Supporting information

Supplementary Methods

Normalization, filtering and statistical analysis of gene expression microarray data: After background correction and $\log_2$ transformation of signal intensities normalization was carried out using quantile normalization by arithmetical mean of the distribution and Tukey's median polish. The acquired CHP files containing background corrected, $\log_2$ transformed and normalized intensity data were then transformed to TXT files using Affymetrix APT-1.15.0 software (Affymetrix, Inc., Santa Clara, CA, USA) and inserted to Bioconductor BRB-ArrayTools 4.3.0 (Richard Simon and Amy Peng Lam, National Cancer Institute, Bethesda, USA). Genes were excluded if less than 20% of expression data have at least a 1.5-fold change in either direction from gene's median value; p-value of the log-ratio variation was greater than 0.05 or at least 50% of data was missing or filtered out. Multiple probes or probe sets were reduced to one per gene symbol by using maximally expressed probe or probe set measured by average intensity across arrays.

Volcano plot filtering and paired t-test with random variance model was applied to reveal differentially expressed genes between primary and metastatic cell lines, considering significant at the nominal 0.05 level with at least 2 fold-change. Exact multivariate permutation test was computed based on 1000 random permutations. The maximum allowed proportion of false-positive genes (FDR) was less than 0.1 and the confidence level of false discovery rate assessment was 80%.

Reverse transcription and qRT-PCR analysis workflow: Reverse transcription was carried out with High Capacity cDNA Reverse Transcription Kit, according to the protocol of the supplier (Life Technologies Corporation, Carlsbad, CA, USA), using 1000 ng total RNA. We used primers in 0.4 μM and UPL probes in 0.2 μM final concentrations per reaction. Each reaction (20 μl in volume) contained 50 ng cDNA and was run in triplicate on LightCycler 480 instrument (Roche Magyarország Kft., Budaörs, Hungary) with the following thermal profile: (1) activation at 95°C for 10 min, (2) amplification (45 cycles): denaturation at 95°C for 10 sec, annealing at 60°C for 30 sec, extension at 72°C for 1 sec, and (3) cooling at 40°C for 30 sec.

Matrigel invasion assay: Before invasion assay the chambers were rehydrated for 2h at 37°C with serum-free RPMI1640 medium. During incubation cell lines with 50-70% confluence were harvested, counted and resuspended with 500μl serum-free medium. The upper chamber of insert was filled with 500μl of cell suspension in serum-free medium ($5x10^4$ cells/well).

Medium supplemented with 10% FBS (750µl) was applied as chemoattractant. Tumor cells were incubated for 24h at 37°C. After non-invading cells have been removed with a sterile cotton swab, the invading cells at the lower layer were fixed with 500µl 100% methanol for 15 min at -20°C and stained with hematoxylin-eosin. The average number of invaded cells was counted using light microscope in 7 different visual fields at 200X magnification. Data were presented as mean±SD of three independent experiments. To select invasive clones, parallel experiments were run with the same conditions. After 24h incubation at 37°C the invading cells at the lower layer were extracted from the membrane using 50µl 0.5% trypsin/0.2% EDTA solution (Sigma-Aldrich Inc., St. Louis, MO, USA) for 1 min at 37°C. Blocking trypsin solution with 1ml serum supplemented medium, cells were incubated for 4h at 37°C till adhesion. Cells were then cultured following standard protocol till 90% confluence and passed to a T25 flask. The selected invasive clones were checked whether they maintain the invasive property comparing to the original cell line and increased invasiveness was observed (Supplementary Figure 3).

Supplementary Figure 1 **–** Pathway analysis of the 413 significantly downregulated genes in metastases. Bonferroni correction was applied with a p-value ≤0.01. **(A)** Altered molecular pathways with at least 5 observations in the selected gene subset. Columns represent the number of genes included in the given pathway, whereas lines represent the significance of the given pathway. **(B)** Molecular functional characterization of the altered genes based on the GO classification.

Supplementary Figure 2 **–** Different expression patterns of selected integrins (ITG) in matched melanoma cell lines: **(A)** WM278 (primary; black dots) and WM1617 (lymph node metastasis; empty dots) cell line pair, **(B)** WM115 (primary; black dots) and WM266-4 (cutaneous metastasis; empty dots) cell line pair, **(C)** WM983A (primary; black dots), $WM983A^{INV}$ (selected invasive primary; gray dots) and WM983B (lymph node metastasis; empty dots) cell line trio, and **(D)** WM793B (primary; black dots), $WM793B^{INV}$ (selected invasive primary; gray dots) and 1205Lu (lung metastasis; empty dots) cell line trio. Sectors of polar charts represent different ITGs. The relative $\log_2$ transformed expression levels increase from the centre towards the outer edge of the chart. Every dot represents the exact relative mRNA level of the corresponding integrin on a $\log_2$ scale. The ITG pattern based on the combined expression changes of the 9 selected ITGs (A2, A3, A4, A6, A9, B1, B3, B5 and B8) was similar in primary samples with regional lymph node and cutaneous metastasis compared to the WM793B forming

distant organ metastasis (lung). GAPDH and ACTB were used as internal control genes, and melanocyte was used as a calibrator sample. Relative mRNA levels are based on the average of three replicates.

Supplementary Figure 3 – Result of matrigel invasion analysis of extracted invasive clones ($WM983A^{INV}$ and $WM793B^{INV}$) comparing to the original cell lines. Both $WM983A^{INV}$ and $WM793B^{INV}$ clones acquire an increased rate of invasion potential. Invasive capacity was approximately 5-fold and 6-fold higher than the original WM983A and WM793B cell lines, respectively. Invasion rate was provided as the ratio of average number of selected clones and original cell lines. Invaded cells were stained with hematoxylin-eosin, and counted using light microscope in 7 different visual fields at 200X magnification. In the table below the diagram, fold-change values of mRNA level of the selected ITGs between invasive primary clones vs. corresponding original primary were represented. Fold-change values were acquired using the following equation: relative ITG mRNA level (invasive clones) divided by the relative ITG mRNA level (original cell line). In case of lymph node metastatic WM983A primary melanoma cell line all the examined ITGs were downregulated in the selected invasive clones, in contrast to the lung metastatic primary melanoma cell line WM793B.

Supplementary Table 1 - Primer sequences and UPL probes used in RT-qPCR experiments.

Supplementary Table 2 - Differentially expressed genes in cell lines derived from melanoma metastases compared to their primary pairs (N=438). Melanocyte was used as a reference array to normalize expression intensity. Genes are significant at the nominal level of $p \leq 0.05$ with a greater than two-fold change in their expression.

Supplementary Table 3 - Biological processes affected in the metastatic melanoma cell lines. Biological processes are significant at the nominal level of $p \leq 0.01$ using Bonferroni correction and containing at least 5 observations in the selected downregulated gene subset.

Supplementary Table 4 - Significantly altered molecular pathways in the metastatic melanoma cell lines. Pathways are significant at the nominal level of $p \leq 0.01$ using Bonferroni correction and containing at least 5 observations in the selected downregulated gene subset.

Supplementary Table 5 - Downregulated genes in the metastatic melanoma cell lines are included in 4 molecular functional groups. Pathways are significant at the nominal level of

$p \leq 0.01$ using Bonferroni correction and containing at least 5 observations in the selected gene subset.

Supplementary Table 6 - Class comparison analysis results of TCGA SKCM RNAseq v2.0 data. A total of 1065 differentially expressed genes were revealed between primary melanomas and at least one of the metastasis subgroups (regional lymph node, (sub)cutaneous metastases or distant organ metastases). A multivariate permutation test was computed based on 1000 random permutations. The maximum allowed proportion of false-positive genes (FDR) was less than 0.1, and the confidence level of false-discovery rate assessment was 80%. The univariate test was a two-sample t-test, considering $p<0.01$ as statistically significant, and the fold-change was greater than 2.

Supplementary Table 7 - Pathway analysis of the 1065 altered genes between primary melanomas and different metastatic subgroups (regional lymph node, (sub)cutaneous metastases or distant organ metastases) based on TCGA RNAseq v2.0 data. Bonferroni correction was applied using a p-value ≤ 0.05, and classes with at least 5 observations in the selected subset were considered as altered pathways.