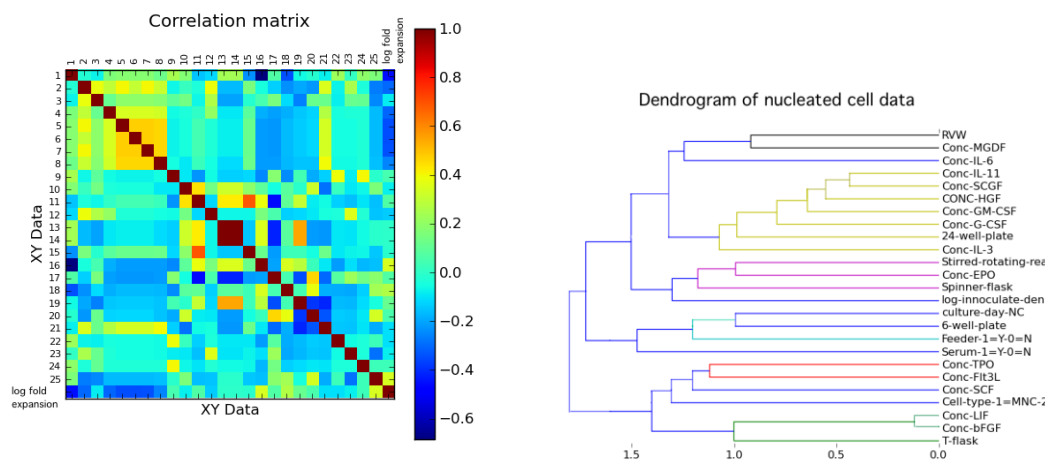


Supplementary information

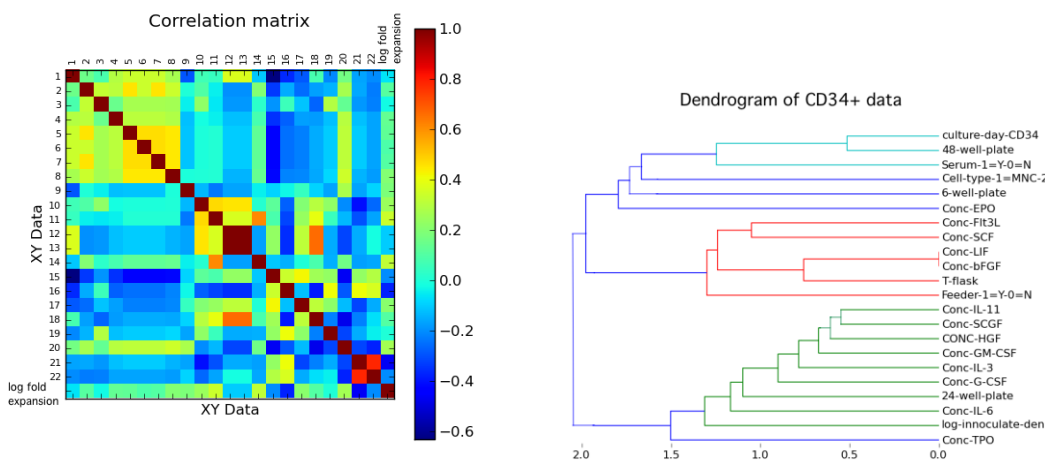
Figure S1 Correlation matrices and dendrograms for independent and dependent variables in models

Nucleated cells (NC)



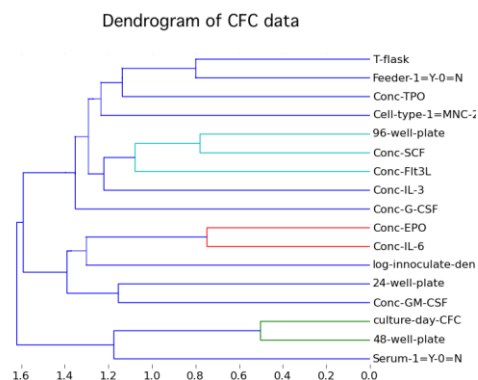
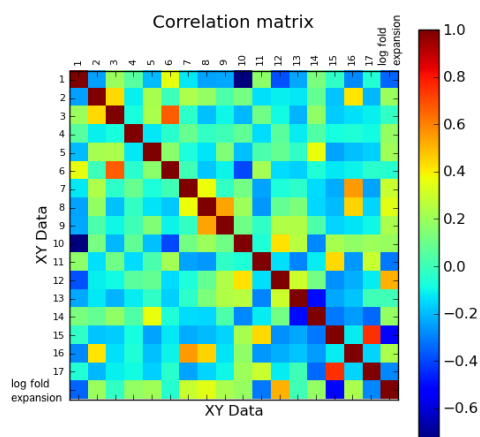
1 logID, 2 [IL-3], 3 [IL-6], 4 [G-CSF], 5 [SCGF], 6 [HGF], 7 [IL-11], 8 [GM-CSF], 9 [EPO], 10 [SCF], 11 [FL], 12 [MGDF], 13 [bFGF], 14 [LIF], 15 [TPO], 16 Cell-type-1=MNC-2=cd34+, 17 Serum-1=Y-0=N, 18 Feeder-1=Y-0=N, 19 T-flask, 20 6-well-plate, 21 24-well-plate, 22 Spinner-flask, 23 RVW, 24 Stirred-rotating-reactor, 25 culture-day

CD34+



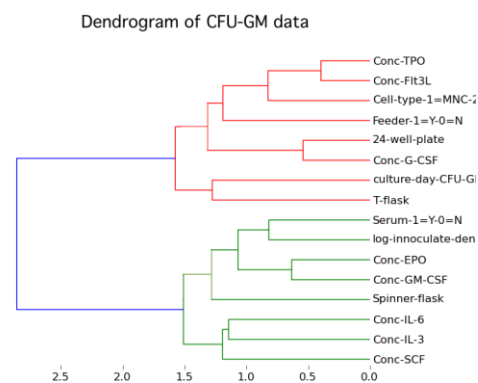
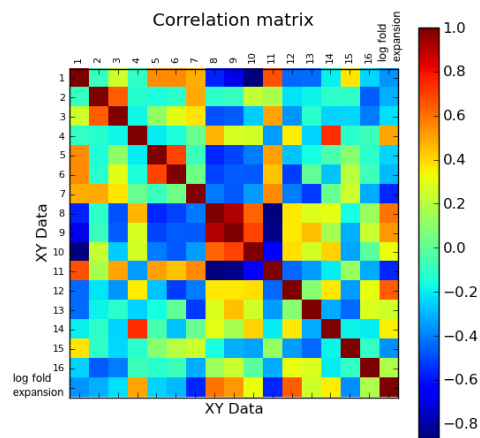
1 logID, 2 [IL-3], 3 [IL-6], 4 [G-CSF], 5 [SCGF], 6 [HGF], 7 [IL-11], 8 [GM-CSF], 9 [EPO], 10 [SCF], 11 [FL], 12 [bFGF], 13 [LIF], 14 [TPO], 15 Cell-type-1=MNC-2=cd34+, 16 Serum-1=Y-0=N, 17 Feeder-1=Y-0=N, 18 T-flask, 19 6-well-plate, 20 24-well-plate, 21 48-well-plate, 22 culture-day

CFC



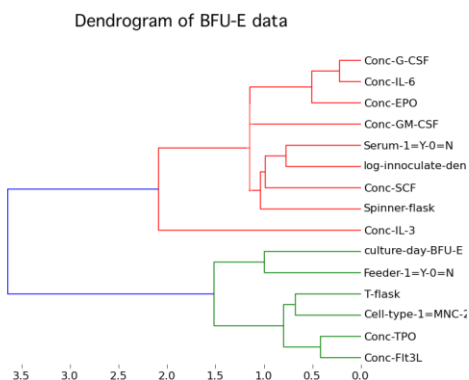
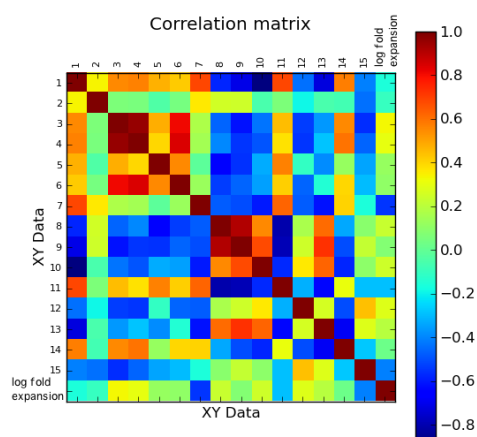
1 logID, 2 [IL-3], 3 [IL-6], 4 [G-CSF], 5 [GM-CSF], 6 [EPO], 7 [SCF], 8 [FL], 9 [TPO], 10 Cell-type-1=MNC-2=cd34+, 11 Serum-1=Y-0=N, 12 Feeder-1=Y-0=N, 13 T-flask, 14 24-well-plate, 15 48-well-plate, 16 96-well-plate, 17 culture-day

CFU-GM



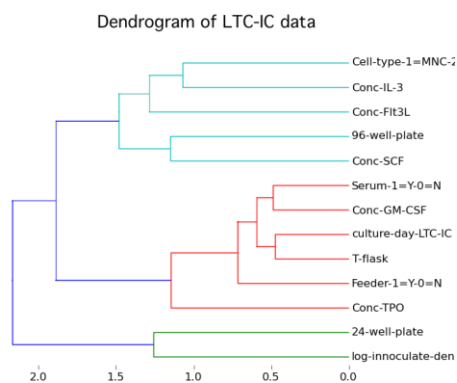
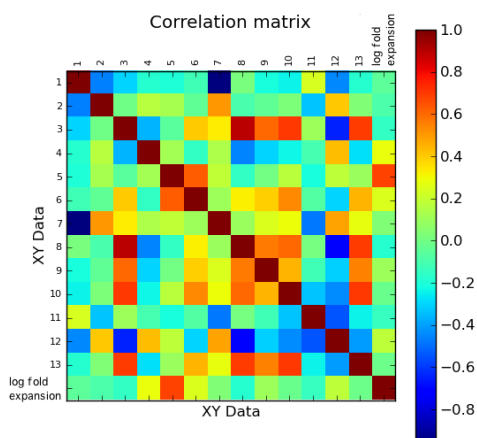
1 logID, 2 [IL-3], 3 [IL-6], 4 [G-CSF], 5 [GM-CSF], 6 [EPO], 7 [SCF], 8 [FL], 9 [TPO], 10 Cell-type-1=MNC-2=cd34+, 11 Serum-1=Y-0=N, 12 Feeder-1=Y-0=N, 13 T-flask, 14 24-well-plate, 15 Spinner-flask, 16 culture-day

BFU-E



1 logID, 2 [IL-3], 3 [IL-6], 4 [G-CSF], 5 [GM-CSF], 6 [EPO], 7 [SCF], 8 [FL], 9 [TPO], 10 Cell-type-1=MNC-2=cd34+, 11 Serum-1=Y-0=N, 12 Feeder-1=Y-0=N, 13 T-flask, 14 Spinner-flask, 15 culture-day

LTC-IC



1 log-D, 2 [IL-3], 3 [GM-CSF], 4 [SCF], 5 [FL], 6 [TPO], 7 Cell-type-1=MNC-2=cd34+, 8 Serum-1=Y-0=N, 9 Feeder-1=Y-0=N, 10 T-flask, 11 24-well-plate, 12 96-well-plate, 13 culture-day

Table S1. Performance of MLR-EM and BRANN bioreactor models (r^2 values) with varying numbers of hidden layer nodes in the BRANN models. The number of experiments for each cell type are shown in brackets.

Cell type	Data set	MLR	BRANN 1 node	BRANN 2 nodes	BRANN 3 nodes
NC (207)	all	0.62	0.59	0.81	0.83
	Training (80%)	0.59	0.52	0.75	0.78
	Test (20%)	0.47	0.60	0.67	0.66
CD34+ (134)	all	0.56	0.48	0.49	0.49
	Training (80%)	0.59	0.50	0.51	0.51
	Test (20%)	0.37	0.38	0.39	0.39
CFC (96)	all	0.66	0.61	0.61	0.71
	Training (80%)	0.70	0.61	0.60	0.62
	Test (20%)	0.47	0.57	0.56	0.55
CFU-GM (44)	all	0.79	0.69	0.70	0.70
	Training (80%)	0.79	0.67	0.68	0.68
	Test (20%)	0.92	0.89	0.89	0.89
BFU-E (27)	all	0.91	0.79	0.81	0.89
	Training (80%)	0.94	0.80	0.84	0.83
	Test (20%)	0.91	0.82	0.87	0.90
LTC-IC (24)	all	0.73	0.50	0.50	0.50
	Training (80%)	0.85	0.43	0.48	0.48
	Test (20%)	0.76	0.71	0.73	0.73

Table S2. Performance of MLR-EM and BRANNLP bioreactor models (r^2 values) with varying degrees of sparsity imposed. The number of parameters in each model is shown in brackets. BRANNLP with no imposed sparsity may still prune less relevant parameters resulting in fewer than the maximum number in the model.

Cell type	Sparsity index	MLR-EM	BRANN 1 node	BRANN 2 nodes	BRANN 3 nodes
NC	0	0.62 (25)	0.61 (21)	0.80 (25)	0.82 (25)
	1	0.60 (20)			
	2	0.57 (15)	0.61 (14)	0.71 (13)	0.84 (17)
	3	0.50 (9)			0.66 (12)
	4	0.41 (5)	0.51 (6)	0.61 (5)	
CD34+	0	0.56 (22)	0.50 (18)	0.50 (18)	0.51 (17)
	1				
	2	0.53 (16)			
	3		0.52 (11)	0.49 (9)	0.59 (10)
	4	0.44 (7)	0.44 (6)	0.48 (5)	0.48 (7)
CFC	0	0.66 (17)	0.62 (13)	0.62 (13)	0.62 (9)
	1	0.62 (12)	0.63 (8)	0.62 (10)	0.63 (8)
	2	0.54 (6)			
	3	0.443 (5)	0.63 (7)	0.63 (8)	0.63 (7)
	4	0.37 (3)			
CFU-GM	0	0.79 (16)	0.74 (11)	0.74 (10)	0.74 (9)
	1	0.75 (8)	0.75 (7)	0.75 (6)	0.75 (7)
	2	0.69 (5)			
	3	0.72 (4)	0.61 (3)	0.72 (4)	0.74 (4)
BFU-E	0	0.91 (15)	0.85 (10)	0.85 (10)	0.84 (11)
	1	0.78 (6)	0.86 (7)	0.84 (6)	
	2	0.73 (5)			0.84 (5)
	3	0.45 (3)	0.62 (4)	0.80 (5)	0.75 (4)
LTC-IC	0	0.73 (13)			
	1	0.70 (10)	0.68 (8)	0.68 (6)	0.58 (4)
	2	0.68 (6)	0.65 (4)	0.65 (4)	0.65 (4)
	3	0.58 (5)	0.48 (1)	0.48 (1)	0.48 (1)