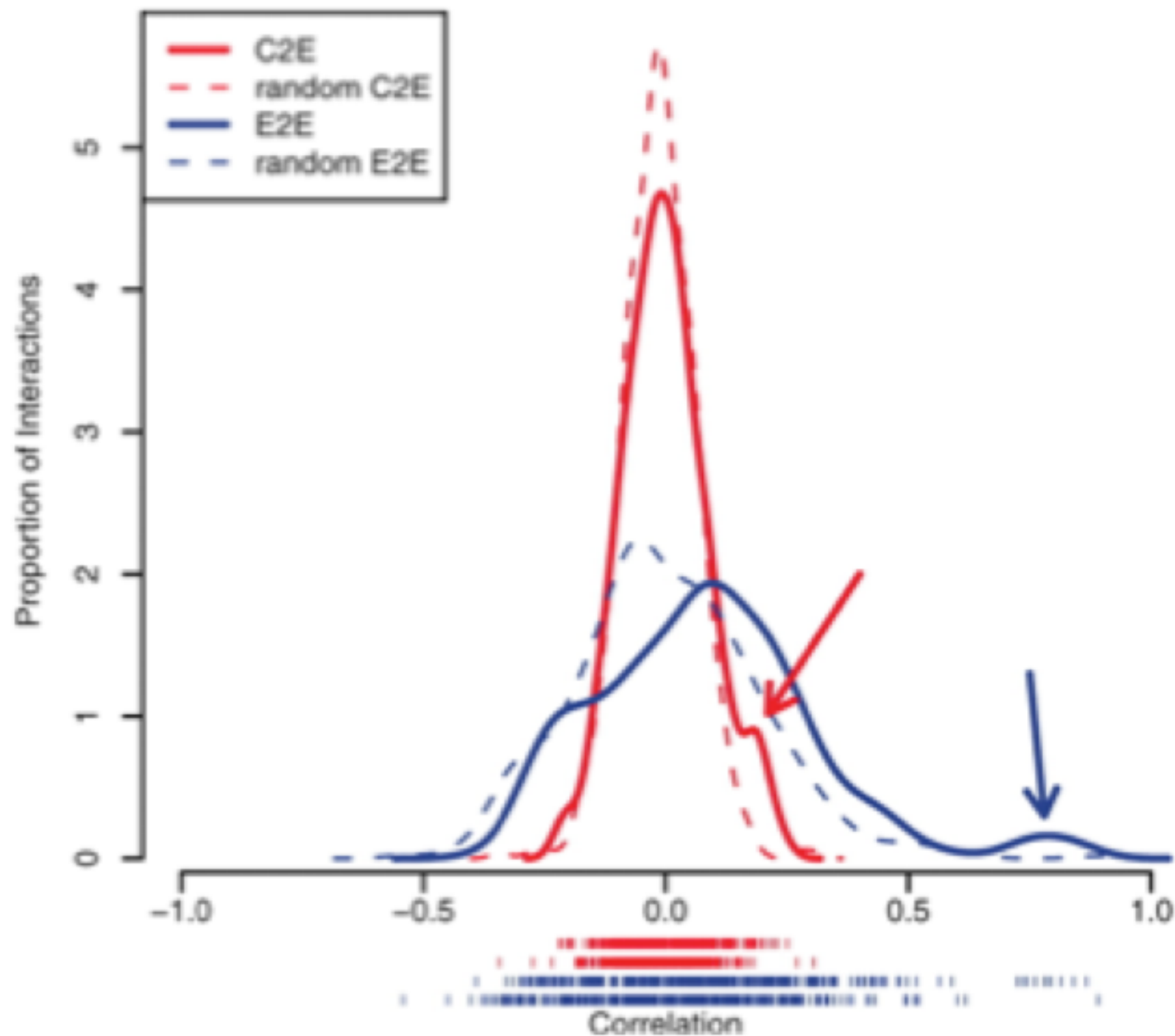


# PARADIGM

Erkin Otles  
CS 838

## NCI Pathway interactions in TCGA GBM data.



Charles J. Vaske et al. *Bioinformatics* 2010;26:i237-i245

# PARADIGM Approach

We developed an approach called PARADIGM (PArthway Recognition Algorithm using Data Integration on Genomic Models) to infer the activities of genetic pathways from integrated patient data.

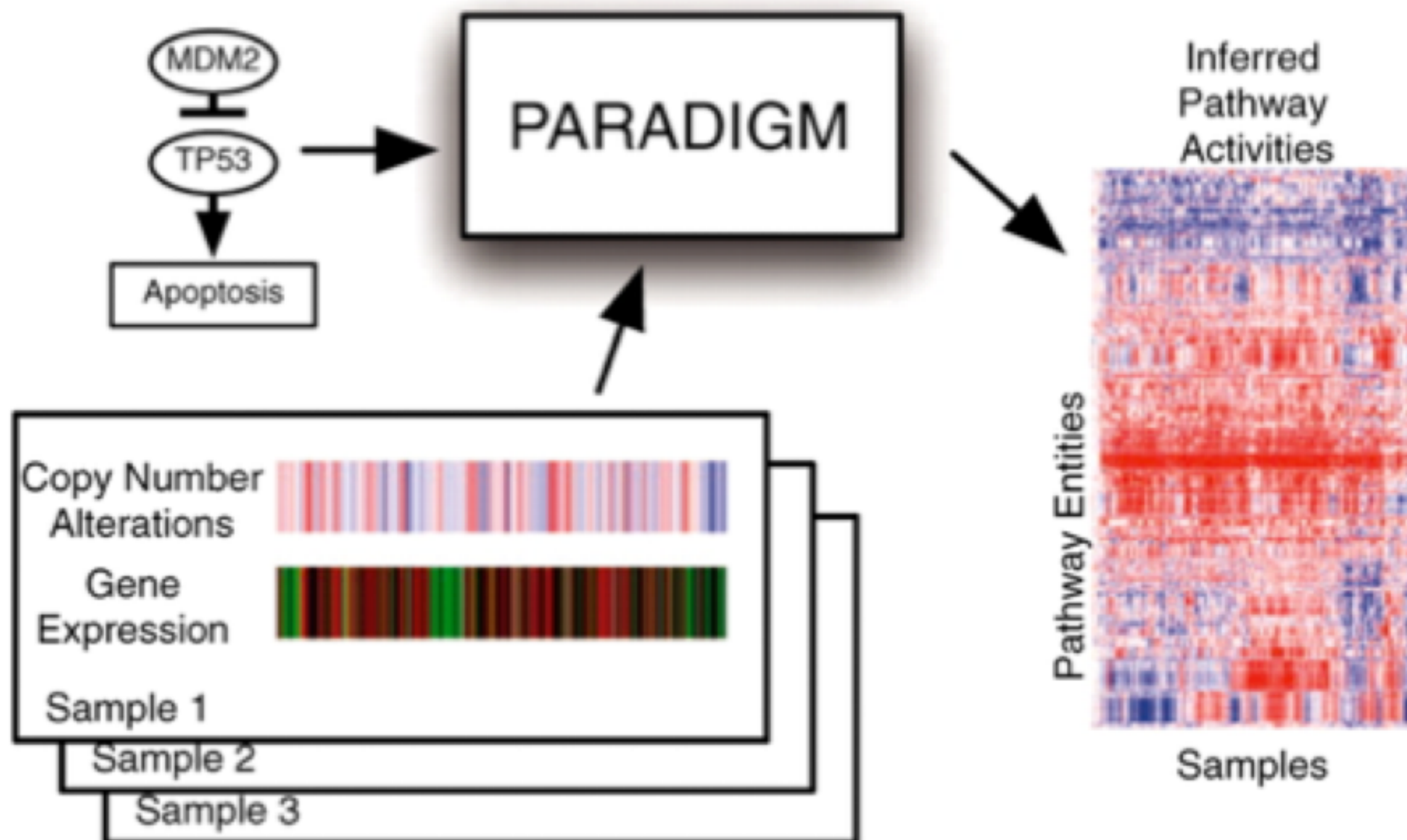
Multiple genome-scale measurements on a single patient sample are combined to infer the activities of genes, products and abstract process inputs and outputs for a single NCI pathway.

# PARADIGM Approach

PARADIGM produces a matrix of integrated pathway activities (IPAs)  $A$  where  $A_{ij}$  represents the inferred activity of entity  $i$  in patient sample  $j$ .



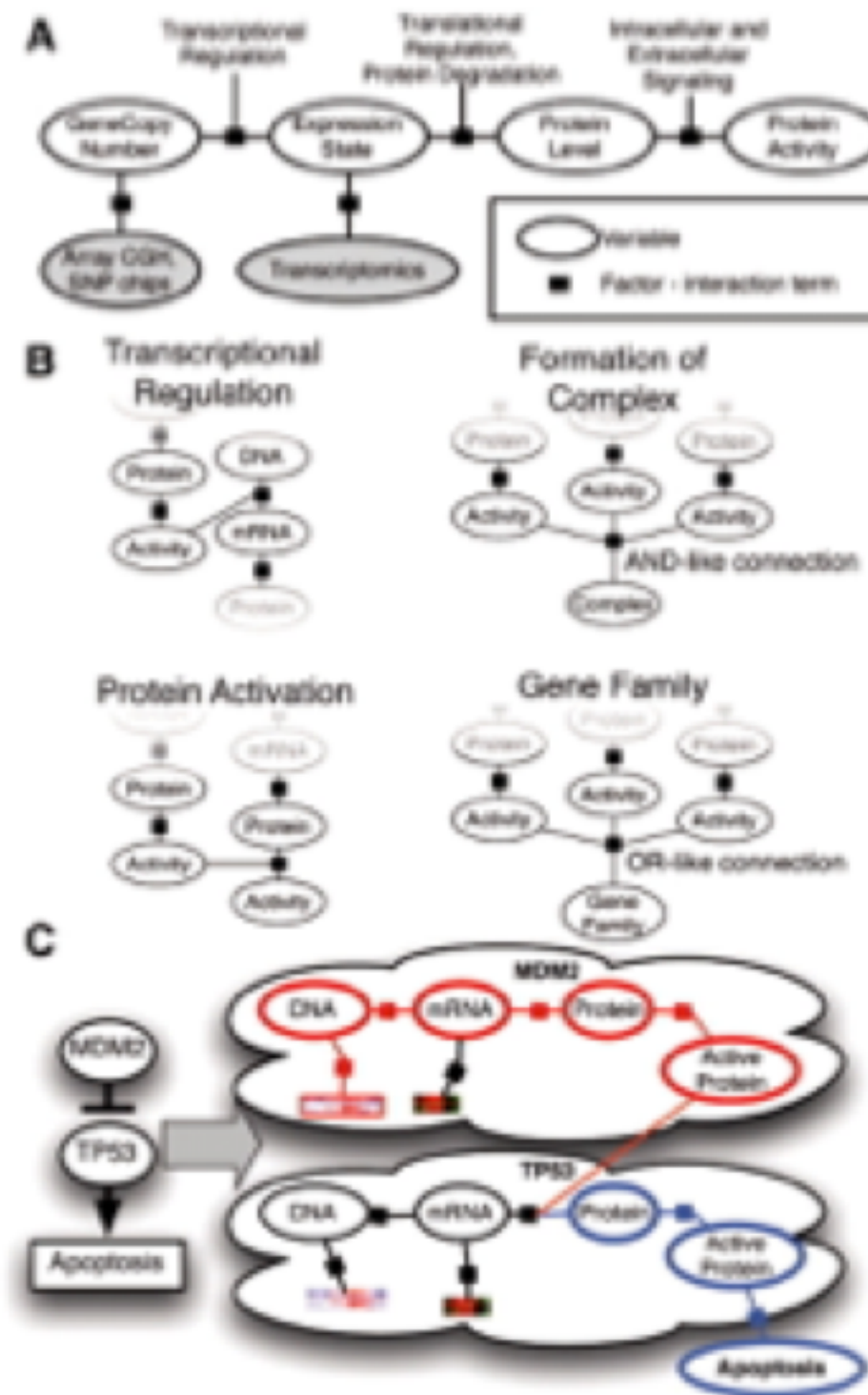
## Overview of the PARADIGM method.



Charles J. Vaske et al. *Bioinformatics* 2010;26:i237-i245



## Conversion of a genetic pathway diagram into a PARADIGM model.



Charles J. Vaske et al. *Bioinformatics* 2010;26:i237-i245

Method

# GOAL!

Make a factor graph that represents the underlying pathway.

Each entity can take on one of three states corresponding to activated, nominal or deactivated relative to a control level (e.g. as measured in normal tissue) and encoded as 1, 0 or  $-1$  respectively.

The states may be interpreted differently depending on the type of entity (e.g. gene, protein, etc)



# Factor Graph Goal

The factor graph encodes the state of a cell using a random variable for each entity  $X=\{x_1, x_2, \dots, x_n\}$  and a set of  $m$  non-negative functions, or factors, that constrain the entities to take on biologically meaningful values as functions of one another.

The  $j$ -th factor  $\phi_j$  defines a probability distribution over a subset of entities  $x_j \subset X$ . The entire graph of entities and factors encodes the joint probability distribution over all of the entities as:

$$P(X) = \frac{1}{Z} \prod_{j=1}^m \phi_j(X_j),$$

where  $Z = \prod_j \sum_{S \subset X_j} \phi_j(S)$  is a normalization constant and  $S \subset X$  denotes that  $S$  is a 'setting' of the variables in  $X$ .

# Construction

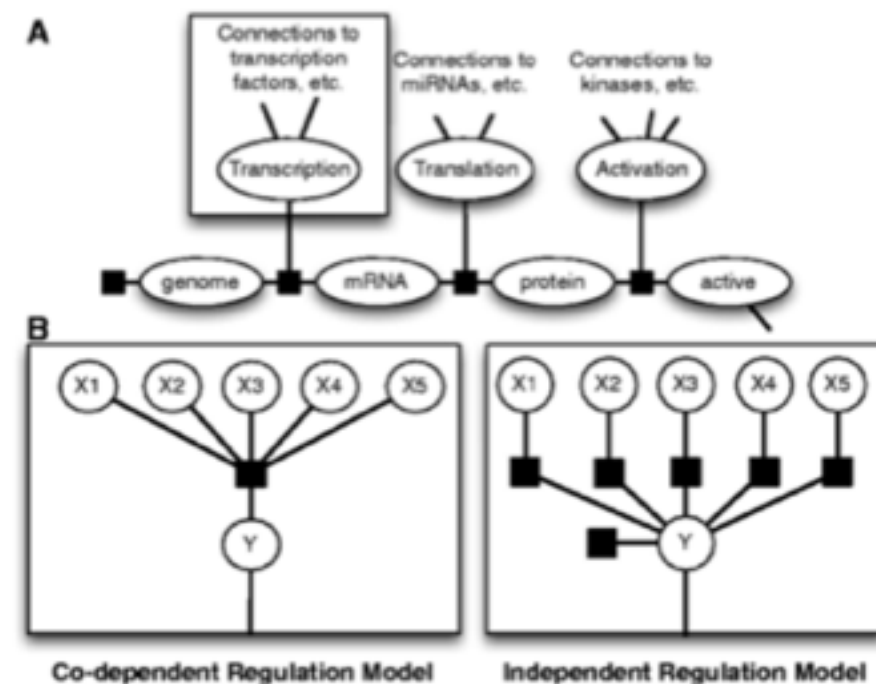
In order to simplify the construction of factors, we first convert the pathway into a directed graph, with each edge in the graph labeled with either positive or negative influence.

Every interaction in the pathway is converted to a single edge in the directed graph.

Using this directed graph, we then construct a list of factors to specify the factor graph.

For every variable  $x_i$ , we add a single factor  $\phi(X_i)$ , where  $X_i = \{x_i\} \cup \{\text{Parents}(x_i)\}$  and  $\text{Parents}(x_i)$  refers to all the parents of  $x_i$  in the directed graph.

Factor graph structures in PARADIGM. (A) Central dogma structure shared by all protein coding genes.



Andrew J. Sedgewick et al. Bioinformatics 2013;29:i62-i70

© The Author 2013. Published by Oxford University Press.

Bioinformatics

# Filling Out the FG

The expected value was set to the majority vote of the parent variables.

$$\phi_i(x_i, \text{Parents}(x_i)) = \begin{cases} 1 - \epsilon & x_i \text{ is the expected state from Parents}(x_i) \\ \frac{\epsilon}{2} & \text{otherwise.} \end{cases}$$

If a parent is connected by a positive edge it contributes a vote of +1 times its own state to the value of the factor. (negative edge, then -1)

The variables connected to  $x_i$  by an edge labeled 'minimum' get a single vote, and that vote's value is the minimum value of these variables, creating an AND-like connection. Similarly the variables connected to  $x_i$  by an edge labeled 'maximum' get a single vote, and that vote's value is the maximum value of these variables, creating an OR-like connection.

Votes of zero are treated as abstained votes. If there are no votes the expected state is zero.

Otherwise, the majority vote is the expected state, and a tie between 1 and -1 results in an expected state of -1 to give more importance to repressors and deletions.

# Inference

Given patient data, we would like to estimate whether a particular hidden entity  $x_i$  is likely to be in state  $a$ .

For example, how likely TP53's protein activity is  $-1$  (inactivated) or 'Apoptosis' is  $+1$  (activated).

To do this, we first compute the prior probability of the event prior to observing the patient's data.

If  $A_i(a)$  represents the singleton assignment set  $\{x_i=a\}$  and  $\Phi$  is the fully specified factor graph, this prior probability is:

$$P(x_i=a|\Phi) = \frac{1}{Z} \prod_{j=1}^m \sum_{\mathbf{s} \sqsubset_{A_j(a)} X_j} \phi_j(\mathbf{s}),$$

# Inference Cont.

The probability that  $x_i$  is in state  $a$  along with all of the observations made for the patient is:

$$P(x_i = a, D | \Phi) = \frac{1}{Z} \prod_{j=1}^m \sum_{\mathbf{S} \sqsubseteq_{A_i(a) \cup D} X_j} \phi_j(\mathbf{S}).$$

For the majority of pathways, we use the junction tree inference algorithm with HUGIN updates to infer the probabilities in equations. For pathways that take longer than 3 s of inference per patient, we use Belief Propagation with sequential updates.

To learn the parameters of the observation factors we use the expectation-maximization (EM) algorithm.

# How to Make IPAs

After inference, we output an IPA for each variable that has an 'active' molecular type.

We compute a log-likelihood ratio using the quantities:

$$\begin{aligned} L(i,a) &= \log\left(\frac{P(D, x_i=a|\Phi)}{P(D, x_i\neq a|\Phi)}\right) - \log\left(\frac{P(x_i=a|\Phi)}{P(x_i\neq a|\Phi)}\right) \\ &= \log\left(\frac{P(D|x_i=a, \Phi)}{P(D|x_i\neq a, \Phi)}\right). \end{aligned}$$

We then compute a single IPA for gene  $i$  based on the log-likelihood ratio as:

$$IPA(i) = \begin{cases} L(i, 1) & L(i, 1) > L(i, -1) \text{ and } L(i, 1) > L(i, 0) \\ -L(i, -1) & L(i, -1) > L(i, 1) \text{ and } L(i, -1) > L(i, 0) \\ 0 & \text{otherwise.} \end{cases}$$





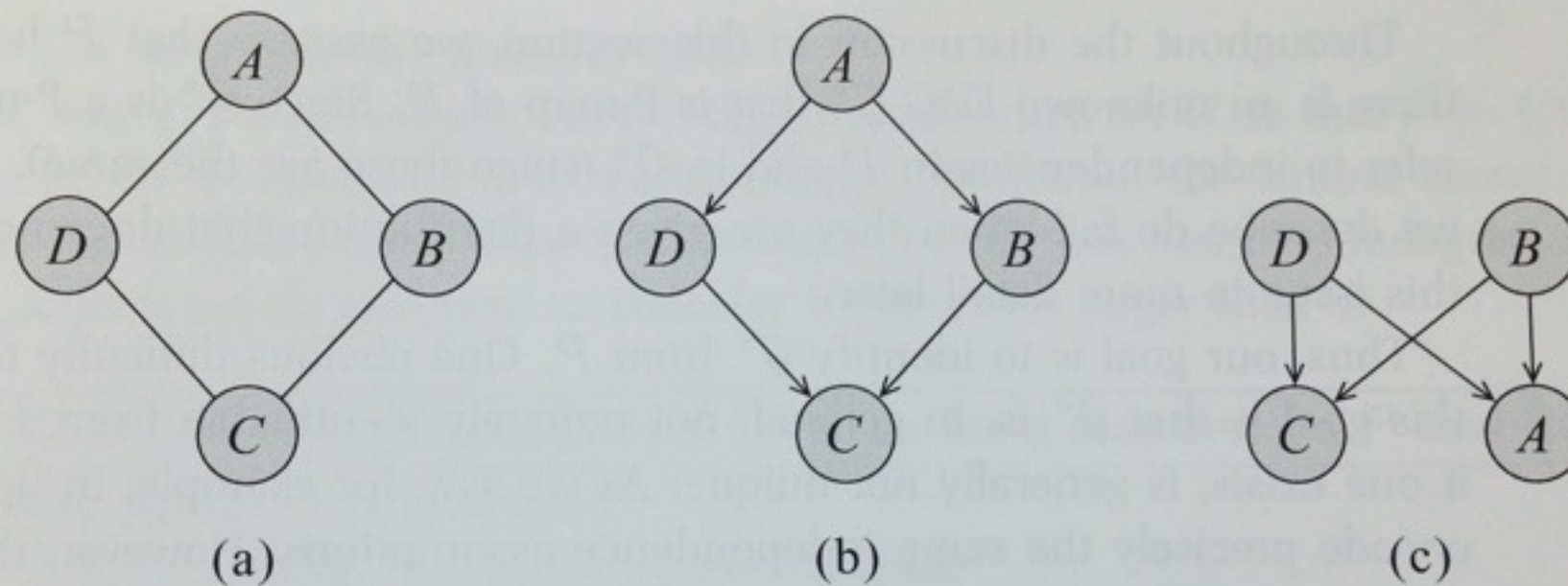
# Aside: Factor Graphs

Consider a scenario where we have four students who get together in pairs to work on the homework for a class. For various reasons, only the following pairs meet: Alice and Bob; Bob and Charles; Charles and Debbie; and Debbie and Alice. (Alice and Charles just can't stand each other, and Bob and Debbie had a relationship that ended badly.) The study pairs are shown in figure 3.10a.

In this example, the professor accidentally misspoke in class, giving rise to a possible misconception among the students in the class. Each of the students in the class may subsequently have figured out the problem, perhaps by thinking about the issue or reading the textbook. In subsequent study pairs, he or she may transmit this newfound understanding to his or her study partners. We therefore have four binary random variables, representing whether the student has the misconception or not. We assume that for each  $X \in \{A, B, C, D\}$ ,  $x^1$  denotes the case where the student has the misconception, and  $x^0$  denotes the case where he or she does not.

Because Alice and Charles never speak to each other directly, we have that  $A$  and  $C$  are conditionally independent given  $B$  and  $D$ . Similarly,  $B$  and  $D$  are conditionally independent given  $A$  and  $C$ . Can we represent this distribution (with these independence properties) using a BN? One attempt is shown in figure 3.10b. Indeed, it encodes the independence assumption that  $(A \perp C \mid \{B, D\})$ . However, it also implies that  $B$  and  $D$  are independent given only  $A$ , but dependent given both  $A$  and  $C$ . Hence, it fails to provide a perfect map for our target distribution. A second attempt, shown in figure 3.10c, is equally unsuccessful. It also implies that  $(A \perp C \mid \{B, D\})$ , but it also implies that  $B$  and  $D$  are marginally independent. It is clear that all other candidate BN structures are also flawed, so that this distribution does not have a perfect map. ■





**Figure 3.10 Attempted Bayesian network models for the Misconception example:** (a) Study pairs over four students. (b) First attempt at a Bayesian network model. (c) Second attempt at a Bayesian network model.

Draw Factor Graph

| $\phi_1(A, B)$ |       |    | $\phi_2(B, C)$ |       |       | $\phi_3(C, D)$ |  |       | $\phi_4(D, A)$ |     |  |       |       |     |
|----------------|-------|----|----------------|-------|-------|----------------|--|-------|----------------|-----|--|-------|-------|-----|
| $a^0$          | $b^0$ | 30 |                | $b^0$ | $c^0$ | 100            |  | $c^0$ | $d^0$          | 1   |  | $d^0$ | $a^0$ | 100 |
| $a^0$          | $b^1$ | 5  |                | $b^0$ | $c^1$ | 1              |  | $c^0$ | $d^1$          | 100 |  | $d^0$ | $a^1$ | 1   |
| $a^1$          | $b^0$ | 1  |                | $b^1$ | $c^0$ | 1              |  | $c^1$ | $d^0$          | 100 |  | $d^1$ | $a^0$ | 1   |
| $a^1$          | $b^1$ | 10 |                | $b^1$ | $c^1$ | 100            |  | $c^1$ | $d^1$          | 1   |  | $d^1$ | $a^1$ | 100 |
| (a)            |       |    |                | (b)   |       |                |  | (c)   |                |     |  | (d)   |       |     |

**Figure 4.1 Factors for the Misconception example**

| Assignment |       |       |       | Unnormalized | Normalized          |
|------------|-------|-------|-------|--------------|---------------------|
| $a^0$      | $b^0$ | $c^0$ | $d^0$ | 300,000      | 0.04                |
| $a^0$      | $b^0$ | $c^0$ | $d^1$ | 300,000      | 0.04                |
| $a^0$      | $b^0$ | $c^1$ | $d^0$ | 300,000      | 0.04                |
| $a^0$      | $b^0$ | $c^1$ | $d^1$ | 30           | $4.1 \cdot 10^{-6}$ |
| $a^0$      | $b^1$ | $c^0$ | $d^0$ | 500          | $6.9 \cdot 10^{-5}$ |
| $a^0$      | $b^1$ | $c^0$ | $d^1$ | 500          | $6.9 \cdot 10^{-5}$ |
| $a^0$      | $b^1$ | $c^1$ | $d^0$ | 5,000,000    | 0.69                |
| $a^0$      | $b^1$ | $c^1$ | $d^1$ | 500          | $6.9 \cdot 10^{-5}$ |
| $a^1$      | $b^0$ | $c^0$ | $d^0$ | 100          | $1.4 \cdot 10^{-5}$ |
| $a^1$      | $b^0$ | $c^0$ | $d^1$ | 1,000,000    | 0.14                |
| $a^1$      | $b^0$ | $c^1$ | $d^0$ | 100          | $1.4 \cdot 10^{-5}$ |
| $a^1$      | $b^0$ | $c^1$ | $d^1$ | 100          | $1.4 \cdot 10^{-5}$ |
| $a^1$      | $b^1$ | $c^0$ | $d^0$ | 10           | $1.4 \cdot 10^{-6}$ |
| $a^1$      | $b^1$ | $c^0$ | $d^1$ | 100,000      | 0.014               |
| $a^1$      | $b^1$ | $c^1$ | $d^0$ | 100,000      | 0.014               |
| $a^1$      | $b^1$ | $c^1$ | $d^1$ | 100,000      | 0.014               |

**Figure 4.2 Joint distribution for the Misconception example.** The unnormalized measure and the normalized joint distribution over  $A, B, C, D$ , obtained from the parameterization of figure 4.1. The value of the partition function in this example is 7,201,840.



# Too Tired to Merge These Slides

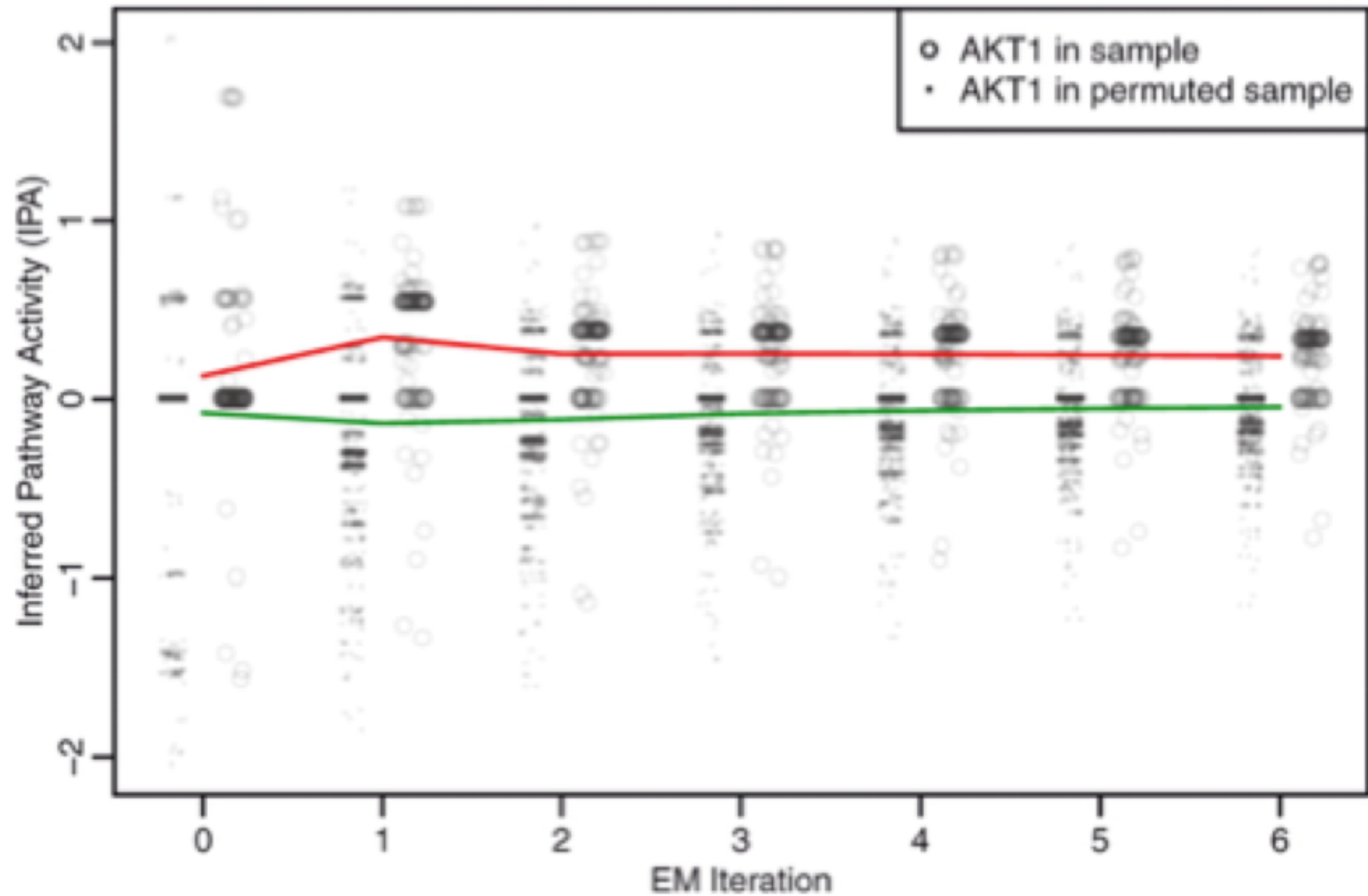
[http://www.cedar.buffalo.edu/~srihari/CSE574/Chap8/  
Ch8-GraphicalModelInference/Ch8.3.2-  
FactorGraphs.pdf](http://www.cedar.buffalo.edu/~srihari/CSE574/Chap8/Ch8-GraphicalModelInference/Ch8.3.2-FactorGraphs.pdf)

[http://disi.unitn.it/~passerini/teaching/2010-2011/  
MachineLearning/slides/09\\_inference\\_in\\_bn/talk.pdf](http://disi.unitn.it/~passerini/teaching/2010-2011/MachineLearning/slides/09_inference_in_bn/talk.pdf)

[http://www.cs.cmu.edu/~sandholm/cs15-780S11/  
slides/19-factor-graphs-mc.pdf](http://www.cs.cmu.edu/~sandholm/cs15-780S11/slides/19-factor-graphs-mc.pdf)

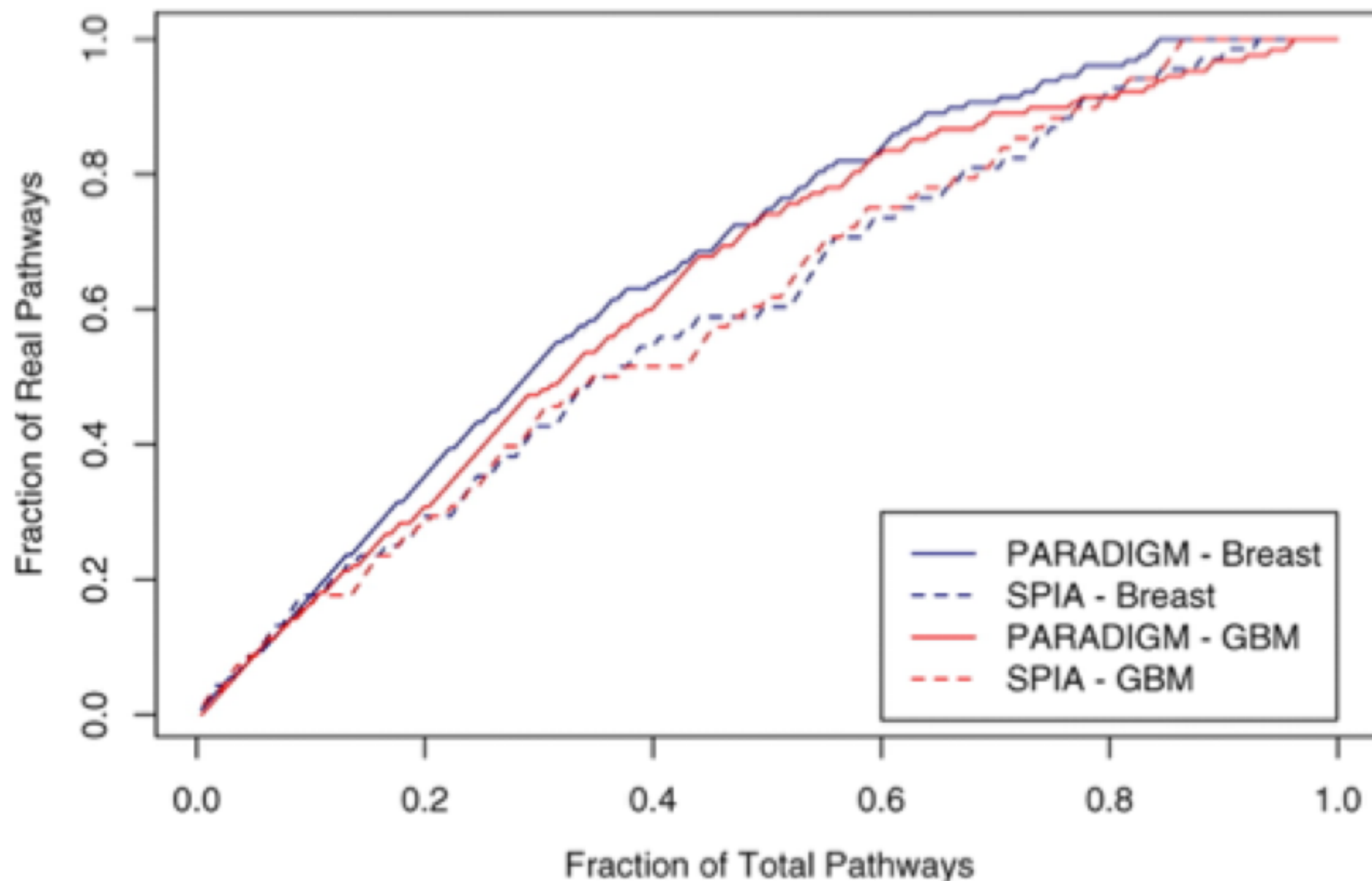
# Results

## Learning parameters for AKT1.



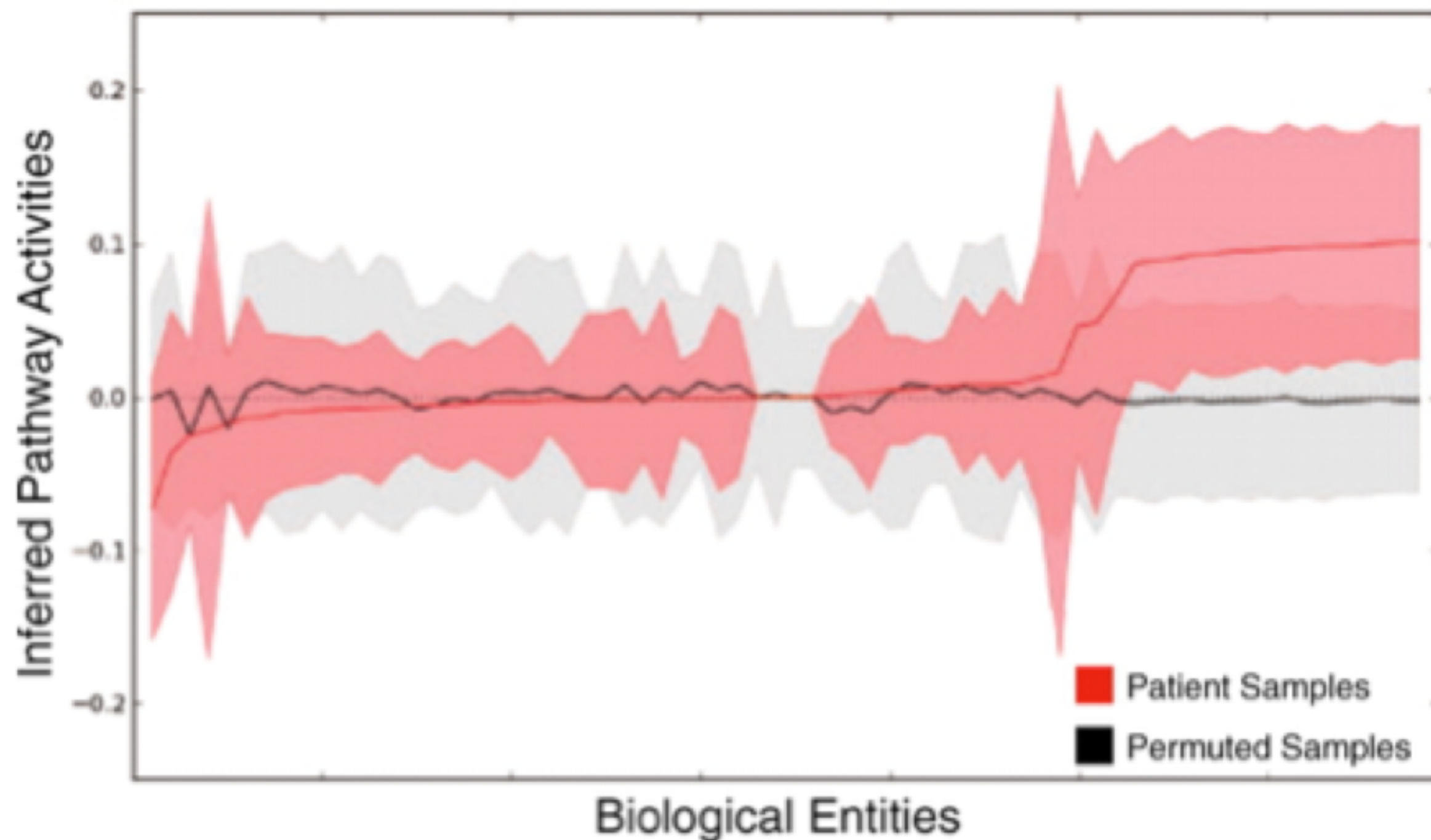
Charles J. Vaske et al. *Bioinformatics* 2010;26:i237-i245

**Distinguishing decoy from real pathways with PARADIGM and SPIA. Decoy pathways were created by assigning a new gene name to each gene in a pathway.**



Charles J. Vaske et al. *Bioinformatics* 2010;26:i237-i245

**Patient sample IPAs compared with 'within' permutations for Class I PI3K signaling events mediated by Akt in breast cancer.**



Charles J. Vaske et al. *Bioinformatics* 2010;26:i237-i245



**Table 1.**  
Top PARADIGM pathways in breast cancer

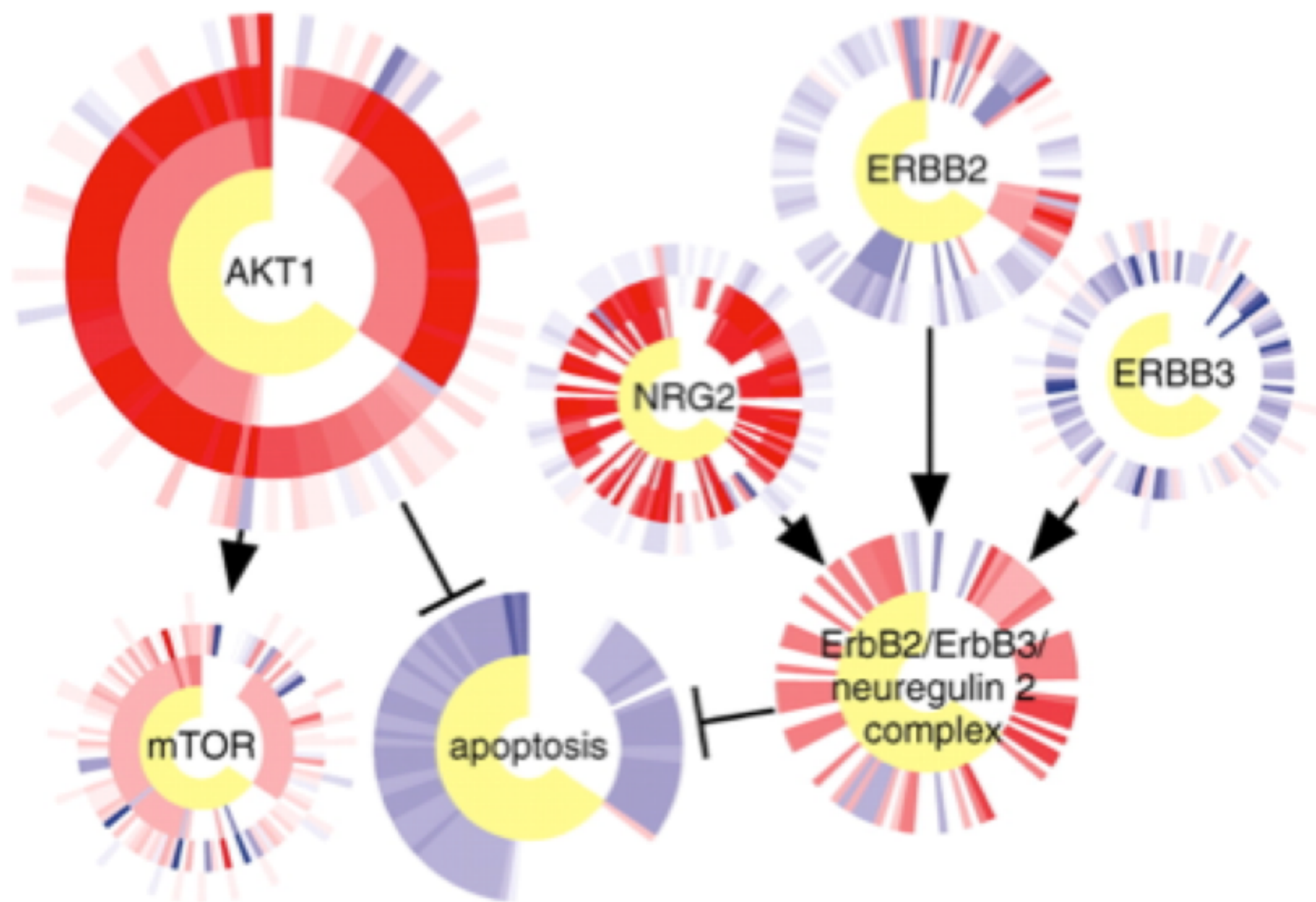
| Rank | Name  | Avg. <sup>a</sup> | SPIA? <sup>b</sup> |
|------|---|-------------------|--------------------|
| 1    | Class I PI3K signaling events mediated by Akt | 20.7              | No                 |
| 2    | Nectin adhesion pathway                       | 14.1              | No                 |
| 3    | Insulin-mediated glucose transport            | 13.8              | No                 |
| 4    | ErbB2/ErbB3 signaling events                  | 12.1              | Yes                |
| 5    | p75(NTR)-mediated signaling                   | 11.5              | No                 |
| 6    | HIF-1-alpha transcription factor network      | 10.7              | No                 |
| 7    | Signaling events mediated by PTP1B            | 10.7              | No                 |
| 8    | Plasma membrane estrogen receptor signaling   | 10.6              | Yes                |
| 9    | TCR signaling in naive CD8+ T cells           | 10.6              | No                 |
| 10   | Angiopoietin receptor Tie2-mediated signaling | 10.1              | No                 |
| 11   | Class IB PI3K non-lipid kinase events         | 10.0              | No                 |
| 13   | Osteopontin-mediated events                   | 9.9               | Yes                |
| 12   | IL4-mediated signaling events                 | 9.8               | No                 |
| 14   | Endothelins                                   | 9.8               | No                 |
| 15   | Neurotrophic factor-mediated Trk signaling    | 9.7               | No                 |

**Table 2.**  
Top PARADIGM pathways in GBM

| Rank | Name   | Avg. <sup>a</sup> | SPIA? <sup>b</sup> |
|------|--|-------------------|--------------------|
| 1    | Signaling by Ret tyrosine kinase             | 46.0              | No                 |
| 2    | Signaling events activated by Hepatocyte GFR | 43.7              | No                 |
| 3    | Endothelins                                  | 42.5              | Yes                |
| 4    | Arf6 downstream pathway                      | 42.3              | No                 |
| 5    | Signaling events mediated by HDAC Class III  | 36.3              | No                 |
| 6    | FOXO1 transcription factor network           | 35.9              | Yes                |
| 7    | IL6-mediated signaling events                | 33.2              | No                 |
| 8    | FoxO family signaling                        | 31.3              | No                 |
| 9    | LPA receptor mediated events                 | 30.7              | Yes                |
| 10   | ErbB2/ErbB3 signaling events                 | 30.1              | No                 |
| 11   | Signaling mediated by p38-alpha and p38-beta | 28.1              | No                 |
| 12   | HIF-1-alpha transcription factor network     | 27.6              | Yes                |
| 13   | Non-genotropic Androgen signaling            | 27.3              | No                 |
| 14   | p38 MAPK signaling pathway                   | 27.2              | No                 |
| 15   | IL2 signaling events mediated by PI3K        | 26.9              | No                 |

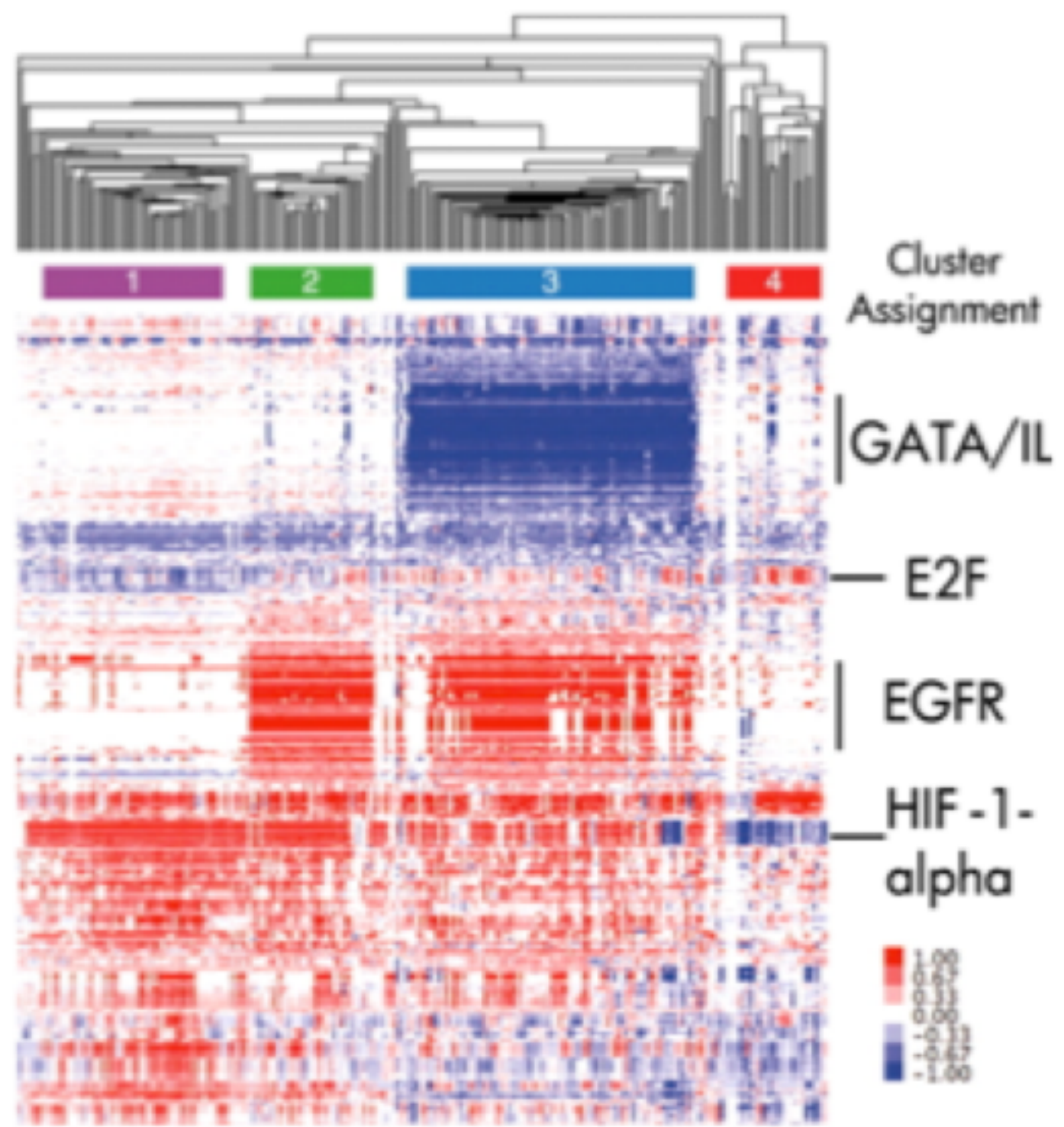


## CircleMap display of the ErbB2 pathway.



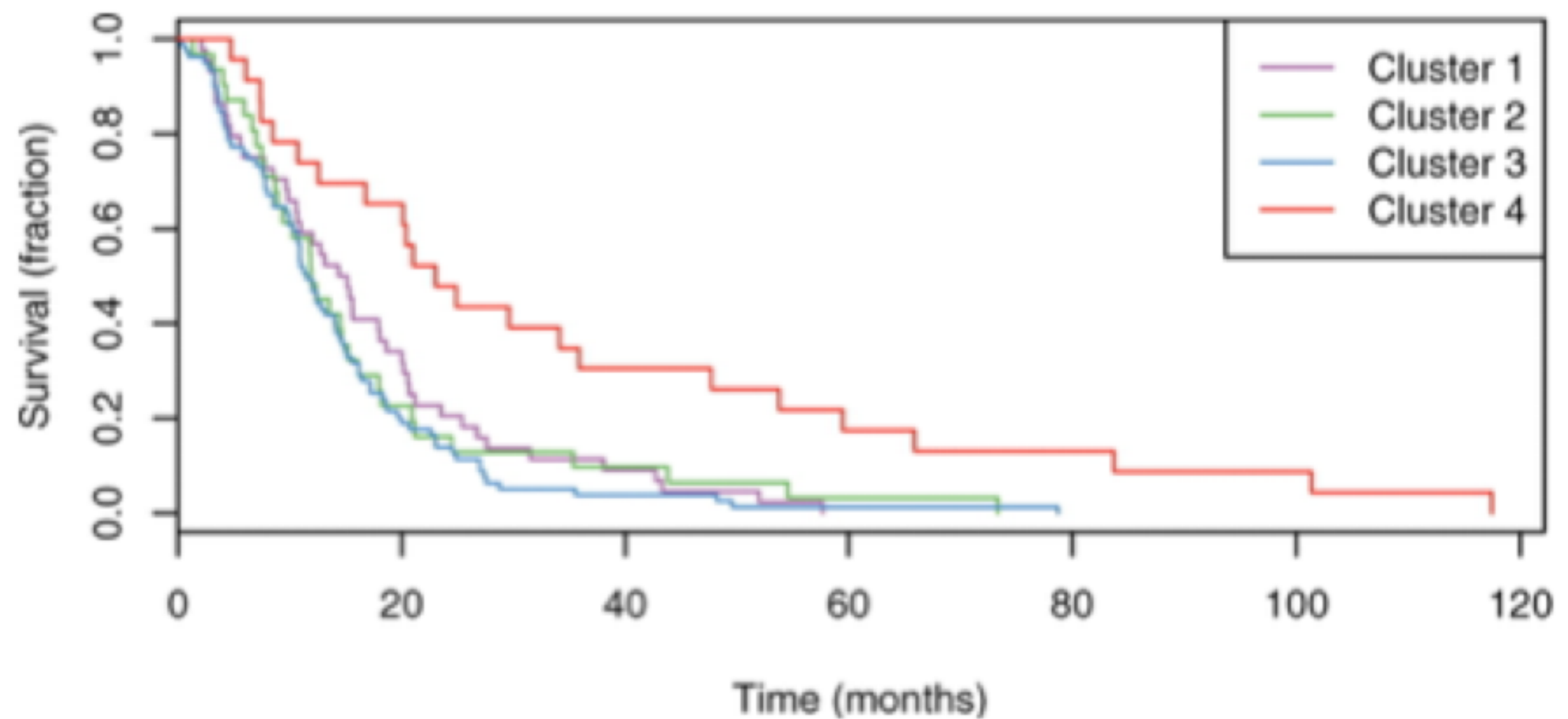
Charles J. Vaske et al. *Bioinformatics* 2010;26:i237-i245

**Clustering of IPAs for TCGA GBM. Each column corresponds to a single sample, and each row to a biomolecular entity.**



Charles J. Vaske et al. *Bioinformatics* 2010;26:i237-i245

### Kaplan-Meier survival plots for the clusters from Figure 8.



Charles J. Vaske et al. *Bioinformatics* 2010;26:i237-i245

# Future Work