CD44 Expression Predicts Local Recurrence after Radiotherapy in Larynx Cancer

Monique C. de Jong[1,2], Jimmy Pramana[1,3], Jacqueline E. van der Wal[10], Martin Lacko[5], Carine J. Peutz-Kootstra[6], Jos M. de Jong[7], Robert P. Takes[8], Johannes H. Kaanders[9], Bernard F. van der Laan[11], Jasper Wachters[10,11], Jeroen C. Jansen[12], Coen R. Rasch[2], Marie-Louise F. van Velthuysen[4], Reidar Grénman[13], Frank J. Hoebers[2], Ed Schuuring[10], Michiel W. van den Brekel[3], and Adrian C. Begg[1]

Abstract

Purpose: To find molecular markers from expression profiling data to predict recurrence of laryngeal cancer after radiotherapy.

Experimental Design: We generated gene expression data on pre-treatment biopsies from 52 larynx cancer patients. Patients developing a local recurrence were matched for T-stage, subsite, treatment, gender and age with non-recurrence patients. Candidate genes were then tested by immunohistochemistry on tumor material from a second series of 76 patients. Both series comprised early stage cancer treated with radiotherapy alone. Finally, gene expression data of eight larynx cancer cell lines with known radiosensitivity were analyzed.

Results: Nineteen patients with a local recurrence were matched with 33 controls. Gene sets for hypoxia, proliferation and intrinsic radiosensitivity did not correlate with recurrence, whereas expression of the putative stem cell marker *CD44* did. In a supervised analysis, probes for all three splice variants of *CD44* on the array appeared in the top 10 most significantly correlated with local recurrence. Immunohistochemical analysis of CD44 expression on the independent validation series confirmed CD44's predictive potential. In 8 larynx cancer cell lines, *CD44* gene expression did not correlate with intrinsic radiosensitivity although it did correlate significantly with plating efficiency, consistent with a relationship with stem cell content.

Conclusions: *CD44* was the only biological factor tested which significantly correlated with response to radiotherapy in early stage larynx cancer patients, both at the mRNA and protein levels. Further studies are needed to confirm this and to assess how general these findings are for other head and neck tumor stages and sites. *Clin Cancer Res; 16(21); 5329–38. ©2010 AACR.*

The incidence of larynx cancer in the United States is around 4.5 cases per 100,000 per year (1). The 5-year relative survival percentage for localized disease has been stable at around 70-80% for the last 20 years (1). In early laryngeal cancer, radiotherapy is an effective treatment modality, with local control rates between 80-90% for T1 tumors (2). Partial laryngectomy or CO2 laser resection are alternative treatments with comparable survival rates, although when used as salvage after a failed radiotherapy course they have a higher complication rate (3). Treatment choice is mainly based on the estimated functional out-

come and the preferences of the clinician. It would therefore be useful to predict beforehand which patients will benefit from radiotherapy. Prediction of resistance is also likely to be increasingly useful in the development of biological modifiers which increase the effects of radiation, providing an alternative treatment for resistant tumors.

Important clinical factors associated with local recurrence after radiotherapy are tumor stage, tumor size, radiotherapy fraction size and year of treatment (4). Treatment choice is now mainly based on T-stage (5), although this is still a relatively poor indicator of survival (6). Since clinical factors

Authors' Affiliations: Departments of [1]Experimental Therapy, [2]Radiation Oncology, [3]Head and Neck Surgery and [4]Pathology, The Netherlands Cancer Institute, The Netherlands; Departments of [5]Otorhinolaryngology and Head and Neck Surgery, [6]Pathology and [7]Radiation Oncology, University Medical Center Maastricht, The Netherlands; Departments of [8]Otorhinolaryngology and Head and Neck Surgery and [9]Radiation Oncology, Radboud University Medical Centre Nijmegen, The Netherlands; Departments of [10]Pathology and [11]Otorhinolaryngology and Head and Neck Surgery, University Medical Center Groningen, University of Groningen, The Netherlands; [12]Department of Otorhinolaryngology and Head and Neck Surgery, Leiden University Medical Center, The Netherlands; and

[13]Department of Otorhinolaryngology - Head and Neck Surgery and Medical Biochemistry, University of Turku, Finland

Translational Relevance

Treatment choice for larynx cancer is based on clinical factors such as T-stage, but these are imprecise indicators of response. Having robust methods to predict outcome of a particular therapy would be extremely valuable, allowing a more rational treatment choice which should lead to greater tumor cell kill and also spare patients from toxic and ineffective therapies. Such predictors should include biological factors as well as clinical factors, given the heterogeneity in tumor biology even for patients presenting with similar sites and stages. The present study employed gene expression profiling in a series of larynx cancers and validated the result in a second similar series using immunohistochemistry. The principle predictor for outcome after radiotherapy was CD44, a putative stem cell marker. In addition, this study sheds light on potential mechanisms of radioresistance, which could lead to the design of targeted drugs for combining with radiation.

cannot provide an accurate prediction, it is likely that recurrence of a tumor can partly be explained by tumor biology. Three biological processes known to influence response to radiotherapy are intrinsic radiosensitivity (7), hypoxia (8) and repopulation (9). For each of these processes, individual markers (mainly immunohistochemical) have been investigated and found to be of predictive value (10–12), although none have been sufficiently validated or are in routine use. Since many genes are involved in each process, in addition to single markers representing these processes, sets of markers (gene sets) for hypoxia (13, 14), intrinsic radiosensitivity (15–17) and repopulation (18) have also been defined. Another factor more recently hypothesized to play a role in response to therapy is the number of stem cells, ultimately determining repopulation of the tumor (19, 20) and so eradication of this subpopulation is of prime importance.

To date, no studies have investigated all these processes simultaneously. Microarrays have been used to measure gene expression (mRNA) on a genome wide scale, and can in principle monitor all the above-mentioned processes concurrently. However, only one microarray study with 14 patients has been carried out for patients treated with radiotherapy alone (21). Several expression profiling studies have been carried out on patients treated with radiotherapy in combination with surgery or chemotherapy (22–26). However, these have often included heterogeneous groups of patients and cannot address the question of factors affecting the response of laryngeal cancer to radiation alone.

Our objective was to find a gene expression profile that will accurately predict local recurrence after radiotherapy in a homogeneous group of patients with early laryngeal carcinoma. We chose to study early stage tumors, since these are likely to be more homogeneous than advanced tumors and

also technically easier to treat, minimizing the chance of geographical misses. Treatment failure is then highly likely to be due to biological rather than technical factors. In addition to giving more insight into the molecular processes underlying treatment failure, accurate prediction would enable treatment to be individualized, leading to increased survival and less unnecessary morbidity. We studied two series of early stage larynx cancer patients treated with radiotherapy alone. The first was a test series of frozen tumor specimens used to study global gene expression to discover predictive markers for local control, which were then validated on a second series by immunohistochemistry.

Materials and Methods

All studies reported here were done with approval of the local Medical Ethics Committees.

Gene expression series

Patients. Fifty two patients were recruited from five different institutes in The Netherlands and were eligible if they had been treated for a T1 or T2 larynx carcinoma (see Supplemental Table 1), and pre-treatment fresh frozen tumor material was available. Patients were treated between 1997 and 2005, and staging was done either clinically or with a CT-scan. Because patients with small tumors did not have a CT-scan, tumor volumes could not be measured for the whole group. Treatment was radiotherapy alone with curative intent, applying fractionation schemes standard in each of the five centers. To compare different radiotherapy schedules, the equivalent dose in 2-Gy fractions (EQD_2) was calculated for every patient with the formula: $EQD_2 = D \times (d + \alpha/\beta)/(2 + \alpha/\beta)$, where D is the total dose, d the given fraction dose, the α/β ratio was assumed to be 10 Gy. Recurrence was defined as a histologically proven local tumor recurrence within two years of the initial treatment, to ensure the analysis of true recurrences rather than second primaries. Since we planned to study a matched series, for every patient with a recurrence we aimed to include two controls, with a recurrence-free follow-up of at least two years and matched for the institute they were treated in, T-stage, subsite, gender and age. There were no significant differences between groups with and without local recurrence in age, gender, subsite, T-stage, total dose, fraction size, tumor percentage or RNA quality (Supplemental Table 1).

RNA isolation. All biopsies were snap frozen in liquid nitrogen. Around 30 slices of 30 μm were deposited in RNA-Bee (Campro scientific). Before and after these 30 slices H&E sections were taken that were subsequently assessed by an experienced pathologist, who scored differentiation and tumor percentage. Only biopsies containing on average more than 50% of tumor cells were included. The tumor material in RNA-Bee was processed using the Qiagen RNeasy mini and RNase-free DNase kits. Total RNA was isolated and DNAse treated using spin columns according to the manufacturers instructions. The Agilent 2100 bioanalyzer was used to assess the integrity (intactness) of the

RNA. Samples with an RNA Integrity Number (RIN) under 6.0 or with no obvious 18S and 28S peaks were discarded.

Gene expression. cDNA was made from one microgram of total RNA and amplified into aRNA with T7-mRNA Superscript-III amplification kit (Invitrogen). Only amplification yields over 1000-fold with a 1 kB smear on a gel were accepted. Hybridization to microarray slides was performed at our Central Microarray Facility (http://microarrays. nki.nl). All samples were hybridized to Illumina bead arrays (v3 Illumina beads) and subsequently scanned using the Illumina scanner. Each Illumina array consists of 3-micron silica beads covered with oligos containing over 48,000 transcript probes per sample, representing around 25,000 known genes. Each transcript probe was represented more than 20-fold per array and final data were averaged for each probe. Fluorescence intensities were measured with the Illumina scanner and averaged per probe.

Data analysis. The dataset was transformed (variance stabilizing method (ref. 27)) and normalized (robust spline method) with the Lumi (28) package for R, version 2.8 (29) (http://www.R-project.org). If, for a specific probe, no patient had a value above background levels, that probe was filtered out. Gene sets for hypoxia, proliferation, radiosensitivity and stem cells were tested (13–16, 18, 30, 31). Unigene identifiers were used to map the genes in a set to the annotations of the Illumina array. For gene sets with known weights contributing to the endpoint (as described in the original publications), Pearson correlations were calculated against the weights of a gene set for each patient. This also allowed assessment of gene sets which included genes both positively and negatively correlating with outcome. For gene sets without weights (each gene assumed to contribute equally), the average expression of the genes in the set was calculated. For these signatures, all genes in the set were correlated in the same direction with outcome.

The Pearson or average values were then used in a logistic regression with local recurrence data. In order to give comparable odds ratios, some Pearson correlations were multiplied by 5 or 10, which does not change the *P*-values but simply provides a better comparison of odds.

In addition to this hypothesis-driven analysis, a data-driven analysis was performed with Biometric Research Branch (BRB) array tools (NIH, http://linus.nci.nih.gov/ brb-arraytools.htm). Genes were first filtered by including probes where at least 20% of samples had a minimum fold change greater than 1.35 and a *P*-value for log-ratio variation under 0.01. The filtered set was entered in a nearest centroid model that finds genes that best predict local recurrence. Genes significantly different between the patients with and without recurrence at the *P* < 0.01 significance level were used for class prediction. The leave-one-out cross-validation method was used to compute mis-classification rates.

Immunohistochemistry series

Patients. Of the patients included in the mRNA expression microarray series, paraffin embedded material for immunohistochemistry (IHC) was used from two of the five institutes (Amsterdam and Groningen). This small subset of 20 cases was used to confirm gene expression values by IHC. A second matched series of 76 patients was used as an independent validation series of our findings from the gene expression study. Paraffin embedded biopsies were used to make cores for a tissue microarray (TMA). The construction of the TMA and the patient characteristics were described previously (10). Briefly, the patients were predominantly male with stage T1 and T2 glottic tumors given a median of 66Gy in 2Gy fractions (Supplemental Table 2).

Immunohistochemistry staining. Sections of 3 μm were cut from either whole tissue blocks or the TMA and mounted on amino-propyl-ethoxy-silan (APES, Sigma-Aldrich,

---

Table 1. Logistic regression of gene sets with local recurrence

Gene set	Range	*P*	OR (95% CI)
Stem cell (CD44)(31)	7.8-9.1	0	20.2 (3.4-172.3)
Stem cell (Glinksky)(45)	5.7-6.9	0.03	6.5 (1.3-42.2)
Acute hypoxia (Chi) ×10(14)	−1.6	0.04	7 (1.1-46.6)
Hypoxia metagene (Winter)(13)	7.2-7.9	0.13	19.3 (0.5-1225.2)
Stem cell genes (various) excluding CD44	6.4-7.4	0.16	7.5 (0.4-129.7)
Radiosensitivity (17), response	−0.5	0.64	0.3 (0.002-46.0)
Proliferation (Shepard*) ×10	−1	0.67	1.6 (0.2-12.2)
Chronic hypoxia (14)	−1.6	0.67	0.7 (0.1-3.8)
Radiosensitivity (15)	6.1-7.2	0.68	0.6 (0.03-9.4)
Proliferation (18)	6.4-7.1	0.95	0.9 (0.03-26.1)

NOTE: Range: lowest to highest value of either Pearson correlations against the weights of a gene set or, for gene sets without weights, the average expression (log2 scale) of the genes in the set. OR: odds ratios with corresponding confidence intervals and *P*-values were generated from a logistic regression of Pearson or average values of the gene sets with local recurrence data. In order to give comparable odds ratios, some Pearson correlations were multiplied by 5 or 10 (not changing the *P*-values).
*From Gene Set Enrichment Analysis molecular signatures database; http://www.broadinstitute.org/gsea/msigdb/.

Diesenhofen Germany)-coated glass slides. Slides were de-paraffinized in xylene and rehydrated in ethanol. Antigen retrieval comprised boiling the slides in a microwave oven in citrate (pH 6.0) for 15 minutes. Endogenous peroxidase was blocked with 0.3% hydrogen peroxidase for 30 minutes. Slides were incubated with a mouse monoclonal antibodies against CD44 (156-3C11; dilution 1:200; Cell Signaling Technology, Danvers, MA) and CD44v6 (clone VFF-18; dilution 1:8000; Bender Medsystems, Vienna, Austria) for 1 h at room temperature. Detection was performed with RAM[HRP] (dilution 1:100) and GAR[HRP] (dilution 1:100), visualized by 3'3-diaminobenzidinetetra-hydrochloride and counterstained with haematoxylin.

Immunohistochemistry scoring. The percentage of tumor cells staining positive for CD44 was scored as well as the intensity of staining (low or high). A CD44 staining score was calculated by adding the percentages of positive low and high intensity cells, weighted by factor of 1 and 2 respectively. This weighted score reflects total CD44 protein better than total percentage positive cells, for better comparison with total mRNA from the microarray analysis. For the set of patients in which concordance between mRNA and IHC was tested, all slides were analyzed independently by two teams, each consisting of a pathologist (MvV and JvdW) and a scientist. Slides scored differently by the two teams were discussed at a conference microscope to reach consensus. Before consensus, the inter-observer correlation for CD44 scores was 0.75 ($P < 0.001$; Supplemental Fig. 1). For the TMA series of 76 patients, scoring was done by one team. Pearson correlations were calculated between mRNA levels and IHC scores. For the TMA analysis, associations between CD44 expression and local recurrence were compared using a logistic regression model. P-values of <0.05 were considered statistically significant. Statistical analysis was performed with SPSS 16.0 for Windows (SPSS Inc., Chicago, IL).

Larynx cancer cell lines

Cell culture. The larynx cancer cell lines UT-SCC-6A,-8,-9,-19A,-19B,-22,-23 and -42A from the University of Turku (Finland) were cultured in DMEM with 10% FBS, 1% NEAA, 1%L-glutamine and 1% penicillin-streptomycin. Information was available on plating efficiency and radiosensitivity for all cell lines (published and unpublished data) (refs. 32–34).

Gene expression. For each cell line, 1×10^6 cells were washed with ice cold PBS at approximately 50% confluence and then collected in RNA-Bee. Illumina microarray data were generated using the same methods and materials as described above for the tumor biopsies.

Results

Gene expression

Gene expression analysis. Exclusion of probes that did not exceed background expression in any patient left 26,454 probes for analysis. Gene expression signatures for hypoxia, intrinsic radiosensitivity, repopulation and stem cells were analyzed in a logistic regression (Table 1). The putative stem cell marker *CD44* was the most significant, with an unrelated stem cell signature as second most significant. A third stem cell signature not including *CD44* (see Supplemental Table 3) was fifth of the 12 signatures tested but was not significant. After multiple testing correction (Bonferroni), only *CD44* expression remained significant ($P = 0.024$). Comparative histograms of *CD44* expression illustrate the higher expression in recurrences versus cures (Fig. 1A). When patients were divided into three groups of low, medium and high *CD44* expression, split so that there were equal numbers of recurrences in each group, the odds of recurrence (number of recurrences divided by number of non-recurrences for each group) was 15.2 fold higher in the highest *CD44* expression group compared with the lowest ($P = 0.003$, Fig. 1B). Expression of acute hypoxia genes was also associated with local recurrence, although significance was lost after correction for multiple testing. Radiosensitivity and proliferation genes showed no relationship with recurrence.
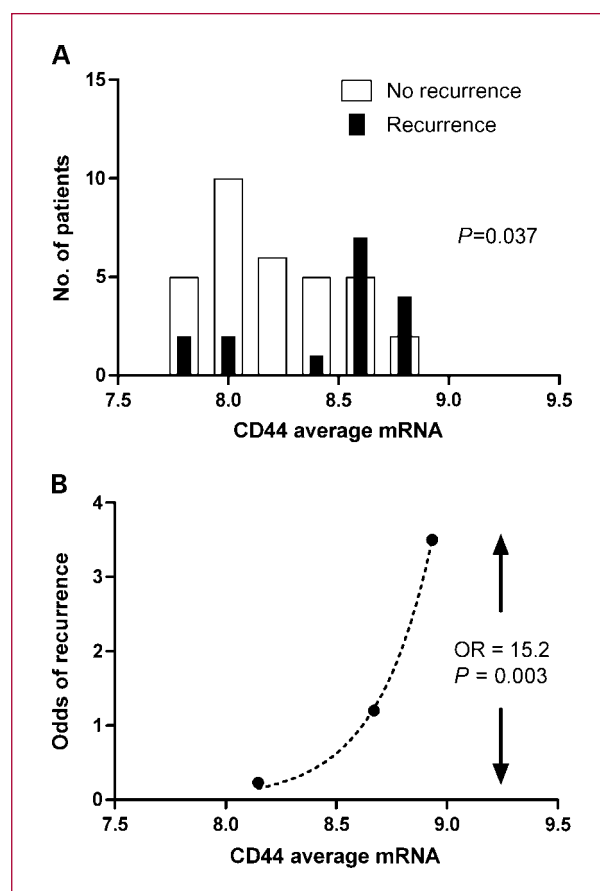

Fig. 1. CD44 expression predicts local recurrence. A, histograms of CD44 mRNA expression for patients subsequently cured (open bars) or those subsequently suffering a recurrence (closed bars). B, odds of recurrence when patients are divided into three groups with increasing mRNA levels, split so that each group contains equal numbers of recurrences. OR: odds ratio of recurrence between highest and lowest CD44 expression groups.

Table 2. Data driven classifier, showing top 34 most significant genes, all with a *P*-value <0.01

Gene symbol	*t*-value	Parametric *P*-value	Probe ID	Description
UP-REGULATED in recurrence				
CD44	−3.49	0.001	ILMN_1803429	CD44 molecule (Indian blood group), transcript variant 4
HSD17B12	−3.17	0.003	ILMN_1702168	hydroxysteroid (17-beta) dehydrogenase 12
BTBD11	−3.09	0.003	ILMN_1705066	BTB (POZ) domain containing 11, transcript variant 1
CHL1	−3.08	0.003	ILMN_1713347	cell adhesion molecule with homology to L1CAM
MGLL	−3	0.004	ILMN_1738589	monoglyceride lipase, transcript variant 1
BNIP3	−2.93	0.005	ILMN_1724658	BCL2/adenovirus E1B 19kDa interacting protein 3
SNX5	−2.92	0.005	ILMN_1673676	sorting nexin 5, transcript variant 1
CD44	−2.85	0.006	ILMN_1778625	CD44 antigen (Indian blood group), transcript variant 1
CAPRIN1	−2.82	0.007	ILMN_1754145	cell cycle associated protein 1, transcript variant 1
CD44	−2.8	0.007	ILMN_2348788	CD44 molecule (Indian blood group), transcript variant 5
CHMP2A	−2.78	0.008	ILMN_1656621	chromatin modifying protein 2A, transcript variant 1
SLC37A4	−2.75	0.008	ILMN_1678678	solute carrier family 37 (glucose-6-phosphate transporter), member 4
CNDP2	−2.72	0.009	ILMN_1726769	CNDP dipeptidase 2 (metallopeptidase M20 family)
ATP1B1	−2.71	0.009	ILMN_1730291	ATPase, Na+/K+ transporting, beta 1 polypeptide, transcript variant 1
HK1	−2.7	0.009	ILMN_1761829	hexokinase 1, transcript variant 1
ACADVL	−2.7	0.009	ILMN_2263466	acyl-Coenzyme A dehydrogenase, very long chain, transcript variant 1
IL1B	−2.69	0.009	ILMN_1775501	interleukin 1, beta
TLE1	−2.69	0.01	ILMN_1751572	transducin-like enhancer of split 1 (E(sp1) homolog, Drosophila)
DOWN-REGULATED in recurrence				
PRSS21	2.7	0.009	ILMN_1774256	protease, serine, 21 (testisin), transcript variant 2
AGPAT4	2.71	0.009	ILMN_1730504	1-acylglycerol-3-phosphate O-acyltransferase 4
RNF7	2.71	0.009	ILMN_1711862	ring finger protein 7, transcript variant 3
LSM1	2.75	0.008	ILMN_2218450	LSM1 homolog, U6 small nuclear RNA associated (S. cerevisiae)
GPATCH2	2.76	0.008	ILMN_1786036	G patch domain containing 2
GOLGA7	2.79	0.007	ILMN_1778673	golgi autoantigen, golgin subfamily a, 7, transcript variant 2
C3orf21	2.83	0.007	ILMN_1671116	chromosome 3 open reading frame 21
TLOC1	2.84	0.006	ILMN_1762003	translocation protein 1
HIST2H2AC	2.86	0.006	ILMN_1768973	histone cluster 2, H2ac
BRF2	2.89	0.006	ILMN_1665554	subunit of RNA polymerase III transcription initiation factor, BRF1-like
SLMO1	2.9	0.005	ILMN_2232157	slowmo homolog 1 (Drosophila)
LYPLAL1	3.05	0.004	ILMN_2142117	lysophospholipase-like 1
MRPL55	3.07	0.003	ILMN_2348090	mitochondrial ribosomal protein L55, transcript variant 5
TERC	3.1	0.003	ILMN_1766573	telomerase RNA component on chromosome 3
MRPS1	3.21	0.002	ILMN_1663664	mitochondrial ribosomal protein S10
KIAA97	3.51	0.001	ILMN_1670752	KIAA0907

After restriction of the dataset to those 8,317 probes that showed significant differences in expression between the tumors, thus removing uninformative probes, we performed a data-driven analysis for which genes best predicted recurrence. When the univariate significance alpha level was set to $P < 0.01$, 34 probes (18 up and 16 down-regulated in tumors subsequently recurring) were found to be predictive (Table 2). The most significant upregulated marker discriminating between cures and recurrences was *CD44* ($P < 0.002$). With the nearest centroid method only 23% of the patients were correctly classified with these 34 genes. In addition, false discovery rates, as calculated by the Benjamini-Hochberg method, were high. However, despite the predictive weakness of the signature as a whole, of note was that all three probes for *CD44* that were present on the array appeared in the top 10 highest ranking upregulated genes. Two of the probes (variants 4 and 5) map to the constant and largest exon (exon 18), while variant 1 maps to the first variable exon (exon 6). Expression of each probe was highly significantly correlated with expression of each of the other probes across the 52 tumors (all *P*-values <0.001; 1 vs 3, 1 vs 5, 3 vs 5).

In addition to *CD44*, the remaining top ranking genes from Table 2 were most highly represented in a pathway relevant to "cell cycle, cellular development, cellular
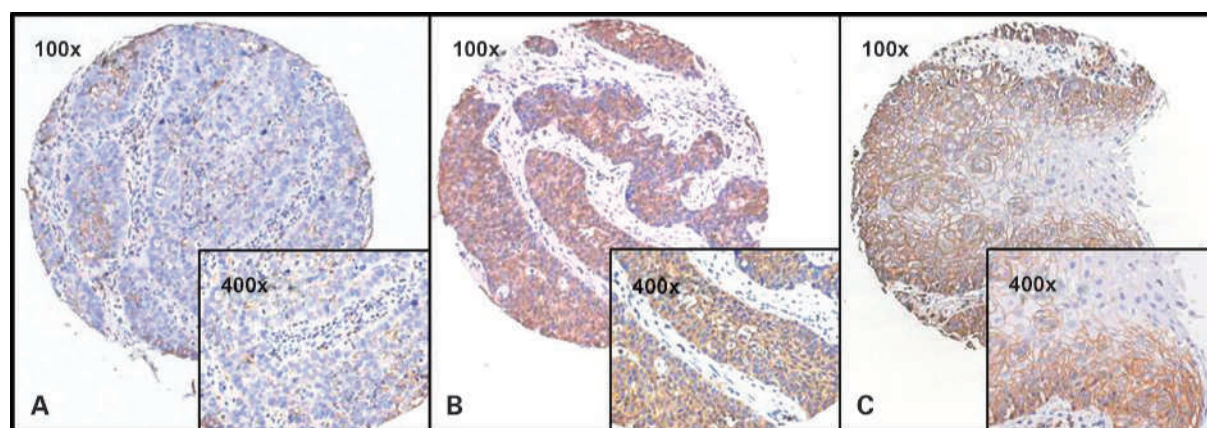
Fig. 2. Examples of CD44 immunohistochemistry. Staining, using antibody 156-3C11, against an epitope common to all CD44 variants, on the tissue microarray for three representative cores at two different magnifications (100× and 400×). Scorings for these cores were: A: 40% intensity I. B: 95% intensity II. C: 80% intensity II.

growth and proliferation" (from Ingenuity Pathway Analysis). This pathway contained *EGF*, *VEGF*, and *HRAS* as hub genes and of 35 genes on the pathway, 11 appeared in list of top ranking genes (Supplemental Fig. 2).

CD44 protein level versus outcome

CD44 mRNA correlates with immunohistochemical expression. Both frozen and paraffin embedded material was readily available from 20 tumors and used to compare RNA and protein expression. Antibodies were tested against an epitope common to all CD44 variants and one specific for the v6 variant. Figure 2 shows examples of CD44 staining. All tumors showed some expression (with on average 24% of tumor cells staining with a low intensity and 52% with a high intensity) although the staining was heterogeneous in all cases. In tumors showing a clear differentiation pattern, the basal cell layers were more intensely stained than the more differentiated cells. Both the CD44 and the CD44v6 immunostaining scores correlated significantly ($P < 0.05$) with the average for all three CD44 mRNA probe levels (Supplemental Fig. 3).

CD44 expression in the validation series. We next tested whether immunohistochemical expression of CD44 correlated with clinical outcome. We used an independent matched series of laryngeal cancers with patient characteristics similar to the test series. Patient characteristics of this validation series, like the 52 patients in the test series, were predominantly male with a T1-2 glottic tumor and treated with radiotherapy alone (Supplemental Table 4). CD44 expression, assessed immunohistochemically for percentage CD44-positive cells weighted according to staining intensity (see Materials and Methods), was significantly associated with clinical outcome. Histograms of the IHC scores showed higher CD44 protein expression in recurrences compared with cures (Fig. 3A). As before, when patients were divided into three groups with low, medium

and high CD44 expression, split to ensure equal numbers of recurrences per group, the odds ratio for recurrence was 6.1 fold higher in the highest group compared with the lowest ($P = 0.005$, Fig. 3B). These data on protein expression thus confirm the mRNA expression data.

Larynx cancer cell lines

In addition to cellular radiosensitivity, the effectiveness of fractionated radiotherapy can be determined by microenvironmental factors such as hypoxia, repopulation rates during therapy, and the fraction of stem cells. As a first step in attempting to dissect the role played by *CD44* on these factors, we studied a series of larynx cancer cell lines under well controlled in vitro conditions. As shown in Fig. 4, *CD44* mRNA levels (average for the three probes) correlated significantly with plating efficiency ($P = 0.03$). Since plating efficiency has been correlated with tumor initiating capacity in several studies, this is consistent with *CD44* being a stem cell marker in this tumor type. In the same experiments, CD44 expression did not correlate with intrinsic radiosensitivity in these 9 larynx cancer cell lines (P-value = 0.71). None of the three *CD44* probes individually showed a correlation with radiosensitivity, while two out of three CD44 probes show a significant correlation with plating efficiency (Supplemental Table 4). These data imply that *CD44* expression is not monitoring intrinsic radiosensitivity but rather the fraction of stem cells.

Discussion

The aim of this study was to find prediction markers for clinical outcome of larynx cancer after radiotherapy using gene expression profiling. We chose to study early stage tumors since these will be inherently less variable than advanced cancer both in terms of genetically different

subpopulations and variability in blood flow and hypoxia. In addition, delivery of the radiotherapy is less complicated with less chance of geographical misses. Any recurrences are therefore likely to be due to inherent resistance of the tumor cells. Secondly, we chose to match recurrent and non-recurrent patients for the most important known clinical variables (T-stage, subsite, treatment, gender and age), so that these would not be confounding factors in the analysis.

In the test series, we studied the expression of several sets of genes monitoring biological processes known to influence the outcome of radiotherapy. We found that *CD44*, chosen as a stem cell marker, showed the most significant correlation with local recurrence. Expression of genes monitoring proliferation and intrinsic radiosensitivity showed no correlation with outcome. A gene set defining acute hypoxia showed a trend, although not significant when corrected for multiple testing. In a separate data-driven analysis including over 8000 genes (after filtering out genes not showing significant expression or significant

variation across the samples), the three probes for *CD44* came out high in the ranking list of genes correlating with recurrence, one of the probes being the most significant of all genes tested. This non-hypothesis-driven approach supported the hypothesis-driven approach, indicating that *CD44* is a good predictor of outcome after radiotherapy in these head and neck squamous cell carcinomas. Furthermore, in an independent validation series, CD44 protein expression measured immunohistochemically correlated significantly with outcome, such that higher CD44 scores were associated with a higher chance of local recurrence. Since both these were matched series, results are independent of the most important clinical predictors.

In a previous expression profiling study from our own institute on a series of 91 HNSCC patients treated with concurrent radiation and cisplatin, *CD44* was higher in tumors from patients which subsequently developed a recurrence, although this did not reach significance ($P = 0.08$) (ref. 25). Kawano et al found CD44s and CD44v6 staining correlated with prognosis in a series of 57 patients treated with surgery and radiotherapy (35). Zhao et al analyzed margins after surgery for 112 HNSCC patients and found that CD44v6 presence in these margins, detected with immunohistochemistry, was predictive of recurrence (36). Wang et al. (37) found that one CD44 isoform (v10) was associated with reduced disease free survival in HNSCC. These, together with the present study, support CD44 expression as a negative predictive factor.

We chose *CD44* as a stem cell marker for HNSCC, since Prince et al. (31) showed that CD44 positive cells in this tumor type were up to an order of magnitude more tumorigenic than CD44 negative cells. These data indicated that CD44 positive cells are enriched in cancer stem cells. However, our and other (38) IHC studies showed a relatively high average percentage of cells staining for CD44, inconsistent with a small minority stem cell fraction. We and others also observed a gradient of CD44 staining, where cells in more basal-like areas stained more positively than cells in the more differentiated areas. Such patterns may reflect more stem like properties of cells in the basal-like areas, analogous to that in normal epithelia.

Assuming that *CD44* has a causal role in determining the chance of recurrence and is not simply an indirect marker for stem cell content or another unknown process, there are several possible explanations for this role in the many functions of CD44. CD44 is a transmembrane glycoprotein with many transcript variants and has hyaluronan, an extracellular matrix protein, as a ligand (39). Various functions of CD44 have been described, including promoting tumorigenesis, cell motility and invasion. CD44, when activated by ligand, can act as a co-receptor for several membrane receptors, triggering various intracellular signalling pathways. In one of these, CD44 acts as co-receptor for the ErbB family which can lead to activation of the *PI3K/AKT* pathway, a pathway known to promote survival after cytotoxic damage, including after irradiation. This suggests a possible link between *CD44* expression and intrinsic radiosensitivity. However, we did not find a
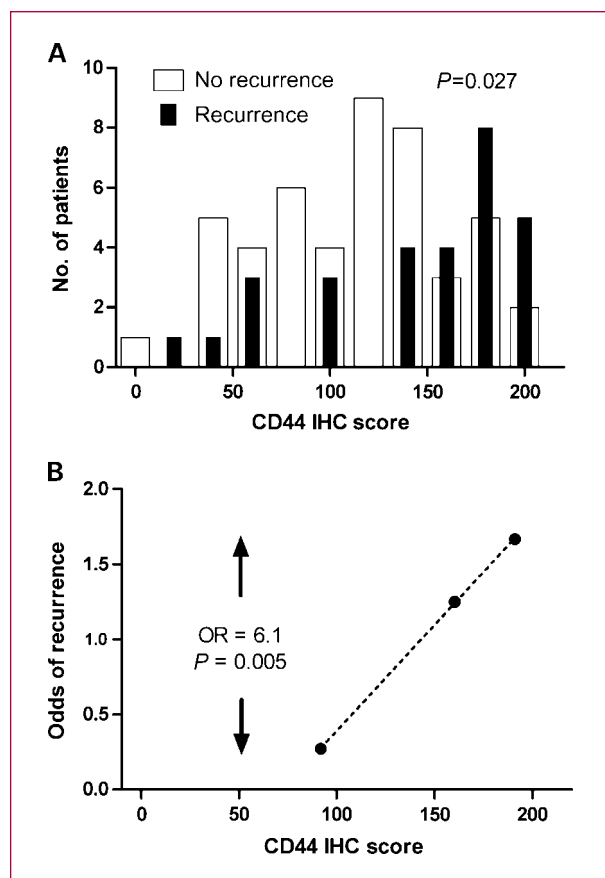

Fig. 3. CD44 IHC predicts local recurrence. A, histograms of CD44 IHC score for patients subsequently cured (open bars) or those subsequently suffering a recurrence (closed bars). B, odds of recurrence when patients are divided into three groups with increasing IHC scores, split so that each group contains equal numbers of recurrences. OR: odds ratio of recurrence between highest and lowest CD44 expression groups.
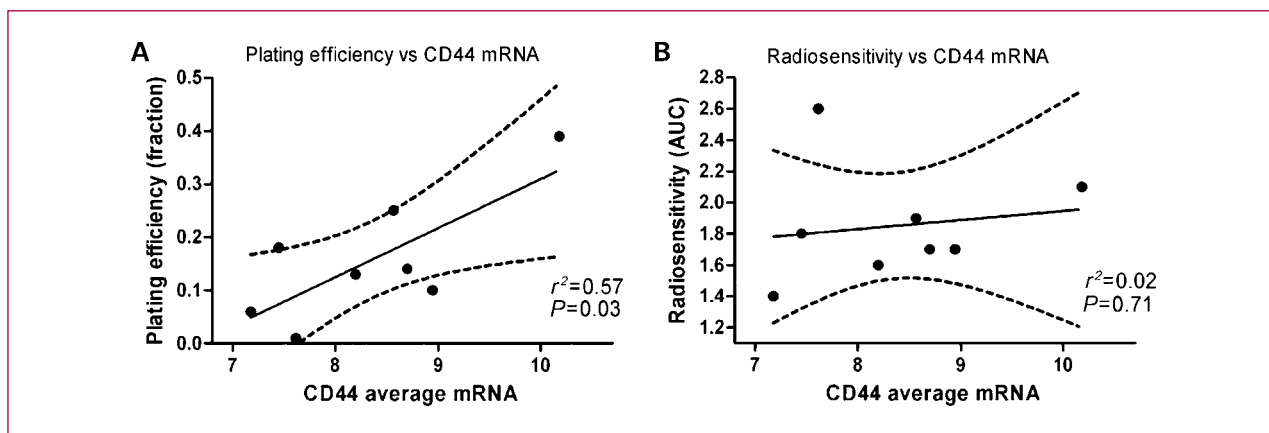
Fig. 4. Correlation of plating efficiency (A) and radiosensitivity (B; as measured by area under the survival curve (AUC)) with CD44 mRNA levels (averaged over the 3 probes).

correlation between *CD44* expression and radiosensitivity in the panel of larynx cancer cell lines. Alternatives therefore need to be sought to explain the relationship between *CD44* expression and radiocurability.

Other possibilities are links with hypoxia or repopulating ability, both known to influence radiotherapy outcome. We found that *CD44* expression correlated with expression of acute hypoxia genes (Supplemental Fig. 4) and a trend ($P = 0.08$) that expression of acute hypoxia genes correlated with chance of recurrence. No significant relationship with expression of chronic hypoxia genes was found. This is consistent with other studies indicating that cells hypoxic for relatively short times are more dangerous than those chronically exposed to hypoxia (40, 41). We found no evidence of a link between *CD44* expression and proliferation associated genes, or in this series between expression of proliferation genes and outcome. This is consistent with our earlier expression profiling studies on advanced head and neck tumors treated with radiotherapy and cisplatin, where proliferation genes were not predictive (25). Whether this is due to relatively slow repopulation rates in these tumors, or because the signatures do not adequately monitor repopulation capacity during fractioned radiotherapy is not known.

A final possible explanation is that *CD44* expression monitors the number of stem or cancer initiating cells. It is unlikely that all CD44-positive cells have stem cell properties, considering the rather ubiquitous expression of *CD44* in normal tissues (www.genecards.org), and the relatively high average fraction of CD44-positive cells in the tumors studied here and elsewhere (37, 38). However, if the cancer stem cells are a constant subfraction of CD44-positive tumor cells, the stem cell fraction (or tumor initiating fraction) will be directly correlated with the CD44-positive fraction. In the current study, this fraction varied by a factor of around 3. Based on Poisson statistics, such a three-fold change in the effective number of cells which need to be killed by radiation would lead to an absolute change in the cure probability of around 30%;

e.g. 1 surviving cell on average would lead to 37% cure probability, whereas 3 surviving cells on average would lead to a 5% cure probability. It is therefore possible that the relationship between cure and CD44 expression is a reflection of the number of cancer initiating cells needed to be killed. This is independent of whether the putative stem cells are more or less radioresistant than bulk tumor cells.

This contention is supported by the cell line data where *CD44* expression correlated significantly with colony forming efficiency of unirradiated cells (and not with radiosensitivity). This suggests a correlation with cancer initiating properties, since several studies have shown a correlation between in vitro plating efficiency and the number of cells required to produce tumors in animals (42–44). In addition, the Glinsky signature (45), a putative stem cell signature, also showed a strong trend with outcome in the test series (Table 1). This *BMI-1*-driven signature was derived by comparing primary and metastatic prostate cancer. We performed an Ingenuity pathway analysis on this 11-gene signature, also including CD44. The only significant pathway resulting from the analysis showed a link between the Glinsky genes and CD44 through an interaction with TGFB1 (supplementary Fig. 2). While not definitive, these data support the notion that *CD44* is in some way monitoring stem cell capacity.

Various CD44 isoforms have been described with different functions (39). In the present study, correlations with outcome were found with mRNA probes for one of the constant regions, and with an antibody against a constantly expressed epitope. Whether variant isoform expression would provide better prediction or understanding of failure needs further study.

Summary and Conclusion

CD44 expression, both at the mRNA and protein levels in independent patient series, correlated with the probability of recurrence after radiotherapy for early stage larynx

cancer. Possible explanations are that *CD44* expression monitors the cancer stem cell fraction or that *CD44* expression monitors the hypoxic fraction. It will be important to distinguish these two possibilities, since interventions to increase cure in patients with high *CD44* expressing tumors will depend on the mechanism (attacking hypoxia, or attacking CD44 itself, or its downstream pathways, or other stem cell specific pathways). Predicting outcome is important partly to spare patients ineffective and toxic therapies. It will be equally or more valuable to provide alternative therapies for patients with resistant tumors. It is likely that CD44 expression, measured with standard immunohistochemical or perhaps PCR-based assays will contribute to better outcome prediction, and the next steps will be to confirm mechanisms and design effective interventions against the consequences of this over-expression. The present data suggest that the association between *CD44* and radioresponse reflects an increased number of cancer initiating cells that are usually resistant to radiation and result in a recurrence.

CD44 might therefore provide a new marker to predict the radiotherapy response in a biopsy of the primary tumor before treatment is initiated.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

References

1. Carvalho AL, Nishimoto IN, Califano JA, Kowalski LP. Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database. Int J Cancer 2005;114:806–16.
2. Sjogren EV, Wiggenraad RG, Le Cessie S, Snijder S, Pomp J, de Jong RJ. Outcome of radiotherapy in T1 glottic carcinoma: a population-based study. Eur Arch Otorhinolaryngol 2009;266:735–44.
3. Dirven R, Swinson BD, Gao K, Clark JR. The assessment of pharyngocutaneous fistula rate in patients treated primarily with definitive radiotherapy followed by salvage surgery of the larynx and hypopharynx. Laryngoscope 2009;119:1691–5.
4. Franchin G, Minatel E, Gobitti C, et al. Radiotherapy for patients with early-stage glottic carcinoma: univariate and multivariate analyses in a group of consecutive, unselected patients. Cancer 2003;98:765–72.
5. Rosenthal DI, Ang KK. Altered radiation therapy fractionation, chemoradiation, and patient selection for the treatment of head and neck squamous carcinoma. Semin Radiat Oncol 2004;14:153–66.
6. Groome PA, Schulze K, Boysen M, et al. A comparison of published head and neck stage groupings in laryngeal cancer using data from two countries. J Clin Epidemiol 2002;55:533–44.
7. Bjork-Eriksson T, West C, Karlsson E, Mercke C. Tumor radiosensitivity (SF2) is a prognostic factor for local control in head and neck cancers. Int J Radiat Oncol Biol Phys 2000;46:13–9.
8. Janssen HL, Haustermans KM, Balm AJ, Begg AC. Hypoxia in head and neck cancer: how much, how important? Head Neck 2005;27:622–38.
9. Kim JJ, Tannock IF. Repopulation of cancer cells during therapy: an important cause of treatment failure. Nat Rev Cancer 2005;5:516–25.
10. Wildeman MA, Gibcus JH, Hauptmann M, et al. Radiotherapy in laryngeal carcinoma: can a panel of 13 markers predict response? Laryngoscope 2009;119:316–22.
11. Buffa FM, Bentzen SM, Daley FM, et al. Molecular marker profiles predict locoregional control of head and neck squamous cell carcinoma in a randomized trial of continuous hyperfractionated accelerated radiotherapy. Clin Cancer Res 2004;10:3745–54.
12. Koukourakis MI, Giatromanolaki A, Danielidis V, Sivridis E. Hypoxia inducible factor (HIF1α and HIF2α) and carbonic anhydrase 9 (CA9) expression and response of head-neck cancer to hypofractionated and accelerated radiotherapy. Int J Radiat Biol 2007;1–6.
13. Winter SC, Buffa FM, Silva P, et al. Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. Cancer Res 2007;67:3441–9.
14. Chi JT, Wang Z, Nuyten DS, et al. Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. PLoS Med 2006;3:e47.
15. Ishigami T, Uzawa K, Higo M, et al. Genes and molecular pathways related to radioresistance of oral squamous cell carcinoma cells. Int J Cancer 2007;120:2262–70.
16. Torres-Roca JF, Eschrich S, Zhao H, et al. Prediction of radiation sensitivity using a gene expression classifier. Cancer Res 2005;65:7169–76.
17. Amundson SA, Do KT, Vinikoor LC, et al. Integrating global gene expression and radiation survival parameters across the 60 cell lines of the National Cancer Institute Anticancer Drug Screen. Cancer Res 2008;68:415–24.
18. Starmans MH, Krishnapuram B, Steck H, et al. Robust prognostic value of a knowledge-based proliferation signature across large patient microarray studies spanning different cancer types. Br J Cancer 2008;99:1884–90.
19. Baumann M, Krause M, Thames H, Trott K, Zips D. Cancer stem cells and radiotherapy. Int J Radiat Biol 2009;85:391–402.
20. Baumann M, Krause M, Hill R. Exploring the role of cancer stem cells in radioresistance. Nat Rev Cancer 2008;8:545–54.
21. Dumur CI, Ladd AC, Wright HV, et al. Genes involved in radiation therapy response in head and neck cancers. Laryngoscope 2009;119:91–101.
22. Chung CH, Parker JS, Karaca G, et al. Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. Cancer Cell 2004;5:489–500.
23. Chung CH, Parker JS, Ely K, et al. Gene expression profiles identify epithelial-to-mesenchymal transition and activation of nuclear factor-κB signaling as characteristics of a high-risk head and neck squamous cell carcinoma. Cancer Res 2006;66:8210–8.
24. Lohavanichbutr P, Houck J, Fan W, et al. Genomewide gene expression profiles of HPV-positive and HPV-negative oropharyngeal cancer: potential implications for treatment choices. Arch Otolaryngol Head Neck Surg 2009;135:180–8.
25. Pramana J, van den Brekel MW, van Velthuysen ML, et al. Gene expression profiling to predict outcome after chemoradiation in head and neck cancer. Int J Radiat Oncol Biol Phys 2007;69:1544–52.
26. Slebos RJ, Yi Y, Ely K, et al. Gene expression differences associated with human papillomavirus status in head and neck squamous cell carcinoma. Clin Cancer Res 2006;12:701–9.
27. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing

transformation for Illumina microarray data. Nucleic Acids Res 2008; 36:e11.

28. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics 2008;24:1547–8.

29. R Development Core Team (2008). R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2009.

30. Amundson SA, Grace MB, McLeland CB, et al. Human *in vivo* radiation-induced biomarkers: gene expression changes in radio-therapy patients. Cancer Res 2004;64:6368–71.

31. Prince ME, Sivanandan R, Kaczorowski A, et al. Identification of a sub-population of cells with cancer stem cell properties in head and neck squamous cell carcinoma. Proc Natl Acad Sci U S A 2007;104:973–8.

32. Grenman R, Burk D, Virolainen E, Wagner JG, Lichter AS, Carey TE. Radiosensitivity of head and neck cancer cells *in vitro*. A 96-well plate clonogenic cell assay for squamous cell carcinoma. Arch Otolaryngol Head Neck Surg 1988;114:427–31.

33. Grenman R, Burk D, Virolainen E, et al. Clonogenic cell assay for anchorage-dependent squamous carcinoma cell lines using limiting dilution. Int J Cancer 1989;44:131–6.

34. Grenman R, Carey TE, McClatchey KD, et al. *In vitro* radiation resistance among cell lines established from patients with squamous cell carcinoma of the head and neck. Cancer 1991;67:2741–7.

35. Kawano T, Nakamura Y, Yanoma S, et al. Expression of E-cadherin, and CD44s and CD44v6 and its association with prognosis in head and neck cancer. Auris Nasus Larynx 2004;31:35–41.

36. Zhao H, Ren J, Zhuo X, Ye H, Zou J, Liu S. Prognostic significance of Survivin and CD44v6 in laryngeal cancer surgical margins. J Cancer Res Clin Oncol 2008;134:1051–8.

37. Wang SJ, Wong G, de Heer AM, Xia W, Bourguignon LY. CD44 variant isoforms in head and neck squamous cell carcinoma progression. Laryngoscope 2009;119:1518–30.

38. Mack B, Gires O. CD44s and CD44v6 expression in head and neck epithelia. PLoS One 2008;3:e3360.

39. Ponta H, Sherman L, Herrlich PA. CD44: from adhesion molecules to signalling regulators. Nat Rev Mol Cell Biol 2003;4:33–45.

40. Martinive P, Defresne F, Bouzin C, et al. Preconditioning of the tumor vasculature and tumor cells by intermittent hypoxia: implications for anticancer therapies. Cancer Res 2006;66:11736–44.

41. Seigneuric R, Starmans MH, Fung G, et al. Impact of supervised gene signatures of early hypoxia on patient survival. Radiother Oncol 2007;83:374–82.

42. Addla SK, Brown MD, Hart CA, Ramani VA, Clarke NW. Character-ization of the Hoechst 33342 side population from normal and malignant human renal epithelial cells. Am J Physiol Renal Physiol 2008; 295:F680–7.

43. Hill RP, Milas L. The proportion of stem cells in murine tumors. Int J Radiat Oncol Biol Phys 1989;16:513–8.

44. Ma S, Chan KW, Hu L, et al. Identification and characterization of tumorigenic liver cancer stem/progenitor cells. Gastroenterology 2007;132:2542–56.

45. Glinsky GV, Berezovska O, Glinskii AB. Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. J Clin Invest 2005;115:1503–21.