parameters

Filtering method:

- (●) Classical filter
- ( ) SUMCOV (automatic filter size detection; Tritchler et al., 2009)

Edge adjacency threshold [0-1]: 0.5

4) E-Mail Notification (optional)

Your e-mail address:

Submit

Cross-Study Normalization

This module provides methods to combine microarray data from different studies and platforms together into a single data set. The input data sets need to be derived from the same tissue type under comparable biological conditions and the genetic probes must overlap. To obtain instructions click **help**.

1) Data Sets

Upload your data:

Please select a zip-file on your computer that contains both input data sets:

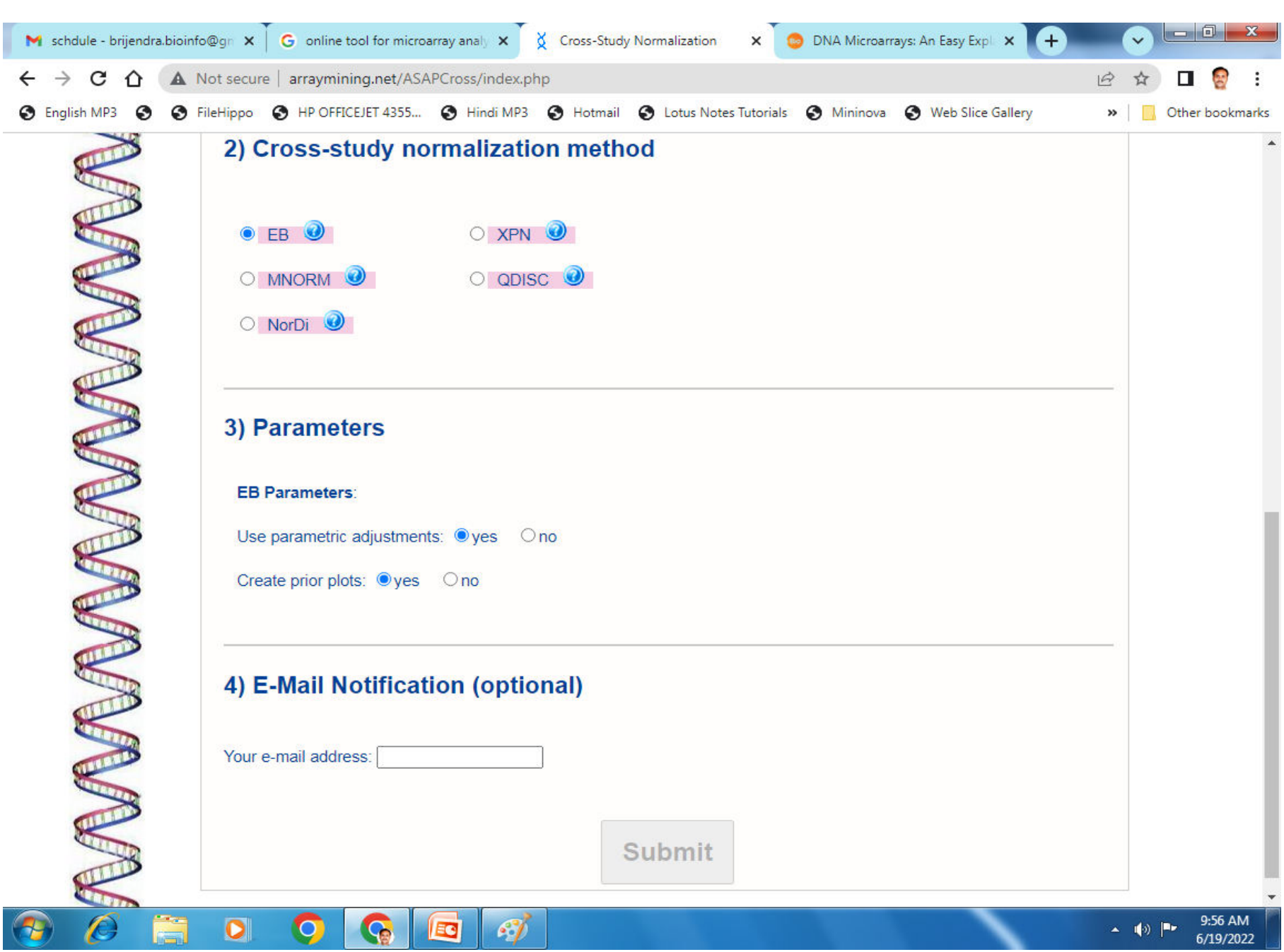
Get help on the input format Download example data

Choose File | No file chosen

Upload

2) Cross-study normalization method

- (•) EB
- ( ) XPN
- ( ) MNORM
- ( ) QDISC


schdule - brijendra.bioinfo@gr
online tool for microarray anal
Cross-Study Normalization
DNA Microarrays: An Easy Expl
Not secure | arraymining.net/ASAPCross/index.php
English MP3
FileHippo
HP OFFICEJET 4355...
Hindi MP3
Hotmail
Lotus Notes Tutorials
Mininova
Web Slice Gallery
Other bookmarks
2) Cross-study normalization method
EB
XPN
MNORM
QDISC
NorDi
3) Parameters
EB Parameters:
Use parametric adjustments: yes no
Create prior plots: yes no
4) E-Mail Notification (optional)
Your e-mail address:
Submit
9:56 AM
6/19/2022

Application of Microarray

- The current scope of microarray applications includes

i) sequencing by hybridization.
ii) Resequencing.
iii) mutation detection.
iv) assessment of gene copy number.
v) comparative genome hybridization.
vi) drug discovery.
vii) Expression analysis, and immunoassay (protein microarrays).
viii) Toxicology.
ix) Protein-protein interaction network and pathway analysis

Conti..

Disadvantages of microarrays -

i) The high cost of a single experiment,

ii) The large number of probe designs based on sequences of low-specificity, as well as the lack of control over the pool of analyzed transcripts.

iii) Analysis only for pre-defined sequences.

iv) Dynamic range limited by scanner.

v) Relies on hybridization

vi) Hybridization potentially non-specific.

Advantages of Microarrays-

i) Well- defined protocols for hybridization

ii) Well-defined analysis pipelines

iii) Standardized approach for data submission

THANK YOU