SUPPLEMENTARY MATERIAL

PLSR Modeling.

An MxN data matrix (**X**) was generated from the Bioplex, xMAP Novagen, quantitative RT-PCR, alkaline phosphatase activity, cell proliferation, and Alizarin Red S Staining assays. Each column corresponded to a different input or output signal: total or phosphorylated kinase signal from days 1, 2, 4, 7, and 14, xMAP Novagen total EGFR and HER2 signals from day 1, 4, 7, and 14, mRNA levels at day 1, 2, 4, and 7, alkaline phosphatase activity at day 7. Each row represented a different surface/medium/inhibitor condition, with a total of eight rows corresponding to control Exp vehicle, tEGF Exp vehicle, control OS vehicle, tEGF OS vehicle, control Exp AG1478, tEGF Exp AG1478, control OS AG1478, and tEGF OS AG1478. An MxP matrix (**Y**) was generated from the cellular output data, with rows corresponding to the same treatment conditions listed above and columns representing 21-day cell proliferation and 21 day matrix mineralization. All data were mean-centered and scaled to unit variance. SIMCA-P solves the PLSR problem with the nonlinear iterative partial least squares (NIPALS) algorithm[71]. Goodness of prediction was tested using a bootstrapping approach. Briefly, cross-validation is performed by omitting an observation from the model development and then using the model to predict **Y**-matrix values without the withheld observation. This procedure is repeated until every observation has been excluded exactly once. To identify the signaling metrics most important for the overall model, a weighted sum of squares (also known as variable importance for projection [VIP]) was calculated as previously described[30]. These steps were repeated during model simplification after reducing the number of input columns at the beginning step.

Alkaline phosphatase enzymatic activity assay

Surfaces with attached cells were moved to new 12-well dishes and rinsed with PBS. Cells were lysed in 200 μL of 0.2% NP-40 in 1 mmol/L $MgCl_2$, scraped with a rubber policeman, and collected. After one minute of sonication in a water bath, equal volumes of lysates were diluted by 10- and 100-fold with lysis buffer. Diluted sample lysate and lysis buffer were placed in a 96 well plate with an all lysis buffer sample used as a background control. A 1:1 solution of 2-Amino-2-methyl-1-propanol, 1.5 mol/L, pH

10.3 at 25°C (Sigma) and stock substrate solution of p-nitrophenyl phosphate, disodium (Sigma) was added to the samples and incubated for 30 minutes at 37°C; sodium hydroxide was added to stop the reaction. Absorbance at 405 nm was read using the SpectraMax plate-reader and background signal from the blank control was subtracted from the readings. Increasing concentrations of p-nitrophenol in sodium hydroxide was used to generate a standard curve to fit the data into Sigma units: that amount of enzyme which catalyses the liberation of 1 micromole p-nitrophenol per minute at 37°C. Total protein was then determined from lysates.

Alizarin Red S staining

At the end of 21 days, polymeric surfaces were transferred to a new 12-well plate, and rinsed with PBS. 10% buffered formalin (Sigma) was used to fix cells at room temperature for 1 hour and discarded. Surfaces were rinsed with DI water, and then 1% w/vol Alizarin Red S (Sigma) was added and incubated for 20 mins at room temperature, before it was rinsed away as well with DI water. Images were captured to visualize staining. Then, the plateswere washed four times with PBS before the addition of 0.1 mlof 10% (wt%) cetylpyridinium chloride for 30 min to releasethe remaining calcium-bound Alizarin Red S. The solution was collected,diluted at a ratio of 1:10 and read at OD_{570} on a SpectraMAX microplatereader (MolecularDevices).

Quantitative reverse transcription polymerase chain reaction

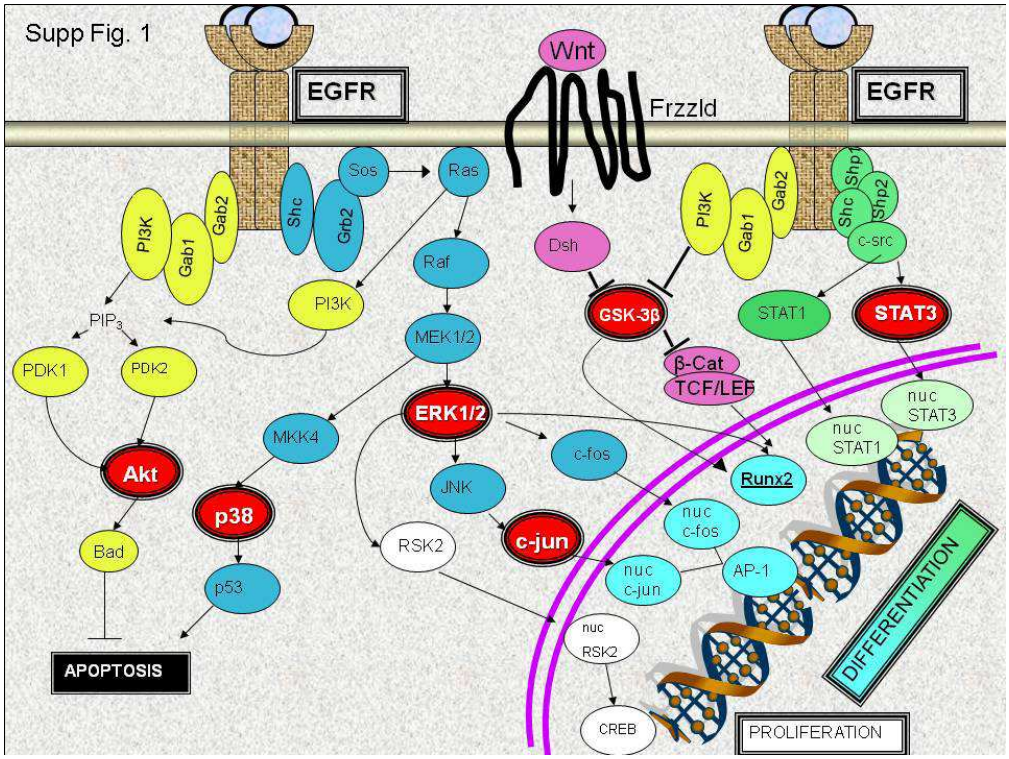
Control and tEGF surfaces with attached cells were moved to new 12 well dishes and rinsed with PBS prior to cell lysis. RNA was isolated using the Qiagen RNEasy kit according to manufacturer's instructions. 200 ng of total RNA was reverse transcribed with Superscript III (Invitrogen) to synthesize first-strand cDNA according to manufacturer's instructions. The cDNA was amplified with SYBR green PCR master mix (Qiagen) on an ABI7500 instrument (Applied Biosystems). Program set for 30 seconds denaturation at 95°C followed by 60 seconds of annealing and extension at 60°C (except for TGFα which was 64°C). Primer sequences are as follows EGF: forward 5'-CAGGGAAGATGACCACCACT-3' reverse 5'-CAGTTCCCACCACTTCAGGT-3' [72]; TGFα forward 5'-TGATACACTGCTGCCAGGTC -3'; reverse: 5'-

ATCTCTGGCAGTGCTGTCCT-3' [72] Alkaline phosphatase Forward 5'-CTTCAAACCGAGATACAAGCAC-3' reverse 5'- CTGGTAGTTGTTGTGAG CATAG-3' [73]. Glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) forward 5'-ACCACAGTCCATGCCATCAC-3' reverse 5'- TCCACCACCCTGTT GCTGTA- 3'. EGFR mRNA was detected using proprietary Taqman® probes (Applied Biosystems) with the same PCR program. Normalization and fold changes were calculated using the ΔΔ Ct method.

Supplementary Figure 1.EGFR binding to tethered EGF activates kinases that may be involved in differentiation programs. The crosstalk and complex nature of the signaling map make it difficult to study one particular pathway linking EGFR activation with osteogenic differentiation since many different mechanisms may be involved. Kinases in red are those measured in this study.

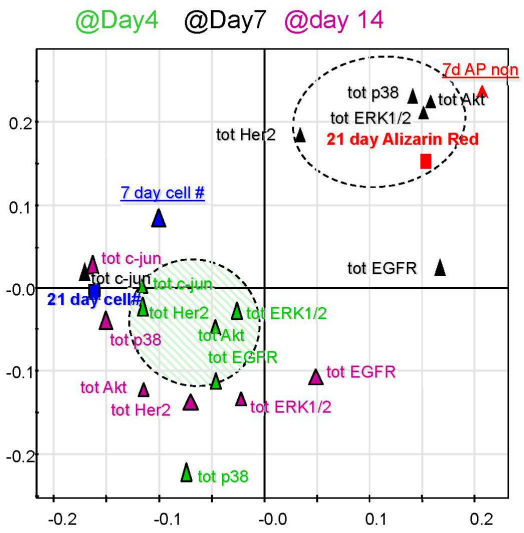
Supplementary Figure 2. Total kinase levels' contribution to an outcome differs during MSC differentiation. Loadings (weight coefficients) of the total kinase levels contribute differently to the 21-day cell response at different time points (Green – day 4; black – day 7; pink – day 14; red – osteogenic differentiation response; blue – cell proliferation response).

Supplementary Table 1. Variable importance of projection (VIP) identifies signals with the greatest contribution to the measured outputs and allow for the reduction of the data needed to successfully predict outcomes. After calculating the VIPs for all of the signals in the fully trained model, 12 out of the top 20 are from day 7, and the top 7 are from day 7. Training the model with only day 7 data accurately predicts (99%) 21-day differentiation responses of MSCs, but requires two lysates and four experimental assays. Further reduction of day 7 model data training to just one lysate and two assays retains 93% predictability of 21 day MSC differentiation responses with EGFR and Akt phosphorylation the most important.



254x190mm (96 x 96 DPI)

Supplementary Fig. 2



254x190mm (200 x 200 DPI)

Supplementary Table 1

Days 1-14: 4 Lysates-7 tests Predictability 99.8%			Day 7: 2 lysate - 4 tests Predictability 99%		Day 7: 1 lysate – 2 tests: Predictabiliity 93%	
Rank	Signal	VIP	Signal	VIP	Signal	VIP
1	tot Akt@day 7	1.69	tot Akt	1.22	p-EGFR	1.20
2	tot ERK1/2@day7	1.60	p-Akt	1.19	p-Akt	1.17
3	tot p38@day7	1.59	tot ERK1/2	1.16	p-ERK1/2	1.08
4	p-Akt@day7	1.56	tot p38	1.14	p-HSP27	1.08
5	p-EGFR@day7	1.54	p-ERK1/2	1.12	p-c-jun	1.02
6	p-ERK1/2@day7	1.47	p-HSP27	1.11	p-GSK-3a/b	0.80
7	p-HSP27@day7	1.44	p-EGFR	1.07	p-p38 MAPK	0.76
8	p-HSP27@day14	1.41	p-c-jun	1.02	p-STAT3	0.76
9	p-STAT3@day14	1.39	TotalEGFR	0.99		
10	AP@Day7	1.34	tot c-jun	0.91		
11	p-c-jun@day7	1.32	p-p38	0.76		
12	p-c-jun@day4	1.32	TotalHer2	0.74		
13	p-c-jun@day2	1.30	p-STAT3	0.72		
14	p-c-jun@day14	1.28	p-GSK-3a/b	0.51		
15	TotalEGFR@Day7	1.28				
16	p-HSP27@day4	1.24				
17	EGFR@Day7	1.22				
18	tot p38@day4	1.21				
19	AP@Day4	1.19				
20	tot c-jun@day 7	1.19				

254x190mm (200 x 200 DPI)