Detailed description of gene signatures

Immune response gene signatures

Four QTC clusters showed a significant association with immune response, QTC1, QTC8, QTC10, and QTC17. QTC1 was highly enriched for GO:0002376 immune system process (p=9.2x10^{-31}) and GO:0046649 lymphocyte activation (7.1x10^{-7}), QTC8 for GO:0002376 immune system processes (p=9.7x10^{-5}), QTC10 for GO:0002474 antigen processing and presentation of peptide antigen via MHC class I (p=2.1x10^{-5}), and QTC17 for the KEGG pathway hsa04060 cytokine-cytokine receptor interaction (p=2.3x10^{-5}), suggesting that these gene clusters represent different aspects of the immune response system. QTC1 and QTC8 were highly correlated (r=0.85) whereas QTC10 and QTC17 appeared more independent. QTC1 and QTC8 were more strongly expressed in the MS2b.1 subtype whereas QTC10 and QTC17 were more strongly expressed in the MS2b2.2 subtype.

QTC1 and QTC8 were merged and subsets of immune-related genes are shown in Figures 1A and 1B. The signature had a prominent activated T-cell theme including key genes ranging from T-cell stimulatory chemokines, T-cell receptor complex genes, T-cell specific markers, as well as signal transducers and effector genes of cytotoxic T-cells. Among the early T-cell activator genes were CCL2, CCL5, IL2RA, IL2RB, and IL7R. CCL2 and CCL5 are known to be produced by tumor cells and induce transendothelial migration and chemotaxis of effector T-cells (Brown et al. 2007, Soria and Ben-Baruch, 2008). CCL5 co-localizes with tumor infiltrating leukocytes and the CCL5 concentrations correlates to CD8+ T-cell infiltration in the epithelial areas of tumors (Negus et al. 1997). IL2RA and IL2RB are receptors for the T-cell stimulatory chemokine IL-2. IL7R is the receptor for IL-7, a survival and proliferation factor of mature T-cells that expands the T-cell population with a bias towards CD8+ cells (Rosenberg et al. 2006). CD3D and CD3G encode the delta and gamma subunits of the CD3 modules that act as stabilizers and signal transducers in the CD3-T-cell receptor complex. CD8A, a cytotoxic T-cell specific antigen with co-receptor function in the T-cell receptor binding to MHC class I antigens. In addition several co-stimulatory factors to T-cell signaling were included e.g., CTLA4 and CD86, several T-cell Intracellular signaling molecules such as FYN, LAT2, and ITK, as well as several effectors genes such as four granzyme (GZM) genes, perforin (PRF1), and Lymphotoxin A and B (LTA and LTB). These findings associate expression of the QTC1 and QTC8 with lymphocyte infiltration. Expression of CD3 was validated at protein level by immunohistochemistry (Figure 2, bottom row for each subtype). For each core, the binned CD3 score was directly compared to the corresponding sample mean mRNA expression of the three genes that encode the core chains of the CD3 molecule, CD3D, CD3E, and CD3G. A strong and highly significant correlation was observed (r = 0.41, p<10^{-17}). Ten representative CD3 stained cores for each subtype are shown in Figure 2 (bottom row for each subtype).

QTC10 was significantly enriched for GO:0002474, antigen processing and presentation of peptide antigen via MHC class I (p=2.0x10^{-4}) and the equivalent KEGG pathway (hsa04612) (p=6.5x10^{-4}). The expression of this gene signature is shown in Figures 1C and 1D. The cluster included HLA class I genes HLA-A, HLA-B, and HLA-F genes, transporter of endogenous peptides genes TAP1 and TAP2, as well as PSMB9 that replaces catalytic subunit 1 (PSMB6) in the immunoproteasome. Hence, the presence of this signature signals an active antigen presenting machinery.

The QTC17 cluster contained several genes related to chemotaxis of the neutrophil/monocyte lineage of myeloid leukocytes (Figures 1E and 1F). The signature included a total of 6 chemokine genes of which the
best characterized are CCL3 and CXCL2, encoding macrophage inflammatory proteins 1α (MIP-1α) and 2α (MIP-2α), respectively. Included were also the CCL3 receptor gene CCR1, expressed in monocytes, macrophages, dendritic cells, NK-cells (Su et al. 1996), as well as the non-chemokine receptor granulocyte CSF receptor gene (CSF3R), critical for recruitment and survival of granulocytes (Touw and Van de Geijn, 2007). This indicates the presence of myeloid cells, particularly of the monocyte/granulocyte lineage, that show high expression of the QTC17 genes.

**Figure 1.** Expression of immune related signatures. (A) T cell activation gene clusters (QTC1 and QTC8) were merged and gene expression is shown for individual samples, and (B) shown as group means. Expression of antigen presenting machinery associated genes (QTC10) shown for (C) individual samples, and (D) shown as group means. Expression of myeloid cell associated gene clusters (QTC17) is shown for (E) individual samples, and (F) shown as group means.
Figure 2. Protein expression of ACTA2 (α-SMA) and CD3. Ten representative cores are shown for each subtype. Strong staining of both markers is seen in MS2b1 cases. MS2b2.2 cases show staining for ACTA2 (α-SMA).
References


ECM and myofibroblast gene signatures

The QTC4 gene signature showed low expression in MS1 and MS2a1 and varying expression in the other MS2 subtypes, with the strongest expression in MS2b1 (Figure 3). GO terms significantly enriched in this cluster included GO:0030198, extracellular matrix organization (p=1.38x10^-9) and GO:0007155 cell adhesion (p=1.92x10^-9), indicating that a major remodeling of the ECM occurs in the MS2b1 class of tumors. The signature included genes for collagens such as collagens 1A2, 3A1, 5A1, 5A2, 6A1, 6A2, 6A3, and 8A1, proteoglycans LUM and DCN, the basal lamina component LAMA4, and SPARC, a modulator of cell-ECM contact. In addition to the general ECM related genes a number of genes known to be specifically expressed in myofibroblasts were included in the cluster, notably ACTA2 and VIM, currently used as markers for myofibroblasts (Eyden et al. 2009). In addition, TGFB3, EDNRA, and PDGFRB encoding signaling molecules capable of inducing a myofibroblast phenotype, were also a part of the signature (Heer et al. 2010, Bostrom et al. 1996, Préfontaine et al. 2008). Furthermore, the MS2b1 class of tumors showed similar increase in CNN1 expression known to be a features of myoepithelial cells (Foschini et al. 2000) although CCN1 was not found in the cluster. The normal urothelium is negative for the common myofibroblast markers ACTA2, VIM, CNN1, and TGFB1 as reported by the Human Protein Atlas (www.proteinatlas.org). Expression of ACTA2 was validated at protein level by immunohistochemistry. For each core, the ACTA2 protein expression score was directly compared to the corresponding sample mRNA expression of the ACTA2 gene. A moderate but highly significant correlation was observed (r = 0.26, p< 10^-17). Ten representative ACTA2 stained cores for each subtype are shown in Figure 2 (top row for each subtype). Taken together, the data indicate a strong upregulation of a wound healing signature not normally expressed in the urothelium.
Figure 3. Expression of ECM/Myofibroblast associated gene cluster (QTC4). (A) Gene expression given for individual samples. (B) Gene expression given as group means. The bottom two panels show expression of myofibroblast marker genes of which ACTA2, VIM, TGFB3, PDGFRB, and EDNRA were included in QTC4, whereas CNN1 was not included in QTC4 but shows similar expression pattern.

References


Angiogenesis gene signature

QTC 13 showed no significant GO term enrichment. However, a closer look at the signature revealed a number of genes associated with endothelial development and maintenance. The expression of this signature is strongest in the MS2b1 group of tumors (Figure 4A and 4B). The signature includes receptors for angiogenic factors such as endothelin (EDNRB), angiopoietin (TEK), and spingosine-1-phosphate (EDG1), as well as the endothelial cadherin (CDH5), and two transcription factors directing lymphangiogenesis and angiogenesis, respectively (SOX17, SOX18). EDG1 expression has been shown to be required for tumor angiogenesis (Chae et al. 2004). Adrenomedullin receptor genes CALCR2 and RAMP3 also provide a link between hypoxia signaling and angiogenesis in this signature as adrenomedullin signaling is hypoxia responsive. (Zudaire et al. 2003). Anti-correlated to this cluster is the gene EGLN3 which is a negative regulator of hypoxia-inducible factor (HIF) at normoxia. We compared our results to that of Croix et al. (2000) that derived an angiogenesis signature by comparing transcript levels (SAGE) in tumor endothelium and normal endothelium. The genes corresponding to the most abundant transcript tags in tumor vasculature have largely the same profile as our angiogenesis signature (Figure 4C and 4D) suggesting that both signatures can be used to detect the angiogenic phenotype in bladder tumor tissue.

**Figure 4.** Expression of angiogenesis associated gene cluster (QTC13). (A) Gene expression given for individual samples. (B) Gene expression given as group means. Expression of the tumor endothelium signature of Croix et al. is given for (C) individual samples, and (D) given as group means.
References


Cell cycle gene signature

QTC3 corresponded to a prominent gene cluster associated with cell cycle regulation showing high expression in the MS2 group of tumors, with the exception of MS2b1 and MS2b2.1. A closer inspection of the QTC signature indicated that the majority genes were associated with the S, G2, and M phases of the cell cycle and did not include genes e.g., CCND1 typical for the G1 phase. We therefore performed a supervised selection of genes with known key regulatory functions of the cell cycle and selected among these the top ANOVA genes (all with Bonferroni adjusted p<10^{-10}) from this list. The resulting 46 genes formed two distinct gene expression patterns with one group predominantly expressed in the MS1 and one in the MS2 tumors (Figure 5A and 5B). CCND1 showed high expression in the MS1 tumors and in the MS2b2.1 subtype. In normal urothelium CCND1 is expressed in the supra-basal cells as reported by the Human Protein Atlas (www.proteinatlas.org). CCND2 showed particular high expression in MS2b1 tumors. This latter finding is in agreement with MS2b1 being infiltrated with immunologic cells and that CCND2 is preferentially expressed in cells derived from the hematological lineages (Lukas et al. 1995, White et al. 2006). Cyclins associated with late G1, S, and the G2 phase, CCNE, CCNA, and CCNB, respectively, all show high expression in MS2 tumors. These cyclin genes show a scattered expression pattern in normal urothelium. The association of MS2 cases with late cell cycle genes was underlined by the high expression of the CCNB activators CDC25A, CDC25B, and CDC25C in MS2, and high expression of the CDC25 inhibitor WEE1 in the MS1 tumors. Even though RB1 did not show a differential expression across the molecular subtypes, RBL2 (p130) showed high expression in MS1 whereas RBL1 (p107) showed expression in MS2. Intriguingly three ID genes, ID1-3, showed an expression pattern similar to CCND1 i.e., high in MS1 tumors. The ID genes are expressed in normal urothelium, but not distinctively localized to a specific cell layer. ID2 is known to interact with RBL2 and may influence the activity of the RBL2-E2F4/F5 complexes that inhibit cell growth in the G0 phase (Zebedee and Hara, 2001). Increased expression of ID1, on the other hand, has been shown to extend the lifespan of epithelial cells by rescuing the cells from senescence (Zebedee and Hara, 2001). In addition to the genes mentioned, MS2 tumors showed increased expression of several genes related to chromosome segregation and cell division such as BUB1, BUB1B, and CDC20, several CENP genes as well as AURKA and AURKB. In normal urothelium these genes usually show a scattered expression of single positive cells (Human Protein Atlas). Taken together, the expression pattern of the cell cycle genes suggests that the cell cycle activity in MS1 tumors is primarily engaged in releasing the cells from G0 to G1 i.e., into actively growing cells, whereas in MS2 the cell cycle has reached maximum activity. The data also suggest that MS1 tumors are still
dependent of the mitogen driven part of the cell cycle i.e., CCND1 activation, whereas MS2 tumors evade the cell cycle restriction point by other means. The exception being the MS2b2.1 group of tumors that did not show as prominent expression of the late cell cycle genes as the MS2a and MS2b2.2 subtypes; the MS2b2.1 subtype showed high CCND1 expression reminiscent of the situation in the MS1 tumors.

Figure 5. Gene expression of the most significant genes involved in cell cycle control identified by ANOVA analysis (Bonferroni adjusted $p<10^{-10}$) between the subtypes. Distinct clusters of early cell cycle genes (top panel), genes up regulated in the infiltrated subtype (middle panel), and Late cell cycle genes (Bottom panel) were observed. (A) Gene expression for individual samples. (B) Gene expression given as group means.

References


Zebedee Z, Hara E. Id proteins in cell cycle control and cellular senescence. Oncogene. 2001 20:8317-8325
Keratin gene expression
A total of 18 keratin genes passed the ANOVA test (Bonferroni corrected p<0.05) and of these 15 are shown in Figure 6A and 6B; KRT12 and KRT81 were omitted from further analysis due to limited information on normal expression patterns. The primary keratins of simple epithelial cells, KRT8/KRT18 and the secondary keratins of simple epithelial cells KRT7/KRT19, known to frequently be co-expressed respectively, were predominantly expressed in the MS1 and MS2a molecular subtypes but also in the MS2b2.1 subtype. All four are general keratins and expressed in all cell layers of normal urothelium (Alonso et al. 2009, Human Protein Atlas www.proteinatlas.org). KRT20, specifically expressed in differentiated umbrella cells, was expressed in MS1b, MS2a1, and MS2a2, but not in the other subtypes. Particularly, MS1a showed a lower expression of KRT20 as compared to the MS1b, MS2a1, and MS2a2, group of tumors. In addition, the established association between KRT20 expression and urothelial cell differentiation was confirmed by the finding that the KRT20 expressing subtypes also expressed the uroplakin family of genes (UPKA1, UPKA2, UPK2, UPK3A, and UPK3B), expressed in terminally differentiated umbrella cells (Figure 6C). KRT13 showed a characteristic expression pattern with high expression in MS1a and MS1b, and lower expression in the MS2 subtypes, except for MS2b2.1. KRT4 showed a less distinct expression pattern but was mainly expressed in MS2b2.1 and MS2b2.2. In the normal urothelium KRT13 is localized to the basal and intermediate cells, but not to superficial umbrella cells. KRT4 expression is scattered but mainly confined to the basal and suprabasal cell layers. The basal keratins KRT5, KRT15, and KRT17 all showed relatively high expression in MS2b2.1 and MS2b2.2. In addition, KRT5 and KRT15 showed high expression in MS1a and KRT17 in MS1a and MS1b. The additional basal keratin KRT14 only showed high expression in MS2b2.2. The KRT6 and KRT16 keratins associated with proliferative activity and not expressed in normal urothelium (Human Protein Atlas www.proteinatlas.org), showed expression in MS2b2.1 and particularly in MS2b2.2. The results for keratin expression are summarized in Figure 6B as group means. This figure indicates that MS1a is associated with the non-proliferation basal keratins, i.e., KRT5/KRT15/KRT13 high, and with a lower level of differentiation, i.e., KRT20 low. The opposite pattern is seen in MS2a1 and MS2a2, whereas MS1b may be considered intermediate between these two phenotypes. MS2b2.2 differed radically from the other subtypes by showing strong expression of KRT6A, 6B, 6C and KRT16 (14, 19, 31, and 4.5 fold higher relative all other subtypes, respectively). A SAM analysis between MS2b2.2 and the remaining groups indicated this subtype to be highly keratinized with additional high expression of KRT14 (19 fold), KRT5 (13 fold) KRT4 (3.5 fold). Of these KRT14 has been associated with squamous differentiation of UC (Chu et al. 2001, Harnden et al. 1997, Tungekar et al. 1988, Chu and Weiss, 2002). To explore this further we used the SSC gene signature of Blaveri et al. (2005) to investigate the expression of bladder SCC specific genes in our data; the MS2b2.2 subtype was found to show a very strong SCC signature signal (Figure 6D). Upon pathological reevaluation of this group, 18 of the 29 cases were shown to have signs of squamous differentiation. In addition, this group showed a different proportion of female/male patients, 15/14 compared to 65/214 (p<1x10⁻³, Chi-2), reminiscent of the 1:1 proportions seen in patients diagnosed with bladder SCC in the Swedish population. MS2b2.1 showed a similar KRT profile as MS2b2.2 with high expression of the basal keratins KRT5, KRT15, and KRT17 as well as of the keratins associated with hyperproliferation, KRT6 and KRT16, albeit with a less extreme expression profile; the KRT6 group of keratins showed 2-5 fold higher expression in MS2b2.2 compared to MS2b2.1. In addition this subtype showed a higher expression of keratins of simple epithelial cells KRT13, KRT7, KRT19, KRT8,
and KRT18, and of KRT20. This latter finding indicate that MS2b2.1 and MS2b2.2 may be of different cellular origin or are governed by different genetic programs. Either way, the KRT signatures indicates MS2b2.1 as a progressed version of the MS1 subtypes.

Figure 6. Expression of cytokeratin genes. (A) Gene expression given for individual samples. (B) Gene expression given as group means. (C) Gene expression of KRT20 compared to the mean expression of UPK1A, UPK1B, UPK2, UPK3A, and UPK3B. (D) Mean expression of the bladder SCC signatures published by Blaveri et al. (2005).

References


FGFR3, PIK3CA, and TP53 mutations, and the FGFR3 gene signature

**FGFR3 gene signature**

QTC cluster 16 contained FGFR3 as a key gene together with an additional 19 genes (Figure 7A and 7B), and showed high expression in both MS1 subtypes as well as in MS2b2.1. The signature contained three genes that may be of particular importance for UC development TP63, IRS1, and WNT7B.

TP63 is a member of the TP53 family of transcriptional regulators that also include TP73. TP63 has a basal/intermediate expression in the normal urothelium (Castillo-Martin et al. 2010, Human Protein Atlas www.proteinatlas.org) and is crucial for normal differentiation of the urothelium. The FGFR3 gene has TP63 responsive promoter elements and is activated by TP63 (Sayan et al. 2010). Hence, a functional link between FGFR3 and TP63 expression may exist. It has been shown that TP63 is significantly downregulated in MI tumors but that a subset of ≥T2 tumors has retained TP63 expression (Urist et al. 2002). This subset of ≥T2 tumors is now identified as being a part of the MS2b2.1 subtype.

IRS1 is upregulated by FGF2 (bFGF) stimulation of MCF-7 breast cancer cells (Lassarre and Ricort, 2003) and may thus be a downstream target of FGFR signaling. In addition, IRS1 may interact with FGFR3 through GRB2. IRS1 specifically bind to various cellular proteins containing SH2 domains and activates phosphatidylinositol 3-kinase (PI3K) when bound to the regulatory p85 subunit (PIK3R). WNT7B is a member of the WNT family of secreted proteins which interacts with the FRZ family of receptors and induces β-catenin signaling. WNT7B has been shown to be four-fold upregulated in non-muscle invasive UC compared to muscle invasive UC or normal urothelial tissue (Bui et al, 1998). These authors suggest that this is an early event in bladder tumor development which is in accordance with this gene being a part of the FGFR3 gene signature.

The FGFR3 signature also contained the gene C16orf74. Decreased expression of C16orf74, a gene of unknown function, has been suggested as a marker for tumor progression (Kim et al, 2010). This is in line with its presence in the FGFR3 signature as FGFR3 mutated/expressing tumors typically has good prognosis. Non-muscle invasive cases with low C16orf74 expression are typically of the MS2a subtype. The FGFR3 gene signature showed a coordinated expression pattern also in the Kim et al. data (Figure 8) even though a MS2b2.1 subtype was not identified in this data set; the coordinated up regulation of the FGFR3 signature in a subset of the Kim MS2b1/MS2b2 cases suggest that MS2b2.1 is split between the MS2b1 and the MS2b2 subtypes in the Kim et al. data.

**FGFR3, PIK3CA, and TP53 mutation data**

FGFR3 mutation analysis clearly revealed two distinct subtypes with respect to mutation frequency, MS1 and MS2a; MS1 showed mutations in 55 %, and MS2a in 7 % of the cases (p<0.0001, Chi-2). MS1a and MS1b, and MS2a1 and MS2a2, respectively, did not differ significantly in this respect. MS1 also showed a higher PIK3CA mutation frequency compared to MS2a, 25% and 8%, respectively, (p=0.002, Chi-2), with no difference between MS1a and MS1b (Figure 7C). A significant association between FGFR3 and PIK3CA mutation was observed in the entire data set as well as within the MS1 subtype (p=0.016, hypergeometric test). This association has been observed previously (Lopez-Knowles et al. 2006, Kompier et al. 2010, Sjödahl et al. 2011) The frequency of TP53 mutations was significantly higher in MS2a (48%) compared to MS1 (11%) (p<0.0001, Chi-2). PTEN, a negative regulator of PIK3CA, showed a significantly lower expression in MS2a tumors (p<0.001, permuted t-test) even when cases where
expression indicated possible homozygous deletions were omitted from the analysis (PTEN_HD_corr in Figure 7D and 7E). This identifies MS1 as FGFR3 and PIK3CA mutated, and MS2a as TP53 mutated and PTEN low. This is in line with the finding that TP53 and PTEN deficiency promotes invasive bladder cancer (Puzio-Kuter et al. 2009) as 85% of the MS2a cases were invasive (≥T1) but only 43% of the MS1 cases (p<0.0001, Chi-2). The MS2b2.1 subtype demonstrated FGFR3 and PIK3CA mutations frequencies similar to MS1 (p>0.25, Chi-2, in both comparisons), and of TP53 similar to MS2 (p>0.6, Chi-2). Furthermore, ten out of twenty MS2b2.1 cases were muscle invasive (≥T2) indicating MS2b2.1 as high risk versions of MS1 (≥T2 in MS1; 8/123 vs. ≥T2 in MS2b2.1; 10/10, p<0.001, Chi-2) that have obtained TP53 mutations. Hence, both the gene expression and the mutation data indicates MS2b2.1 as an evolved version of MS1 tumors.

Figure 7. Expression of the FGFR3 associated gene cluster (QTC13), distribution of FGFR3, PIK3CA, and TP53 mutations, and expression of PTEN. (A) Gene expression of the FGFR3 associated gene cluster (QTC13) for individual samples (B) Gene expression of the FGFR3 associated gene cluster (QTC13) given as group means. (C) Mutation distribution in the molecular subtypes, black indicates mutation, white indicates wild type, thin grey lines indicates the four samples for which no data was available. (D) Gene expression for PTEN and PTEN HD-corrected shown as subtype means together with gene mutation frequencies of FGFR3, PIK3CA, and TP53 rescaled to be represented as heat maps (red, high frequency; green, low frequency). (E) Expression of PTEN shown for individual samples.
Figure 8. Expression of the FGFR3 associated gene cluster (QTC13) in the Kim et al. data set. Gene expression for individual samples. (B) Gene expression given as group means.
Expression of the FGFR3 gene signature in FGFR3 mutated cases

To evaluate the relationship between FGFR3 mutation and FGFR3 gene signature further, we selected all mutated cases in the data set and calculated the mean expression level of the FGFR3 gene signature for each case. The cases were then rank ordered with respect to mean signature expression level and the MS types indicated (Figure 9). This analysis shows a significant drop in signature expression in FGFR3 mutated MS2a and MS2b2.2 cases.

Figure 9. Expression of the FGFR3 associated gene cluster in FGFR3 mutated cases. Mean expression of the FGFR3 associated gene cluster (QTC13) sorted by rank. Each dot represents a FGFR3 mutated case, and the colors represent tumor clusters. MS1 and MS2b2.1 cases show high mean FGFR3 signature expression, whereas MS2a1, MS2a2, and MS2b2.2 show a drop in mean FGFR3 signature expression. For clarity, Urobasal A (MS1a, MS1b), green; Genomically Unstable (MS2a1, MS2a2), blue; Urobasal B (MS2b2.1), orange; SCC-like (MS2b2.2), red.

References


**Receptor Tyrosine Kinases**

Receptor tyrosine kinase (RTK) genes were identified by the GO term GO:0004714, transmembrane receptor protein tyrosine kinase activity. Of the identified genes, twenty-seven showed MS specific expression (ANOVA) (Figure 10A and 10B). We noticed four main RTK expression patterns; FGFR3, EPHB6, MST1R, EPHA1, EPHA2, and RYK expressed in MS1 and MS2b2.1; ERBB2 and ERBB3 predominantly expressed in MS2a; CSF1R, PDGFRB, AXL, DDR2, ROR2, NRP1, TEK, PDGFRα, EPHA3, and TIE1 expressed in MS2b1; and EPHB3, EPHB4, MET, EGFR, and MUSK expressed in MS2b2.2.

The receptors typical for MS1 tumors, FGFR3, EPHB6, MST1R, EPHA1, and EPHA2 are all expressed uniformly throughout normal urothelium (basal, intermediate, and umbrella cells) as determined by Human Protein Atlas (www.proteinatlas.org). Three of the RTKs associated with MS1 belong to the Ephrin receptor family. These receptors mediate their function by binding to cell surface bound ephrin ligands on neighboring cells. The effects of EPH signaling can be bidirectional and may act as both tumor promoters and suppressors depending on context (Genander and Frisén 2010).

**ERBB2** and **ERBB3** showed high expression in MS2a tumors, less uniformly in MS1 tumors, and low expression in MS2b2 subtypes. **ERBB2** is a known oncogene and is amplified in several tumor types as well as in UC (Lae et al, 2010, Tapia et al, 2007, Coogan et al, 2004). We identified a number of RTKs highly expressed in the immune-cell infiltrated subtype, MS2b1. Indeed many of these receptors, including **CSF1R, AXL, ROR2, NRP1, TEK, PDGFRα, PDGFRB**, and **TIE1**, are known to be expressed in immune-cells or endothelial cells (Tallquist et al. 1999, Neubauer et al. 1994, Yano et al. 1997, Favre et al. 2003, Radzun et al. 1988).
Five receptor genes were highly specific for MS2b2.2, EPHB3, EPHB4, MET, EGFR, and MUSK. EGFR is a known oncogene and has been linked to tumors with SCC histology in both breast and bladder cancer (Guo et al. 2009, Bossyut et al. 2005). Expression of both EGFR and MET protein can be detected in normal urothelium (Human protein Atlas www.proteinatlas.org). MUSK is involved in signaling between nerve and muscle cells and has not been studied in relation to bladder cancer.

Four RTK genes, FGFR3, ERBB2, EGFR, and MET showed a subtype specific expression pattern in which FGFR3 expression was seen in the MS1 and the MS2b2.1 subtypes, ERBB2 in MS2a subtypes, and EGFR and MET in the MS2b2.2 subtype (Figure 11A). These expression patterns cut across tumor stage and are hence specific for the respective MS subtypes (Figure 11B). This is particularly evident for FGFR3 and ERBB2 that are expressed in MS1 and MS2a tumors, respectively, irrespective of their stage (Ta, T1, or ≥T2).

**Figure 10.** Expression of RTK genes. (A) Gene expression for individual samples. (B) Gene expression given as group means.
Figure 11. Mean expression of the subtype specific RTK genes. (A) Mean expression of the subtype specific RTK genes using all samples. (B) Mean expression of the subtype specific RTK genes stratified by stage group. MS1a and MS1b, and MS2a1 and MS2a2 were merged to avoid that the stratified groups became too small to calculate a relevant mean expression level. Grey squares indicate that the stage/subtype group was excluded because it contained fewer than 5 samples.

References


Subtype specific expression of potential drug targets

We downloaded drug targets from the Drugbank database (Knox et al. 2011, http://www.drugbank.ca) with a described or potential use in cancer treatment, and limited the result to one compound per target gene. Of 60 genes that were targetable, 39 that were expressed in a subtype specific pattern (ANOVA, Bonferroni corrected p<0.05) are shown in Figure 12. Among the identified drugs many were in clinical use for cancer treatment including trastuzumab (anti-ERBB2, recombinant antibody), erlotinib (EGFR inhibitor), and sirolimus (rapamycin, mTOR inhibitor). We also searched The Cochrane Central Register of Controlled Trials (The Cochrane Library, http://www.thecochranelibrary.com) for compounds tested as potential cancer drugs. We limited the search to the time period 2000 to 2011 and to the search term “cancer” anywhere in the text and either of the terms “targeted” or “inhibitor” in the record title. We then refined the resulting list to include only drugs with a described use or potential use in cancer, and limited the result to one compound per target gene. This resulted in 46 compound-target pairs, of which 37 subtype specific (ANOVA, Bonferroni corrected p<0.05) are shown in Figure 13. Of these, seven targets were also present in the Drugbank gene list.

![Figure 12. Expression of drug target genes from the Drugbank database. (A) Gene expression for individual samples. (B)](image-url)
Figure 13. Expression of drug target genes from the Cochrane Central Register of Controlled Trials. (A) Gene expression for individual samples. (B) Gene expression given as group means.

References
