Supplementary Table 1: Patients databases used for analysis

Reference	Source of data	N° of samples	Technological platform	N° of probe sets	N° of samples used
van't Veer et al., Nature 2002	http://www.rii.com/publications/2002/vantveer.html	117	Agilent Hu25K	25K	117
van de Vijver et al., NEJM 2002	http://microarray-pubs.stanford.edu/wound_NKI/	295	Agilent Hu25K	25K	254
Expression Project for Oncology (expO), 2005	https://expo.intgen.org/geo GEO: GSE2109	348	Affymetrix U133 Plus 2.0	54K	348
Wang Y et al., Lancet 2005	GEO: GSE2034	286	Affymetrix U133A	22K	286
Minn AJ et al., Nature 2005	GEO: GSE2603	99	Affymetrix U133A	22K	99
Farmer P et al., Oncogene 2005	GEO: GSE1561	49	Affymetrix U133A	22K	49
Ivshina et al., Cancer Res 2006	GEO: GSE4922, GSE1456	448	Affymetrix U133 A+B	2x22K	448
Sotiriou C et al., J Natl Cancer Inst 2006	GEO: GSE2990	189	Affymetrix U133A	22K	80
Hess KR et al., J Clin Oncol 2006	MDA133	133	Affymetrix U133A	22K	133
Desmedt C etal., Clin Cancer Res 2007	GEO: GSE7390	198	Affymetrix U133A	22K	154
Bonnefoi et al., Lancet Oncol 2007	GEO: GSE6861, GSE4779	161	Affymetrix X3P	61K	125
Klein A et al., Int J Cancer 2007	GEO: GSE6596	26	Affymetrix U133A	22K	24
Schmidt M etal., Cancer Res 2008	GEO: GSE11121	200	Affymetrix U133A	22K	200
Yu K et al., PLoS Genet 2008	GEO: GSE5364	196	Affymetrix U133A	22K	183
Merritt WM et al., N Engl J Med 2008	Array Express: E-MTAB-158	130	Affymetrix U133AAofAv2	23K	130
Marty et al., Breast Cancer Res 2008	GEO: GSE13787	23	Affymetrix U133 Plus 2.0	54K	23
Bos et al., Nature 2009	GEO: GSE12276	204	Affymetrix U133 Plus 2.0	54K	204
Zhang Y et al., Breast Cancer Res Treat 2009	GEO: GSE12093	136	Affymetrix U133A	22K	136
Hoeflich et al., Clin Cancer Res 2009	GEO: GSE12763	30	Affymetrix U133 Plus 2.0	54K	30
Jonsson et al., BCR 2010	GEO: GSE22133	359	Swegene H_v2.1.1 55K	55K	346

Supplementary Table 1: Patients databases used for analysis

Reference	Source of data	N° of samples	Technological platform	N° of probe sets	N° of samples used
Prat A et al., Breast Cancer Res 2010	GEO: GSE18229	337	Agilent Hu25K	25K	264
Popovici V et al., Breast Cancer Res 2010	GEO: GSE20194	278	Affymetrix U133A	22K	91
Chen et al., Breast Cancer Res Treat 2010	GEO: GSE10780	185	Affymetrix U133 Plus 2.0	54K	42
Tabchy A et al., Clin Cancer Res 2010	GEO: GSE20271	178	Affymetrix U133A	22K	178
Miller WR et al., Breast Cancer Res 2010	GEO: GSE5462	116	Affymetrix U133A	22K	116
Silver et al., J Clin Oncol 2010	GEO: GSE18864	84	Affymetrix U133 Plus 2.0	54K	84
Korde et al., Breast Cancer Res Treat 2010	GEO: GSE18728	61	Affymetrix U133 Plus 2.0	54K	61
Barry et al., J Clin Oncol 2010	GEO: GSE23593	50	Affymetrix U133 Plus 2.0	54K	50
Van der Auwera et al., Br J Cancer 2010	Array Express: E-MTAB-1006	96	Affymetrix U133 Plus 2.0	54K	96
Guedj et al., Oncogene 2011	Array Express: E-MTAB-365	537	Affymetrix U133 Plus 2.0	54K	452
Hatzis C et al., JAMA 2011	GEO: GSE25066	508	Affymetrix U133A	22K	508
Sabatier R et al., (our IPC series) PLoS One 2011	GEO: GSE31448	353	Affymetrix U133 Plus 2.0	54K	352
Iwamoto T et al., J Natl Cancer Inst 2011	GEO: GSE22093, GSE23988	164	Affymetrix U133A	22K	100
Desmedt et al., J Clin Oncol 2011	GEO: GSE16446	120	Affymetrix U133 Plus 2.0	54K	120
Pusztai et al., Breast Cancer Res Treat 2011	GEO: GSE22597	83	Affymetrix U133A	22K	83
Curtis et al., Nature 2012	EGA: EGAS00000000083	2136	Illumina HT 12	49K	1974
TCGA, Nature 2012	TCGA Data Portal - BRCA -	1215	Illumina, RNAseq V2	20K	1095
Ellis et al., Nature 2012	GEO: GSE29442, GSE35186	201	Agilent-014850 4x44K	44K	201
TOTAL		10329			9236

Supplementary Table 2: Clinicopathological characteristics of sample

Characteristics		N (%)
Age at diagnosis (years)	≤50	2650 (36%)
	>50	4630 (64%)
AJCC Stage	1	1006 (36%)
	2	1130 (40%)
	3	599 (21%)
	4	82 (3%)
Pathological type	ductal	4184 (79%)
	lobular	514 (10%)
	other	580 (11%)
Pathological grade	1	739 (11%)
	2	2633 (40%)
	3	3151 (48%)
Pathological lymph node status (pN)	negative	3667 (57%)
	positive	2789 (43%)
Pathological tumor size (pT)	pT1	2117 (37%)
	pT2	2926 (52%)
	pT3	605 (11%)
ER status (mRNA)	negative	2867 (31%)
	positive	6369 (69%)
PR status (mRNA)	negative	4818 (52%)
	positive	4361 (48%)
HER2 status (mRNA)	negative	8087 (88%)
	positive	1149 (12%)
Molecular subtypes	HR+/HER2-	6061 (66%)
	HER2+	1149 (12%)
	TN	2007 (22%)
Follow-up, median (months)		57.2 (0-299.24)
Metastasic relapse		899 (24%)
5-year MFS [95%CI]		74% [72-75]

Supplementary Table 3: Histo-clinical variables analyses

logistic regression HC grp 'II' vs. 'I'	Univariate		
	N	Odds-ratio [CI95]	p-value
age.g, >50 vs. <= 50 years	7280	0.77 [0.71-0.84]	**1.10E-07**
Grade, 2 vs. 1	6523	1.57 [1.36-1.81]	**2.79E-07**
Grade, 3 vs. 1	6523	3.36 [2.92-3.88]	**1.84E-44**
pN, 1 vs. 0	6252	1.2 [1.1-1.3]	**3.78E-04**
pT, pT2 vs. pT1	5648	1 [0.91-1.1]	0.965
pT, pT3 vs. pT1	5648	0.98 [0.85-1.14]	0.858
Histology, ILC vs. IDC	5278	1.04 [0.89-1.21]	0.664
Histology, other vs. IDC	5278	0.85 [0.74-0.98]	0.069
Mol. subtype, HER2+ vs. HR+/HER2-	9217	2.29 [2.05-2.55]	**5.41E-36**
Mol. subtype, TN vs. HR+/HER2-	9217	3.8 [3.46-4.17]	**1.85E-123**
RT adj, 1 vs. 0	2102	1.59 [1.37-1.84]	**2.42E-07**
CT adj, 1 vs. 0	6260	1.87 [1.67-2.08]	**4.78E-21**
HT adj, 1 vs. 0	4912	0.65 [0.59-0.72]	**1.47E-13**

Supplementary Table 4: Univariate and multivariate analyses of correlation between overexpression of C(X)C genes with CSC-associated profiles.

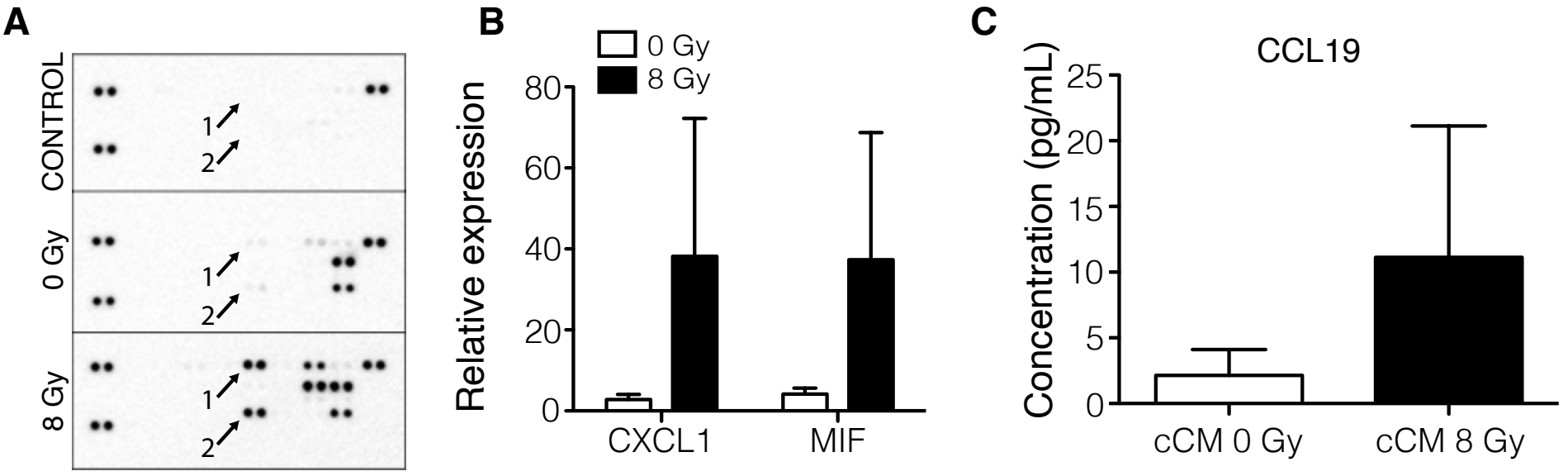
logistic regression HC grp 'II' vs. 'I'	Univariate		
	N	Odds-ratio [CI95]	p-value
Prat, Claudin-low vs. non claudin-low	9236	3.93 [3.38-4.55]	**7.60E-52**
Lim, luminal prog vs. mature luminal	9236	8.69 [7.88-9.59]	**2.80E-288**
Lim, MaSC vs. mature luminal	9236	4.41 [4.02-4.83]	**1.32E-156**
Prat differentiation, pL->mL vs. MaSC -> pL	9236	0.27 [0.25-0.29]	**2.13E-188**
Charafe, ALDH1pos-like vs. ALDH1neg-like	9236	1.29 [1.21-1.39]	**6.45E-10**
Creighton, CD44+/CD24- like yes vs. no	9236	0.14 [0.12-0.16]	**1.56E-98**

logistic regression HC grp 'II' vs. 'I'	**Multivariate ***		
	N	Odds-ratio [CI95]	p-value
Prat, Claudin-low vs. non claudin-low	2017	5.29 [1.88-14.84]	**7.95E-03**
Lim, luminal prog vs. mature luminal	2017	7.29 [5.74-9.24]	**5.26E-43**
Lim, MaSC vs. mature luminal	2017	4.54 [3.69-5.59]	**5.20E-33**
Prat differentiation, pL->mL vs. MaSC -> pL	2017	0.37 [0.3-0.44]	**3.57E-19**
Charafe, ALDH1pos-like vs. ALDH1neg-like	2017	1.00 [0.85-1.17]	0.959
Creighton, CD44+/CD24- like yes vs. no	2017	0.09 [0.06-0.13]	**4.25E-20**

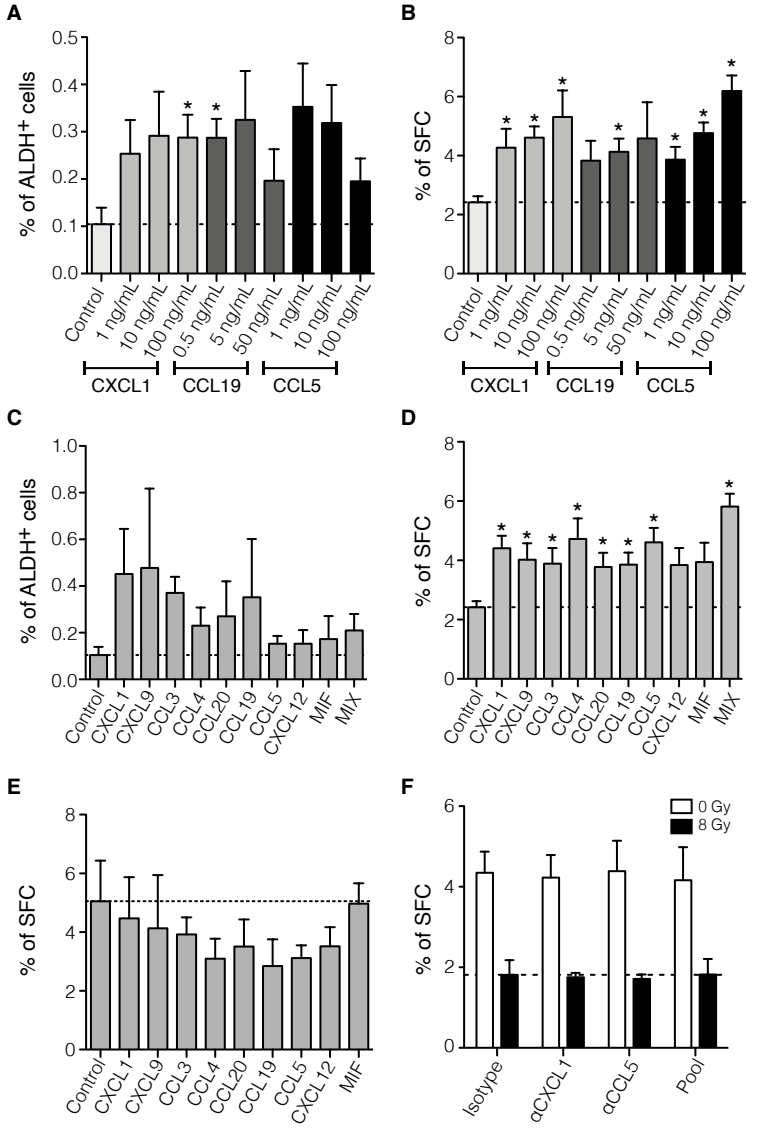
** multivariate vars (p5% UV) : Age Diag, Grade, pN, moleclular subtype, RT, HT & CT*

Supplementary Table 5: Chemokines receptors membrane expression in SUM159PT cells. SUM159PT were immunostained with specific antibodies and analysed by flow cytometry for the membrane expression of various chemokines receptors expression. Flow panels are available in Supplementary Figure 3.

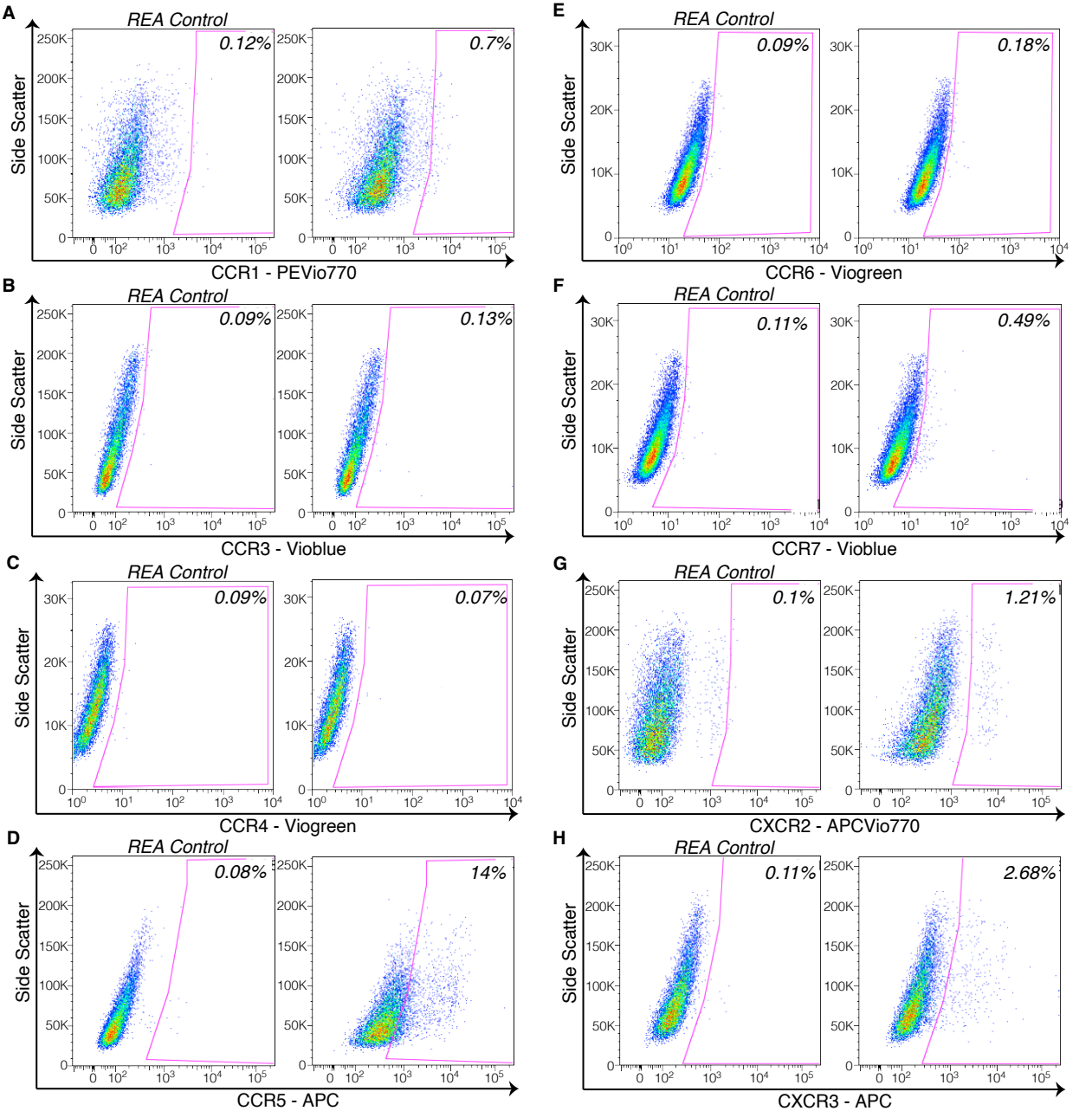
Receptors	Ligands	Expression
CCR1	CCL3,4,**5**,7,14,15,16,23	0.5-1%
CCR3	CCL**5**,7,11,13,15,24,26,28	0.1 %
CCR4	CCL2,3,**5**,17,22	0.1%
CCR5	CCL3,4,**5**,8	10-15%
CCR6	CCL20	<0.5%
CCR7	CCL19,21	0.5 %
CXCR2	CXCL**1**,2,3,5,6,7,8	1-2%
CXCR3	CXCL9,10,11	2-5%



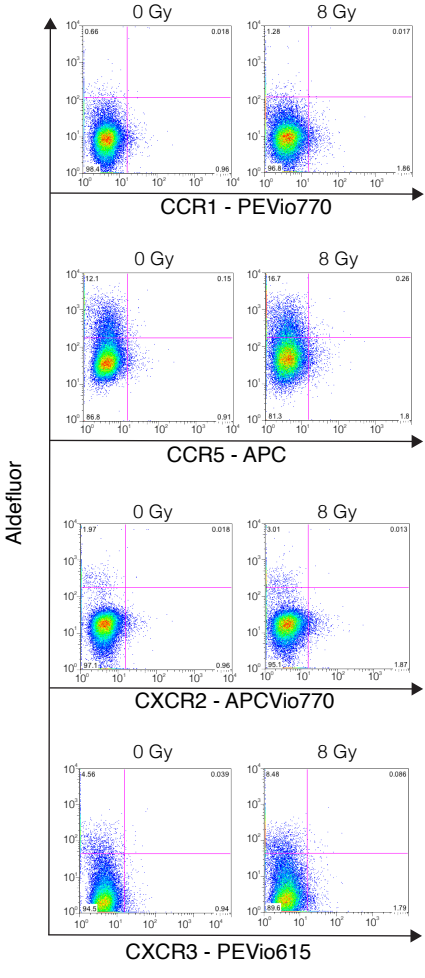
Supplementary Figure 1: Radiation-induced secretion of cytokines. (a) ALDH-/low non-CSCs were sorted and seeded in monolayer. Twenty-four hours after seeding, the cells were irradiated at 0 or 8 Gy. Five days after irradiation, CM was collected and analyzed by a chemokine array. The control consisted of fresh medium incubated for 6 days without cells at 37°C and 5% CO2. CXCL1 (1) and MIF (2) are indicated by arrows. **(b)** Relative expression of cytokines that was modified after irradiation. **(c)** Relative quantification of CCL19 by ELISA in 5-fold concentrated CM (cCM). As described before, CM was collected five days after irradiation and concentrated with an Amicon Ultra 10k. Data are represented by means ± SEM. n = 3.



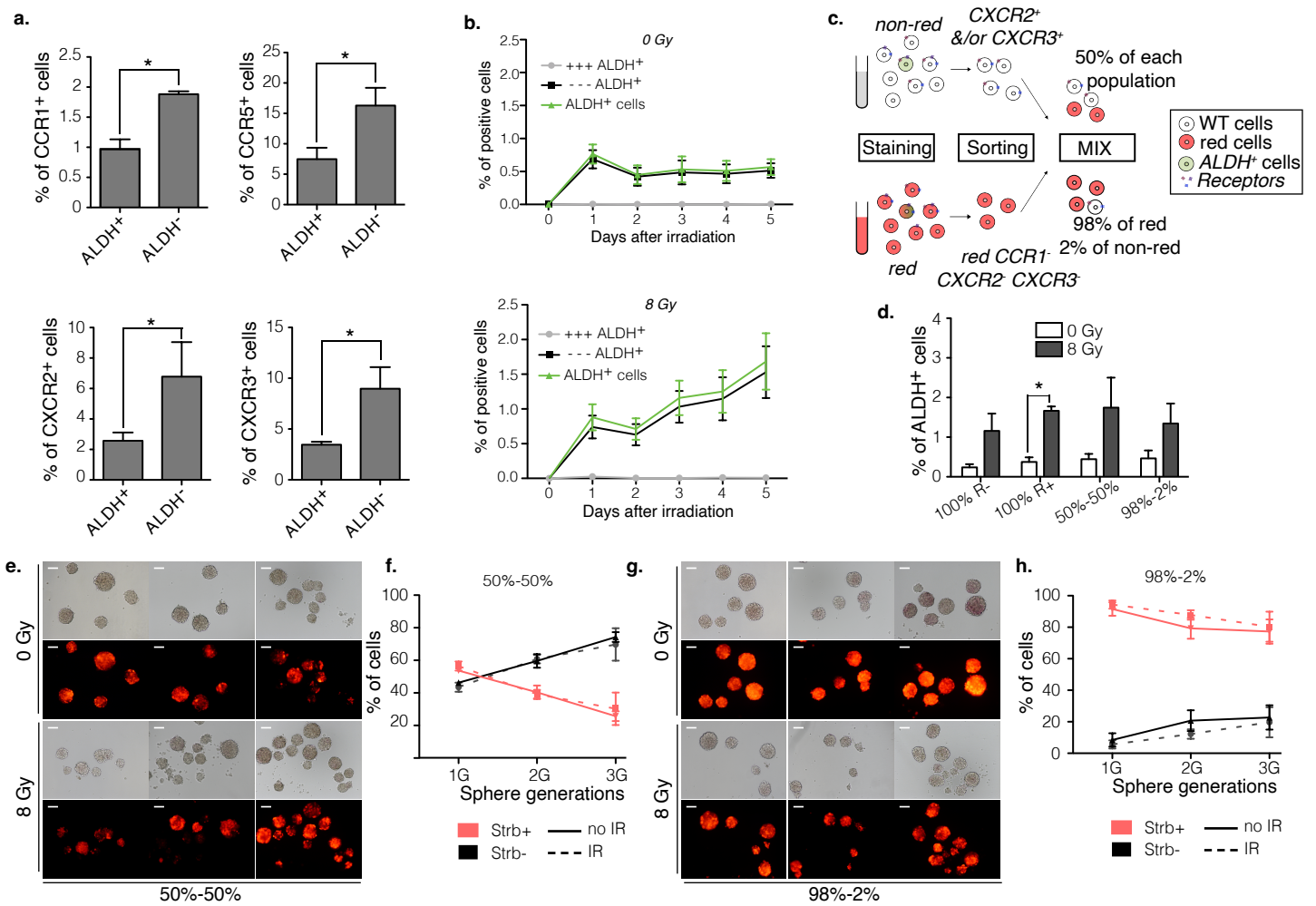
Supplementary Figure 2: Chemokines are involved in radiation-induced reprogramming. (a, b) Freshly sorted non-CSCs were seeded in a monolayer. Twenty-four hours after seeding, the cells were treated with different concentrations of chemokines CXCL1, CCL19 and CCL5. The control cells were treated with PBS 0.1% BSA. Five days after the treatments, Aldefluor assay was performed **(a)**, and the sphere-forming capacity was evaluated **(b)**. Data are represented by means ± SEM. *p<0.05, t-test, n = 3. **(c, d)** Individual chemokine treatments or mixed treatment (concentrations in Materials and Methods) were performed using freshly sorted non-CSC 24 hours after sorting. The control cells were treated with PBS 0.1% BSA. Five days after the treatments, Aldefluor assay was performed **(c)**, and the sphere-forming capacity was evaluated **(d)**. Data are represented by means ± SEM. *p<0.05, t-test, n = 3. **(e)** The same experiment was performed on bulk SUM159PT cells without sorting. Cells were plated in monolayer conditions. Twenty-four hours later, cells were treated with chemokines at the usual concentrations. Five days after the treatments, the sphere-forming capacity was evaluated. Data are represented by means ± SEM. *p<0.05, t test, n = 3. **(f)** Freshly sorted non-CSCs were treated with CM from irradiated non-CSCs supplemented with neutralizing antibodies targeting CXCL1 and CCL5 or isotype controls. Five days after treatment, the cells were seeded under low adherence conditions, and the sphere-forming capacity of the treated cells was analyzed. Data are represented by means ± SEM. n = 3.



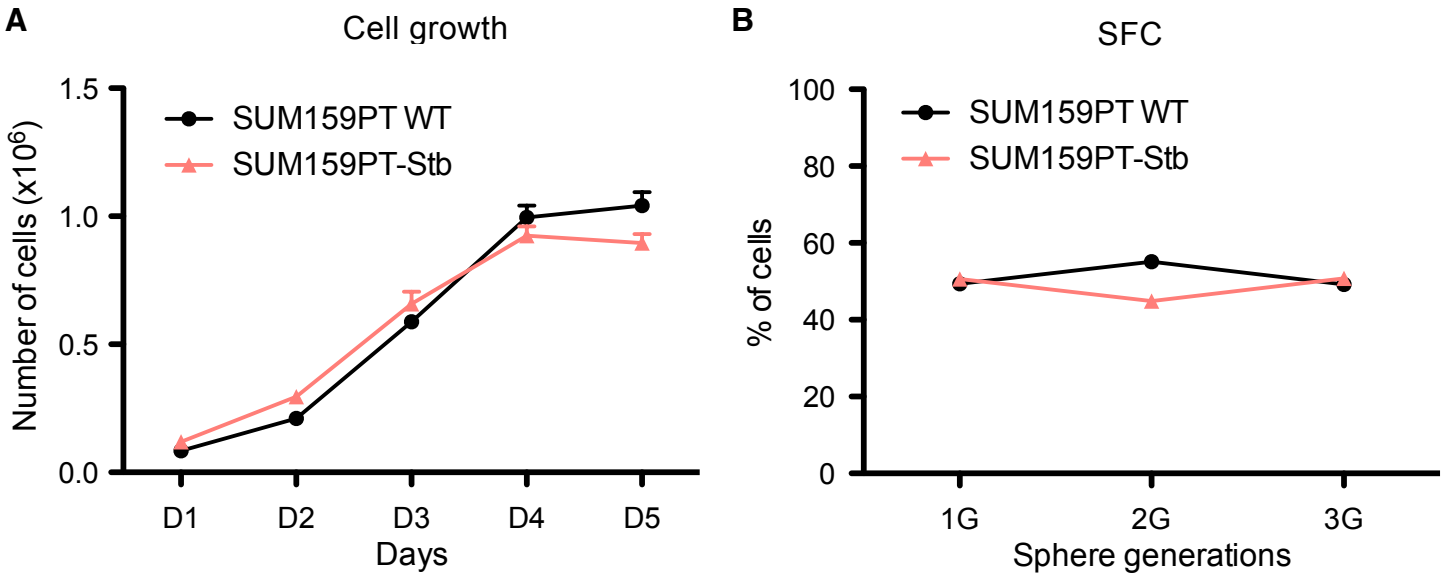
Supplementary Figure 3: Chemokines receptors membrane expression in SUM159PT cells. SUM159PT were immunostained with antiCCR1-PEVio770 **(a)**, antiCCR3-Vioblue **(b)**, antiCCR4-Biotin and antiBiotin-Viogreen **(c)**, antiCCR5-APC **(d)**, antiCCR6-Biotin and antiBiotin-Viogreen **(e)**, antiCCR7-Vioblue **(f)**, antiCXCR2-APCVio770 **(g)** and antiCXCR3-APC **(h)**. A REA control staining was performed for each condition (left figure on each panel), to define the positive population. Cells were analyzed by flow cytometry.



Supplementary Figure 4: Membrane expression of chemokine receptors after irradiation. Expression of CCR1, CCR5, CXCR2, and CXCR3 in Aldefluor stained cells by flow cytometry. A representative illustration of ALDH assay and individual receptor expression analyzed by flow

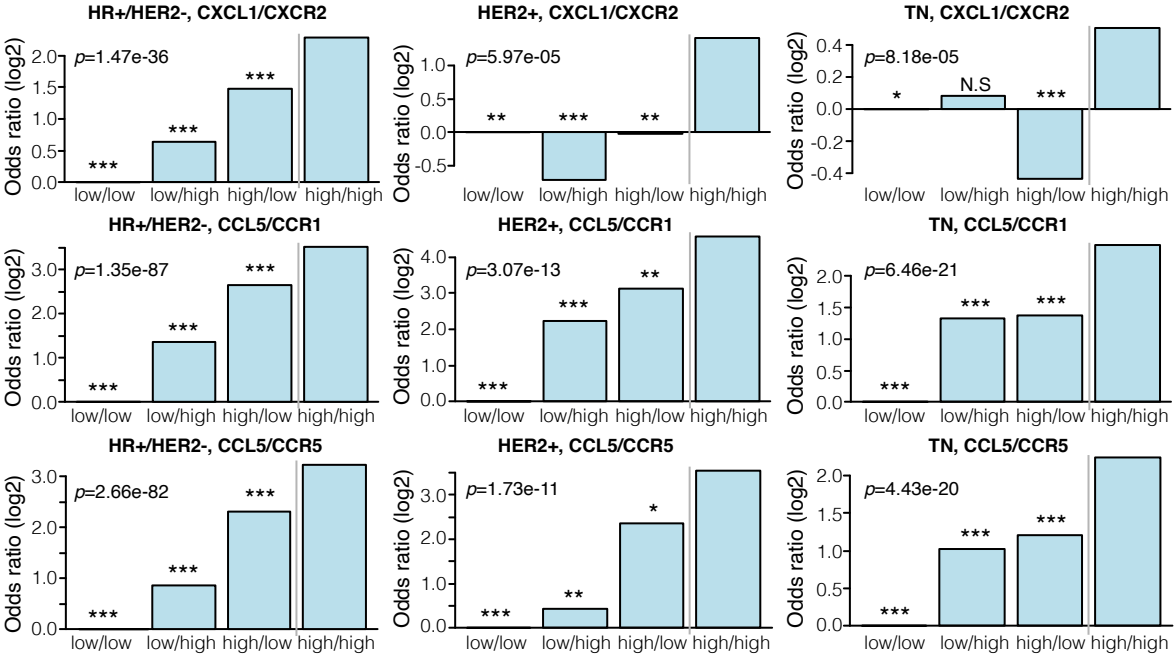


Supplementary Figure 5: Dynamique of expression and cell population expression CCR/CXCR. Membrane expression of chemokine receptors after irradiation and reprogrammable potential of receptors positive cells. (a) Expression of CCR1, CCR5, CXCR2, and CXCR3 in Aldefluor stained cells by flow cytometry. Data are represented by means ± SEM. *p<0.008, t-test, n = 3. (b) ALDH-/low non-CSCs were sorted for their receptors expression, and CCR1- CXCR2- CXCR3- non-CSCs were seeded and irradiated. Every day after irradiation, the cells were analyzed by flow cytometry for ALDH status and receptor expression, in particular cells expressing CCR1, CXCR2,,and CXCR3. The light gray line represents CCR1+ CXCR2+ CXCR3+ ALDH+ cells; the dark line shows CCR1- CXCR2- CXCR3- ALDH+ cells, The green line illustrates ALDH+ cells. Triple-positive and triple-negative cell populations were compared to bulk population. Data are represented by means ± SEM, two-way ANOVA test (comparing 0 vs 8 Gy for each sub-conditions: Bulk, ALDH-, ALDH+), n = 3. (c) Sorting protocol for SUM159PT WT and SUM159PT Strawberry+ cells (red). Each cell line was sorted based on ALDH staining and their receptor expression (either "ALDH- CCR1- CXCR2- CXCR3-" or "ALDH- CXCR2+ and/or CXCR3+"). Cells were mixed as follows: 50% of WT "ALDH- CXCR2+ and/or CXCR3+" and 50% of Strawberry+ "ALDH- CCR1- CXCR2- CXCR3-"; 2% of WT "ALDH- CXCR2+ and/or CXCR3+" and 98% of Strawberry+ "ALDH- CCR1- CXCR2- CXCR3-". Cells were then irradiated as usual. (d) Aldefluor assay of the different populations 5 days after irradiation. Data are represented by means ± SEM. *p<0.002, t-test, n = 3. (e, f, g, h) Mixed populations were maintained as spheres up to 3 generations, and for each generation, (e, g) Representative photographs of spheres from mixed populations of different generations. Photos were obtained with a Nikon Eclipse Ti microscope, 10x lens. Scale corresponds to 100 µm. (f, h) represent the percentage of Stb+ cells determined by flow cytometry. Data are represented by means ± SEM. *p<0.004, two-way ANOVA test, n = 4

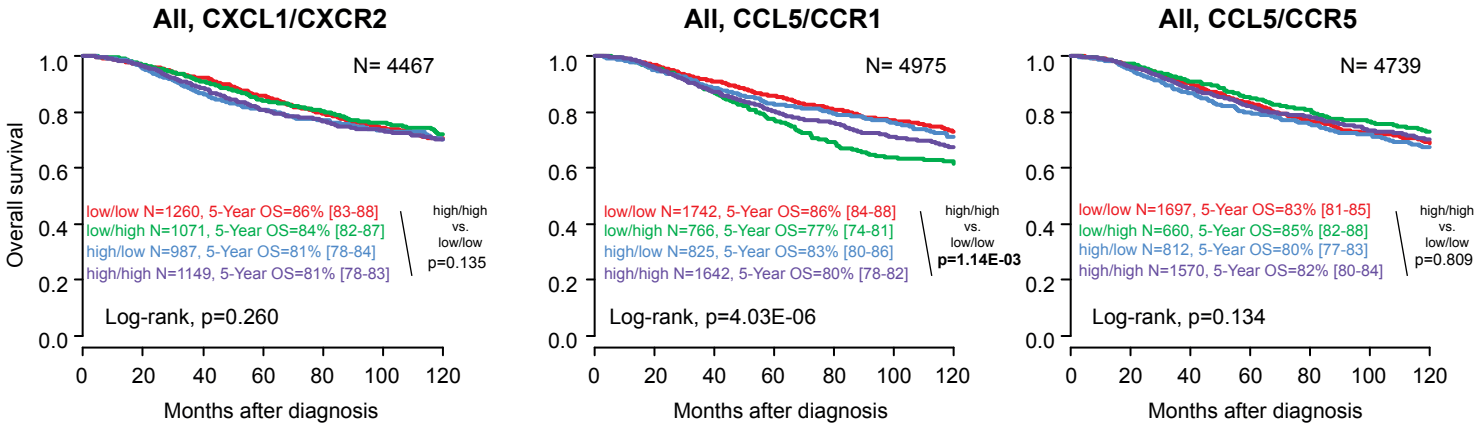


Supplementary Figure 6: SUM159PT-Strawberry controls. (a, b) SUM159PT WT and SUM159PT-Strawberry cells were seeded in monolayer **(a)** or equally mixed and seeded in suspension **(b)**. **(a)** The number of cells was evaluated every day for 5 days by cell counting. n = 2. **(b)** For 3 generations of spheres, cells were dissociated as previously described, and the proportion of each population (WT or Strawberry) was assessed by flow cytometry. n = 1.

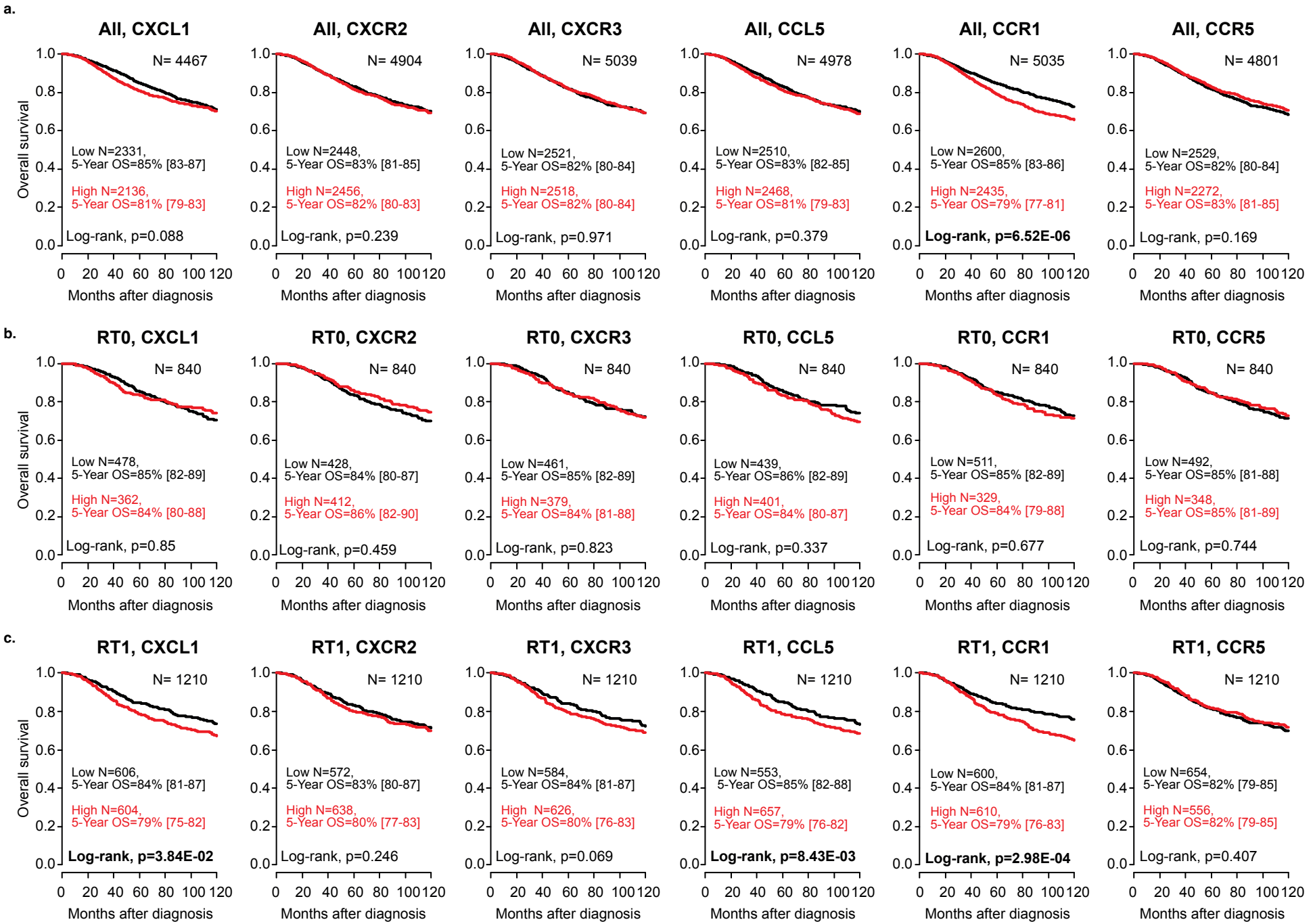
Bailleul *et al.*, Supplementary Figure 6



Supplementary Figure 7: Correlation between stem-cell phenotype and chemokine expression within BC subtypes. We evaluated the correlation by logistic regression with classification according to the stem cell signature as variable and to be explained. Four levels of expression were used: high/high, high/low, low/high and low/low, and low/low as a reference. The modality coefficient was plotted, and ANOVA was applied to evaluate significant correlations. *, corresponds to a significant t-test between each level of expression compared with the high/high condition.



Supplementary Figure 8: Overall survival (OS) of patients with breast cancer stratified according to C(X)CR and C(X)CL pair expression. The Kaplan-Meier curves are shown. log-rank test. N = 4467 to 4975. Individual curves have been compared using a one-way ANOVA test.



Supplementary Figure 9: Overall survival (OS) of patients with breast cancer stratified according to C(X)CR or C(X)CL pair expression and cluster by treatment. (a) Unclustered patients according to their treatment, **(b)** cluster of patients who didn't receive radiotherapy and **(c)** cluster of patients who have been treated with radiotherapy. Overall survival was followed over 120 months. log-rank test. N=840 to 5039.