Supporting Information

Rapid phenotyping of cancer stem cells using multichannel nanosensor arrays

Yingying Geng[a], Hira L. Goel, PhD[b], Ngoc B. Le, PhD[c], Tatsuyuki Yoshii, PhD[c, 1], Rubul Mout, PhD[c], Gulen Y. Tonga, PhD[c, 2], John J. Amante[b], Arthur M. Mercurio, PhD[b], Vincent M. Rotello, PhD[c, *]

[a]*Molecular and Cellular Biology Program, University of Massachusetts, Amherst, MA 01003, U.S.A.*
[b]*Department of Molecular, Cell and Cancer Biology, University of Massachusetts Medical School, Worcester, MA 01605, U.S.A.*
[c]*Department of Chemistry, University of Massachusetts, Amherst, MA 01003, U.S.A.*

* Correspondence should be addressed to V.R. (rotello@chem.umass.edu)

Table of contents

[1] Department of Life Science of Applied Chemistry, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan
[2] Department of Anaesthesia, Boston Children's Hospital, 300 Longwood Ave., Boston MA 02115, U.S.A.

1. Benzyl ligand-functionalized gold nanoparticles (BenzNPs)

Reported procedures in literature were followed to synthesize Benzyl ligand and BenzNPs (1-4). Briefly, Brust-Schiffrin two-phase synthesis method was used to synthesize pentanethiol-protected AuNPs (C5-AuNPs) with core diameter ca. 2 nm (5,6). To obtain benzyl ligand functionalized NPs (BenzNPs), Murray place-exchange method (7) was followed. C5-AuNPs (20 mg) and thiol ligand (60 mg) were dissolved in a mixture of dry DCM (6 mL) and methanol (2 mL) and stirred under N2 atmosphere for 3 days at room temperature. After place exchange, solvents were removed under reduced pressure and the resulting precipitate was washed with hexane (15 mL) three times, hexane:DCM (v/v, 1/1, 16 mL) three times, and DCM (15 mL) twice. After that, the precipitate was dissolved in distilled water (~ 8 mL) and dialyzed for three days (membrane molecular weight cut-off =10,000, volume of the dialysis bucket is ~5 L) to remove excess ligands, pentanethiol, acetic acid, and other salts present in the nanoparticle solution. After dialysis, the particle was lyophilized to yield a solid product. The particles were then re-dispersed in deionized water. ^{1}H NMR-spectra in D2O showed substantial broadening of the proton peaks with no sign of free ligands. The presence of ligands on AuNPs was also confirmed by mass spectroscopy.

2. Characterization of BenzNP

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). MALDI-MS has been performed to characterize the surface ligand on the BenzNPs (8). A saturated α-Cyano-4-hydroxycinnamic acid (α-CHCA) stock solution was prepared in 70% acetonitrile, 30% H_2O, and 0.1% trifluoroacetic acid. An equal volume of 2 μM BenzNP solution was added to the matrix stock solution. 2.5 L of this mixture was applied to the sample carrier, and then the MALDI-MS analysis was performed on a Bruker Autoflex III mass spectrometer. The molecular ions of Benzyl ligand were detected at m/z 498, and the disulfide ion formed between the benzyl ligand and the original pentanethiol was also detected at m/z 600.

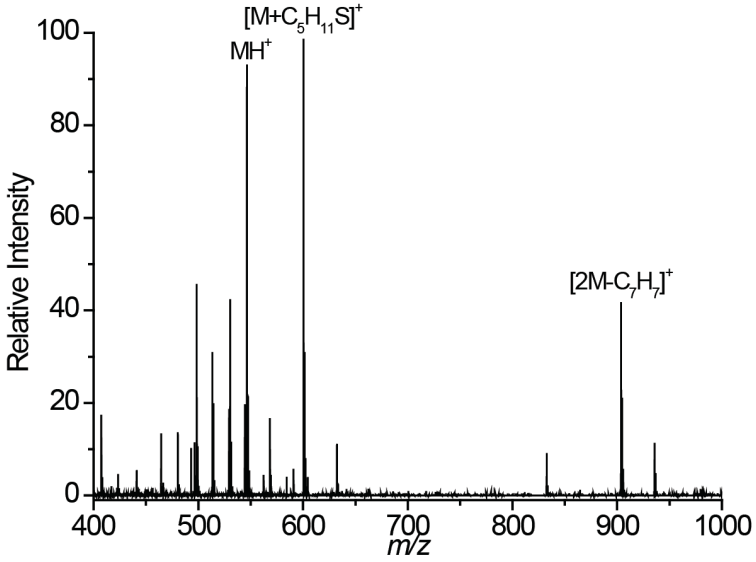

Figure S1. MALDI-MS spectrum of BenzNPs.

^{1}H NMR spectrum

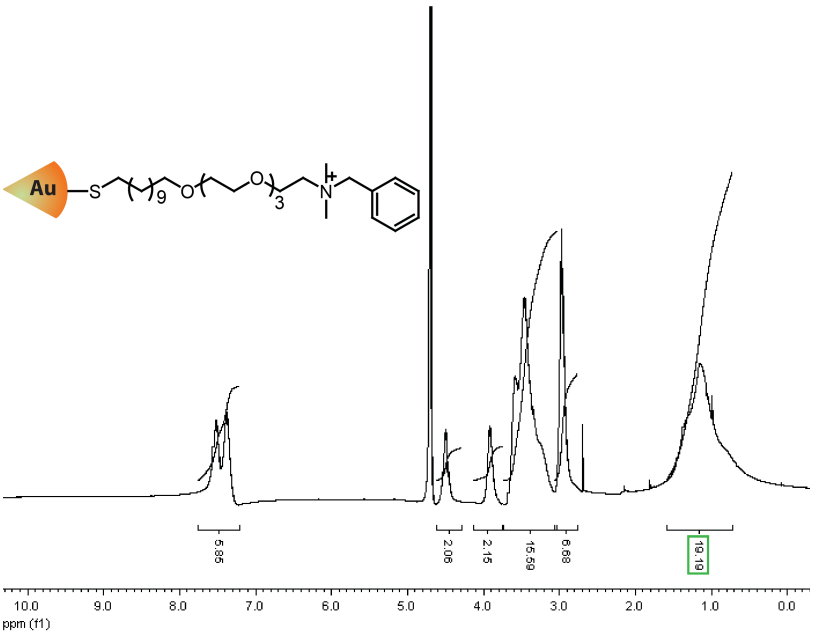

Figure S2. ^{1}H NMR spectrum (400 MHz) of BenzNP in D_2O (D, 99.8%).

Hydrodynamic size and zeta potential of BenzNP

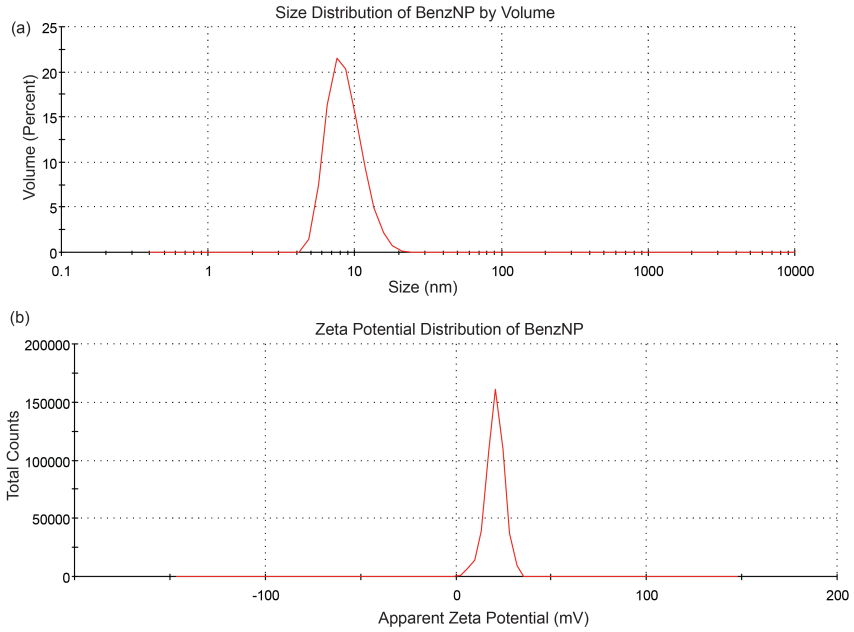

Figure S3. Characterization of the BenzNPs. (a) Size (diameter) of BenzNP is measured by DLS in 5 mM PB. DLS measurement shows that size of the BenzNP is 8.7 ± 2.6 from three independent replicates. (b) Zeta potential of BenzNP was measured by DLS in 5 mM PB. The overall charge of BenzNP is measured as 20.8 ± 6.5 mV from three independent replicates.

3. Fluorescence bar graphs

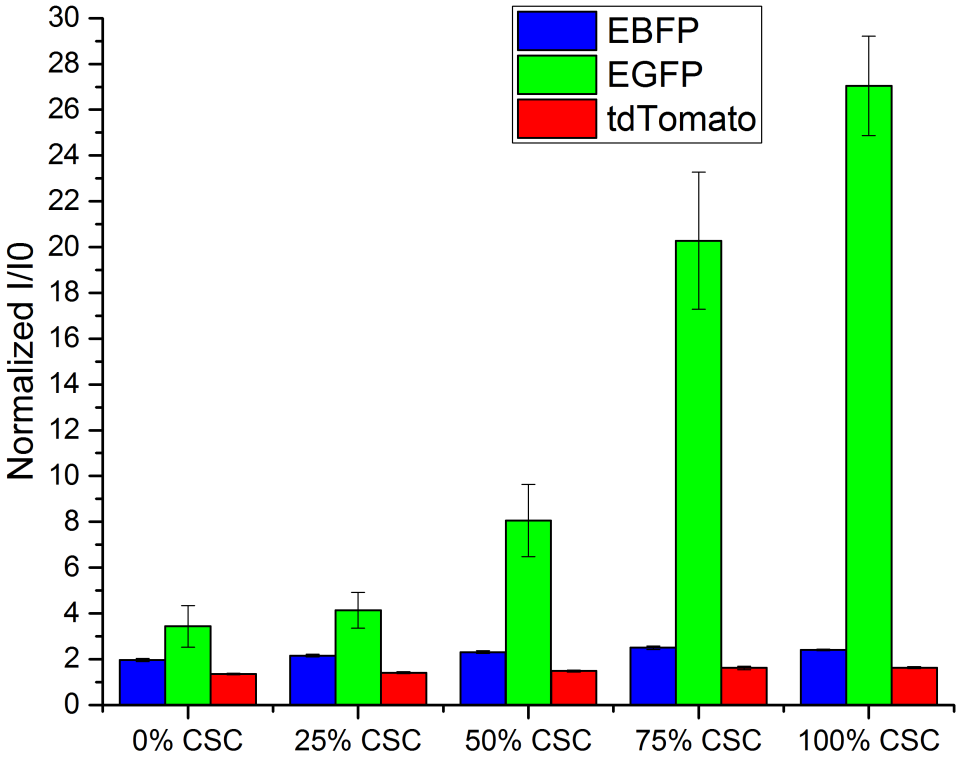

Figure S4. Fluorescent responses of CSC and non-CSC mixture from the engineered S1 system using BenzNP-FP nanosensor. Normalized fluorescent intensities against sensor only (I/I_0) were obtained with BenzNP-FP nanosensor against different ratios of the CSCs and non-CSCs from S1 cells (n = 8).

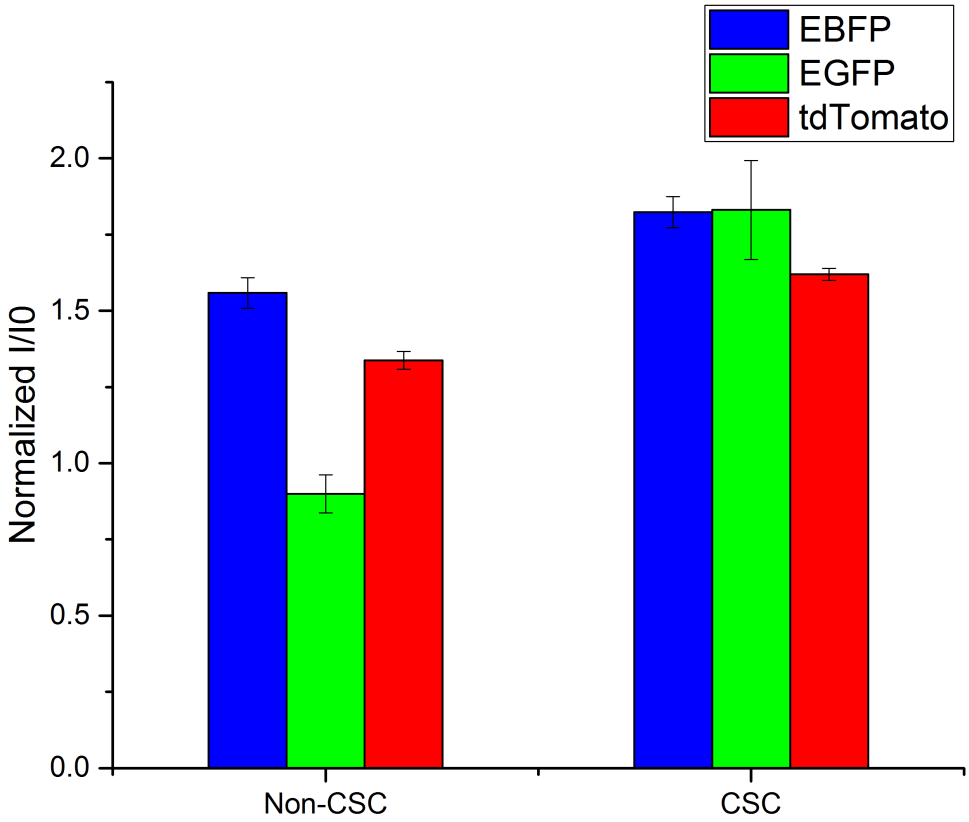

Figure S5. Fluorescent responses of non-CSC and CSC group isolated from the patient-derived xenografts using BenzNP-FP nanosensor. Normalized fluorescent intensities against sensor only (I/I_0) were obtained (n = 8).

4. Normalized fluorescence responses and LDA output data

Fluorescence intensity of each channel was normalized against sensor only (I/I_0) and used for LDA analysis. In LDA, measured fluorescent responses were transformed to generate canonical discriminants that best maximize the distance between classes relative to the variation within the classes. Each data input in the analysis is then reduced to a single score using the discriminants and plotted in a new space, called LDA solution space. At the end of analysis, LDA separates classes of objects into distinct groups, which are shown in the canonical plot. More details of the analysis could be found in the literature (9).

Table S1. Normalized fluorescence responses and LDA output for CSC and non-CSC mixture from the MCF-10A ER-Src system. Score (1) and score (2) correspond to Fig. 2 in the main text.

Sample name	I/I_0			LDA output		
	EBFP	EGFP	tdTomato	Score (1)	Score (2)	Score (3)
0% CSC	1.817865	2.571088	1.543197	6.035518	0.815081	-1.116892

0% CSC	1.877957	3.597334	1.576064	4.875273	1.723847	-0.591839
0% CSC	1.853246	3.187297	1.627913	4.645387	-1.272182	-1.037974
0% CSC	1.893202	3.119675	1.603582	5.493806	0.879185	-0.248070
0% CSC	1.937026	3.486669	1.660884	4.872620	0.044842	0.162784
0% CSC	1.937280	2.923962	1.652830	5.810772	0.144158	0.405962
0% CSC	1.891855	2.577243	1.659455	5.707937	-1.630782	-0.260530
0% CSC	1.951027	2.886411	1.686218	5.684131	-0.800036	0.538873
25% CSC	1.815277	3.455055	1.608589	3.979070	-1.526568	-1.718327
25% CSC	1.879651	3.264141	1.655565	4.561543	-1.558069	-0.712575
25% CSC	1.951227	3.755545	1.635729	4.905596	1.593743	0.390670
25% CSC	1.971403	4.419992	1.693186	3.543490	0.164896	0.286667
25% CSC	1.916584	3.435186	1.673797	4.563710	-1.110001	-0.210669
25% CSC	2.030952	4.048030	1.763659	4.092707	-1.030763	1.197015
25% CSC	2.057281	4.972833	1.794968	2.688262	-1.116267	1.190464
25% CSC	2.128230	6.493353	1.797440	1.234702	1.544462	1.820022
50% CSC	1.934003	7.297329	1.678799	-1.110540	0.805232	-1.379783
50% CSC	1.998554	7.126077	1.725433	-0.551983	0.800948	-0.377194
50% CSC	2.010963	7.587152	1.736926	-1.217612	0.902178	-0.378474
50% CSC	2.007442	6.653028	1.732638	0.195721	0.583459	-0.073290
50% CSC	2.037541	6.693281	1.763118	0.182639	0.279864	0.318988
50% CSC	2.097015	7.492101	1.822550	-0.921556	0.008526	0.827073
50% CSC	2.117329	8.760666	1.876474	-3.157455	-1.025044	0.511043
50% CSC	2.142849	8.762654	1.865720	-2.734278	0.175270	0.982079
75% CSC	1.991160	8.282214	1.721532	-2.348245	1.212182	-0.923264
75% CSC	1.968108	9.248002	1.723889	-4.114830	0.822621	-1.686247
75% CSC	2.057312	8.594579	1.761937	-2.435092	1.705772	-0.053289
75% CSC	1.950658	8.308941	1.725859	-2.930746	-0.168442	-1.638059
75% CSC	2.130475	7.321027	1.827359	-0.303591	0.750681	1.444512
75% CSC	2.115066	9.305848	1.833830	-3.559438	0.845743	0.415294
75% CSC	2.148668	8.292792	1.853676	-1.826109	0.640594	1.299018
75% CSC	2.149504	7.879390	1.870226	-1.365711	-0.169966	1.411146
100% CSC	1.982202	10.215631	1.758995	-5.773980	0.234332	-1.930358
100% CSC	2.036678	9.964893	1.801886	-5.179397	0.044779	-1.056620

100% CSC	2.160530	10.619534	1.876826	-5.440018	1.026159	0.548438
100% CSC	2.077684	9.621597	1.857655	-4.745843	-1.105569	-0.422079
100% CSC	2.167733	10.309326	1.913087	-5.265232	-0.342715	0.662224
100% CSC	2.151665	10.722718	1.943721	-6.409637	-1.886629	0.127727
100% CSC	2.169244	10.260940	1.955830	-5.624084	-2.035223	0.558507
100% CSC	2.187815	10.751325	1.948257	-6.057505	-0.970295	0.717029

Table S2. Normalized fluorescence responses and LDA output for CSC, non-CSC, and de-CSC groups from the S1 system. Score (1), and score (2) correspond to Fig. 4a in the main text.

Sample name	I/I_0			LDA output	
	EBFP	EGFP	tdTomato	Score (1)	Score (2)
Non-CSC	2.12363	2.89472	1.44655	4.92438	0.22478
Non-CSC	2.12705	3.03057	1.42358	5.07213	0.58905
Non-CSC	2.01352	2.18519	1.40060	5.43533	-1.20956
Non-CSC	2.23732	4.05547	1.51209	3.70867	1.43119
Non-CSC	2.14893	3.02788	1.55671	3.65838	-0.78265
Non-CSC	2.16583	2.44019	1.46911	5.47335	0.76117
Non-CSC	2.18345	2.46521	1.50439	5.15827	0.62180
Non-CSC	2.24590	3.65487	1.49484	4.42022	1.85637
CSC	2.44737	14.37155	1.90014	-10.74314	-0.52859
CSC	2.48668	15.77496	1.81416	-10.98282	1.27598
CSC	2.38455	14.10910	1.81482	-9.90788	-0.56266
CSC	2.35731	13.49822	1.83447	-9.66795	-1.30262
CSC	2.53000	12.57462	1.93750	-8.58412	0.68002
CSC	2.46262	12.26964	1.89071	-8.19300	0.04364
CSC	2.57623	13.52768	1.97121	-9.69438	1.04024
CSC	2.53310	15.38998	2.03550	-12.83783	-0.78499
De-CSC	2.22196	3.54636	1.51302	4.14647	1.16237
De-CSC	2.09805	2.52856	1.39384	5.76450	0.47157
De-CSC	2.01011	2.49832	1.40419	5.02097	-1.34601
De-CSC	1.93891	2.00929	1.34375	5.74932	-1.85816
De-CSC	2.12216	2.05062	1.43482	5.98629	0.41631

De-CSC	2.07704	1.86363	1.46871	5.45543	-0.88791
De-CSC	2.12702	1.96935	1.49530	5.39476	-0.29813
De-CSC	2.10250	1.73142	1.51493	5.24263	-1.01323

Table S3. Normalized fluorescence responses and LDA output for CSC and non-CSC mixture from the S1 system. Score (1), score (2) and score (3) correspond to Fig. 4b in the main text.

Sample name	**I/I_0**			**LDA output**		
	EBFP	**EGFP**	**tdTomato**	**Score (1)**	**Score (2)**	**Score (3)**
0% CSC	1.927829	2.280813	1.382625	-6.254980	-3.264813	-1.160907
0% CSC	2.015818	3.814123	1.405952	-5.1804	-2.04917	-0.8603
0% CSC	1.894175	3.006865	1.325646	-5.92037	-3.7247	0.16316
0% CSC	1.976269	3.100487	1.337187	-5.55608	-2.14282	0.491984
0% CSC	1.934918	3.012791	1.353656	-5.8052	-3.09893	-0.25028
0% CSC	1.99233	3.705166	1.356275	-5.23015	-2.12417	0.219186
0% CSC	1.884859	3.209683	1.317727	-5.84293	-3.91262	0.329
0% CSC	2.070647	5.336487	1.376016	-4.14016	-1.108	0.555414
25% CSC	2.211207	5.213129	1.436473	-3.73965	1.367433	0.031804
25% CSC	2.225454	4.909753	1.497276	-3.95448	1.280287	-1.46671
25% CSC	2.167979	3.963315	1.396236	-4.45346	1.127627	0.541268
25% CSC	2.227722	4.703979	1.413107	-3.87409	2.02816	0.67106
25% CSC	2.099328	4.208876	1.371428	-4.56737	-0.16873	0.705794
25% CSC	2.056839	2.876339	1.350324	-5.35808	-0.51681	0.717962
25% CSC	2.119984	3.542361	1.433016	-4.93618	-0.02909	-0.82317
25% CSC	2.132413	3.672445	1.398216	-4.74901	0.45805	0.180288
50% CSC	2.294439	8.244357	1.50777	-2.04453	1.703754	-0.69143
50% CSC	2.351831	9.255812	1.564005	-1.42206	2.179603	-1.54318
50% CSC	2.28559	5.912176	1.45326	-3.11968	2.582553	0.265733
50% CSC	2.33778	8.906347	1.468003	-1.45622	2.720587	0.753353
50% CSC	2.307862	10.63157	1.426227	-0.64452	1.945414	1.875474
50% CSC	2.343071	7.369459	1.467797	-2.19159	3.256174	0.552753
50% CSC	2.178894	6.185306	1.448612	-3.41968	0.338295	-0.36353
50% CSC	2.362341	7.898417	1.522657	-1.96292	3.088279	-0.62368
75% CSC	2.61412	17.98517	1.700067	3.695498	4.145011	-1.68606

75% CSC	2.548706	24.12151	1.643194	6.565398	1.533054	0.262522
75% CSC	2.51909	22.35307	1.613422	5.631098	1.637716	0.522504
75% CSC	2.499048	22.1025	1.556952	5.539775	1.725043	1.778787
75% CSC	2.370701	14.40664	1.463052	1.403079	1.918667	2.001623
75% CSC	2.438336	20.17128	1.662395	4.118778	0.197447	-1.6779
75% CSC	2.498401	19.97131	1.641533	4.312975	1.653326	-0.73023
75% CSC	2.522878	21.0975	1.674161	4.903347	1.598116	-1.20385
100% CSC	2.436318	24.53319	1.623508	6.340826	-0.75203	-0.0021
100% CSC	2.438839	22.94076	1.59714	5.620131	-0.05808	0.437348
100% CSC	2.361345	27.03728	1.590033	7.332001	-2.73678	0.697516
100% CSC	2.3543	28.79826	1.566125	8.219954	-3.186	1.537888
100% CSC	2.41953	27.79243	1.643489	7.837091	-2.15244	-0.11811
100% CSC	2.393368	27.89005	1.648199	7.76666	-2.75597	-0.4172
100% CSC	2.413503	28.65629	1.680892	8.161339	-2.80191	-0.98212
100% CSC	2.458712	28.71857	1.68304	8.37586	-1.90154	-0.69167

Table S4. Normalized fluorescence responses and LDA output for CSC and non-CSC groups isolated from the PDX samples. Score (1) and score (2) correspond to Fig. 5c in the main text. Different sample sizes (n = 8 for non-CSC and n = 6 for CSC) are due to the scarcity of isolated CSCs from PDX samples.

Sample name	I/I_0			LDA output	
	EBFP	EGFP	tdTomato	Score (1)	Score (2)
Non-CSC	1.636849	0.92968	1.359301	5.291044	0.487366
Non-CSC	1.619845	0.999395	1.344677	4.623209	0.389741
Non-CSC	1.587549	0.978464	1.325545	4.807328	0.071918
Non-CSC	1.577691	0.943424	1.347194	4.184204	-0.39109
Non-CSC	1.506717	0.863826	1.288606	5.584475	-0.87542
Non-CSC	1.477156	0.806908	1.291248	5.448338	-1.43668
Non-CSC	1.567925	0.857513	1.321228	5.899989	-0.23062
Non-CSC	1.528191	0.848237	1.362723	3.546835	-1.47735
Non-CSC	1.553219	0.857804	1.365521	3.922945	-1.0732
Non-CSC	1.527163	0.908992	1.365989	2.701327	-1.53118
CSC	1.866588	1.747248	1.635223	-8.87459	0.984683
CSC	1.853689	2.050124	1.62956	-12.4664	0.869076

CSC	1.86197	1.951221	1.63178	-11.2114	0.973441
CSC	1.774863	1.664615	1.612887	-9.25993	-0.34286
CSC	1.76247	1.741178	1.587472	-9.50174	-0.21629

5. **Jackknifed analysis of the fluorescence responses**

Jackknifed classification or leave-one-out cross validation is performed to assess the quality and reliability of the sensor array. The procedure begins by leaving out one sample at a time and uses the rest of the data as a training set to generate the linear discriminant function. The function is then used to categorize the excluded sample into the correct cluster. This process is repeated for all data points. The results from the Jackknifed analysis are listed below (Table S5 – S6).

Table S5. Percentage of accurate classification among different CSC mixtures from the MCF-10A ER-Src system using Jackknifed analysis. The results show an overall 100% correct classification.

	0% CSC	50% CSC	100% CSC	% correct
0% CSC	8	0	0	100
50% CSC	0	8	0	100
100% CSC	0	0	8	100
Total	8	8	8	100

Table S6. Percentage of accurate classification among different CSC mixtures from the S1 cell model using Jackknifed analysis. The results show no overlap between the analyzed groups, indicating that all cases are classified into the correct groups.

	0% CSC	50% CSC	100% CSC	% correct
0% CSC	8	0	0	100
50% CSC	0	8	0	100
100% CSC	0	0	8	100
Total	8	8	8	100

6. Hierarchical clustering analysis

To ensure the accuracy of the data clustering, we employed another method, Hierarchical clustering analysis (HCA), on our data set. HCA of the average data set was performed using the *hclust* function of the stats package of R assuming a complete linkage method (10). *hclust* begins with each case serving as its own cluster. During each processing step, the two most similar cases or clusters are joined. The step-wise analysis iterates until all cases fall into a single cluster. The results from the HCA approach are listed below (Figure S6-S8).

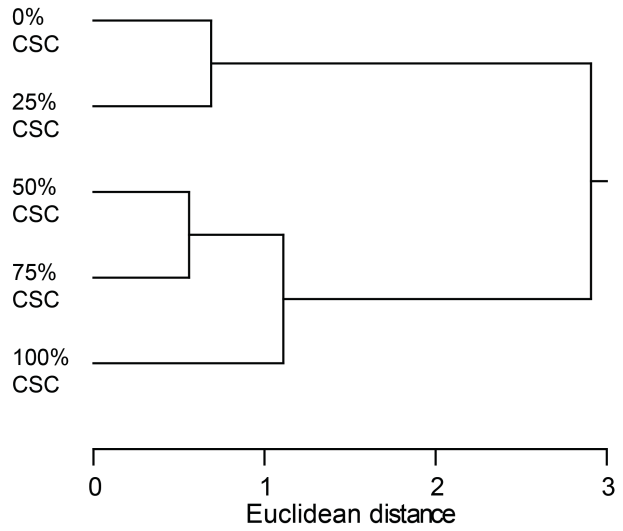

Figure S6. The dendrogram derived from unsupervised HCA using average fluorescence responses in five different MCF-10A ER-Src mixture samples (n=8). Analysis was carried out using average linkage method, where the distance metric is Euclidean distance. A similar trend is observed as the one in the LDA plot from Figure 2 in the main text.

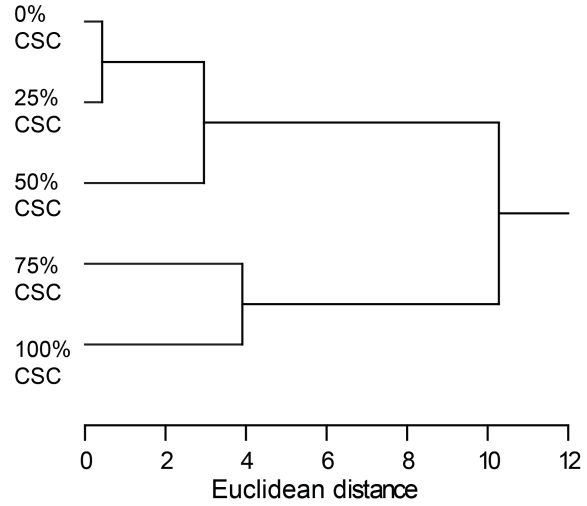

Figure S7. The dendrogram derived from unsupervised HCA using average fluorescence responses in five different S1 mixture samples (n=8). Analysis was carried out using average linkage method, where the distance metric is Euclidean distance. A similar trend is observed as the one in the LDA plot from Figure 4b in the main text.

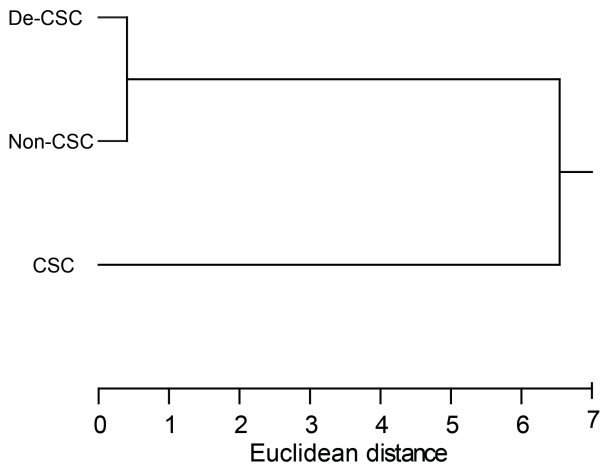

Figure S8. The dendrogram derived from unsupervised HCA using average fluorescence responses in three S1 cell lines (n=8). Analysis was carried out using average linkage method, where the distance metric is Euclidean distance. A similar trend is observed as the one in the LDA plot from Figure 4a in the main text.

7. Identification of unknown samples

Unknown or blinded samples were identified first by converting their normalized fluorescence responses to canonical scores using the discriminant functions established from the reference set. Then, Mahalanobis distance (11,12) of that case to the centroid of each training cluster in the LDA space was computed. Blinded cases were predicted to belong to the closest group, defined by the shortest Mahalanobis distance. The results from the unknown identification are listed below (Table S7 – S10).

Table S7. Unknown identification of three blinded CSC mixtures in MCF-10A ER-Src system from the training set (Figure 2 and Table S1). The results show an overall 96% correct unknown identification with only one case from 0% CSC misidentified as 50% CSC.

Unknown sample #	I/I_0			True ID	Identified as	Correct prediction
	EBFP	EGFP	tdTomato			
1	1.754519	3.423045	1.501593	0% CSC	0% CSC	yes
2	1.76912	4.862428	1.535601	0% CSC	50% CSC	no
3	1.848319	4.453536	1.547369	0% CSC	0% CSC	yes
4	1.878453	4.930182	1.619786	0% CSC	0% CSC	yes
5	1.913332	3.773694	1.602403	0% CSC	0% CSC	yes
6	1.944757	4.835558	1.616316	0% CSC	0% CSC	yes
7	2.001841	4.817133	1.658191	0% CSC	0% CSC	yes
8	1.98959	4.835699	1.676257	0% CSC	0% CSC	yes
9	1.810397	5.443651	1.571686	50% CSC	50% CSC	yes
10	1.819309	4.933177	1.541244	50% CSC	50% CSC	yes
11	1.90041	5.164846	1.590407	50% CSC	50% CSC	yes
12	1.891537	6.44706	1.677087	50% CSC	50% CSC	yes
13	1.953976	6.660632	1.74388	50% CSC	50% CSC	yes

14	1.939982	6.803571	1.728521	50% CSC	50% CSC	yes
15	1.96229	6.756974	1.757715	50% CSC	50% CSC	yes
16	2.006735	6.426179	1.747898	50% CSC	50% CSC	yes
17	1.805842	11.94979	1.849768	100% CSC	100% CSC	yes
18	1.901518	11.80875	1.919809	100% CSC	100% CSC	yes
19	1.888191	10.17209	1.863975	100% CSC	100% CSC	yes
20	1.950687	12.00072	1.946979	100% CSC	100% CSC	yes
21	2.118642	11.57802	1.962532	100% CSC	100% CSC	yes
22	1.957026	11.80334	1.99206	100% CSC	100% CSC	yes
23	1.971589	11.67191	2.014595	100% CSC	100% CSC	yes
24	1.991975	11.60949	2.035304	100% CSC	100% CSC	yes

Table S8. Unknown identification of blinded CSCs, Non-CSCs and De-CSCs in S1 system from the training set (Figure 4a and Table S2). Due to the similar nature between Non-CSCs and De-CSCs, they are named as one group Non/De-CSC for unknown identification studies. The results show an overall 100% accurate unknown identification.

Unknown sample #	I/I_0			True ID	Identified as	Correct prediction
	EBFP	EGFP	tdTomato			
1	1.983416	2.448272	1.399256	Non/De-CSC	Non/De-CSC	yes
2	1.949518	2.514552	1.337073	Non/De-CSC	Non/De-CSC	yes
3	1.901489	2.164027	1.296893	Non/De-CSC	Non/De-CSC	yes
4	1.858121	1.761826	1.266455	Non/De-CSC	Non/De-CSC	yes
5	1.749786	2.117801	1.244752	Non/De-CSC	Non/De-CSC	yes
6	1.819328	2.050408	1.299469	Non/De-CSC	Non/De-CSC	yes
7	1.798689	2.011638	1.312134	Non/De-CSC	Non/De-CSC	yes
8	1.892374	1.921637	1.348044	Non/De-CSC	Non/De-CSC	yes
9	1.900544	1.544735	1.320328	Non/De-CSC	Non/De-CSC	yes
10	1.834241	1.480186	1.268752	Non/De-CSC	Non/De-CSC	yes
11	1.775655	1.232072	1.228482	Non/De-CSC	Non/De-CSC	yes
12	1.713579	1.409896	1.223249	Non/De-CSC	Non/De-CSC	yes
13	1.728734	1.688588	1.212232	Non/De-CSC	Non/De-CSC	yes
14	1.680008	1.203106	1.238495	Non/De-CSC	Non/De-CSC	yes
15	1.845183	1.523259	1.344059	Non/De-CSC	Non/De-CSC	yes
16	1.864388	1.679419	1.366456	Non/De-CSC	Non/De-CSC	yes
17	2.339016	17.37641	1.668717	CSC	CSC	yes
18	2.204321	10.25706	1.532161	CSC	CSC	yes
19	2.183013	11.51956	1.562858	CSC	CSC	yes
20	2.130075	9.966903	1.530473	CSC	CSC	yes
21	2.184406	10.11168	1.545874	CSC	CSC	yes
22	2.097712	8.286319	1.626988	CSC	CSC	yes

23	2.22251	8.583179	1.659878	CSC	CSC	yes
24	2.277516	13.40781	1.727029	CSC	CSC	yes

Table S9. Identification of three unknown CSC mixtures in S1 system from the training set (Figure 4b and Table S3). The results show an overall 100% correct unknown identification.

Unknown sample #	I/I_0					Correct prediction
	EBFP	EGFP	tdTomato	True ID	Identified as	
1	1.803439	1.562091	1.352891	0% CSC	0% CSC	yes
2	1.79796	1.503799	1.363111	0% CSC	0% CSC	yes
3	1.778555	1.421646	1.273746	0% CSC	0% CSC	yes
4	1.764604	1.506237	1.287965	0% CSC	0% CSC	yes
5	1.851735	1.399424	1.289832	0% CSC	0% CSC	yes
6	1.74576	1.334019	1.319181	0% CSC	0% CSC	yes
7	1.842817	1.58972	1.356863	0% CSC	0% CSC	yes
8	1.887588	1.710484	1.326241	0% CSC	0% CSC	yes
9	2.57129	10.09375	1.634484	50% CSC	50% CSC	yes
10	2.515968	10.89496	1.599549	50% CSC	50% CSC	yes
11	2.558933	9.360174	1.599581	50% CSC	50% CSC	yes
12	2.516355	8.433984	1.592995	50% CSC	50% CSC	yes
13	2.420688	8.567795	1.648032	50% CSC	50% CSC	yes
14	2.435266	10.51007	1.540313	50% CSC	50% CSC	yes
15	2.459633	10.61531	1.573778	50% CSC	50% CSC	yes
16	2.481606	10.32127	1.569332	50% CSC	50% CSC	yes
17	2.513698	21.0672	1.642772	100% CSC	100% CSC	yes
18	2.516317	20.6477	1.629466	100% CSC	100% CSC	yes
19	2.435207	21.76076	1.622545	100% CSC	100% CSC	yes
20	2.36229	22.52546	1.518622	100% CSC	100% CSC	yes
21	2.459088	22.3944	1.623563	100% CSC	100% CSC	yes
22	2.41154	21.15326	1.605775	100% CSC	100% CSC	yes
23	2.438052	22.59211	1.625857	100% CSC	100% CSC	yes
24	2.524149	21.07948	1.667968	100% CSC	100% CSC	yes

Table S10. Identification of unknown PDX samples from the training set (Figure 5c and Table S4) using BenzNP-FP nanosensor. The results show 100% accurate unknown identification.

Unknown sample #	I/I_0					Correct prediction
	EBFP	EGFP	tdTomato	True ID	Identified as	
1	1.566597	0.849864	1.321603	Non-CSC	Non-CSC	yes
2	1.551777	0.914188	1.290197	Non-CSC	Non-CSC	yes
3	1.537542	0.939296	1.2961	Non-CSC	Non-CSC	yes
4	1.488406	0.893524	1.287482	Non-CSC	Non-CSC	yes

5	1.566868	0.888473	1.300348	Non-CSC	Non-CSC	yes
6	1.613927	0.883811	1.323813	Non-CSC	Non-CSC	yes
7	1.703125	0.979484	1.374356	Non-CSC	Non-CSC	yes
8	1.667582	0.961879	1.375677	Non-CSC	Non-CSC	yes
9	1.784828	1.09145	1.406615	Non-CSC	Non-CSC	yes
10	1.738162	1.153031	1.358105	Non-CSC	Non-CSC	yes
11	1.724078	1.617118	1.566202	CSC	CSC	yes
12	1.792467	1.799601	1.626918	CSC	CSC	yes
13	1.710386	1.515836	1.649519	CSC	CSC	yes
14	1.846764	1.683434	1.700884	CSC	CSC	yes
15	1.812834	2.002943	1.553019	CSC	CSC	yes

8. Importance of each fluorescent channel

To further evaluate the importance of each FP channel in our sensor, identification of unknowns using individual FP were carried out. As shown in Figure S9-10, each channel contributes to the identification to certain extent. However, when combined all together, it reaches the highest identification accuracy.

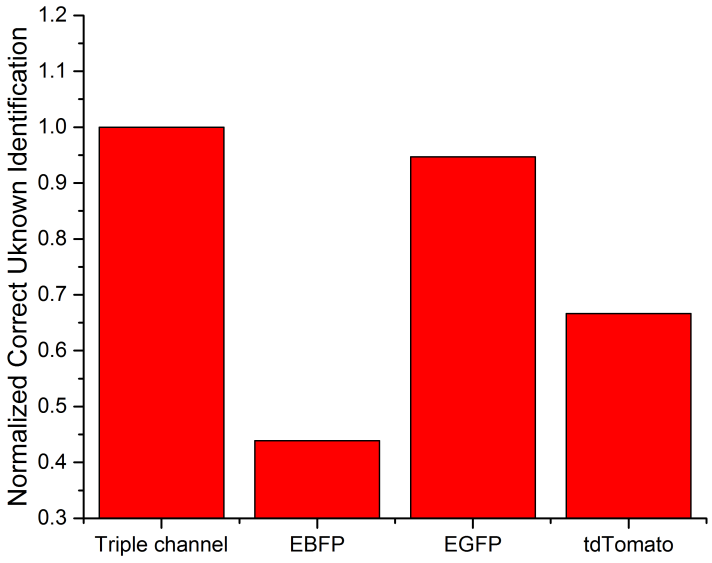

Figure S9. Correct unknown identification (CUI) of blinded cell mixture samples from the MCF-10A ER-Src system using either three FP or individual FP channel. Mixtures were in 0, 20, 50, 75 or 100% CSC. CUI was normalized to the percentage from the triple channel, which provides the highest identification accuracy.

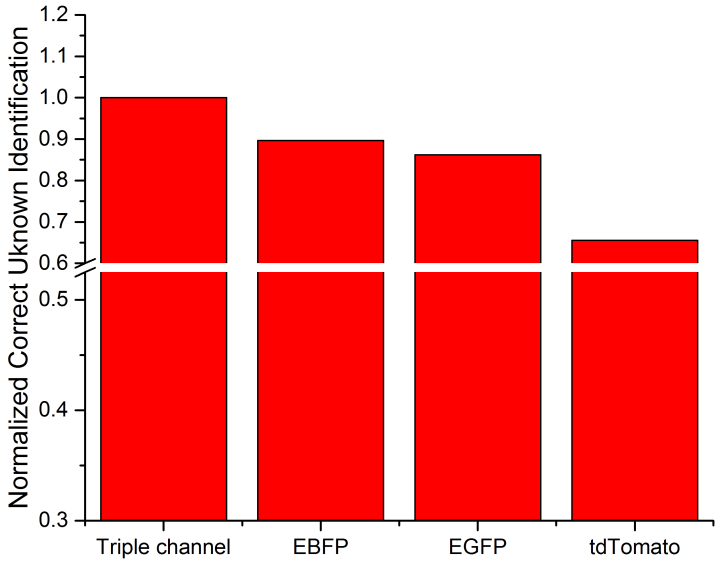

Figure S10. Correct unknown identification (CUI) of blinded cell mixture samples from the S1 system using either three FP or individual FP. Mixtures were in 0, 20, 50, 75 or 100% CSC. CUI was normalized to the percentage from the triple channel, which provides the highest identification accuracy.

References

1. Tonga, G. Y., Jeong, Y., Duncan, B., Mizuhara, T., Mout, R., Das, R., Kim, S. T., Yeh, Y-C., Yan, B., Hou, S., and Rotello, V. M. (2015) Supramolecular regulation of bioorthogonal catalysis in cells using nanoparticle-embedded transition metal catalysts. *Nat. Chem.***7**, 597–603
2. Tonga, G. Y., Mizuhara, T., Saha, K., Jiang, Z., Hou, S., Das, R., and Rotello, V. M. (2015) Binding studies of cucurbit[7]uril with gold nanoparticles bearing different surface functionalities. *Tetrahedron Lett.* **56**, 3653–3657
3. Kim, C., Tonga, G. Y., Yan, B., Kim, C. S., Kim, S. T., Park, M. H., Zhu, Z., Duncan, B., Creran, B., and Rotello, V. M. (2015) Regulating exocytosis of nanoparticles via host-guest chemistry. *Org. Biomol. Chem.* **13**, 2474-2479
4. Rana, S., Le, N. D. B., Mout, R., Saha, K., Tonga, G. Y., Bain, R. E. S., Miranda, O. R., Rotello, C. M., and Rotello, V. M. (2015) A multichannel nanosensor for instantaneous readout of cancer drug mechanisms. *Nat. Nanotechnol.* **10**, 65-69
5. Kanaras, A. G., Kamounah, F. S., Schaumburg, K., Kiely, C. J., and Brust, M. (2002) Thioalkylated tetraethylene glycol: a new ligand for water soluble monolayer protected gold clusters. *Chem. Commun.* ***20,*** 2294–2295
6. Brust, M., Walker, M., Bethell, D., Schiffrin, D. J., and Whyman, R. (1994) Synthesis of thiol derivatised gold nanoparticles in a two-phase Liquid–Liquid system. *J. Chem. Soc. Chem. Commun.* 801-802
7. Templeton, A. C., Wuelfing, W. P., and Murray, R. W. (2000) Monolayer-protected cluster molecules. *Acc. Chem. Res.* **33**, 27–36

8. Yan, B., Zhu, Z. J., Miranda, O. R., Chompoosor, A., Rotello, V. M., and Vachet, R. W. (2009) Laser desorption/ionization mass spectrometry analysis of monolayer-protected gold nanoparticles. *Anal. Bioanal. Chem.* ***396***, 1025–1035
9. Rana, S., Singla, A. K., Bajaj, A., Elci, S. G., Miranda, O. R., Mout, R., Yan, B., Jirik, F. R., and Rotello, V. M. (2012) Array-based sensing of metastatic cells and tissues using nanoparticle-fluorescent protein conjugates. *ACS Nano.* **6**, 8233–8240
10. R Development Core Team (2010). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.Rproject.org.
11. Mahalanobis, P. C. (1936) On the generalized distance in statistics. *Proc. Natl. Inst. Sci. India.* **2**, 49–55
12. Gnanadesikan, R., and Kettenring, J. R. (1972) Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics* **28**, 81–124