Tumor Classification by Partial Least Squares Using Microarray Gene Expression Data

Supplementary Color Figures

Danh V. Nguyen
Center for Image Processing and Integrated Computing
and
David M. Rocke
Department of Applied Science

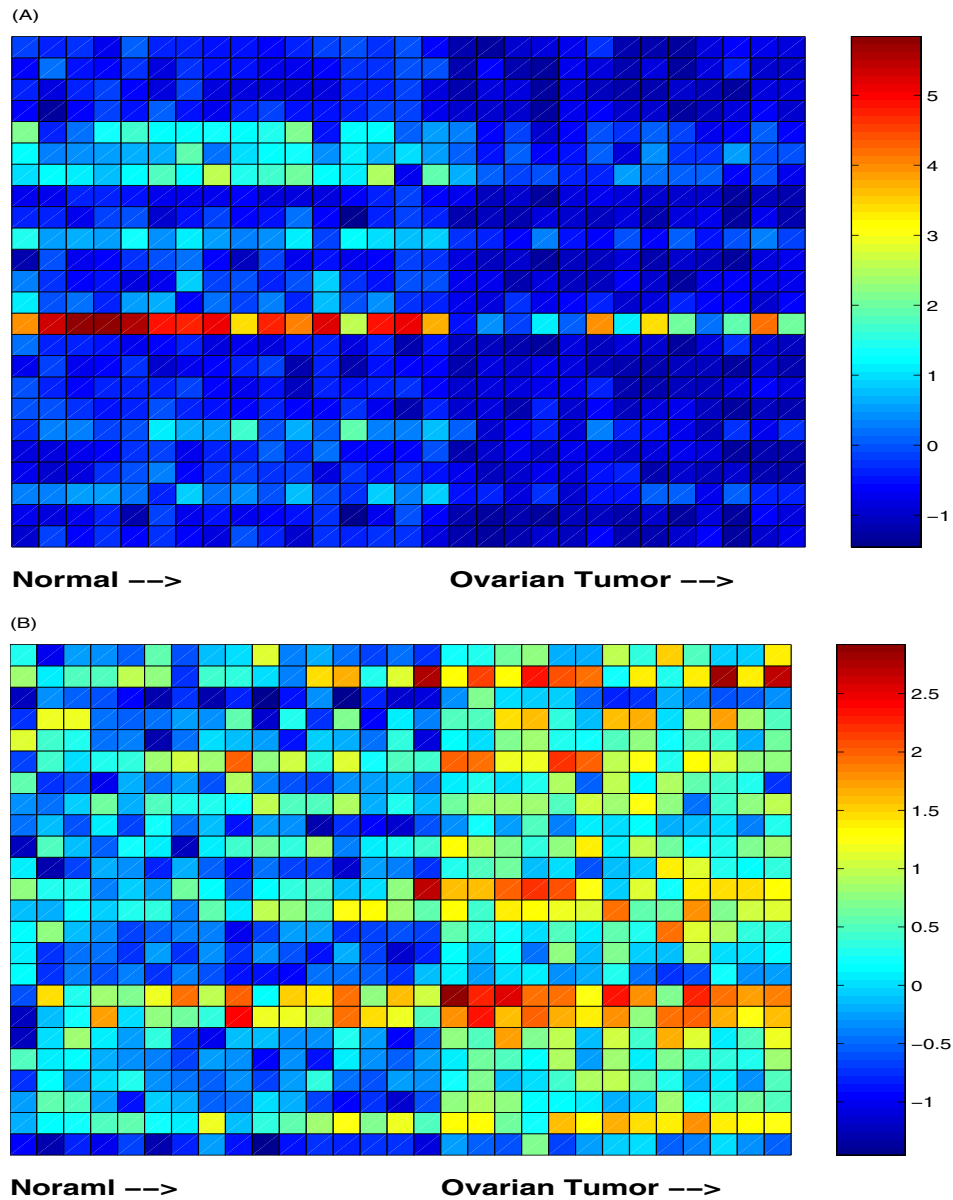University of California, Davis
Davis, CA 95616

July, 2001

(A)

Normal ––> **Ovarian Tumor ––>**

(B)

Noraml ––> **Ovarian Tumor ––>**

Figure 1: **Ovarian data**. Displayed are expression intensities of genes distinguishing normal from ovarian samples based on t-statistics. (A) Top 25 genes highly expressed in normal relative to ovarian tumor samples. (B) Top 25 genes highly expressed in ovarian tumor relative to normal samples. Expressions values here are rescaled signal − background and log transformed measurements. Expressions are standardized so that the mean is 0 and the standard deviation is 1.

(A)

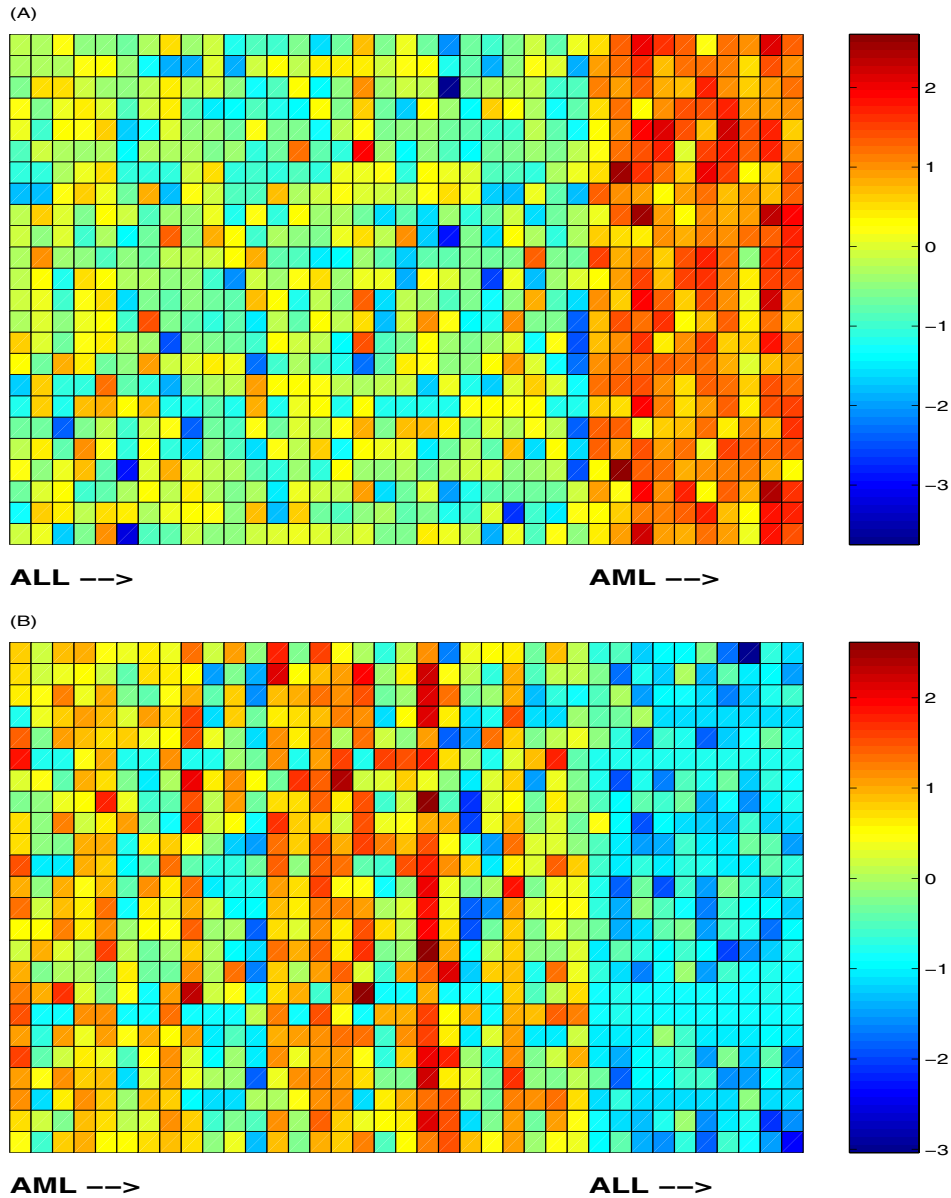ALL --> AML -->

(B)

AML --> ALL -->

Figure 2: **Leukemia data**. Displayed are expression intensities of genes distinguishing ALL from AML samples based on t-statistics. (A) Top 25 genes highly expressed in AML relative to ALL samples. (B) Top 25 genes highly expressed in ALL relative to AML samples. Expressions are log transformed and standardized so that the mean is 0 and the standard deviation is 1.
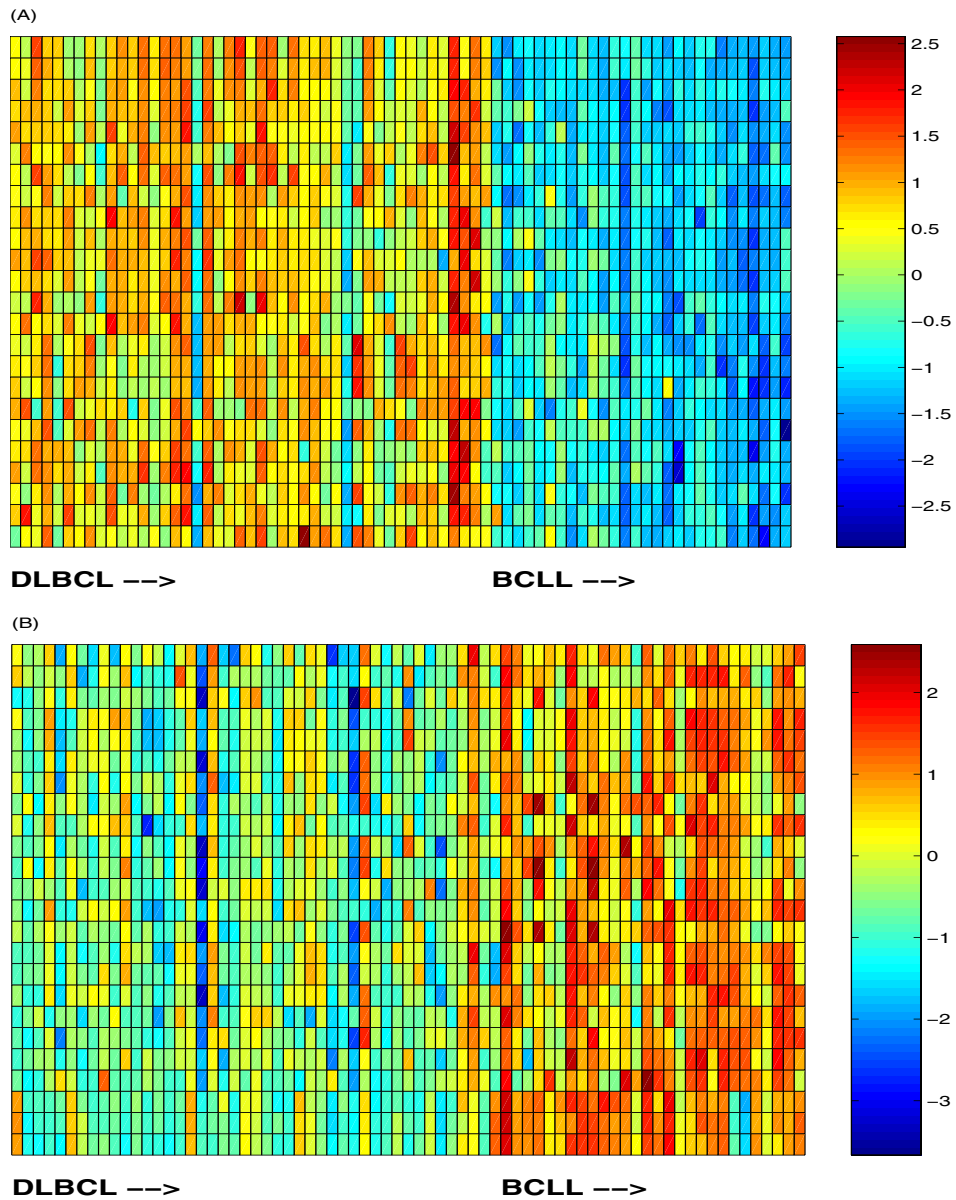
3

(A)

DLBCL --->   BCLL --->

(B)

DLBCL --->   BCLL --->

Figure 3: **Lymphoma data**. Displayed are the relative expression intensities of genes distinguishing DLBCL from BCLL samples based on t-statistics. (A) Top 25 genes highly expressed in DLBCL relative to BCLL samples. (B) Top 25 genes highly expressed in BCLL relative to DLBCL samples. Relative expression values here are $\log(\text{Cy5}/\text{Cy3})$ values and Cy5 and Cy3 values are signal $-$ background measurements. Expressions are standardized so that the mean is 0 and the standard deviation is 1.

(A)

Normal -->   *   **Colon -->**   *   *   *   *

(B)

Normal -->   *   **Colon -->**   *   *   *   *
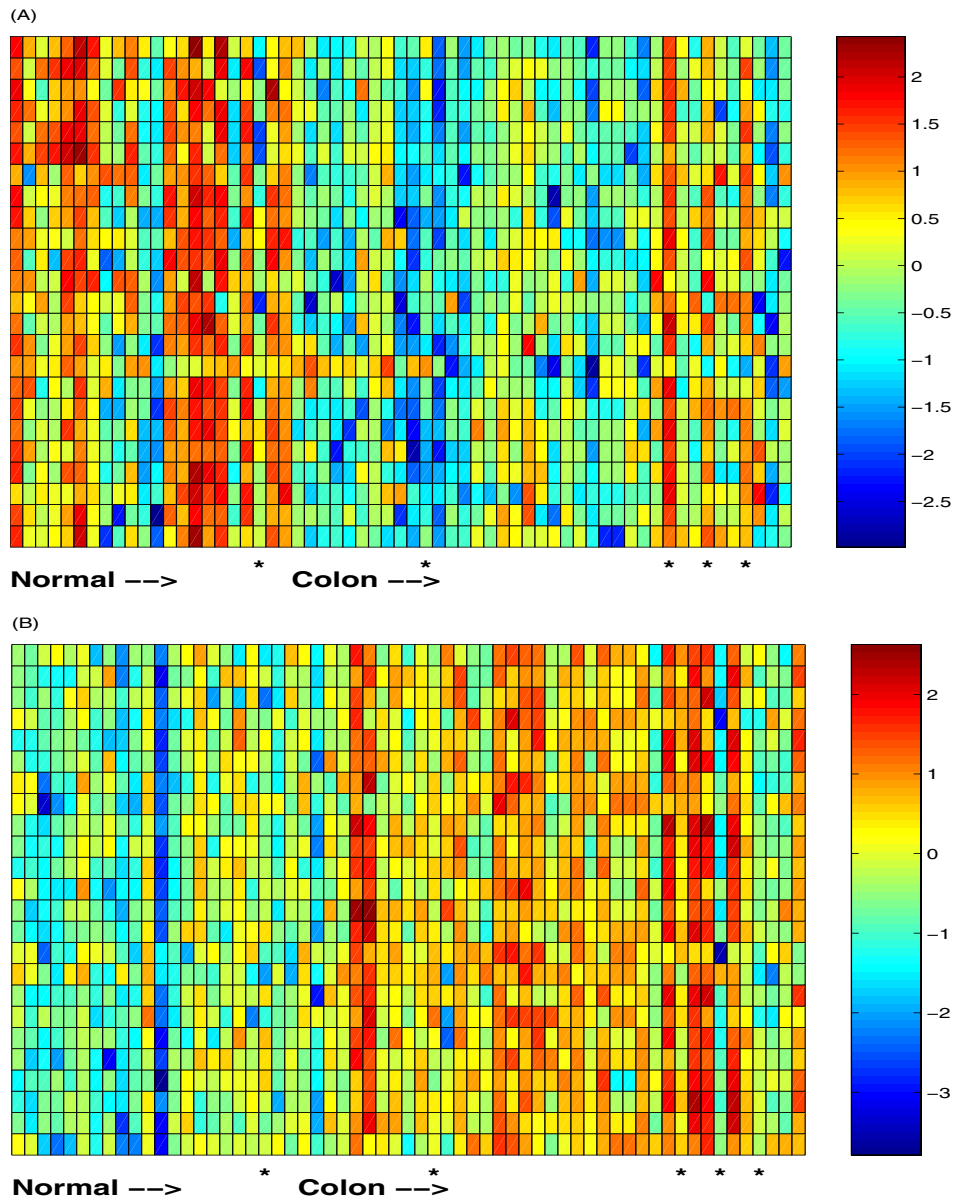
Figure 4: **Colon data**. Displayed are the expression intensities of genes distinguishing tumor from normal colon samples based on t-statistics. (A) Top 25 genes highly expressed in normal relative to colon tissue samples. (B) Top 25 genes highly expressed in colon relative to normal tissue samples. Expressions are log transformed and standardized so that the mean is 0 and the standard deviation is 1. Some samples misclassified by various methods are marked with an *.
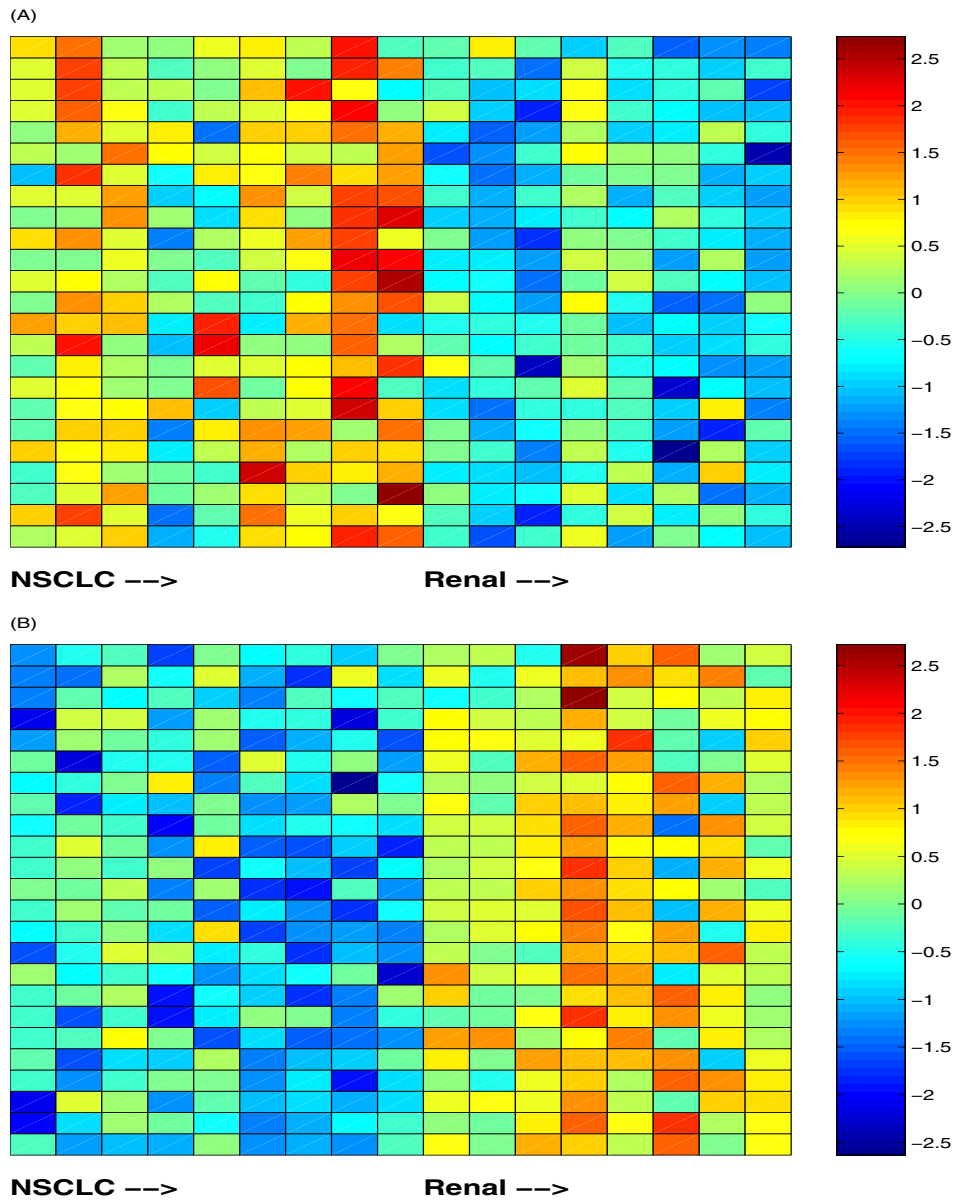
Figure 5: **NCI60 data**. Displayed are the relative expression intensities of genes distinguishing tumor from NSCLC and renal cell lines based on t-statistics. (A) Top 25 genes highly expressed in NSCLC relative to renal samples. (B) Top 25 genes highly expressed in renal relative to NSCLC samples. Relative expression values here are $\log(Cy5/Cy3)$ values and $Cy5$ and $Cy3$ values are $signal - background$ measurements. Expressions are standardized so that the mean is 0 and the standard deviation is 1.