

Supplementary

Table S1 Oligonucleotide sequences used for Q-PCR analysis

Name	Sequences
E-Cadherin	F 5'- ATTTTCCCTCGACACCCGAT -3' R 5'- TCCCAGGCGTAGACCAAGA -3'
N-Cadherin	F 5'- TGCGGTACAGTGTAAGTGGG -3' R 5'- GAAACCGGGCTATCTGCTCG -3'
DDR1	F 5'- AAGGGACATTTTGATCCTGCC -3' R 5'- CCTTGGGAAACACCGACCC -3'
Snail	F: 5'-CGCCTGACTGAGCAACTGG-3' R: 5'-GGAAGCTGGCGGTAAGTGT-3'
Slug	F: 5'-CGAACTGGACACACATACAGTG-3' R: 5'-CTGAGGATCTCTGGTTGTGGT-3'
MMP9	F: 5'- AGACCTGGGCAGATTCCAAAC-3' R: 5'- CGGCAAGTCTTCCGAGTAGT-3'
Vimentin	F: 5'- AGTCCACTGAGTACCGGAGAC-3' R: 5'- CATTTCACGCATCTGGCGTTC-3'
GAPDH	R 5'-ACAACCTTGGTATCGTGGAAGG-3' F 5'-GCCATCAGCCACAGTTTC-3'

Table S2 Statistics of the three microarray databases derived from the GEO database

Dataset ID	HCC	Normal	Total number
GSE99807	4	4	8
GSE101685	24	8	32
GSE121248	70	47	107