

The GCM data set is as previously described (Ramaswamy *et al.*, 2001). The gene expression measures were calculated using Affymetrix genechip analysis software MAS4 (S. Ramaswamy, personal communication). In addition to NCI60, GCM data set (Ramaswamy *et al.*, 2001) has also been extensively utilized as a benchmark set to compare multiclass cancer classification algorithms (Ross *et al.*, 2000; Ramaswamy *et al.*, 2001; Yeang *et al.*, 2001; Ooi and Tan, 2003; Peng *et al.*, 2003). As shown in the following Table, the GA/SVM method achieved either comparable or superior LOOCV predicting accuracies to those previously reported. The GA/SVM based method also significantly reduced the number of required predicting features from 16,063 to 40. Repeating the GA/MLHD approach on this data set resulted in an LOOCV accuracy of 77.08%, similar to the previously reported 79.33% accuracy (Ooi and Tan, 2003). This correspondence supports the idea that the discrepancy between the GA/MLHD studies seen in the NCI60 classification was due to erroneous NCI60 tumor sample labeling in the previous study.

Table. Comparison of GA/SVM with other previously described multiclass tumor distinction methods.

	NCI 60 Dataset		GCM Dataset		
Classification Method	LOOCV %	Number of features (genes)	LOOCV %	Number of features (genes)	Reference
Hierarchical Clustering	81	6831	-	-	Ross <i>et al.</i>
OVA/SVM	-	-	78	16063	Ramaswamy <i>et al.</i>
OVA/SVM	-	-	81.25	16063	Yeang <i>et al.</i>
OVA/KNN	-	-	72.92	16063	Yeang <i>et al.</i>
GA/MLHD	85.37	13	79.33	32	Ooi <i>et al.</i>
GA/MLHD	70.73	12	77.08	15	This study
GA/SVM	88.52	40	80.99	40	This study