**Supplementary Figure S1**. Clustering of gene expression data

**

*Hierarchical clustering of the 34 CTC-specific genes in the 63 mCRPC patients. Red indicates above median and green below median expression level of a gene over all samples.*

**Supplementary Figure S2.** PFS and OS in relation to the clustering groups



*Kaplan-Meier curves of (A) progression-free survival (PFS) and (B) overall survival (OS) in relation to the groups divided by the cluster that was created in Supplementary Figure S1.*

**Supplementary Figure S3.** PFS and OS in the 5-gene expression profile at baseline

C:\Users\482024\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.Outlook\P98SK0C7\Supp Figure 3 - OS and PFS vs Risk preds2.tif

*Kaplan-Meier curves of (A) progression-free survival (PFS) and (B) overall survival (OS) in relation to the 5-gene expression profile. The groups are divided into two categories based on their low (1) and high (2) predicted risk based on this 5-gene profile by BRB Array tools.*

**Supplementary Figure S4**. Cut off epithelial profile

****

*Pearson correlation between the epithelial profile that has been determined on 12 genes in metastatic breast cancer patients and the epithelial profile used in this metastatic prostate cancer group based on 6 out of those 12 genes (A). In B) a receiving operating characteristic (ROC) curve with an area under the curve (AUC) value of this threshold is shown.*

**Supplementary table S1.** Genes used for gene-expression profile

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **A** | **All 96 genes in the gene-expression panel (including 3 reference genes)** | | | | |
|  | *ABCC4* | *CDK12* | *FZD7* | *NTS* | *STYK1* |
|  | *ACPP* | *CDK4* | *GUCY1A3* | *PCGEM1* | *TACC2* |
|  | *ADRA2C* | *CEACAM5* | *HNF1A* | *PIK3CA* | *TBX3* |
|  | *AGR2* | *COL18A1* | *IGFBP4* | *PIK3CB* | *TCF7L2* |
|  | *AKR1C3* | *CREB3L4* | *IGFBP5* | *PLAU* | *TMPRSS2* |
|  | *APC* | *CRISP3* | *ITGA6* | *PMEPA1* | *TMPRSS2\_ERG* |
|  | *APP* | *CYP11A1* | *KIT* | *PPAP2A* | *UGT2B15* |
|  | *AR* | *CYP17A1* | *KLK2* | *PPARG* | *WNT11* |
|  | *AR\_V1* | *DDC* | *KLK3* | *PRKACB* | *WNT4* |
|  | *AR\_V3* | *DHCR24* | *KRT8* | *PTEN* | *WNT5A* |
|  | *AR\_V7* | *DKK1* | *LAD1* | *PTPN1* | *ZBTB16* |
|  | *AR\_V9* | *DRG1* | *LEF1* | *RAD51B* |  |
|  | *AR\_WT\_fl* | *EAF2* | *LRIG1* | *RAD51C* | Reference genes |
|  | *AXIN2* | *ELL2* | *MALAT1* | *RB1* | *GUSB* |
|  | *BMP7* | *EPCAM* | *MED12* | *S100A16* | *HMBS* |
|  | *BRCA2* | *EZH2* | *MKI67* | *SLC45A3* | *HPRT1* |
|  | *CAV1* | *FAT1* | *MSMB* | *SLCO1B3* |  |
|  | *CCND1* | *FKBP5* | *MYC* | *SOX9* |  |
|  | *CDH1* | *FOLH1* | *NDRG1* | *SP5* |  |
|  | *CDH10* | *FOXA1* | *NKX3\_1* | *SRD5A2* |  |
|  |  |  |  |  |  |
| **B** | **CTC-specific genes** | | | | |
|  | *AGR2* | *BMP7* | *FKBP5* | *MSMB* | *TACC2* |
|  | *AR* | *CCND1* | *FOLH1* | *NKX3\_1* | *TBX3* |
|  | *AR-V1* | *CDH1* | *FOXA1* | *PPAP2A* | *TMPRSS2* |
|  | *AR-V3* | *CRISP3* | *KLK2* | *PPARG* | *TMPRSS2\_ERG* |
|  | *AR-V7* | *DKK1* | *KLK3* | *SOX9* | *UGT2B15* |
|  | *AR-V9* | *EPCAM* | *KRT8* | *SP5* | *WNT5A* |
|  | *AR\_WT\_fl* | *FAT1* | *LAD1* | *STYK1* |  |

*(A) All 96 genes that were measured in the samples with available RNA. (B) The 34 genes with higher expression in CTC samples compared to the background of leukocytes.*

**Supplementary Table S2**. Univariate regression analysis of PFS

|  |  |  |  |
| --- | --- | --- | --- |
| **Univariate analysis PFS** | |  |  |
|  |  |  |  |
| **Gene** | **Exp(B)** | **95% CI** | **sig** |
| AGR2 | 1.087 | 1.02-1.16 | 0.008 |
| AR | 1.069 | 1.0-1.15 | 0.059 |
| AR\_V1 | 1.012 | 0.96-1.07 | 0.653 |
| AR\_V3 | 1.034 | 0.96-1.12 | 0.393 |
| AR\_V7 | 1.013 | 0.93-1.10 | 0.768 |
| AR\_V9 | 0.996 | 0.94-1.06 | 0.905 |
| AR\_WT\_fl | 1.066 | 1.00-1.14 | 0.071 |
| BMP7 | 1.042 | 0.99-1.10 | 0.141 |
| CCND1 | 1.031 | 0.92-1.16 | 0.607 |
| CDH1 | 1.037 | 0.96-1.13 | 0.391 |
| CRISP3 | 0.962 | 0.86-1.08 | 0.498 |
| DKK1 | 1.072 | 1.01-1.14 | 0.022 |
| EPCAM | 1.056 | 0.94-1.18 | 0.341 |
| FAT1 | 1.068 | 1.01-1.13 | 0.030 |
| FKBP5 | 1.004 | 0.84-1.21 | 0.969 |
| FOLH1 | 1.047 | 1.00-1.10 | 0.074 |
| FOXA1 | 1.078 | 1.01-1.15 | 0.024 |
| KLK2 | 1.055 | 0.99-1.13 | 0.115 |
| KLK3 | 1.034 | 0.97-1.10 | 0.310 |
| KRT8 | 1.065 | 0.96-1.18 | 0.223 |
| LAD1 | 1.036 | 0.99-1.09 | 0.146 |
| MSMB | 1.028 | 0.98-1.08 | 0.261 |
| NKX3\_1 | 1.012 | 0.78-1.31 | 0.928 |
| PPAP2A | 1.058 | 0.93-1.20 | 0.376 |
| PPARG | 0.985 | 0.85-1.14 | 0.832 |
| SOX9 | 1.030 | 0.98-1.09 | 0.283 |
| SP5 | 1.041 | 0.99-1.10 | 0.115 |
| STYK1 | 1.013 | 0.95-1.08 | 0.699 |
| TACC2 | 1.040 | 0.99-1.09 | 0.129 |
| TBX3 | 1.050 | 0.99-1.11 | 0.092 |
| TMPRSS2 | 1.058 | 1.00-1.12 | 0.037 |
| TMPRSS2\_ERG | 1.026 | 0.98-1.08 | 0.318 |
| UGT2B15 | 1.052 | 0.98-1.13 | 0.173 |
| WNT5A | 1.033 | 0.96-1.12 | 0.407 |

*Univariate Cox Regression analysis of all 34 genes in comparison with PFS in 63 patients with available gene expression data.*