

# IsoGeneGUI: a graphical user interface for analyzing dose-response studies in microarray experiments

Research team:

Setia Pramana<sup>1</sup>, Dan Lin<sup>1</sup>, Ziv Shkedy<sup>1</sup>, Tobias Verbeke<sup>2</sup>,  
Philippe Haldermans<sup>1</sup>, Willem Talloen<sup>3</sup>.

1. IBioStat, Universiteit Hasselt, Belgium

2. OpenAnalytics, Belgium

3. Johnson & Johnson Pharmaceutical R&D, Belgium

useR! 2010 Conference, 21 July 2010, Gaithersburg, Maryland,  
USA

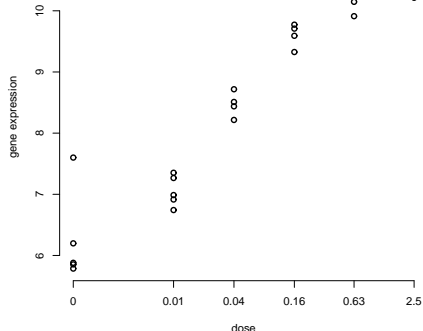


# Overview

- Background: Dose-response Studies
- Test for Monotone Trend
- The Software
- Analysis using IsoGeneGUI Package
  - Data exploration
  - Likelihood Ratio Test  $E^2$  analysis
  - Resampling-based analysis
  - Significance Analysis of Microarrays (SAM)
  - User defined graphical display

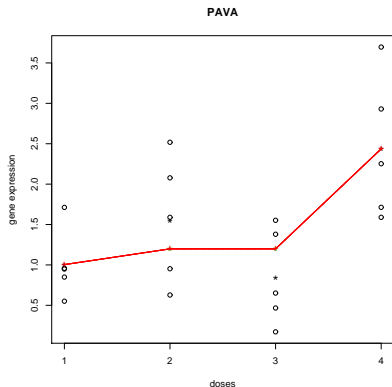
# Dose-response studies

- The fundamental objective of drug development is to find a dose or a range of dose that efficacious and safe.
- Dose-response study investigates the dependence of the response on doses.
- Within a microarray setting the response of interest is gene expression.



# Dose-response studies in Microarray Experiments

- The aim is to identify a subset of genes with a monotone trend.
- In general, increasing the dose will result in a proportional increase in the response.
- Estimation: Isotonic regression with **Pool-Adjacent-Violator-Algorithm (PAVA)**.



# Testing for Monotonic Trend

- Hypotheses: For gene  $i$  ( $i = 1, \dots, m$ ) with  $K + 1$  doses ( $j = 0, \dots, K$ )

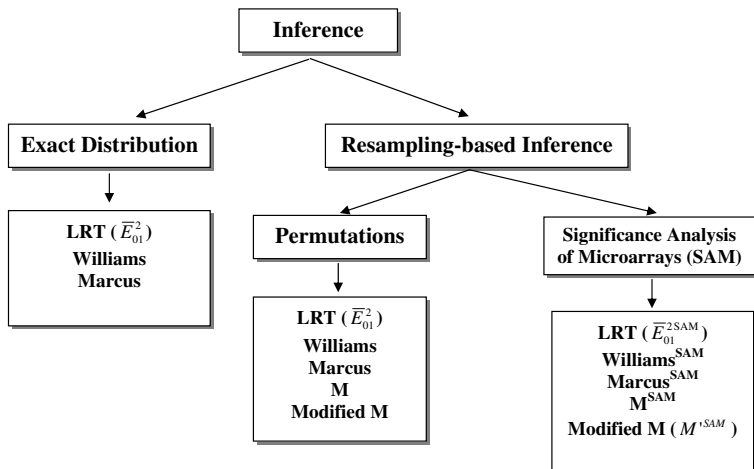
$$H_0 : \quad \mu(d_0) = \mu(d_1) = \dots = \mu(d_K) \quad (\text{No Dose Effect})$$

$$H_1^{Up} : \quad \mu(d_0) \leq \mu(d_1) \leq \dots \leq \mu(d_K) \quad (\text{Ordered alternative})$$

- The likelihood ratio test ( $\bar{E}_{01}^2$ ).
- The  $t$ -type tests: Williams, Marcus,  $M$  and  $M'$

Lin et.al 2007

# Testing for Monotonic Trend: Inference



Both approaches will be discussed in the next slides

# Resampling-based Inference: Permutations

- Observed test statistics  $E^2$ :

$$\bar{E}_1^2, \bar{E}_2^2, \dots, \bar{E}_m^2$$

- Resampling-based Inference, the permutation matrix  $T$ :

$$T = \begin{pmatrix} \bar{E}_{11}^2 & \bar{E}_{11}^2 & \dots & \bar{E}_{1B}^2 \\ \bar{E}_{21}^2 & \bar{E}_{22}^2 & \dots & \bar{E}_{2B}^2 \\ \vdots & & \vdots & \\ \bar{E}_{m1}^2 & \bar{E}_{m2}^2 & \dots & \bar{E}_{mB}^2 \end{pmatrix}$$

- $B$  is the number of permutations.
- Raw  $p$ -values:  $P_i = \frac{\#\{b: \bar{E}_{ib}^2 \geq E_i^2\}}{B}$ .

## Significance Analysis of Microarrays (SAM, Tusher et al., 2001) for microarrays under order restriction

- A modified score is assigned to each gene.
- For the  $t$ -type test statistics (i.e., Williams, Marcus, the  $M$ , and the  $M'$ ):

Original:

$$M' = \frac{\hat{\mu}_K^* - \hat{\mu}_0^*}{\hat{s}'}$$

Modified:

$$M'^{SAM} = \frac{\hat{\mu}_K^* - \hat{\mu}_0^*}{\hat{s}' + \hat{s}_0}$$

- $\hat{s}_0$  is the fudge factor added to protect against inflated statistic due to small standard errors.



# Software

Software for dose-response in microarrays analysis:

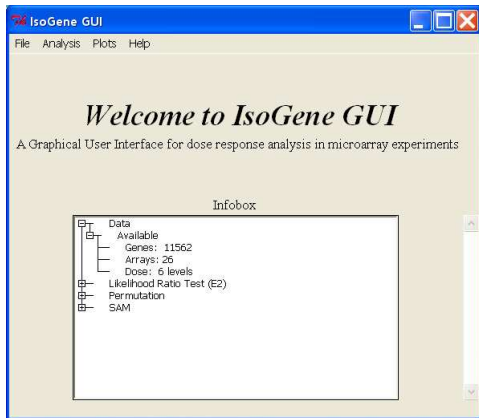
- R package: [IsoGene](http://cran.ii.uib.no/web/packages/IsoGene/), available in CRAN:  
<http://cran.ii.uib.no/web/packages/IsoGene/>

Pramana et.al 2010, The R Journal

- R GUI package: [IsoGeneGUI](http://bioconductor.org/packages/2.7/bioc/html/IsoGeneGUI.html), available in the Bioconductor project:  
<http://bioconductor.org/packages/2.7/bioc/html/IsoGeneGUI.html>

# IsoGeneGUI: Main Window

- Load Data
- Data Exploration
- Analysis:
  - $\bar{E}_{01}^2$  exact distribution.
  - Resampling-based Inference
  - Multiplicity adjustment (FWER and FDR)
  - Significant Analysis of Microarrays (SAM)
- Output & Graphical Displays



# IsoGeneGUI: Data Format

The package can read data in the following formats:

- R workspace (\*.RData files)
- Excel file (\*.xls files)
- text file (\*.txt files)

# IsoGeneGUI: Data Exploration

## Exploratory Data Analysis

**Isotonic Regression Plot**

Insert Gene name/ row name you want to draw:

Gene Name(s), e.g., gene1, gene3

Row Number(s), e.g., 2, 3, 4

Range of Row Number From: To

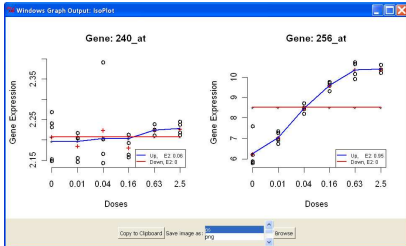
Dose as ordinal

Show isotonic regression curve for both directions

Show summary of the data

OK Cancel

## The Isotonic regression plots



## The summary statistic

**Summary Statistics for gene**

Summary Statistics for gene: 240\_at

Dose	N	Observed Mean	S.e	IsoMean Up	IsoMean Down
0.00	5	2.208	0.0108	2.196	2.209
0.01	5	2.185	0.0064	2.196	2.209
0.04	4	2.225	0.0223	2.203	2.209
0.16	4	2.181	0.0051	2.203	2.209
0.63	4	2.224	0.0022	2.224	2.209
2.50	4	2.228	0.0030	2.228	2.209

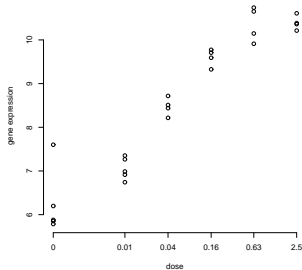
E2 Up : 0.06 | E2 Down : 0.00

Summary Statistics for gene: 256\_at

Dose	N	Observed Mean	S.e	IsoMean Up	IsoMean Down
0.00	5	6.264	0.1498	6.264	8.69
0.01	5	7.053	0.0495	7.053	8.69
0.04	4	8.470	0.0406	8.470	8.69
0.16	4	9.600	0.0387	9.600	8.69
0.63	4	10.362	0.0780	10.362	8.69
2.50	4	10.392	0.0320	10.392	8.69

# IsoGeneGUI: The $\bar{E}_{01}^2$ Analysis

- Antipsychotic study.
- 6 dose levels
- 4-5 samples at each dose level.
- 11,565 genes.
- The dataset is available inside the package.



- The  $\bar{E}_{01}^2$  test and BH-FDR (0.05) are specified:

The screenshot shows the 'Global Likelihood Ratio Test (E2) Analysis' window. The main area contains the following settings:

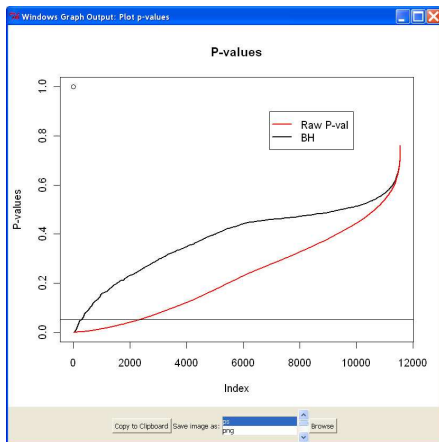
- Calculate the p-values for E2
- Select the genes:  All Genes
- Genes Range: From [ ] To [ ]
- Calculate button
- P-value adjustment:
  - BH
  - BY
  - SidakSS
  - SidakSD
  - Holm
  - Hochberg
  - Bonferroni
- Overall Significant Level: 0.05
- Display the significant genes
- Calculate button
- Save significant genes button
- Save results for all genes button
- Plots:
  - Plot E2 values vs raw p-values
  - Plot Fold change vs E2
  - Ranking Plot (Adjust the p-values first)
- Produce plots button
- Exit the analysis button

# IsoGeneGUI: $\bar{E}_{01}^2$ Result

- $\bar{E}_{01}^2$  with BH-FDR (0.05).
- 298 genes with significant monotone trend

```

Global Likelihood Ratio Test (E2)
Results for E2 using exact distribution p-value
Number of Genes analyzed : 11562
Overall significant level : 0.05
P-value adjustment: BH
Number Significant Genes : 298
  
```



# IsoGeneGUI: List of Significant Genes

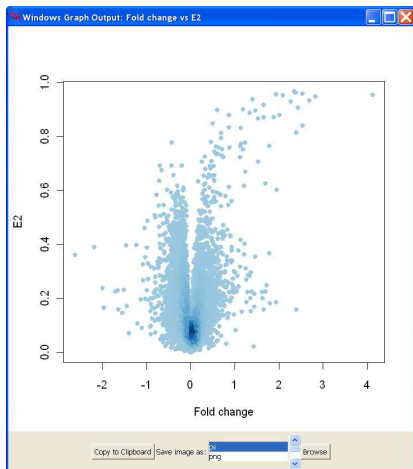
- List significant genes.
- The list be saved as an R object or/and an excel file

List of Significant Genes based on Asymptotic P-values, FDR: 0.05 Number of Genes analyzed : 11562

	ProbeID	row.num	Mu.Diff	Direction	E2Val	pval.Asymp	BH.p.val	BY.p.val
1	6415_at	6215	2.34951303	u	0.9654130	2.378309e-16	0.000000	0.000000
2	1803_at	1603	2.38403091	u	0.9628712	4.870726e-16	0.000000	0.000000
3	4242_at	4042	2.53672430	u	0.9584245	1.623438e-15	0.000000	0.000000
4	4833_at	4633	1.93891924	u	0.9548713	3.818261e-15	0.000000	0.000000
5	256_at	56	4.12847733	u	0.9528423	6.030189e-15	0.000000	0.000000
6	380_at	180	2.09541507	u	0.9507085	9.554572e-15	0.000000	0.000000
7	1360_at	1160	2.82725604	u	0.9468426	2.105868e-14	0.000000	0.000000
8	3941_at	3741	1.40675989	u	0.9381753	1.018548e-13	0.000000	0.000000
9	924_at	724	2.69715769	u	0.9339285	2.039269e-13	0.000000	0.000000
10	4848_at	4648	2.27824021	u	0.9285061	4.653657e-13	0.000000	0.000000
11	3746_at	3546	1.65736950	u	0.9163942	2.397639e-12	0.000000	0.000000
12	896_at	696	2.44329292	u	0.9064505	7.796396e-12	0.000000	0.000000
13	7541_at	7341	1.19362811	u	0.8988077	1.778998e-11	0.000000	0.000000
14	5099_at	4899	0.60531374	u	0.8968981	2.165298e-11	0.000000	0.000000
15	1359_at	1159	1.47442793	u	0.8956025	2.469078e-11	0.000000	0.000000
16	1568_at	1368	1.34435519	u	0.8878167	5.261434e-11	0.000000	0.000000
17	1269_at	1069	0.86672056	u	0.8798123	1.086907e-10	0.000000	0.000001
18	9835_at	9635	2.01765002	u	0.8787428	1.193201e-10	0.000000	0.000001
19	8754_at	8554	1.88526389	u	0.8713131	2.232660e-10	0.000000	0.000001
20	9931_at	9731	1.66185945	u	0.8711961	2.254151e-10	0.000000	0.000001
21	11124_at	10924	1.53998549	u	0.8665769	3.268060e-10	0.000000	0.000002
22	11448_at	11248	0.50159972	u	0.8486610	1.234955e-09	0.000001	0.000006
23	6992_at	6792	2.54487370	u	0.8386370	2.431316e-09	0.000001	0.000012
24	6922_at	6722	0.86925494	u	0.8334940	3.387152e-09	0.000002	0.000016
25	9421_at	9221	1.13472792	u	0.8305570	4.074771e-09	0.000002	0.000019
26	7917_at	7717	2.40088293	u	0.8139082	1.098002e-08	0.000005	0.000048
27	283_at	83	0.69466377	u	0.8115807	1.252357e-08	0.000005	0.000053
28	5684_at	5484	0.69700485	u	0.7937990	3.253788e-08	0.000013	0.000133
29	10058_at	9858	0.46128782	u	0.7795599	6.601657e-08	0.000026	0.000261
30	477_at	277	1.24433243	u	0.7778047	7.180537e-08	0.000028	0.000275

# IsoGeneGUI: Volcano Plot

- $\bar{E}_{01}^2$  vs. Fold change
- The graph can be copied into clipboard and be saved into different image types.





# IsoGeneGUI: Resampling-based Inference

- Required specification: number of permutations, test statistic and multiplicity adjustment.
- The permutation result can be saved and used later.
- List of significant genes using the selected statistic(s).
- Graphical displays.

Permutation Analysis - IsoGene GUI

New Permutation

Select the genes:

All Genes

Genes Range From  To

Number of Permutation:

Save the permutations result:

Obtain the raw p-value from a file:

Statistic:

M  E2  Marcus  Williams  M'

Controlling FDR Procedure:

BH

Controlling FWER Procedure:

SidakSS  SidakSD  Holm  Hochberg  Bonferroni

Overall Significant Level:

Display the significant genes

Plots:

Specify the statistic that you want to plot:

M  E2  Marcus  Williams  M'

Specify the plot(s):

Plot of the specified statistic vs its raw p-values

Plot of fold change vs the specified statistic

Rank plot (Adjust the p-value first!)

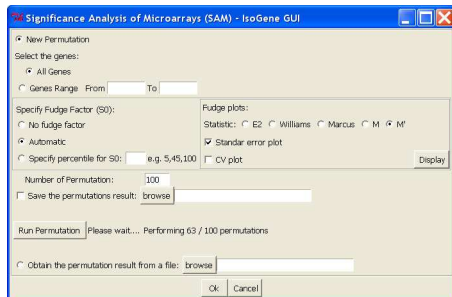
# IsoGeneGUI: SAM (specification)

- Required : Number of permutations and the fudge factor.
- SAM regularized test statistic, e.g.,:

$$M^{SAM} = \frac{\hat{\mu}_K^* - \hat{\mu}_0^*}{\hat{S}' + \hat{S}_0}$$

- Delta ( $\Delta$ ) is chosen with estimation of the FDR.

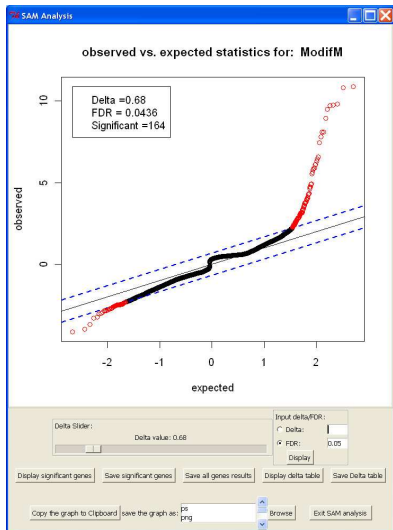
## SAM permutation dialog box



# IsoGeneGUI: SAM (Results)

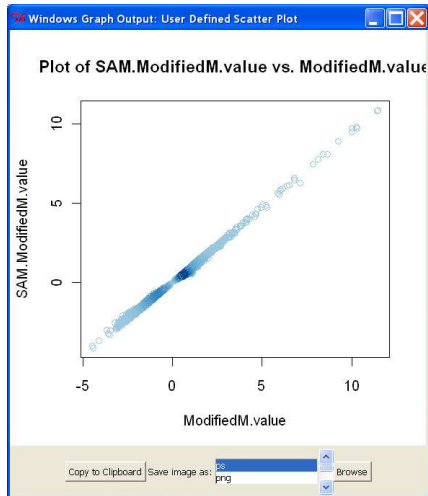
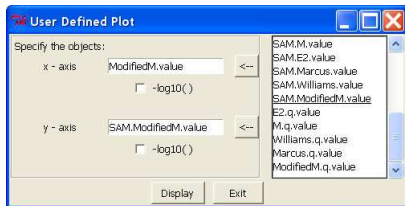
- Threshold/Delta: 0.68
- FDR: 0.0436
- Number of significant genes: 164
- Delta table:

	Delta	FalsePositive50	FalsePositive90	Called	FDR50	FDR90
1	0.00	10332.4241	10335.1060	11562	0.8937	0.8939
2	0.04	10332.4241	10335.1060	11562	0.8937	0.8939
3	0.08	10330.6362	10333.3181	11559	0.8937	0.8940
4	0.12	10317.6738	10325.2724	11545	0.8937	0.8944
5	0.16	10285.0441	10302.0294	11534	0.8917	0.8932
6	0.20	10199.2237	10245.9779	11476	0.8887	0.8928
7	0.24	9996.2941	10100.2620	11317	0.8833	0.8925
8	0.28	9488.9701	9748.5769	10971	0.8649	0.8886
9	0.32	5082.1796	7423.3792	7383	0.6884	1.0055
10	0.36	389.3209	805.9970	1526	0.2551	0.5282
11	0.40	228.8545	492.9312	1094	0.2092	0.4506
12	0.44	136.7763	312.9764	831	0.1646	0.3766
13	0.48	75.9869	190.8611	570	0.1333	0.3348
14	0.52	45.5921	130.9656	462	0.0987	0.2835
15	0.56	31.2887	98.4253	368	0.0850	0.2675
16	0.60	19.2202	62.8456	274	0.0701	0.2294
17	0.64	12.0685	44.6981	222	0.0544	0.2013
18	0.68	7.1517	29.5902	164	0.0436	0.1804
19	0.72	6.2577	26.0143	152	0.0412	0.1711
20	0.76	3.5759	21.6339	121	0.0296	0.1788
21	0.80	2.6819	17.0747	113	0.0237	0.1511
22	0.84	1.7879	12.6049	100	0.0179	0.1260
23	0.88	0.8940	11.6215	96	0.0093	0.1211
24	0.92	0.8940	9.0290	93	0.0096	0.0971
25	0.96	0.8940	7.2411	90	0.0099	0.0805
26	1.00	0.0000	63.3471	83	0.0000	0.0765
27	1.04	0.0000	3.7546	80	0.0000	0.0469
28	1.08	0.0000	3.6652	76	0.0000	0.0482
29	1.12	0.0000	3.5759	74	0.0000	0.0483
30	1.16	0.0000	2.7713	73	0.0000	0.0380



# User Defined Graphical Display

User can define their own figures:



# Help

## Html help with screen shots



## IsoGeneGUI: A Graphical User Interface for the IsoGene

---

### User Guides

1. [About the IsoGeneGUI](#)
2. [The statistics for testing for monotonic trend](#)
3. [IsoGeneGUI input files](#)
4. [Data exploration](#)
5. Analysis:
  - a. [Likelihood ratio test statistic \(E2\)](#)
  - b. [Permutation](#)
  - c. [Significance Analysis of Microarrays \(SAM\)](#)
6. [Example Data: Dopamine Study](#)

# More detail...

- Book of Lin, D., Shkedy, Z., Yekutieli, D., Amaratunga, D., and Bijmens., L., editors. **Modeling Dose-response Microarray Data in Early Drug Development Experiments Using R.** Springer. To be published in 2010.
- IsoGene Project site:  
<http://www.ibiostat.be/software/IsoGeneGUI/index.html>

The screenshot shows a web browser window with the URL <http://www.ibiostat.be/software/IsoGeneGUI/index.html>. The page title is "The IsoGene Project". Below the title is a navigation menu with links for Home, IsoGene, IsoGeneGUI, Examples, and Publications. The "Background" section contains the following text: "The IsoGene Project is a project aimed for developing software to perform analysis of dose-response studies in screening experiments. Two packages in R are developed in this project: IsoGene and IsoGeneGUI. The IsoGene package is a command based package in R. The Graphical User Interface (IsoGeneGUI) is a user friendly interface of the IsoGene package which is developed for the user with no or limited knowledge about R programming on his/her can perform the analysis of dose-response in screening setting, easily. The IsoGeneGUI was developed using Shiny." The "Project developers" section lists: "Dan Lin, Seta Prasanna, Philippe Haldemann, Ziv Shkedy, Tobias Verbeke, Tamas Buraykovski, Dorit Yekutieli, Koen Adetayo-Harish Ghisvann, An De Brandt, Wilkes Talson, Luc Bijmens, Dharmika Amaratunga." The "Quick Start" section lists: "1. IsoGene package: • Download IsoGene source code (.tar.gz) • Download Windows (IsoGene) installer (.zip) • Download user manual 2. IsoGeneGUI as an R package: • Download IsoGeneGUI source code (.tar.gz) • Download Windows (IsoGeneGUI) installer (.zip)"

# Thank you for your attention...