

Supplementary Material

FGP Ttitle: Predicting anti-EGFR therapy in patient-derived xenografts using machine learning and integrative omics analysis

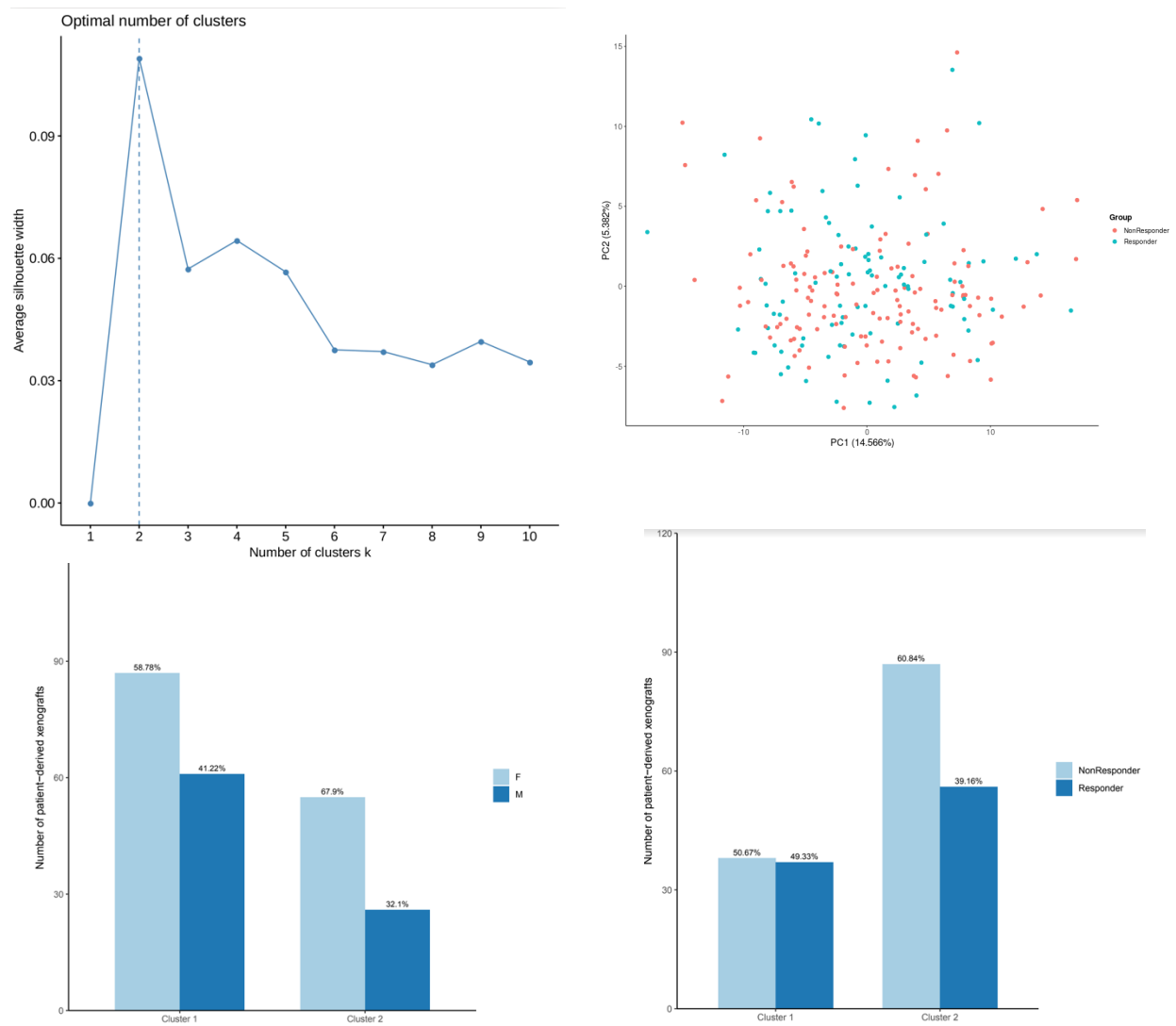
Author: Aina Montalbán

Scientific Director: Manuel Castro de Moura and David Piñeyro

Content

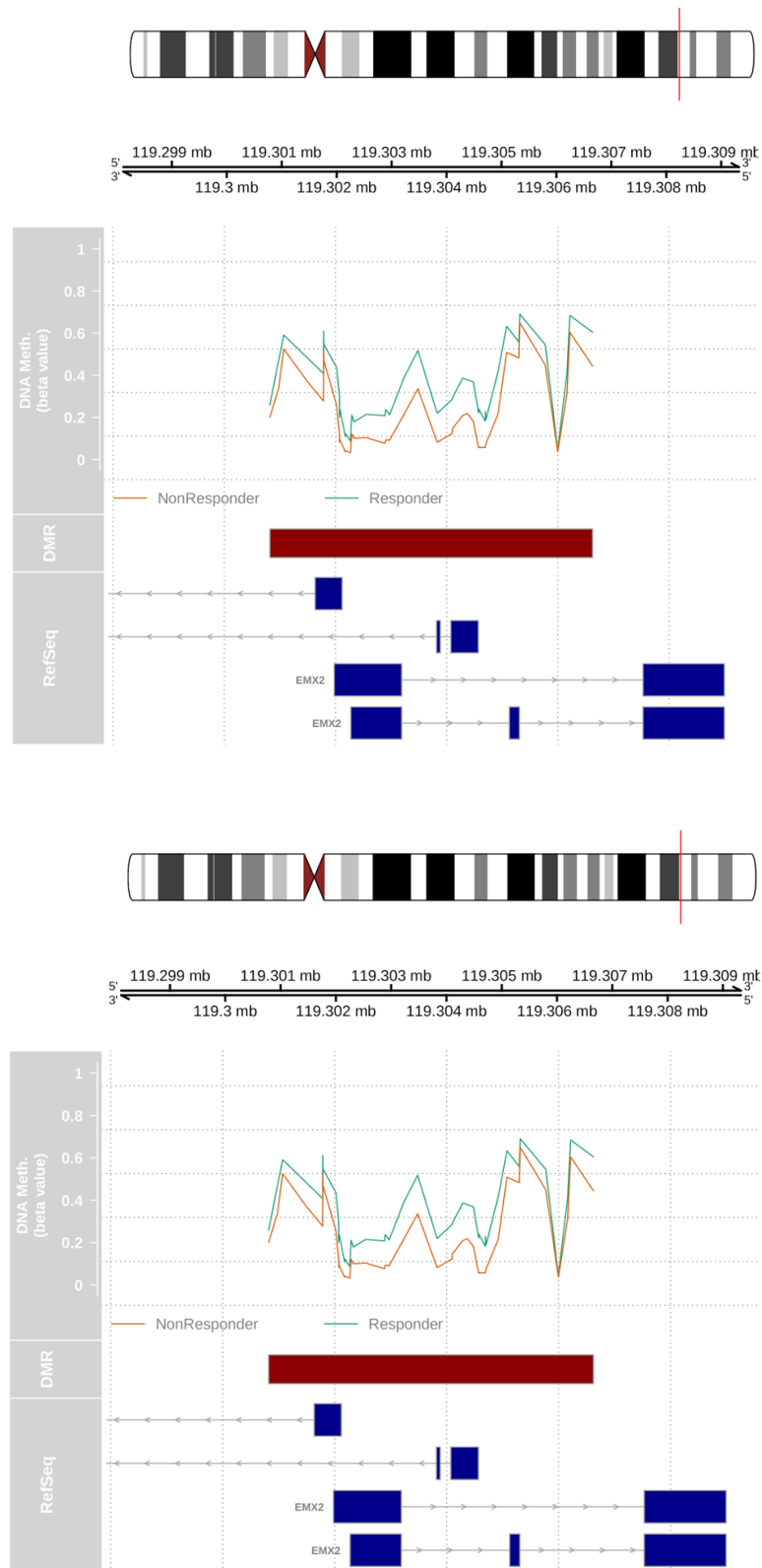
Supplementary Figures	2
Supplementary Figure 1.....	2
Supplementary Figure 2.....	3
Supplementary Tables	4
Supplementary Table 1	4
Supplementary Table 2	6
Supplementary Table 3	7
Supplementary Table 4	10
Supplementary Table 5	11
Supplementary Table 6	12
Supplementary Table 7.	14
Supplementary Table 8	15
Supplementary Table 9	16
Supplementary Table 11.....	17
Supplementary Table 12.	18

Supplementary Figures



Supplementary Figure 1. a. Optimal number of clusters using silhouette width b. Principal component analysis by drug response. c. Distribution of responders and nonresponders in the two clusters. d. Distribution of females and males in the two clusters.

Supplementary Figure 2. Top differentially methylated regions.



Supplementary Tables

Supplementary Table 1. Top 50 differentially methylated CpGs between Responders and NonResponders.

Name	chr	Gene Name	Gene Group	Diff. Meth
cg27616660	chr6	LAMA2	Body	0.32
cg03280299	chr7	ST7	Body	0.23
cg23184766	chr7			0.24
cg02767104	chr10	PARD3	Body	0.27
cg17848679	chr1	LINC01364	TSS200	0.28
cg10526370	chr8	TOX	Body	0.29
cg01427849	chr8	TOX	Body	0.31
cg20149359	chr1	HIVEP3	5'UTR	0.26
cg20443946	chr6	MAP7	Body;5'UTR	0.28
cg17661686	chr8	C8orf84	Body	0.28
cg13868872	chr3			0.31
cg16336556	chr2	LTBP1	Body	0.28
cg01704609	chr13	TNFRSF19	Body	0.26
cg02313670	chr2	MAP4K4	Body	0.26
cg06346126	chr19	MYH14	Body	0.25
cg05593985	chr8			0.29
cg08315399	chr14	TTLL5	Body	0.29
cg04206986	chr2			0.26
cg11664516	chr15	TLN2	ExonBnd;Body	0.26
cg10974746	chr4	MRPL1	Body	0.26
cg03227850	chr9			0.27
cg06586469	chr8			0.23
cg14315508	chr22	SFI1	Body	0.24
cg18621366	chr2			0.28
cg09164377	chr8	TOX	Body	0.28
cg13364311	chr2	EPHA4	Body	0.28

cg15139034	chr7			0.21
cg02324558	chr8	TOX	Body	0.26
cg21582831	chr3	NBEAL2	Body	0.26
cg10758454	chr8	TOX	Body	0.24
cg04774751	chr6			0.26
cg02819582	chr1	HIVEP3	5'UTR	0.26
cg00569313	chr8			0.30
cg00531786	chr3	PHLDB2	Body	0.26
cg08935242	chr3			0.23
cg12777178	chr11			0.30
cg02889177	chr3	FNDC3B	Body	0.22
cg22946082	chr6	SLC22A23	5'UTR;Body	0.21
cg03350316	chr9	LRRC8A	Body	0.24
cg15058210	chr2	HDAC4	Body	0.31
cg05878460	chr1	HIVEP3	5'UTR	0.24
cg03796321	chr15	AKAP13	Body	0.26
cg05532733	chr8			0.23
cg00068038	chr2	CRIM1	Body	0.23
cg26872527	chr4			0.25
cg26829990	chr12	GLTP	Body	0.23
cg02992224	chr11	HEPHL1	Body	0.22
cg18677548	chr13			0.25
cg01741041	chr7	ACTB	Body	0.22
cg17990814	chr2	SLC25A12	Body	0.22
cg23920047	chr8	TOX	Body	0.25
cg03709892	chr1	CAPZB	Body	0.21

Name: ;Chr.; Gene Name: Gene Group: Mean differential methylation:

Supplementary Table 2. Most significantly enriched GO terms for biological processes in hypermethylated genes in non-responders.

Term	Overlap	Adjusted P.value	Combined Score
extracellular matrix organization (GO:0030198)	132/229	0.0040	19.4289
calcium ion transport (GO:0006816)	75/118	0.0034	20.5910
neuron projection morphogenesis (GO:0048812)	97/163	0.0057	17.9886
inorganic cation transmembrane transport (GO:0098662)	85/140	0.0055	17.9704
ion transport (GO:0006811)	158/289	0.0055	15.8869
modulation of excitatory postsynaptic potential (GO:0098815)	21/25	0.0150	22.0391
positive regulation of excitatory postsynaptic potential (GO:2000463)	17/19	0.0160	23.0010
regulation of cation channel activity (GO:2001257)	47/71	0.0170	16.7087
peptidyl-tyrosine modification (GO:0018212)	36/51	0.0161	17.7103
monovalent inorganic cation transport (GO:0015672)	64/104	0.0179	15.1256
sensory perception of pain (GO:0019233)	20/24	0.0166	20.4513
metal ion transport (GO:0030001)	62/101	0.0219	14.5264
positive regulation of phosphatidylinositol 3-kinase signaling (GO:0014068)	36/53	0.0405	14.9417
cell morphogenesis involved in neuron differentiation (GO:0048667)	59/97	0.0409	13.2580
chemical synaptic transmission (GO:0007268)	152/289	0.0385	11.4529

regulation of glutamate receptor signaling pathway (GO:1900449)	24/32	0.0438	15.9841
modulation of chemical synaptic transmission (GO:0050804)	51/82	0.0434	13.1797
cardiac muscle cell development (GO:0055013)	21/27	0.0427	16.4037
divalent metal ion transport (GO:0070838)	54/88	0.0418	12.8955
calcium-dependent cell-cell adhesion via plasma membrane cell adhesion molecules (GO:0016339)	29/41	0.0412	14.8004
positive regulation of synaptic transmission (GO:0050806)	43/67	0.0414	13.3455
membrane depolarization (GO:0051899)	30/43	0.0430	14.3699
collagen fibril organization (GO:0030199)	22/29	0.0440	15.4996
sodium ion transmembrane transport (GO:0035725)	26/36	0.0436	14.6976

Supplementary Table 3. Most significantly enriched GO terms for biological processes in hypomethylated genes in non-responders.

Term	Overlap	Adjusted P.value	Combined Score
phosphorylation (GO:0016310)	187/386	0.0000	50.6861
protein phosphorylation (GO:0006468)	216/470	0.0000	44.5892
regulation of apoptotic process (GO:0042981)	325/815	0.0000	27.5193
vesicle-mediated transport (GO:0016192)	181/410	0.0000	30.3721
transmembrane receptor protein tyrosine kinase signaling pathway (GO:0007169)	175/396	0.0000	29.6115
vascular endothelial growth factor receptor signaling pathway (GO:0048010)	45/70	0.0000	41.4806
viral process (GO:0016032)	105/220	0.0000	27.5864

protein autophosphorylation (GO:0046777)	87/175	0.0000	28.0843
cellular protein modification process (GO:0006464)	379/1001	0.0000	21.1621
ephrin receptor signaling pathway (GO:0048013)	48/82	0.0000	31.6103
regulation of epidermal growth factor receptor signaling pathway (GO:0042058)	37/58	0.0001	33.7695
regulation of transcription from RNA polymerase II promoter (GO:0006357)	534/1478	0.0001	18.8074
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay (GO:0000184)	60/112	0.0001	27.5513
stress-activated protein kinase signaling cascade (GO:0031098)	42/70	0.0001	30.6207
protein dephosphorylation (GO:0006470)	65/125	0.0001	26.2166
peptidyl-serine phosphorylation (GO:0018105)	73/145	0.0001	25.3539
regulation of small GTPase mediated signal transduction (GO:0051056)	71/140	0.0001	25.5252
positive regulation of transcription, DNA- templated (GO:0045893)	414/1120	0.0001	18.5538
nuclear-transcribed mRNA catabolic process (GO:0000956)	84/174	0.0001	23.8178
regulation of mitotic cell cycle (GO:0007346)	80/165	0.0001	23.2629
regulation of cell cycle (GO:0051726)	98/215	0.0003	20.6888
positive regulation of gene expression (GO:0010628)	292/771	0.0004	16.7217
peptidyl-serine modification (GO:0018209)	80/170	0.0005	20.2385
positive regulation of transcription from RNA polymerase II promoter (GO:0045944)	316/848	0.0007	15.6215
epidermal growth factor receptor signaling pathway (GO:0007173)	34/57	0.0007	24.8176
viral gene expression (GO:0019080)	56/110	0.0008	21.0255
glycerophospholipid biosynthetic process (GO:0046474)	77/164	0.0008	19.3605
cotranslational protein targeting to membrane (GO:0006613)	49/93	0.0008	21.5310
regulation of translation (GO:0006417)	95/213	0.0009	18.0015

ubiquitin-dependent protein catabolic process (GO:0006511)	141/341	0.0010	16.4988
intracellular protein transport (GO:0006886)	143/347	0.0011	16.3555
SRP-dependent cotranslational protein targeting to membrane (GO:0006614)	47/89	0.0010	20.9497
cellular macromolecule biosynthetic process (GO:0034645)	150/367	0.0010	16.1808
protein targeting to ER (GO:0045047)	50/97	0.0012	20.0513
actin cytoskeleton reorganization (GO:0031532)	35/61	0.0013	22.1655
regulation of Ras protein signal transduction (GO:0046578)	47/90	0.0014	20.0321
regulation of cell-matrix adhesion (GO:0001952)	31/52	0.0013	22.8240
positive regulation of cell-substrate adhesion (GO:0010811)	31/52	0.0013	22.8240
regulation of programmed cell death (GO:0043067)	114/268	0.0013	16.2595
viral transcription (GO:0019083)	56/113	0.0014	18.7478
activation of protein kinase activity (GO:0032147)	101/233	0.0015	16.2904
regulation of epithelial cell migration (GO:0010632)	29/48	0.0016	22.5409
protein transport (GO:0015031)	134/326	0.0017	15.2141
Ras protein signal transduction (GO:0007265)	97/223	0.0017	16.0348
mitotic cell cycle phase transition (GO:0044772)	96/221	0.0020	15.7759
negative regulation of gene expression (GO:0010629)	234/618	0.0020	13.7348
endocytic recycling (GO:0032456)	19/27	0.0020	25.4091
gene expression (GO:0010467)	163/411	0.0021	14.2276
actin filament organization (GO:0007015)	58/120	0.0021	17.3126
negative regulation of epidermal growth factor receptor signaling pathway (GO:0042059)	25/40	0.0023	22.1847
ERBB signaling pathway (GO:0038127)	42/80	0.0024	18.5589
translation (GO:0006412)	99/232	0.0031	14.6710

positive regulation of protein kinase activity (GO:0045860)	77/172	0.0032	15.3162
negative regulation of ERBB signaling pathway (GO:1901185)	26/43	0.0034	20.5379
positive regulation of apoptotic process (GO:0043065)	125/307	0.0042	13.5070
DNA metabolic process (GO:0006259)	127/314	0.0052	13.1180
protein ubiquitination involved in ubiquitin-dependent protein catabolic process (GO:0042787)	37/70	0.0053	17.0919
protein polyubiquitination (GO:0000209)	116/283	0.0053	13.2235
regulation of autophagy (GO:0010506)	87/203	0.0066	13.4751
negative regulation of transcription from RNA polymerase II promoter (GO:0000122)	212/565	0.0070	11.7200

Supplementary Table 4. Top 10 differentially methylated regions between Responders and NonResponders.

chr	start	end	width	n°	Fisher	maxdiff	mean diff	overlapping genes
chr1 0	11930078 3	119306659	5877	42	1.00E-46	0.2010	0.1223	EMX2, EMX2OS
chr1 0	8092775	8097689	4915	57	1.01E-37	0.1907	0.1043	RP11-379F12.4, GATA3, GATA3-AS1, RP11-379F12.3
chr6	29570008	29575145	5138	28	1.56E-33	0.2104	0.1232	GABBR1, SNORA20
chr6	29585579	29593913	8335	38	1.91E-32	0.2191	0.1073	GABBR1, SNORA20
chr2 0	13199787	13202225	2439	30	2.69E-30	0.1940	0.0975	NA
chr1 5	76627594	76636200	8607	41	1.36E-29	0.2043	0.0956	ISL2, RP11-685G9.4, RP11-685G9.2
chr1 2	54653065	54655026	1962	10	3.27E-28	-0.2504	-0.1713	CBX5, snoMe28S-Am2634
chr1 2	22093960	22095330	1371	16	6.36E-28	0.2401	0.1748	ABCC9, snoMe28S-Am2634
chr9	23820979	23823125	2147	17	1.06E-27	0.2500	0.1646	SNORA30, ELAVL2

Supplementary Table 5. Most significantly enriched GO terms for biological processes in overlapping genes of differentially methylated regions.

Term	Overlap	P.value	Adjusted P.value	Combined Score
regulation of endothelial cell migration (GO:0010594)	41/69	0.0001	0.3023	16.0911
positive regulation of mesonephros development (GO:0061213)	13/15	0.0001	0.1979	22.8191
positive regulation of neuron differentiation (GO:0045666)	53/98	0.0002	0.3070	12.9690
renal system development (GO:0072001)	35/59	0.0002	0.2697	13.9646
negative regulation of cellular response to growth factor stimulus (GO:0090288)	33/55	0.0002	0.2411	13.9393
aorta morphogenesis (GO:0035909)	13/16	0.0003	0.2352	18.5199
urogenital system development (GO:0001655)	16/22	0.0005	0.3573	15.4191
regulation of developmental growth (GO:0048638)	20/30	0.0006	0.3845	13.7504
regulation of glutamate receptor signaling pathway (GO:1900449)	21/32	0.0006	0.3437	13.5253

Supplementary Table 6. Top differentially expressed genes between responders and non-responders.

ID	geneSymbol	logFC	adj.P.Val
ILMN_1781198	PPP1R3D	-0.4615	0.0000
ILMN_1657679	VAV3	-0.8731	0.0000
ILMN_2283196	ZSWIM3	-0.4263	0.0000
ILMN_1755303	ZNF217	-0.4809	0.0000
ILMN_1676759	DDX27	-0.5020	0.0000
ILMN_3251550	PHLDA1	0.7092	0.0000
ILMN_2049727	OSER1	-0.4420	0.0000
ILMN_2080080	MAP7D2	-1.0228	0.0000
ILMN_2386179	ZMYND8	-0.4985	0.0000
ILMN_2415421	SLC30A2	-0.9375	0.0000
ILMN_2391400	PITX2	-0.5988	0.0000
ILMN_3245659	ERI1	0.4777	0.0000
ILMN_1696295	LY6G6D	-1.6491	0.0000
ILMN_1687978	PHLDA1	1.0423	0.0000
ILMN_1808707	FSCN1	1.2862	0.0000
ILMN_1787705	ATP6V1B2	0.5411	0.0000
ILMN_1815656	SERINC3	-0.4922	0.0000
ILMN_1721128	TOMM34	-0.5015	0.0000
ILMN_1736796	RB1CC1	-0.4835	0.0000
ILMN_2198239	HGD	0.6389	0.0000
ILMN_1652407	ZMYND8	-0.4831	0.0000
ILMN_1731948	PCK1	-0.6154	0.0000
ILMN_1706275	C8orf33	-0.5973	0.0000
ILMN_1679600	ACOT8	-0.3983	0.0000
ILMN_1691436	BLVRA	0.8466	0.0000
ILMN_1837428	RAB27B	0.8438	0.0000
ILMN_3243461	RPRD1B	-0.3491	0.0000
ILMN_1812312	NDUFS4	0.4027	0.0000

ILMN_1788961	PPP2R2A	0.5876	0.0000
ILMN_1755047	LRRC2	-0.4444	0.0000
ILMN_2349124	OSBPL2	-0.4931	0.0000
ILMN_1703487	LMO4	0.7529	0.0000
ILMN_1693428	TNNC2	-0.7606	0.0000
ILMN_1808245	SBSPO	1.2371	0.0000
ILMN_1799887	CTSE	1.1467	0.0000
ILMN_1768282	SNX21	-0.3918	0.0000
ILMN_1702279	KIF3B	-0.3779	0.0000
ILMN_1796847	PITX2	-0.3488	0.0000
ILMN_2364131	TTPAL	-0.4070	0.0000
ILMN_1776076	POFUT1	-0.5029	0.0000
ILMN_1730998	TSPAN6	-0.6144	0.0000
ILMN_1766652	MOCS3	-0.3266	0.0000
ILMN_1868866	NA	0.9011	0.0000
ILMN_2169383	REG4	1.6888	0.0000
ILMN_2387224	CTSE	0.9389	0.0000
ILMN_2117171	LMO4	0.6915	0.0000
ILMN_1700204	ZMYND8	-0.2826	0.0000
ILMN_1667893	TNS3	-0.5527	0.0000
ILMN_1814247	TCFL5	-0.5852	0.0000
ILMN_1790309	PINX1	0.3424	0.0000
ILMN_1686116	THBS1	-0.8192	0.0000
ILMN_1778226	EXTL3	0.3806	0.0000
ILMN_1744949	RHOBTB3	1.1092	0.0000
ILMN_1738657	SATB2	-0.8121	0.0000
ILMN_1719286	CTSA	-0.5824	0.0000
ILMN_1768469	TCN1	1.2468	0.0000
ILMN_1747020	NA	-0.3908	0.0000
ILMN_1722489	TFF1	1.3288	0.0000

Supplementary Table 7. Most significantly enriched GO terms for biological processes in upregulated genes in non-responders.

Term	Overlap	Adjusted P.value	Combined Score
respiratory electron transport chain (GO:0022904)	18/94	0.0012	63.4802
mitochondrial ATP synthesis coupled electron transport (GO:0042775)	16/85	0.0036	55.1445
negative regulation of peptidase activity (GO:0010466)	13/62	0.0066	56.7988
mitochondrial electron transport, NADH to ubiquinone (GO:0006120)	11/46	0.0071	62.8838
positive regulation of epithelial cell proliferation (GO:0050679)	17/107	0.0076	40.7605
negative regulation of viral genome replication (GO:0045071)	11/50	0.0112	53.7345
negative regulation of viral life cycle (GO:1903901)	12/61	0.0132	46.6949
neutrophil degranulation (GO:0043312)	43/479	0.0151	20.7885
neutrophil activation involved in immune response (GO:0002283)	43/483	0.0163	20.2332
regulation of protein phosphorylation (GO:0001932)	28/261	0.0149	24.3499
neutrophil mediated immunity (GO:0002446)	43/487	0.0162	19.6932
cellular protein modification process (GO:0006464)	74/1001	0.0158	16.3919
regulation of cell proliferation (GO:0042127)	58/740	0.0212	16.7392
regulation of endopeptidase activity (GO:0052548)	12/68	0.0205	37.5355
retinal ganglion cell axon guidance (GO:0031290)	4/6	0.0211	140.4245
negative regulation of apoptotic process (GO:0043066)	42/485	0.0214	18.0944
positive regulation of cell proliferation (GO:0008284)	38/424	0.0216	18.5896
negative regulation of endopeptidase activity (GO:0010951)	13/83	0.0287	31.3227
regulation of viral genome replication (GO:0045069)	11/63	0.0339	34.0808
cellular protein metabolic process (GO:0044267)	41/484	0.0336	16.4566

mitochondrial respiratory chain complex assembly (GO:0033108)	14/97	0.0334	27.9047
NADH dehydrogenase complex assembly (GO:0010257)	11/64	0.0339	33.0001
mitochondrial respiratory chain complex I biogenesis (GO:0097031)	11/64	0.0324	33.0001
mitochondrial respiratory chain complex I assembly (GO:0032981)	11/64	0.0310	33.0001
positive regulation of protein phosphorylation (GO:0001934)	36/412	0.0372	16.3568
response to cytokine (GO:0034097)	17/138	0.0401	22.7488
positive regulation of interleukin-2 production (GO:0032743)	6/20	0.0392	55.3153
regulation of apoptotic process (GO:0042981)	60/815	0.0411	13.4374
cellular respiration (GO:0045333)	10/57	0.0429	31.7292
regulation of endothelial cell proliferation (GO:0001936)	12/79	0.0424	27.3985

Supplementary Table 8. Most significantly enriched GO terms for biological processes in downregulated genes in non-responders.

Term	Overlap	Adjusted P.value	Combined Score
ERBB signaling pathway (GO:0038127)	16/80	0.0097	52.42
ERBB2 signaling pathway (GO:0038128)	10/39	0.0420	56.21
epidermal growth factor receptor signaling pathway (GO:0007173)	12/57	0.0359	45.11

Supplementary Table 9. Most significantly enriched GO terms for biological processes in DMP-DEG pairs.

Term	Overlap	P.value	Adjusted P.value	Combined Score
ubiquitin-like protein ligase binding (GO:0044389)	27/297	0.0022	1.0000	11.0594
sphingolipid transporter activity (GO:0046624)	3/6	0.0023	1.0000	60.6365
ubiquitin protein ligase binding (GO:0031625)	26/284	0.0024	0.9214	10.9889
transcriptional repressor activity, RNA polymerase II core promoter proximal region sequence-specific binding (GO:0001078)	12/97	0.0034	0.9719	14.0098
ubiquitin protein ligase activity (GO:0061630)	19/192	0.0038	0.8722	10.9803
N6-methyladenosine-containing RNA binding (GO:1990247)	3/7	0.0038	0.7293	47.5254
adrenergic receptor binding (GO:0031690)	4/15	0.0055	0.9113	27.5707
secondary active transmembrane transporter activity (GO:0015291)	9/66	0.0056	0.7985	14.0947
intronic transcription regulatory region sequence-specific DNA binding (GO:0001161)	3/8	0.0059	0.7489	38.3601

Supplementary Table 10. Correlation table between DMP-DEGs.

ID	cpg_names	Gene Symbol	r	Pval	Relation	Diff Meth	logFC Exp.
ILMN_1714444	cg03356288	KLF12	0.793	2.39E-37	hyper-upregulated	0.231	0.401
ILMN_1684461	cg14377278	CADPS2	0.652	1.26E-21	hyper-upregulated	0.204	0.345
ILMN_1686555	cg17546878	FYN	0.633	4.06E-20	hyper-upregulated	0.225	0.671
ILMN_1686555	cg03244438	FYN	0.630	7.09E-20	hyper-upregulated	0.259	0.671
ILMN_1714444	cg14043104	KLF12	0.628	1.05E-19	hyper-upregulated	0.260	0.401

ILMN_1686555	cg25529557	FYN	0.621	3.34E-19	hyper-upregulated	0.238	0.671
ILMN_1789400	cg15691862	"FOXD2	-0.625	1.76E-19	hyper-downregulated	0.260	-0.717
ILMN_1789400	cg14566738	FOXD2	-0.663	1.51E-22	hyper-downregulated	0.236	-0.717
ILMN_1724658	cg26685590	BNIP3	-0.796	9.46E-38	hyper-downregulated	0.208	-0.832
ILMN_1724658	"cg07288471"	BNIP3	-0.796	8.95E-38	hyper-downregulated	0.208	-0.832
ILMN_1786429	cg16071219	LPAR6	-0.798	4.15E-38	hypo-upregulated	0.206	0.581
ILMN_1786429	cg03646329	LPAR6	-0.818	1.88E-41	hypo-upregulated	0.213	0.581
ILMN_1813641	cg27064692	TOR4A	-0.831	8.43E-44	hypo-upregulated	0.204	0.479
ILMN_1813641	cg09899914	TOR4A	-0.831	6.60E-44	hypo-upregulated	0.205	0.479

Supplementary Table 11. Subset of DNA methylation features selected by Random Forest.

Name	Importance	UCSC_RefGene_Name	UCSC_RefGene_Group	Relation to Island
cg02767104	6.622	PARD3	Body	OpenSea
cg27616660	5.958	LAMA2	Body	OpenSea
cg14663517	5.007	C5orf66	5'UTR	OpenSea
cg00136547	4.944	DST	Body;TSS200	OpenSea
cg23184766	4.800			OpenSea
cg17848679	4.4912	LINC01364	TSS200	OpenSea
cg03922648	4.476	KLRC1	TSS1500;TSS200	OpenSea
cg15139034	4.4147			OpenSea
cg04206986	4.318			OpenSea
cg21582831	4.295	NBEAL2	Body	OpenSea
cg15058210	4.189	HDAC4	Body	Island
cg16336556	4.068	LTBP1	Body	OpenSea

cg07703756	3.752	RIMBP2	Body	OpenSea
cg07936950	3.676	DLX1;DLX1	3'UTR	Island
cg17385190	3.4963	SEC14L2	Body	OpenSea
cg18865824	3.439	NRF1	Body	OpenSea
cg15637573	3.434	NR2F2-AS1	Body	OpenSea
cg22674412	3.417	RPRML	3'UTR;1stExon	Island
cg03687070	3.366			OpenSea
cg04181261	3.0388	TMEM59L	TSS200	Island

Supplementary Table 12. Subset of gene expression features selected by Random Forest.

Probe_Id	Importance	Symbol	Definition
ILMN_1731948	5.5275	PCK1	Homo sapiens phosphoenolpyruvate carboxykinase 1 (soluble) (PCK1), mRNA.
ILMN_3251550	4.51484	PHLDA1	Homo sapiens pleckstrin homology-like domain, family A, member 1 (PHLDA1), mRNA.
ILMN_1686555	4.436789	FYN	Homo sapiens FYN oncogene related to SRC, FGR, YES (FYN), transcript variant 2, mRNA.
ILMN_1713496	4.32306	ST3GAL5	Homo sapiens ST3 beta-galactoside alpha-2,3-sialyltransferase 5 (ST3GAL5), transcript variant 2, mRNA.
ILMN_2337655	3.9759	WARS	Homo sapiens tryptophanyl-tRNA synthetase (WARS), transcript variant 1, mRNA.
ILMN_1767665	3.7663	LOC493869	Homo sapiens similar to RIKEN cDNA 2310016C16 (LOC493869), mRNA.
ILMN_1685415	3.5945	HBP1	Homo sapiens HMG-box transcription factor 1 (HBP1), mRNA.
ILMN_1719797	3.2969	MESDC2	Homo sapiens mesoderm development candidate 2 (MESDC2), mRNA.
ILMN_2277099	3.2658	YWHAB	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide (YWHAB), transcript variant 1, mRNA.
ILMN_1815719	3.2159	PLCG2	Homo sapiens phospholipase C, gamma 2 (phosphatidylinositol-specific) (PLCG2), mRNA.
ILMN_1702279	3.1514	KIF3B	Homo sapiens kinesin family member 3B (KIF3B), mRNA.
ILMN_2227573	3.1033	GSTO1	Homo sapiens glutathione S-transferase omega 1 (GSTO1), mRNA.
ILMN_2349124	3.0769	OSBPL2	Homo sapiens oxysterol binding protein-like 2 (OSBPL2), transcript variant 2, mRNA.
ILMN_1683250	2.8713	LOC440731	PREDICTED: Homo sapiens hypothetical LOC440731, transcript variant 2 (LOC440731), mRNA.
ILMN_1789244	2.8485	SOX8	Homo sapiens SRY (sex determining region Y)-box 8 (SOX8), mRNA.

ILMN_173218 7	2.7516	TMEM143	Homo sapiens transmembrane protein 143 (TMEM143), mRNA.
ILMN_167108 7	2.7103	C6orf222	Homo sapiens chromosome 6 open reading frame 222 (C6orf222), mRNA.
ILMN_174188 1	2.3501	C9orf72	Homo sapiens chromosome 9 open reading frame 72 (C9orf72), transcript variant 1, mRNA.
ILMN_327002 1	2.2997	LOC100128319	PREDICTED: Homo sapiens similar to putative protein STRF7 (LOC100128319), mRNA.
ILMN_228228 2	1.8882	MLH3	Homo sapiens mutL homolog 3 (E. coli) (MLH3), transcript variant 1, mRNA.