Supplementary Table 3 Cross-species comparison of mouse HCC and human ovarian cancer

	CCP		1NN		3NN		NC		SVM		LDA	
Predicted subclass	A	B	A	B	A	B	A	B	A	B	A	B
human ovarian cancer												
Cluster A (*n* = 19)	17	2	16	3	14	5	16	3	14	5	17	2
Cluster B (*n* = 25)	6	19	3	22	3	22	6	19	4	21	6	19
percent correctly classified *	82%		86%		82%		80%		80%		82%	
Mouse HCC												
DENA (*n* = 3)	0	3	0	3	0	3	0	3	0	3	0	3
Myc (*n* = 8)	6	2	6	2	8	0	6	2	6	2	6	2
E2f1 (*n* = 10)	3	7	5	5	3	7	3	7	3	7	3	7
Myc/E2f1 (*n* = 9)	1	8	5	4	2	7	1	8	2	7	1	8
Myc/Tgfa (*n* = 9)	5	4	4	5	4	5	5	4	4	5	5	4

*Percentage for correct prediction during leave-one-out cross validation
CCP, Compound Covariate Predictor; **1NN**, 1 Nearest Neighbor; **3NN**, 3 Nearest Neighbor; **NC**, Nearest Centroid; **SVM**, Support Vector Machines; **LDA**, Linear Discriminator Analysis.

Analysis procedure

Previous microarray-based gene expression profiling study identified two distinct subgroups of human ovarian cancer; cluster A and B (Reference 24 from main text). Cluster A group showed relatively higher proliferative gene expression signature than cluster B. Out of 44 Ovarian cancer examined, 19 and 25 tumors were assigned to cluster A and B respectively based on their hierarchical clustering analysis. Gene expression data were downloaded from Stanford microarray database (http://genome-www5.stanford.edu) to compare gene expression pattern with mouse HCC. Before combining two independent data, orthologous genes were selected in both microarray platform and genes were filtered by applying same criteria used previously (genes with less than 30% of missing expression data and a expression ratios of 2 or more in at least 10% of samples were only selected), expression ratios were then standardized as described previously. Total 800 orthologous genes were used in prediction methods. Human ovarian cancer data were used for training prediction methods and mouse data were assigned to prediction set. Outcome of prediction is shown in table above.