Data Analysis Approaches for Modern Cell Biology Technologies

Amir A. Handzel & Andy Liaw

Merck Research Laboratories

Non-Clinical Biostatistics Conference 2009

In Collaboration with…

North Wales, PA

Francesca Santini

Frank Stefano

Yaping Liu

Angela Jaramillo

Laszlo Kiss

Boston, MA

Alex Szewczak

Grace Ka Chan

Ilona Kariv

Cem Elbi

Julie Gorenstein

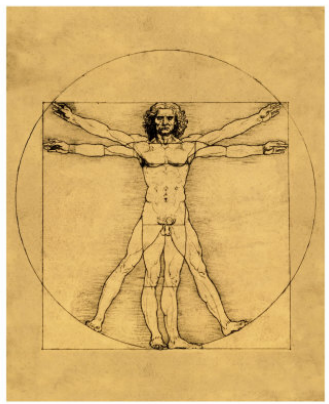
Rahway, NJ

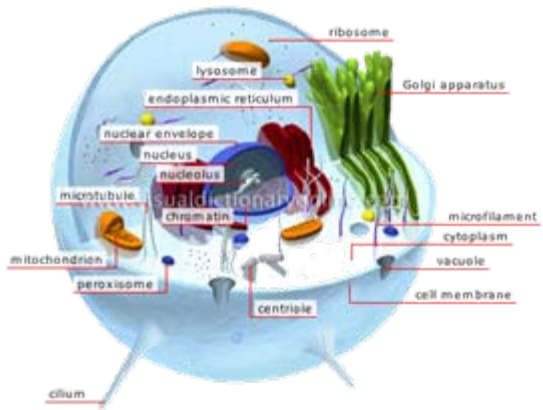
Anka G. Ehrhardt

Richard Wnek

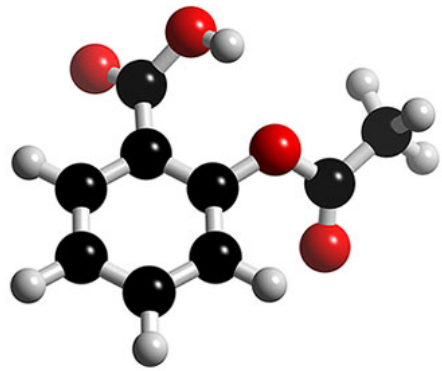
Paul Fischer

Cell Biology in Discovery and Early Development





Simplicity, through-put

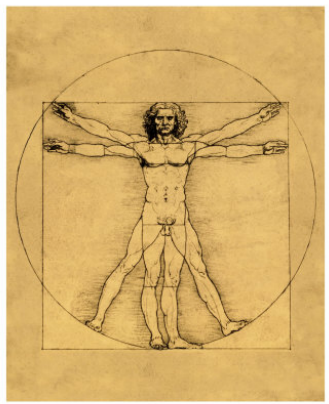


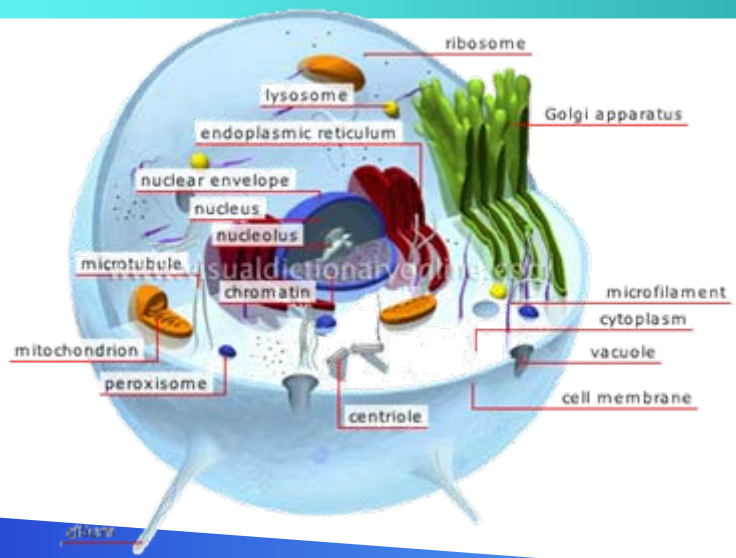
Human-like, complexity, price

MERCK

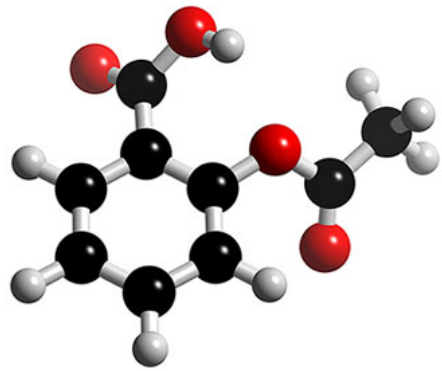
Cell Biology in Discovery and Early Development

- Perturbation (chemical, genomic, other) of living cells and detailed study of effect on phenotype





Simplicity, through-put

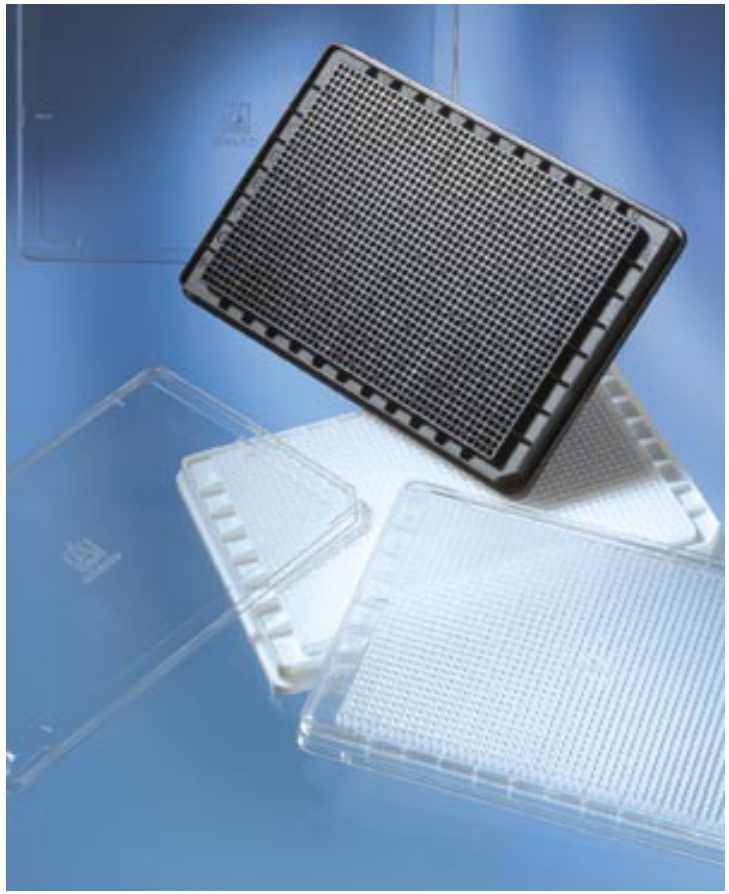


Human-like, complexity, price

MERCK

High Throughput Assays

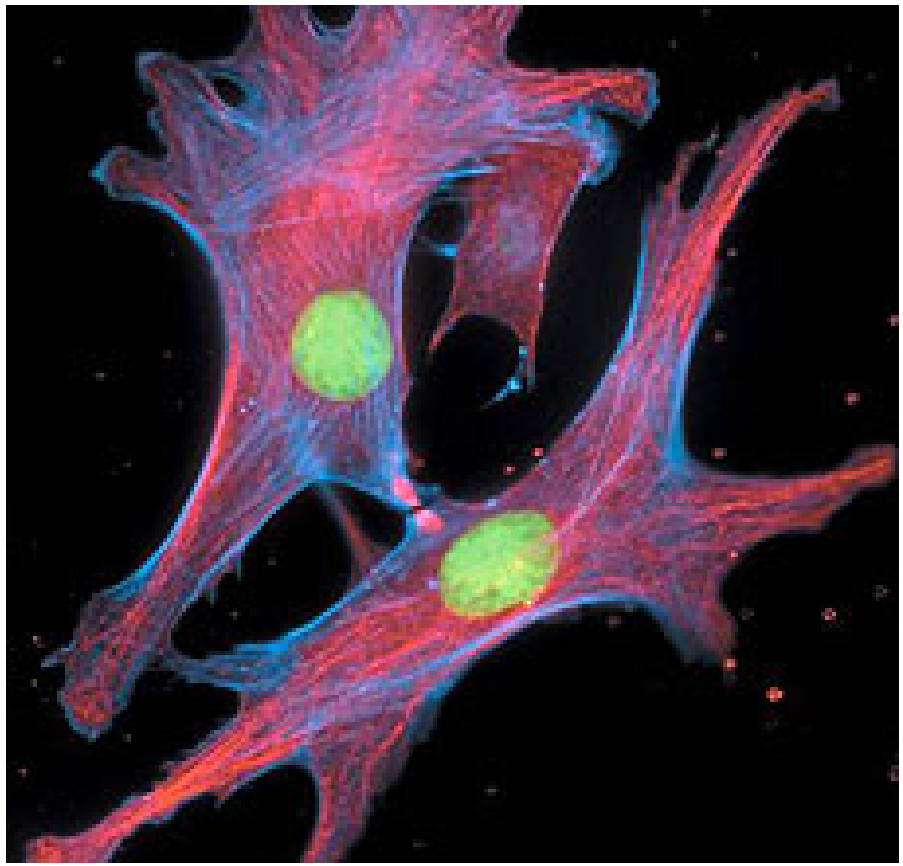
- High and UH density plates (96 / 384 / 1536 wells)
- Cell culture in each well receives specific treatment
- Traditional large scale assays measure single output from each well – gross average of cell population



MERCK

Microscopy: Studying Cell Phenotypes

- Cell morphology
- Nuclear morphology
- Organelles and internal structure
- Cell membrane
- Quantification of protein abundance
- Location and translocation
- ….



→ Measured features are *diverse* and may be *numerous* (100s)

MERCK

High Content Analysis/Screening

- HCA/S data are rich in:
  1. breadth – multiple diverse features of each cell

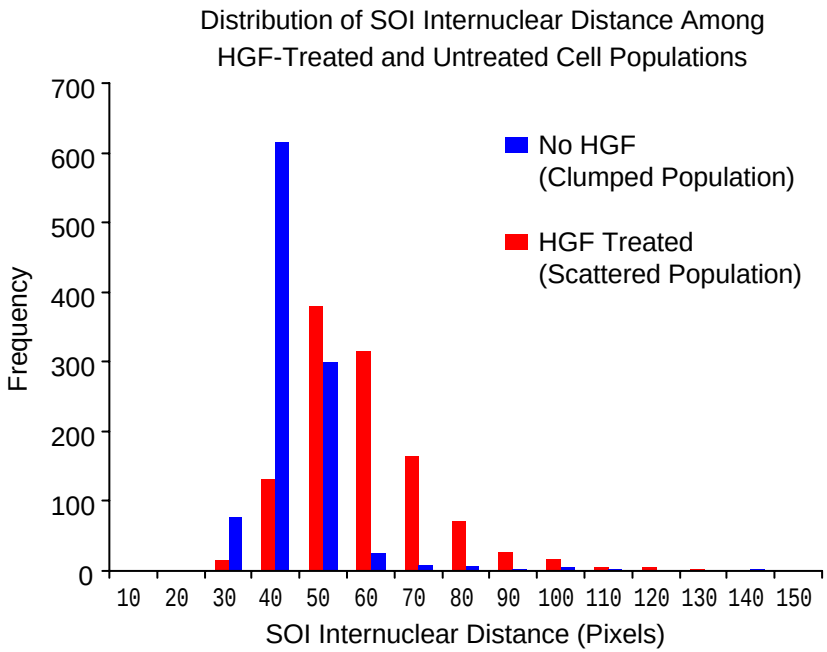
  2. depth – allows simultaneous examination of numerous individual cells

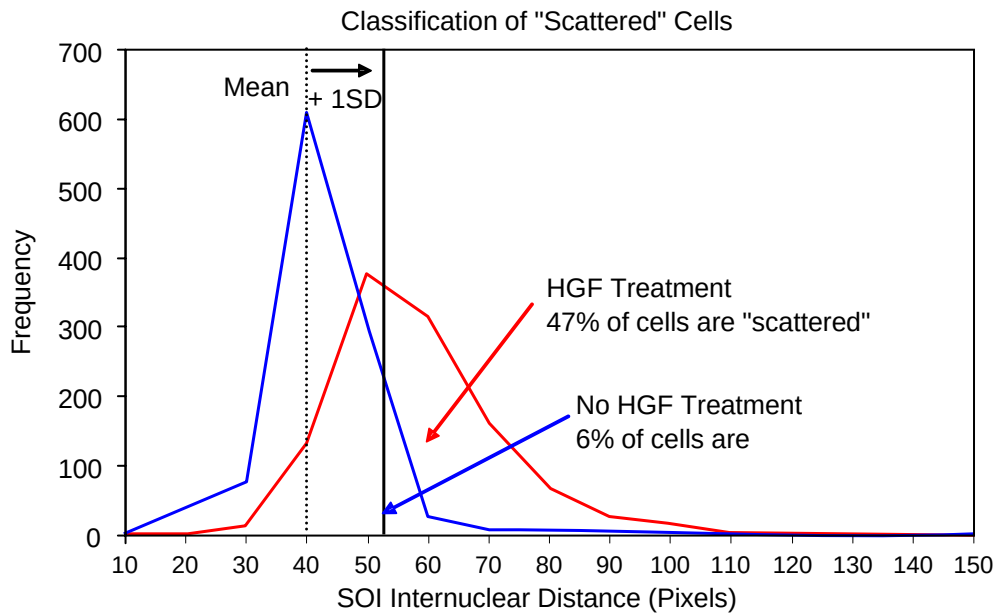
- Used in compound screening to lead optimization, as well as target discovery, target validation and safety

	A	B	C	D	E	F	G	H	I	J	K	
1	ROW	COL	CELL_NO	ObjectAreaCh1	ObjectPerimC	ObjectShapeF	ObjectShapeLW	ObjectShapeBFF	ObjectLengthCh1	ObjectWidthCh1	ObjectAngleCh1	O
2	A	1	1	123.6634313	54.33094207	1.862702176	2.324118079	0.526717512	23.3593558	10.05084725	24.93872293	2
3	A	1	2	163.0818	56.37660658	1.529917556	1.898584251	0.694518785	21.11425133	11.1210505	88.31965025	1
4	A	1	3	167.4500625	76.1447164	2.678859352	3.375265195	0.505320036	33.44362524	9.908443726	35.35729099	3
5	A	1	4	133.0239938	58.37643234	2.015762513	2.920158227	0.704656215	23.47899698	8.040316707	176.5031488	
6	A	1	5	415.0889438	110.6617099	2.32440956	2.893365813	0.579834189	45.51141582	15.72957544	86.9865542	4
7	A	1	6	439.2183938	124.4493295	2.758042815	3.077109069	0.535882547	50.22002198	16.32052061	124.3718504	
8	A	1	7	199.275975	77.04901946	2.348596777	3.418396499	0.643319692	32.54057824	9.519252156	0.660812852	3
9	A	1	8	238.382325	83.83887766	2.305190764	1.960370896	0.449686423	32.23675073	16.44421002	144.7725198	3
10	A	1	9	254.7113063	73.19244531	1.655771036	1.40694924	0.543612827	25.67547449	18.24904108	133.4968607	2
11	A	1	10	247.9509	68.01345829	1.461021572	1.717818214	0.671902925	25.17782325	14.65686127	128.7414975	2
12	A	1	11	218.3091188	83.3084745	2.46669506	2.785191308	0.495197045	35.04083379	12.58112277	154.3004944	3
13	A	1	12	217.7890875	108.2270908	4.114762676	6.455893585	0.573430157	49.51718811	7.670075019	127.431366	4
14	A	1	13	132.8159813	49.32744531	1.428225919	1.694466531	0.640206353	18.74915396	11.06493024	51.92911322	1
15	A	1	14	233.1820125	69.58011955	1.627707962	1.382045737	0.576307318	23.64730558	17.11036397	62.68309805	2
16	A	1	15	444.4187063	146.7650264	3.767018861	4.005900252	0.436905631	63.8340515	15.93500773	140.2540887	6
17	A	1	16	174.6264938	59.55576784	1.6005917	1.418809231	0.654039881	19.46323176	13.71800475	83.27151419	2
18	A	1	17	391.3755188	123.2107294	3.03904459	1.683675383	0.472931523	37.32732585	22.1701441	108.5360473	5
19	A	1	18	349.2529875	86.34062604	1.678810277	1.696851158	0.615806095	31.02201102	18.28210499	45.2718786	3
20	A	1	19	190.2274313	79.91747775	2.591717725	3.879831901	0.583099657	35.57718922	9.169775941	130.5109764	3
21	A	1	20	324.1874813	95.55004865	2.209003082	1.704107376	0.591366603	30.56456373	17.93582033	140.0682848	3
22	A	1	21	251.4871125	77.92321315	1.887786122	2.263171323	0.56106639	31.85000144	14.07317295	118.1442175	3
23	A	1	22	722.0113875	133.2216552	1.937353173	1.176022224	0.631432803	36.67044821	31.18176465	135.2825441	
24	A	1	23	109.6225875	46.35618721	1.53230914	1.675888172	0.601225438	17.48050072	10.43058899	37.34166698	1
25	A	1	24	307.2344625	85.87504865	1.881273867	2.599557325	0.614194024	36.0605105	13.87178892	28.48547961	3
26	A	1	25	272.08035	73.97496153	1.575530884	1.441668177	0.635541484	24.84330053	17.23232913	33.0822559	2
27	A	1	26	219.5571938	82.4935454	2.418110709	1.960038909	0.458965129	30.62078616	15.62253995	158.4541208	
28	A	1	27	223.3014188	72.98454216	1.864413833	1.47903285	0.529160829	24.98278251	16.89129657	78.73041006	2
29	A	1	28	233.9100563	80.82178081	2.194467448	3.226086307	0.664788969	33.69151463	10.44346351	176.5940337	3
30	A	1	29	261.679725	66.43337117	1.323580204	1.521378542	0.63421729	25.05443663	16.46824635	123.3974622	1
31	A	1	30	327.2036625	80.50321315	1.554900519	1.966830263	0.641353075	31.67699152	16.10560511	37.52195304	2
32	A	1	31	505.7823938	134.1852229	2.802815238	3.541407678	0.620780139	53.7156591	15.1678835	166.4371183	5
33	A	1	32	502.6622063	120.4188939	2.269225843	2.14941623	0.531476055	45.0875386	20.97664378	83.66100798	5
34	A	1	33	326.6836313	108.6428971	2.822171463	2.54135393	0.453081517	42.8063394	16.84391099	48.83619612	4
495	A	2	1	238.1743125	82.32754865	2.220926477	1.876575436	0.633565687	26.56040233	14.15365555	11.15454857	3
496	A	2	2	215.2929375	63.7197873	1.476149701	1.178505605	0.721174865	18.75687253	15.91581105	44.05841442	2
497	A	2	3	170.0502188	54.37678081	1.359987382	1.457993459	0.710302734	18.68291518	12.81412825	139.1816889	
498	A	2	4	236.3022	67.16611631	1.499103208	1.016658583	0.571488917	20.50300615	20.16705164	130.8521782	2
499	A	2	5	182.634975	51.82919369	1.15402624	1.108495154	0.72206336	16.74446285	15.10558057	161.058657	1
500	A	2	6	495.1737563	91.29828405	1.330741824	1.6138956	0.796785463	31.66983696	19.62322529	178.8499059	2
501	A	2	7	188.7713438	56.53310982	1.329341791	1.585899156	0.723040796	20.34812592	12.83065562	96.71682196	1

	A	B	C	D	E	F	G	H	I	J	K	
1	ROW	COL	CELL_NO	ObjectAreaCh1	ObjectPerimC	ObjectShapeF	ObjectShapeLW	ObjectShapeBFF	ObjectLengthCh1	ObjectWidthCh1	ObjectAngleCh1	O
2	A	1	1	123.6634313	54.33094207	1.862702176	2.324118079	0.526717512	23.3593558	10.05084725	24.93872293	2
3	A	1	2	163.0818	56.37660658	1.529917556	1.898584251	0.694518785	21.11425133	11.1210505	88.31965025	1
4	A	1	3	167.4500625	76.1447164	2.678859352	3.375265195	0.505320036	33.44362524	9.908443726	35.35729099	3
5	A	1	4	133.0239938	58.37643234	2.015762513	2.920158227	0.704656215	23.47899698	8.040316707	176.5031488	
6	A	1	5	415.0889438	110.6617099	2.32440956	2.893365813	0.579834189	45.51141582	15.72957544	86.9865542	4
7	A	1	6	439.2183938	124.4493295	2.758042815	3.077109069	0.535882547	50.22002198	16.32052061	124.3718504	
8	A	1	7	199.275975	77.04901946	2.348596777	3.418396499	0.643319692	32.54057824	9.519252156	0.660812852	3
9	A	1	8	238.382325	83.83887766	2.305190764	1.960370896	0.449686423	32.23675073	16.44421002	144.7725198	3
10	A	1	9	254.7113063	73.19244531	1.655771036	1.40694924	0.543612827	25.67547449	18.24904108	133.4968607	2
11	A	1	10	247.9509	68.01345829	1.461021572	1.717818214	0.671902925	25.17782325	14.65686127	128.7414975	2
12	A	1	11	218.3091188	83.3084745	2.46669506	2.785191308	0.495197045	35.04083379	12.58112277	154.3004944	3
13	A	1	12	217.7890875	108.2270908	4.114762676	6.455893585	0.573430157	49.51718811	7.670075019	127.431366	4
14	A	1	13	132.8159813	49.32744531	1.428225919	1.694466531	0.640206353	18.74915396	11.06493024	51.92911322	1
15	A	1	14	233.1820125	69.58011955	1.627707962	1.382045737	0.576307318	23.64730558	17.11036397	62.68309805	2
16	A	1	15	444.4187063	146.7650264	3.767018861	4.005900252	0.436905631	63.8340515	15.93500773	140.2540887	6
17	A	1	16	174.6264938	59.55576784	1.6005917	[illegible]	0.654039881	19.46323176	13.71800475	83.27151419	2
18	A	1	17	391.3755188	123.2107294	3.03904459	203.1	0.472931523	37.32732585	22.1701441	108.5360473	5
19	A	1	18	349.2529875	86.34062604	1.678810277		0.615806095	31.02201102	18.28210499	45.2718786	3
20	A	1	19	190.2274313	79.91747775	2.591717725	3.879831901	0.583099657	35.57718922	9.169775941	130.5109764	3
21	A	1	20	324.1874813	95.55004865	2.209003082	1.704107376	0.591366603	30.56456373	17.93582033	140.0682848	3
22	A	1	21	251.4871125	77.92321315	1.887786122	2.263171323	0.56106639	31.85000144	14.07317295	118.1442175	3
23	A	1	22	722.0113875	133.2216552	1.937353173	1.176022224	0.631432803	36.67044821	31.18176465	135.2825441	
24	A	1	23	109.6225875	46.35618721	1.53230914	1.675888172	0.601225438	17.48050072	10.43058899	37.34166698	1
25	A	1	24	307.2344625	85.87504865	1.881273867	2.599557325	0.614194024	36.0605105	13.87178892	28.48547961	3
26	A	1	25	272.08035	73.97496153	1.575530884	1.441668177	0.635541484	24.84330053	17.23232913	33.0822559	2
27	A	1	26	219.5571938	82.4935454	2.418110709	1.960038909	0.458965129	30.62078616	15.62253995	158.4541208	
28	A	1	27	223.3014188	72.98454216	1.864413833	1.47903285	0.529160829	24.98278251	16.89129657	78.73041006	2
29	A	1	28	233.9100563	80.82178081	2.194467448	3.226086307	0.664788969	33.69151463	10.44346351	176.5940337	3
30	A	1	29	261.679725	66.43337117	1.323580204	1.521378542	0.63421729	25.05443663	16.46824635	123.3974622	1
31	A	1	30	327.2036625	80.50321315	1.554900519	1.966830263	0.641353075	31.67699152	16.10560511	37.52195304	2
32	A	1	31	505.7823938	134.1852229	2.802815238	3.541407678	0.620780139	53.7156591	15.1678835	166.4371183	5
33	A	1	32	502.6622063	120.4188939	2.269225843	2.14941623	0.531476055	45.0875386	20.97664378	83.66100798	5
34	A	1	33	326.6836313	108.6428971	2.822171463	2.54135393	0.453081517	42.8063394	16.84391099	48.83619612	4
495	A	2	1	238.1743125	82.32754865	2.220926477	1.876575436	0.633565687	26.56040233	14.15365555	11.15454857	3
496	A	2	2	215.2929375	63.7197873	1.476149701	1.178505605	0.721174865	18.75687253	15.91581105	44.05841442	2
497	A	2	3	170.0502188	54.37678081	1.359987382	1.457993459	0.710302734	18.68291518	12.81412825	139.1816889	
498	A	2	4	236.3022	67.16611631	1.499103208	1.016658583	0.571488917	20.50300615	20.16705164	130.8521782	2
499	A	2	5	182.634975	51.82919369	1.15402624	1.108495154	0.72206336	16.74446285	15.10558057	161.058657	1
500	A	2	6	495.1737563	91.29828405	1.330741824	1.6138956	0.796785463	31.66983696	19.62322529	178.8499059	2
501	A	2	7	188.7713438	56.53310982	1.329341791	1.585899156	0.723040796	20.34812592	12.83065562	96.71682196	1

Cell Population-based Analysis





	Spacing SOI Mean (pixels)	Spacing SOI Std. Dev. (pixels)	% Scattered Cells[1]
Control	39	16	5.9
HGF treated	55	15	47

[1]Scattered cells calculated using SOI population-based method.

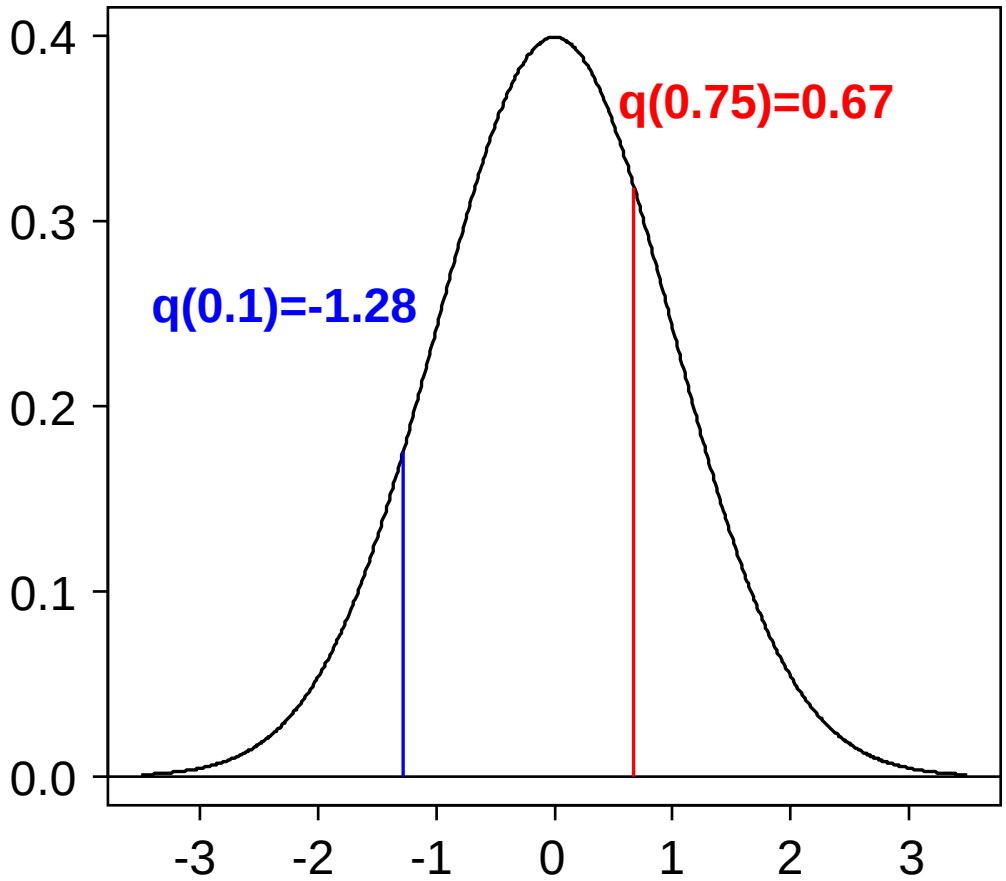
Comparison of internuclear distance mean, standard deviation, and calculated percentage of scattered cells in control and HGF-treated wells.

Chan et al. *J. Biomol. Screen*. (2008) 13:847-854

MERCK

Characterising Distributions with Quantiles

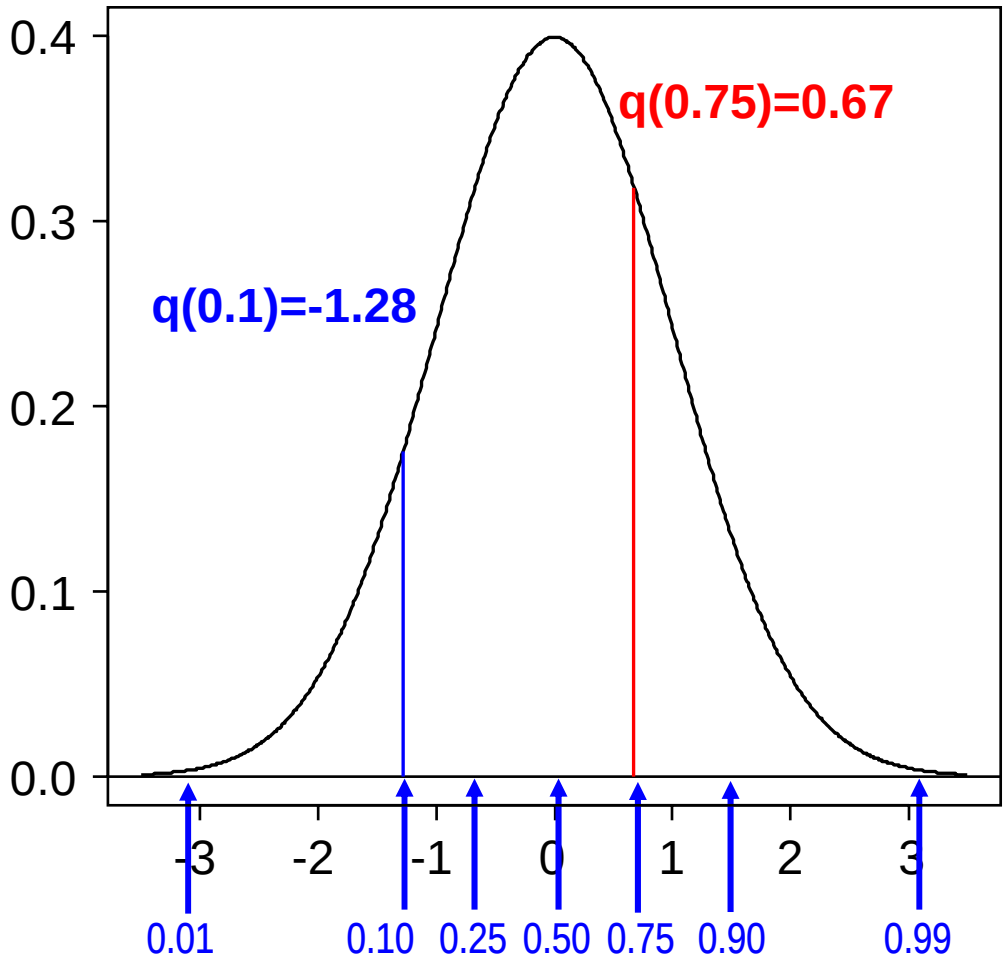
- Summary description of distribution by sequence of quantiles and
- Spread: d = q(90) – q(10)

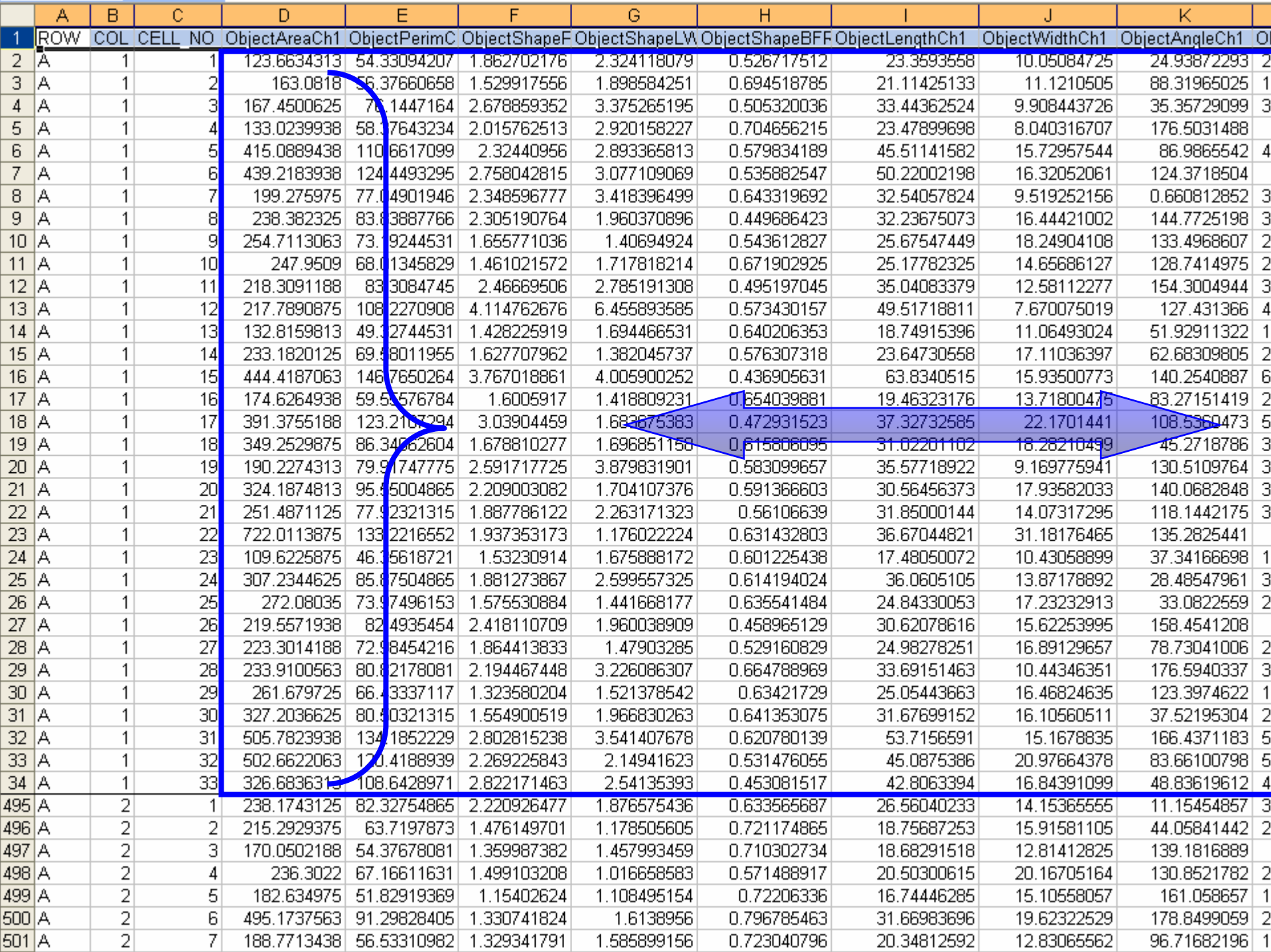


Characterising Distributions with Quantiles

- Summary description of distribution by sequence of quantiles and
- Spread: d = q(90) – q(10)

feature –>
{ feature01, feature10,…
…feature99,featured10 }

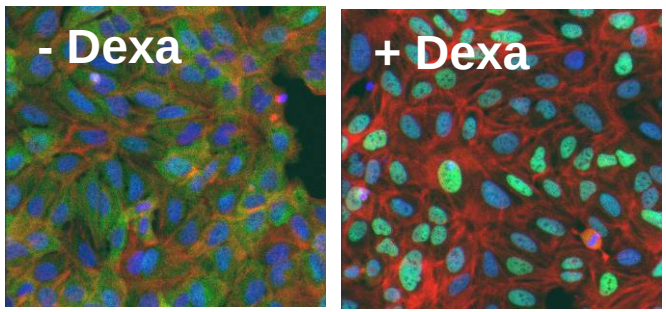




	A	B	C	D	E	F	G	H	I	J	K	
1	ROW	COL	CELL_NO	ObjectAreaCh1	ObjectPerimC	ObjectShapeF	ObjectShapeLW	ObjectShapeBFF	ObjectLengthCh1	ObjectWidthCh1	ObjectAngleCh1	O
2	A	1	1	123.6634313	54.33094207	1.862702176	2.324118079	0.526717512	23.3593558	10.05084725	24.93872293	2
3	A	1	2	163.0818	56.37660658	1.529917556	1.898584251	0.694518785	21.11425133	11.1210505	88.31965025	1
4	A	1	3	167.4500625	76.1447164	2.678859352	3.375265195	0.505320036	33.44362524	9.908443726	35.35729099	3
5	A	1	4	133.0239938	58.37643234	2.015762513	2.920158227	0.704656215	23.47899698	8.040316707	176.5031488	
6	A	1	5	415.0889438	110.6617099	2.32440956	2.893365813	0.579834189	45.51141582	15.72957544	86.9865542	4
7	A	1	6	439.2183938	124.4493295	2.758042815	3.077109069	0.535882547	50.22002198	16.32052061	124.3718504	
8	A	1	7	199.275975	77.04901946	2.348596777	3.418396499	0.643319692	32.54057824	9.519252156	0.660812852	3
9	A	1	8	238.382325	83.83887766	2.305190764	1.960370896	0.449686423	32.23675073	16.44421002	144.7725198	3
10	A	1	9	254.7113063	73.19244531	1.655771036	1.40694924	0.543612827	25.67547449	18.24904108	133.4968607	2
11	A	1	10	247.9509	68.01345829	1.461021572	1.717818214	0.671902925	25.17782325	14.65686127	128.7414975	2
12	A	1	11	218.3091188	83.3084745	2.46669506	2.785191308	0.495197045	35.04083379	12.58112277	154.3004944	3
13	A	1	12	217.7890875	108.2270908	4.114762676	6.455893585	0.573430157	49.51718811	7.670075019	127.431366	4
14	A	1	13	132.8159813	49.12744531	1.428225919	1.694466531	0.640206353	18.74915396	11.06493024	51.92911322	1
15	A	1	14	233.1820125	69.58011955	1.627707962	1.382045737	0.576307318	23.64730558	17.11036397	62.68309805	2
16	A	1	15	444.4187063	146.7650264	3.767018861	4.005900252	0.436905631	63.8340515	15.93500773	140.2540887	6
17	A	1	16	174.6264938	59.55576784	1.6005917	1.418809231	0.654039881	19.46323176	13.71800475	83.27151419	2
18	A	1	17	391.3755188	123.2107294	3.03904459	1.683875383	0.472931523	37.32732585	22.1701441	108.5369473	5
19	A	1	18	349.2529875	86.34962604	1.678810277	1.696851158	0.615806095	31.02201102	18.28210499	45.2718786	3
20	A	1	19	190.2274313	79.91747775	2.591717725	3.879831901	0.583099657	35.57718922	9.169775941	130.5109764	3
21	A	1	20	324.1874813	95.55004865	2.209003082	1.704107376	0.591366603	30.56456373	17.93582033	140.0682848	3
22	A	1	21	251.4871125	77.92321315	1.887786122	2.263171323	0.56106639	31.85000144	14.07317295	118.1442175	3
23	A	1	22	722.0113875	133.2216552	1.937353173	1.176022224	0.631432803	36.67044821	31.18176465	135.2825441	
24	A	1	23	109.6225875	46.15618721	1.53230914	1.675888172	0.601225438	17.48050072	10.43058899	37.34166698	1
25	A	1	24	307.2344625	85.87504865	1.881273867	2.599557325	0.614194024	36.0605105	13.87178892	28.48547961	3
26	A	1	25	272.08035	73.97496153	1.575530884	1.441668177	0.635541484	24.84330053	17.23232913	33.0822559	2
27	A	1	26	219.5571938	82.4935454	2.418110709	1.960038909	0.458965129	30.62078616	15.62253995	158.4541208	
28	A	1	27	223.3014188	72.98454216	1.864413833	1.47903285	0.529160829	24.98278251	16.89129657	78.73041006	2
29	A	1	28	233.9100563	80.82178081	2.194467448	3.226086307	0.664788969	33.69151463	10.44346351	176.5940337	3
30	A	1	29	261.679725	66.43337117	1.323580204	1.521378542	0.63421729	25.05443663	16.46824635	123.3974622	1
31	A	1	30	327.2036625	80.90321315	1.554900519	1.966830263	0.641353075	31.67699152	16.10560511	37.52195304	2
32	A	1	31	505.7823938	134.1852229	2.802815238	3.541407678	0.620780139	53.7156591	15.1678835	166.4371183	5
33	A	1	32	502.6622063	120.4188939	2.269225843	2.14941623	0.531476055	45.0875386	20.97664378	83.66100798	5
34	A	1	33	326.6836313	108.6428971	2.822171463	2.54135393	0.453081517	42.8063394	16.84391099	48.83619612	4
495	A	2	1	238.1743125	82.32754865	2.220926477	1.876575436	0.633565687	26.56040233	14.15365555	11.15454857	3
496	A	2	2	215.2929375	63.7197873	1.476149701	1.178505605	0.721174865	18.75687253	15.91581105	44.05841442	2
497	A	2	3	170.0502188	54.37678081	1.359987382	1.457993459	0.710302734	18.68291518	12.81412825	139.1816889	
498	A	2	4	236.3022	67.16611631	1.499103208	1.016658583	0.571488917	20.50300615	20.16705164	130.8521782	2
499	A	2	5	182.634975	51.82919369	1.15402624	1.108495154	0.72206336	16.74446285	15.10558057	161.058657	1
500	A	2	6	495.1737563	91.29828405	1.330741824	1.6138956	0.796785463	31.66983696	19.62322529	178.8499059	2
501	A	2	7	188.7713438	56.53310982	1.329341791	1.585899156	0.723040796	20.34812592	12.83065562	96.71682196	1

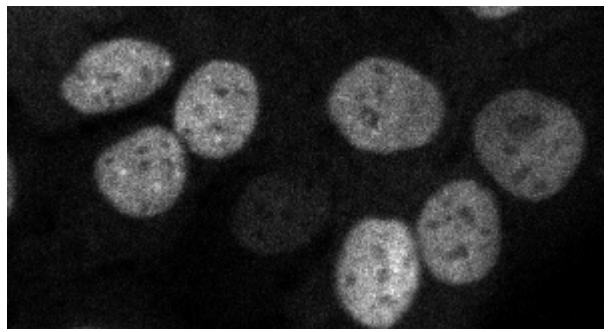
Example: Small Molecule Screen for Receptor Translocation

GR-GFP nuclear translocation
(primary effect)



GR-GFP
DNA
F-actin

Nuclear foci formation
(secondary effect)



- 1536 well format imaging assay

MERCK

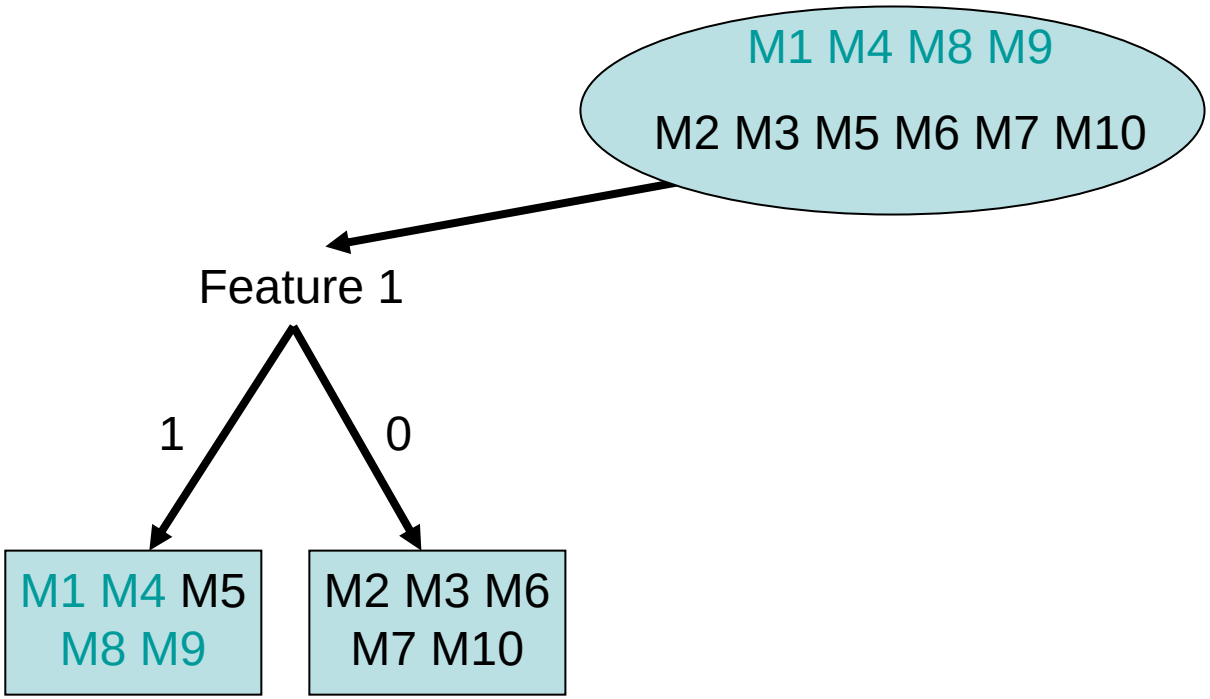
Small Molecule Screen: Goals

- Develop hit selection method based on known control compounds
- Find the most relevant features

 supervised classification

Classification Trees

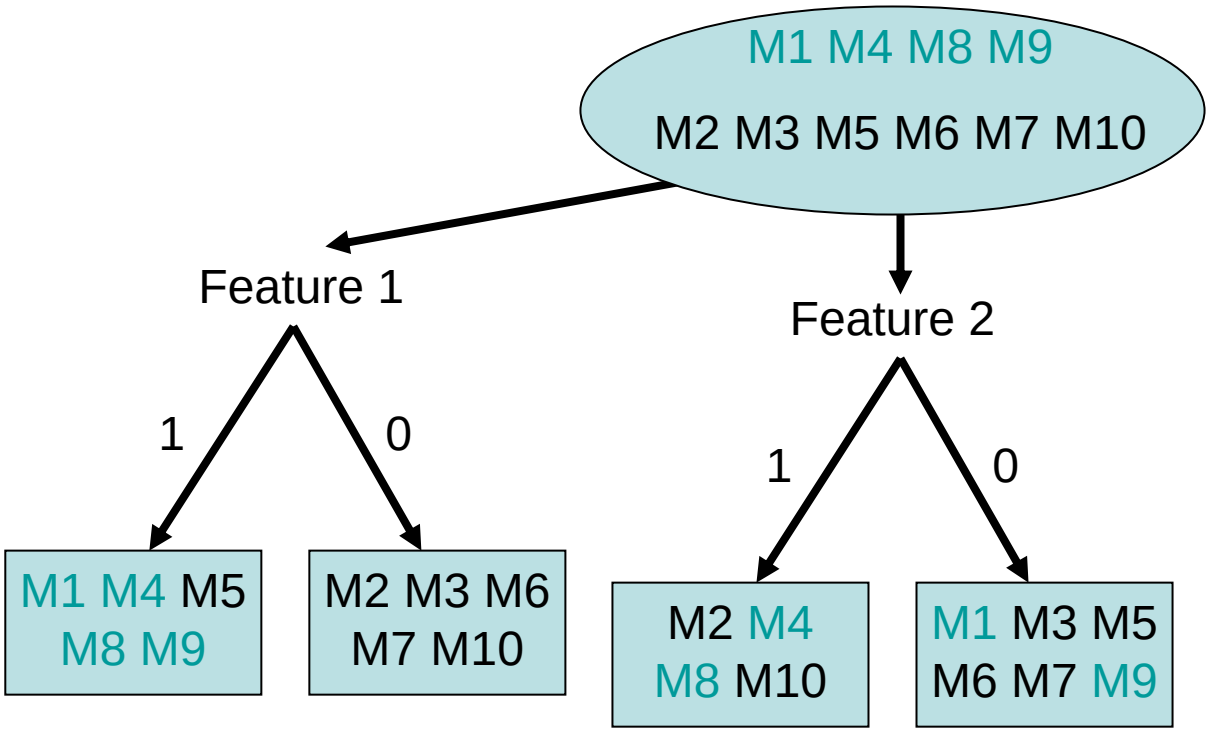
	Activity	Feature 1	Feature 2	…
Molecule 1	active	1	0	…
Molecule 2	inactive	0	1	…
Molecule 3	inactive	0	0	…
…	…	…	…	



MERCK

Classification Trees

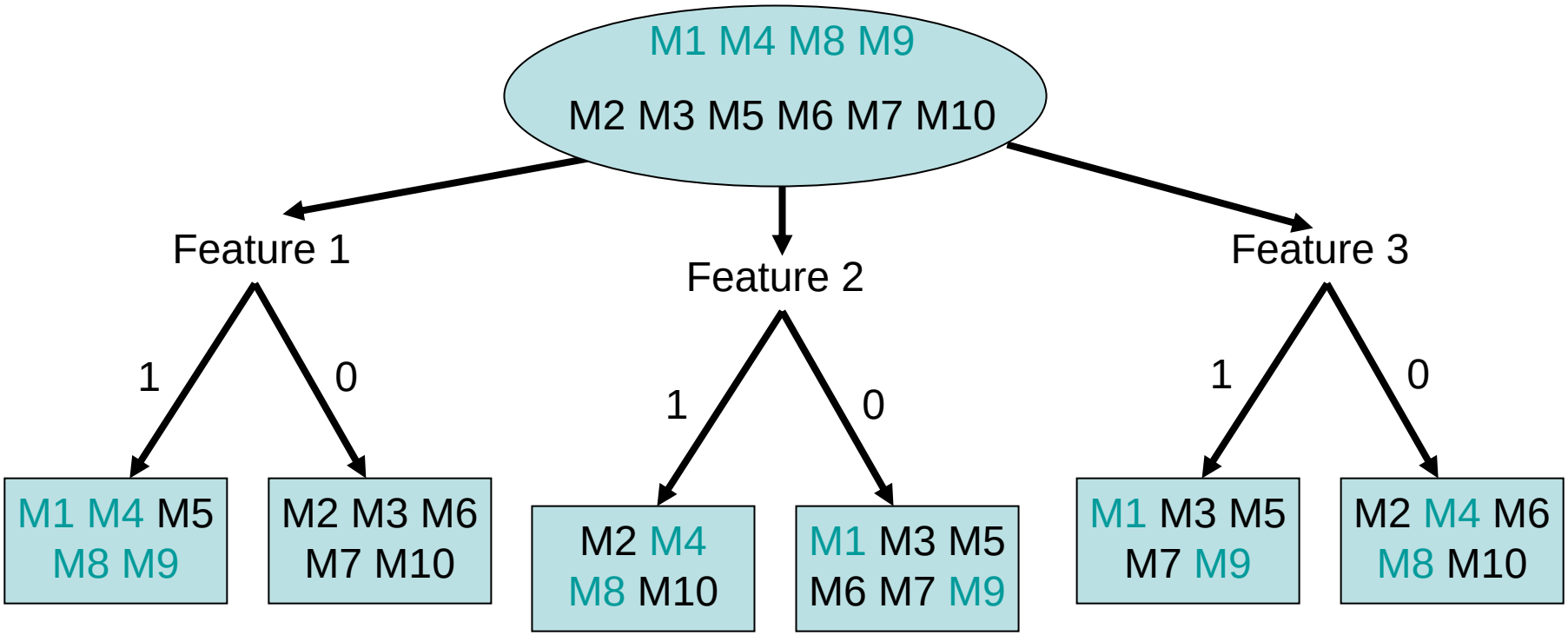
	Activity	Feature 1	Feature 2	…
Molecule 1	active	1	0	…
Molecule 2	inactive	0	1	…
Molecule 3	inactive	0	0	…
…	…	…	…	



MERCK

Classification Trees

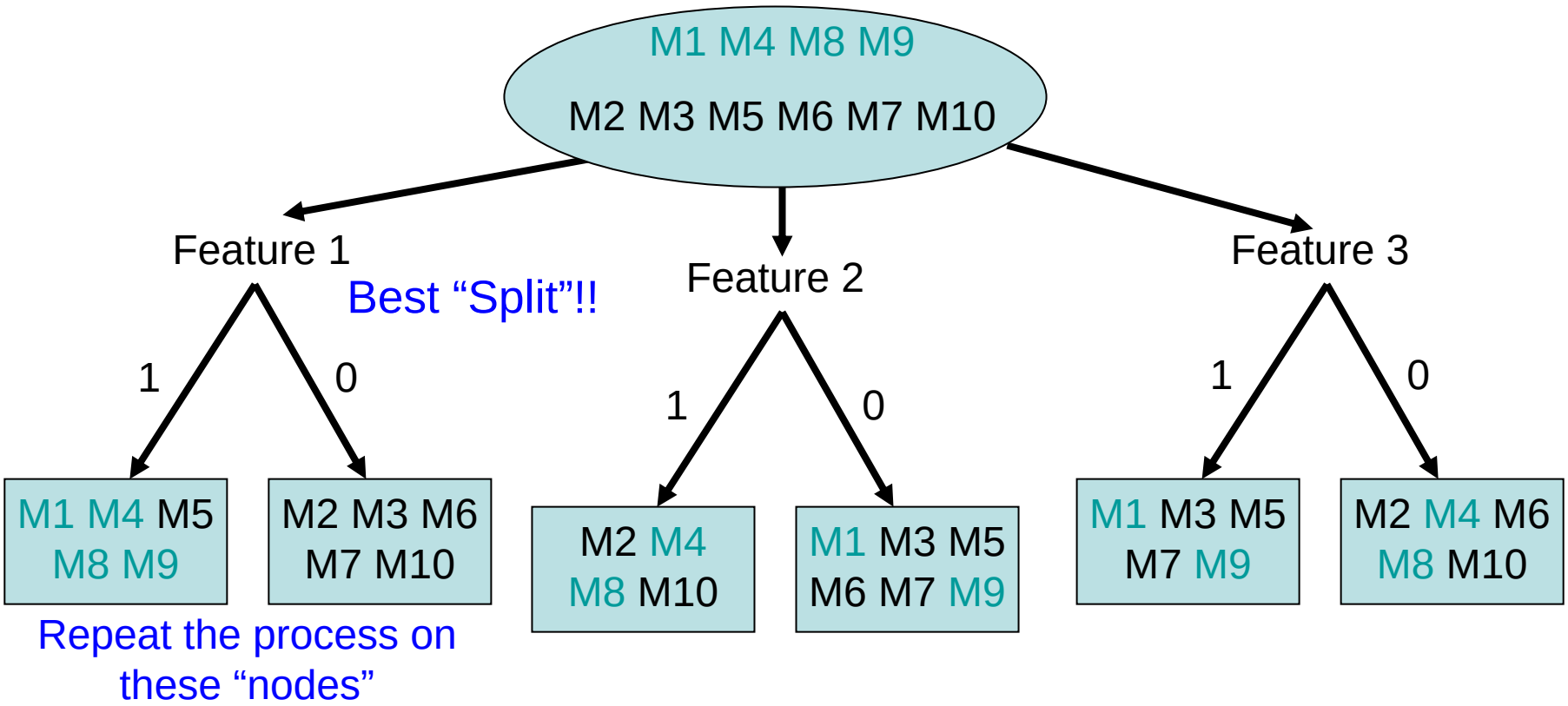
	Activity	Feature 1	Feature 2	…
Molecule 1	active	1	0	…
Molecule 2	inactive	0	1	…
Molecule 3	inactive	0	0	…
…	…	…	…	



MERCK

Classification Trees

	Activity	Feature 1	Feature 2	…
Molecule 1	active	1	0	…
Molecule 2	inactive	0	1	…
Molecule 3	inactive	0	0	…
…	…	…	…	



MERCK

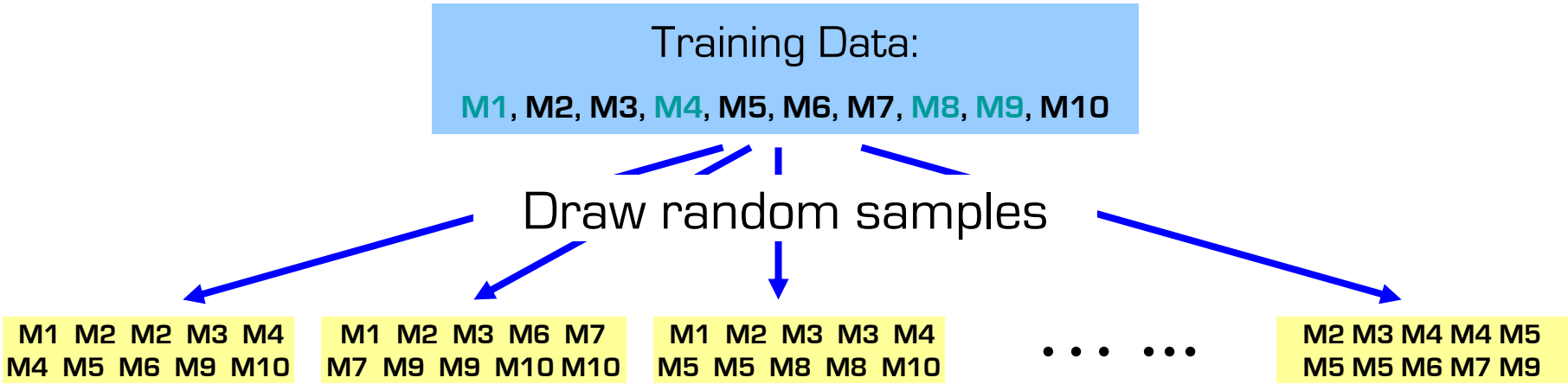
Growing a (Random) Forest

Training Data:

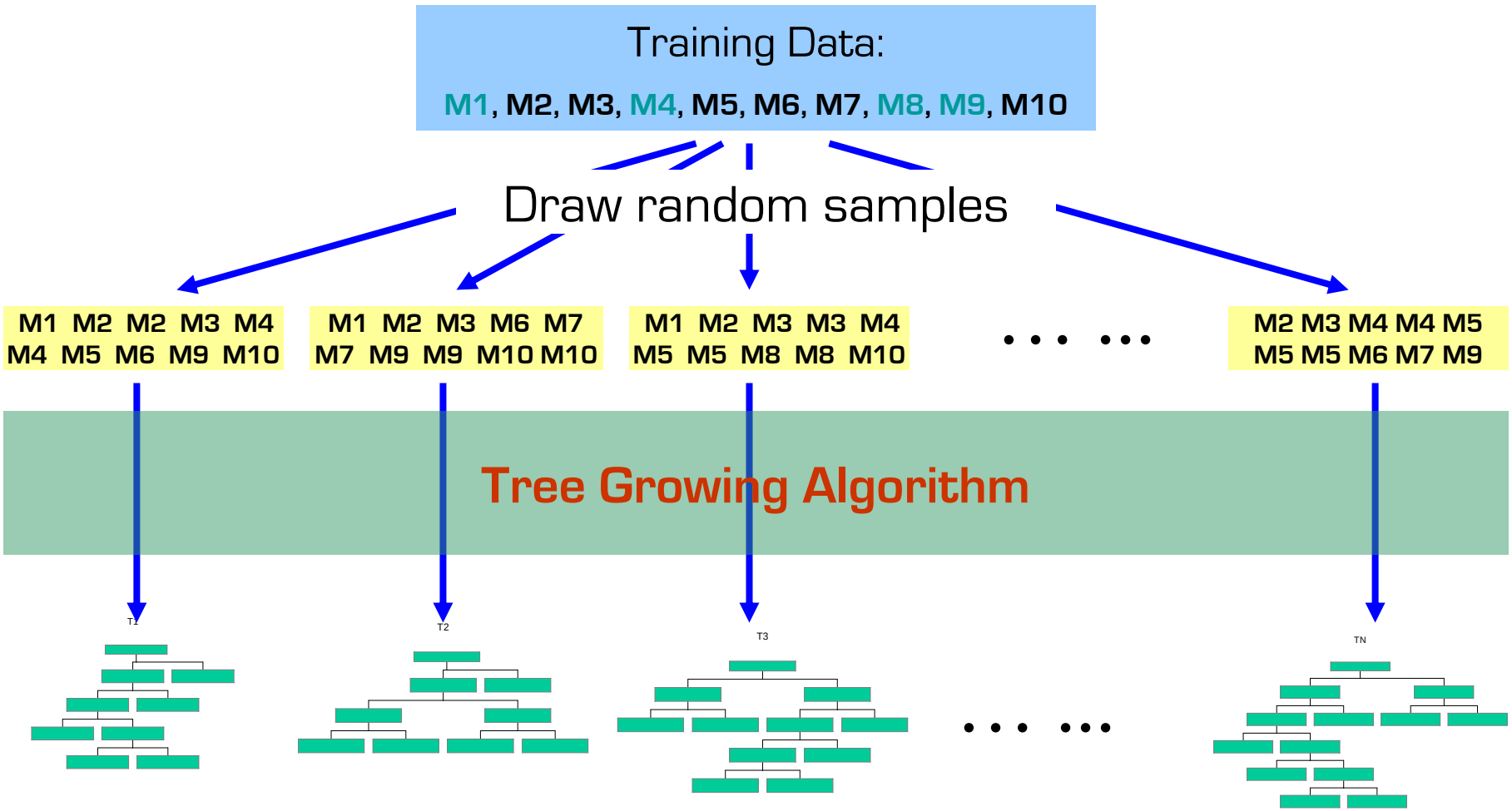
M1, M2, M3, M4, M5, M6, M7, M8, M9, M10

MERCK

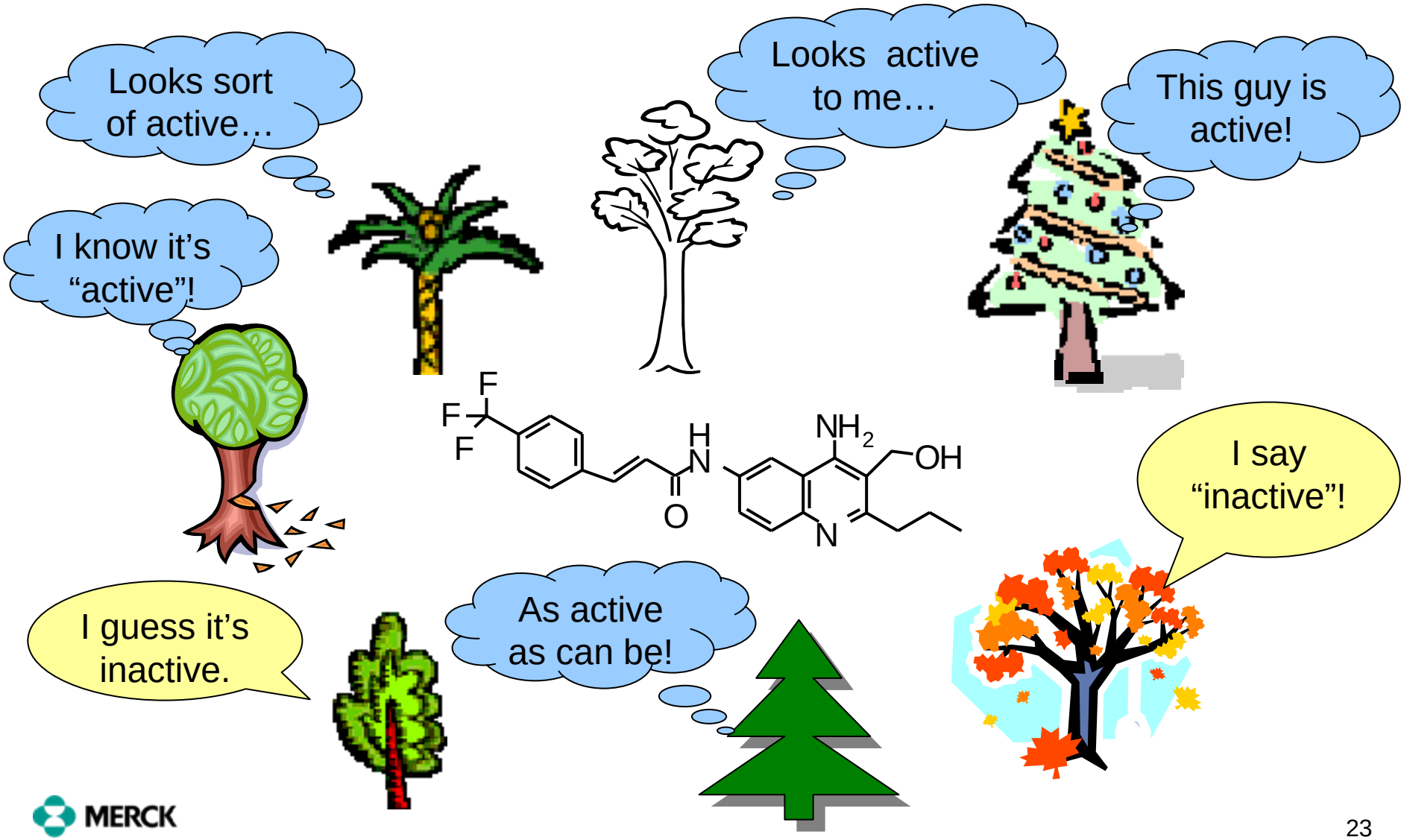
Growing a (Random) Forest



Growing a (Random) Forest



Ensemble Prediction with a Forest

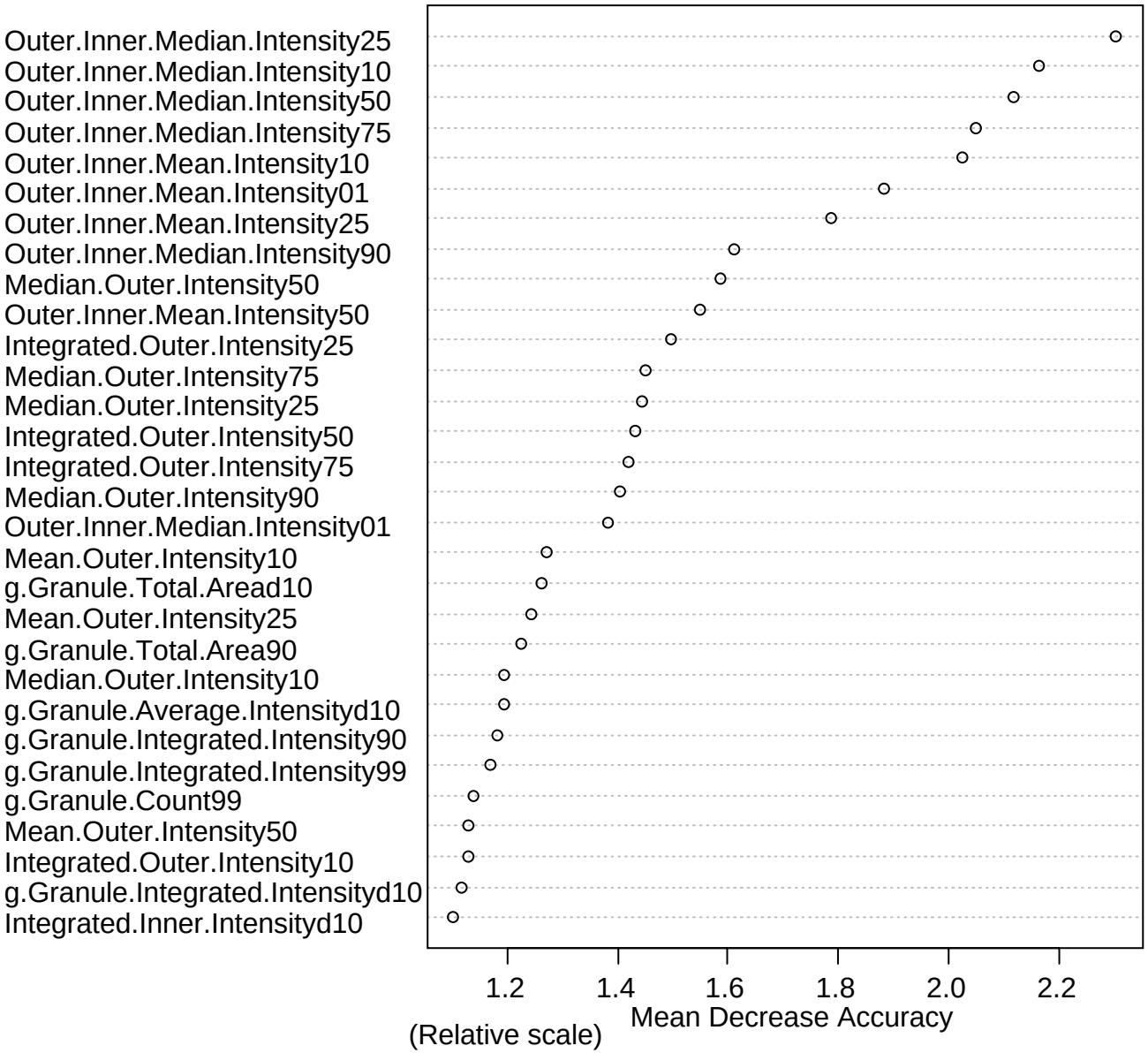


MERCK

Random Forests: Our Swiss Army Knife

- The Random Forest classifier
  - is a powerful and flexible classifier that has excellent prediction performance on data from diverse areas
  - has internal feature ranking/selection for the task at hand
  - handles well large number of variables even when correlated
  - provides probability estimates of class membership

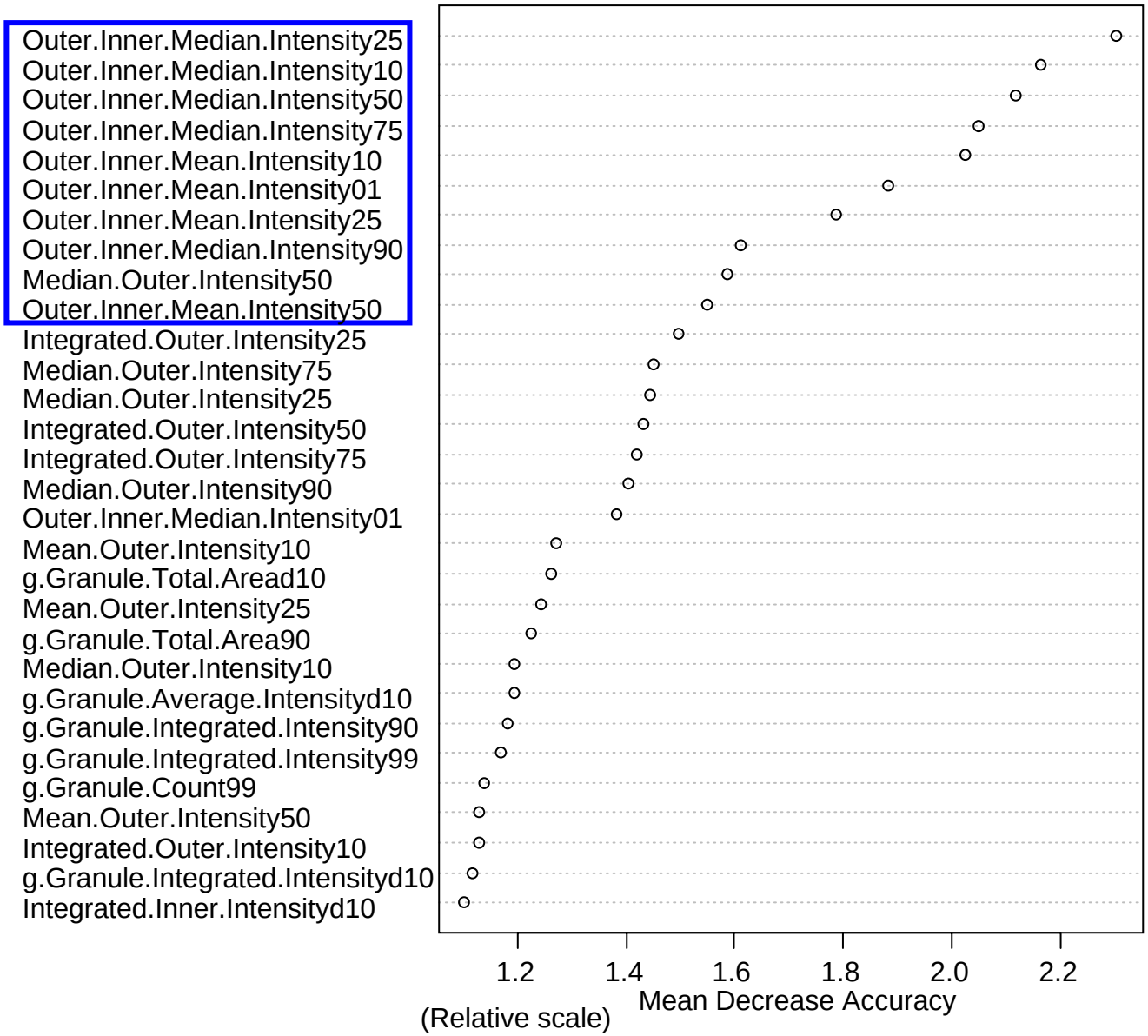
Example: Feature Importance (RF)



MERCK

Example: Feature Importance (RF)

Strongly separating feature seen across the distribution



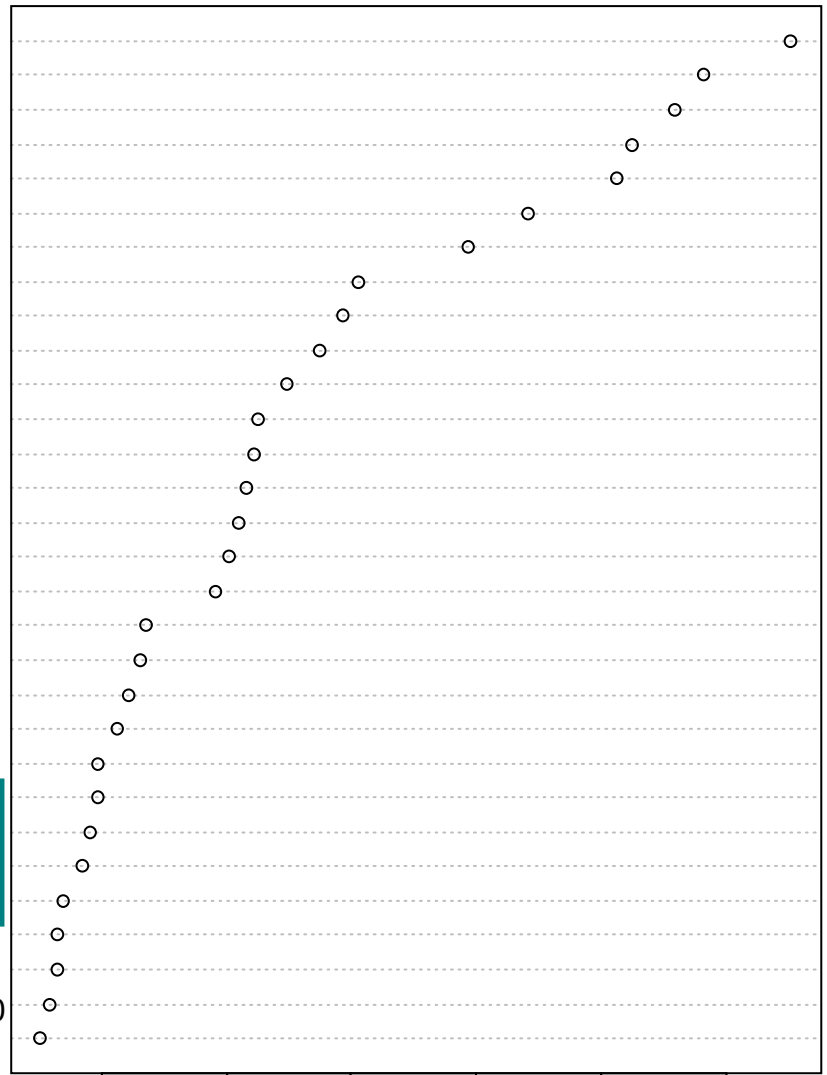
MERCK

Example: Feature Importance (RF)

Strongly separating feature seen across the distribution

Feature separating
In the upper tail

Outer.Inner.Median.Intensity25
Outer.Inner.Median.Intensity10
Outer.Inner.Median.Intensity50
Outer.Inner.Median.Intensity75
Outer.Inner.Mean.Intensity10
Outer.Inner.Mean.Intensity01
Outer.Inner.Mean.Intensity25
Outer.Inner.Median.Intensity90
Median.Outer.Intensity50
Outer.Inner.Mean.Intensity50
Integrated.Outer.Intensity25
Median.Outer.Intensity75
Median.Outer.Intensity25
Integrated.Outer.Intensity50
Integrated.Outer.Intensity75
Median.Outer.Intensity90
Outer.Inner.Median.Intensity01
Mean.Outer.Intensity10
g.Granule.Total.Aread10
Mean.Outer.Intensity25
g.Granule.Total.Area90
Median.Outer.Intensity10
g.Granule.Average.Intensityd10
g.Granule.Integrated.Intensity90
g.Granule.Integrated.Intensity99
g.Granule.Count99
Mean.Outer.Intensity50
Integrated.Outer.Intensity10
g.Granule.Integrated.Intensityd10
Integrated.Inner.Intensityd10

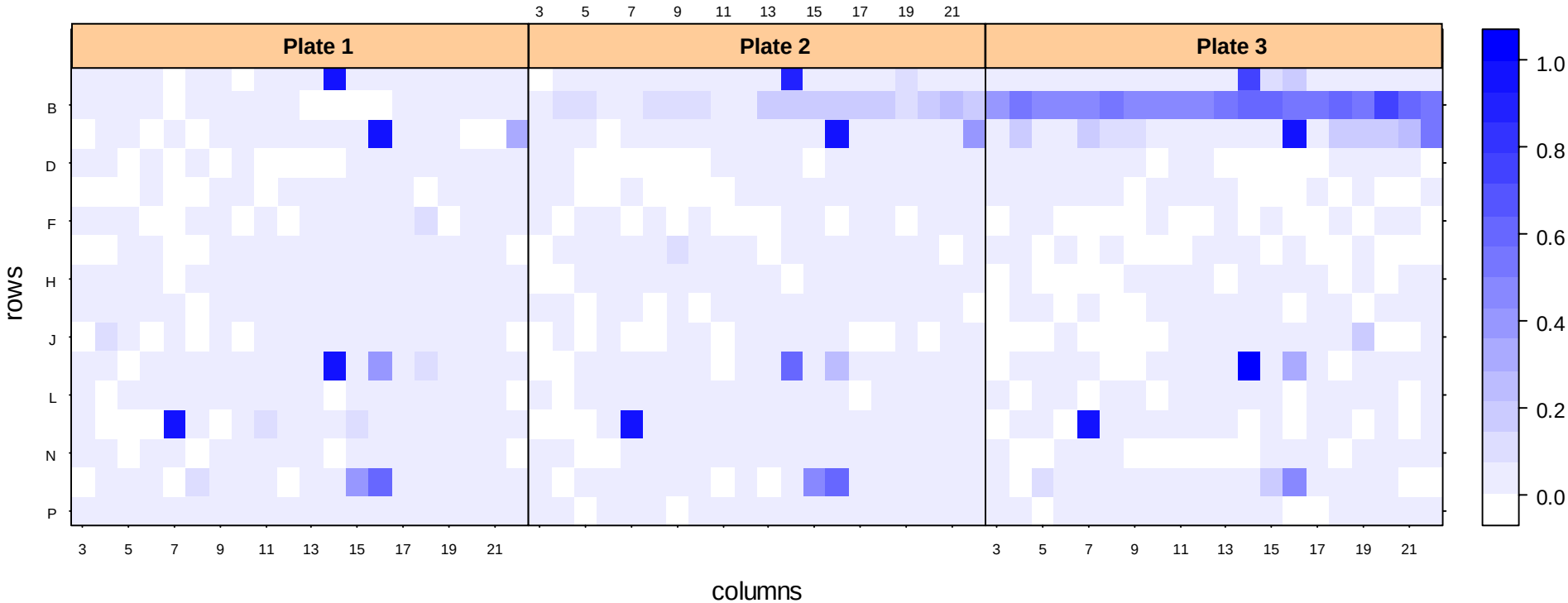


MERCK

Small Molecule Screen: Detecting Non-Background Signal

Based on full multivariate feature set

Probability of *NOT* being background (DMSO-like)



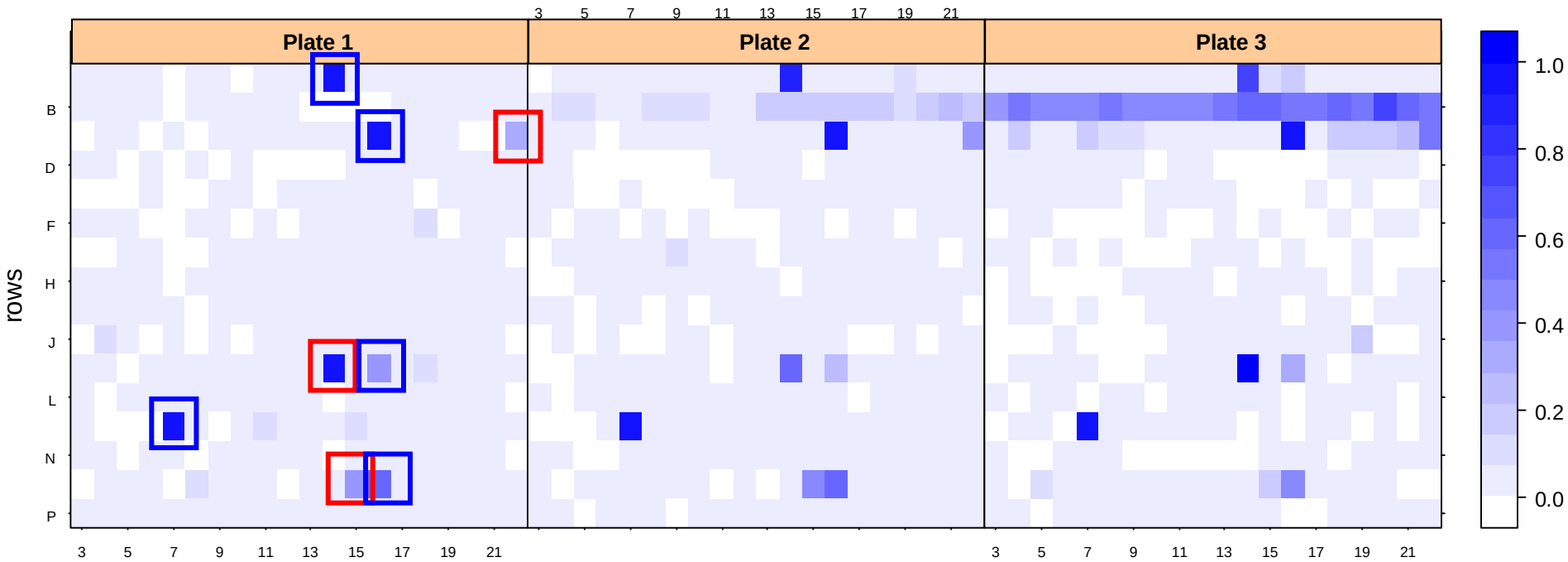
Small Molecule Screen: Detecting Non-Background Signal

Candidates:

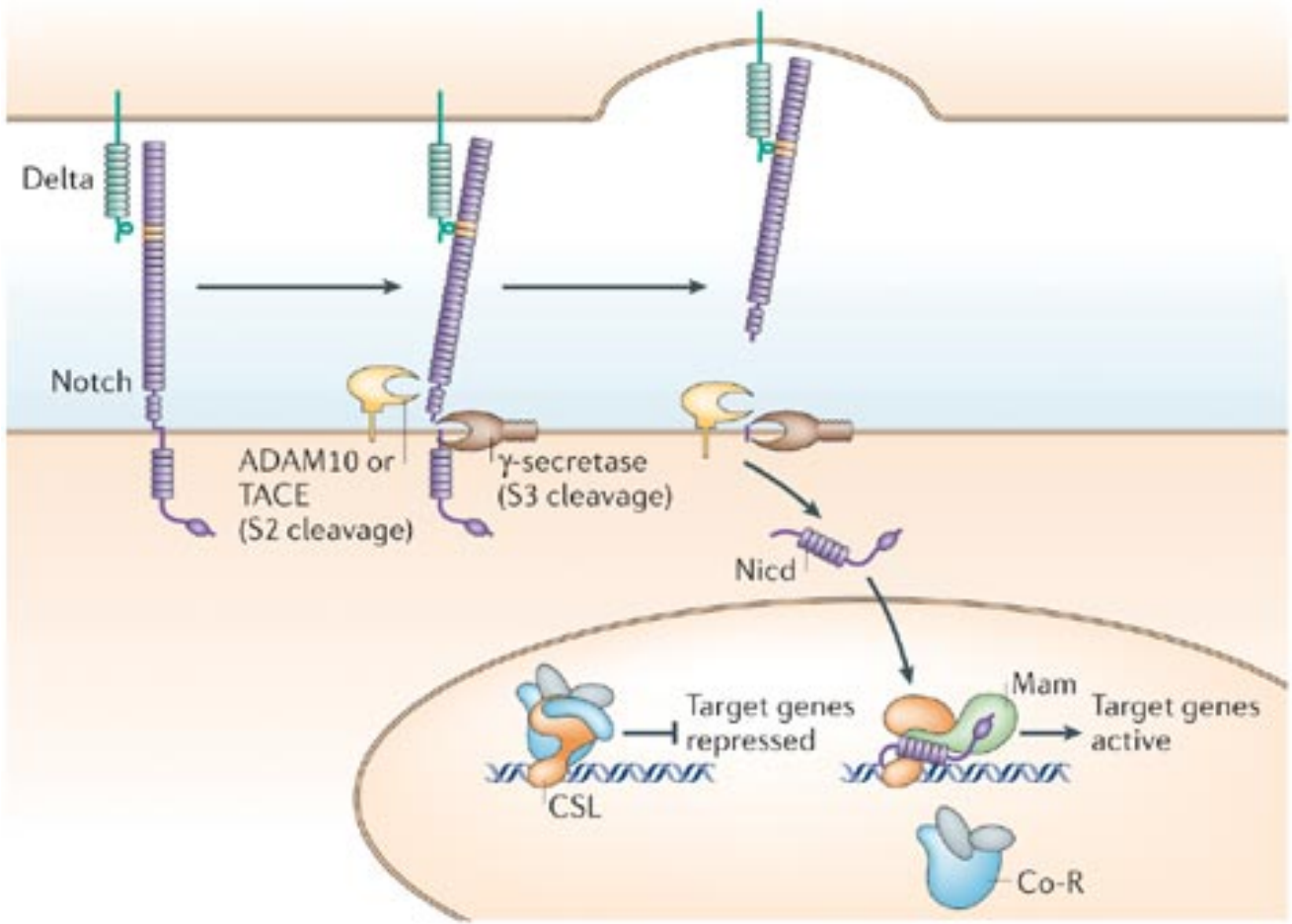
Hits

Toxic compounds

Probability of *NOT* being background (DMSO-like)

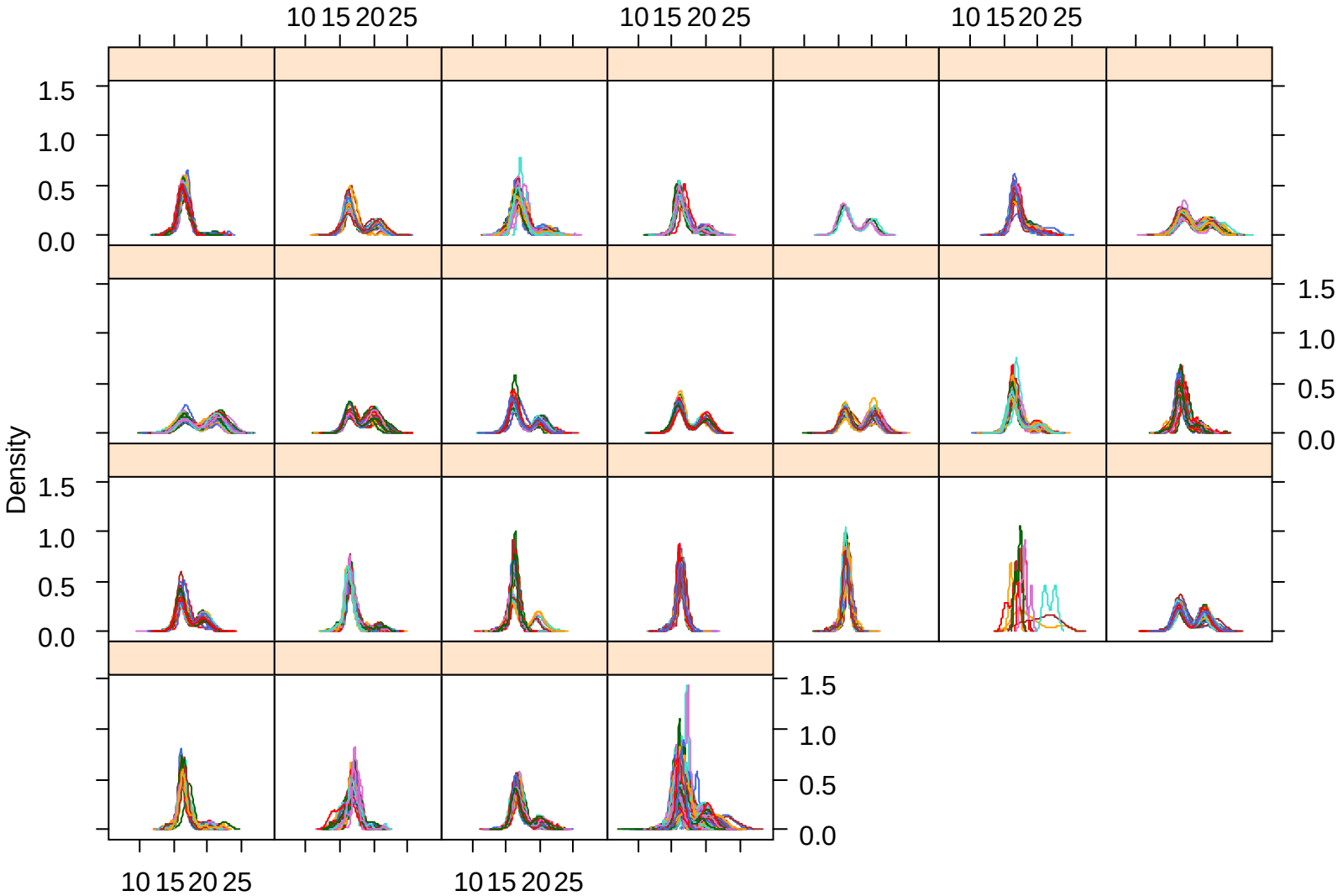


Example: Exploratory Study of Cell Signaling Pathway

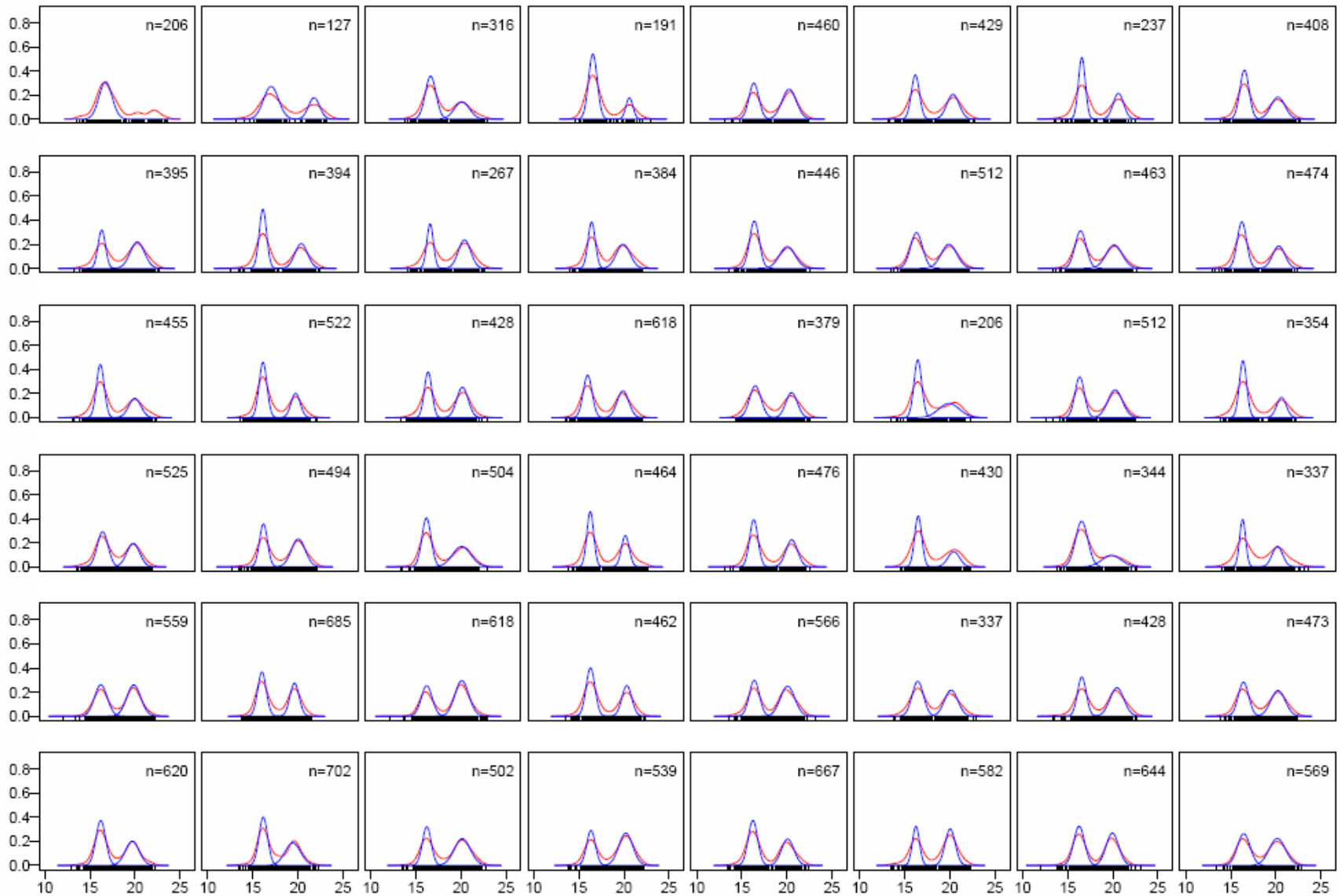


Bray *Nature Reviews Molecular Cell Biology* **7**, 678–689 (September 2006) | doi:10.1038/nrm2009

Cell Signaling Pathway: Cell Sub-populations

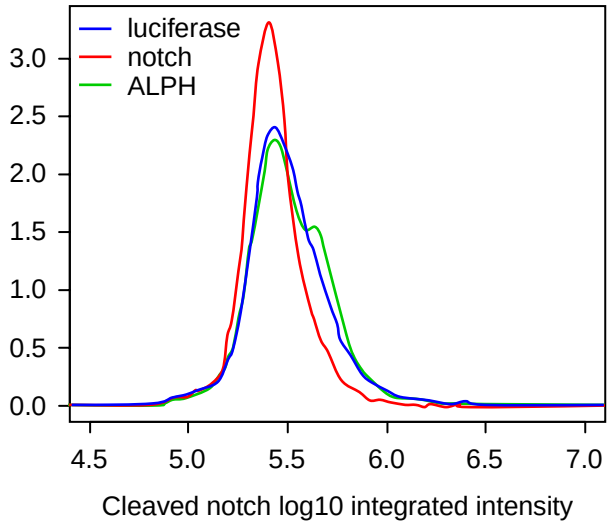


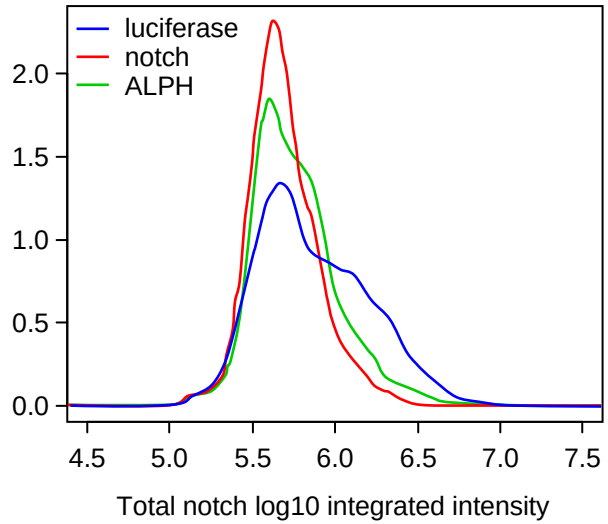
Capturing Cell Sub-populations with Mixture Models



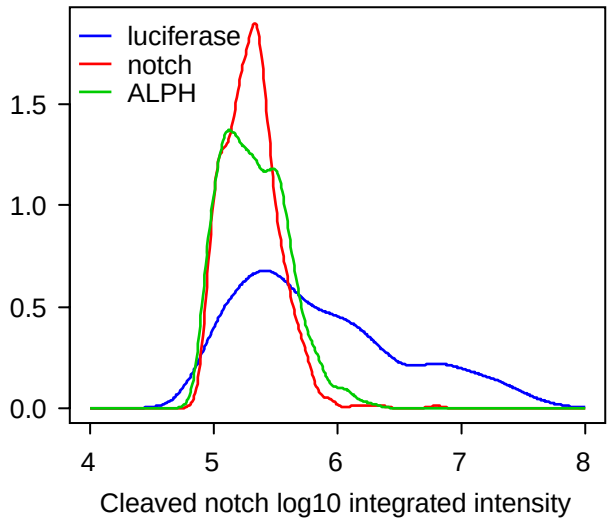
Composite Q-Scores for Multiple Controls and Channels

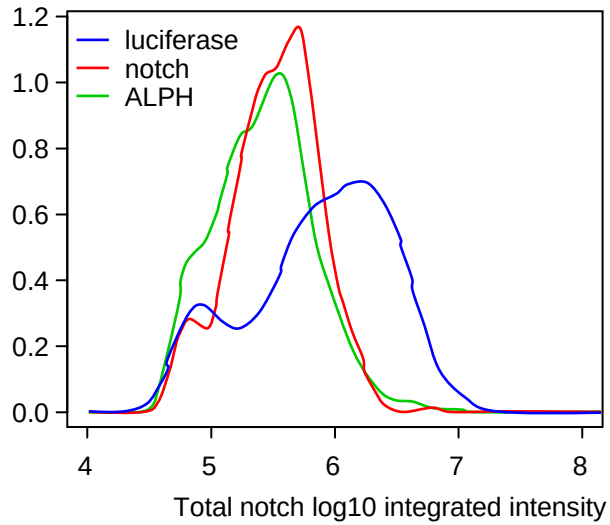
A plate with little separation between negative and positive controls





A plate with good separation between negative and positive controls





MERCK

Composite Q-Scores for Multiple Controls and Channels

- Compute the *q* quantile (e.g. 90th percentile for effects in right tail of distribution) for the controls and each sample well on a plate

$$q1 = \{q_{sample} - q_{pos1}\} / \{q_{neg} - q_{pos1}\}$$

$$q2 = \{q_{sample} - q_{pos2}\} / \{q_{neg} - q_{pos2}\}$$

- Compute q-scores for each of the fluorescence channels and combine according to goal of study

What have we used?

- Summary of distributions using quantiles enabling comparison and additional analysis
- Mixture modelling of cell sub-populations for extraction of relevant parameters
- Supervised classification based on stochastic ensemble training and prediction for hit selection
- Composite scores that correspond to biological processes

Take Home Message

- No two studies are the same: need array of statistical methods, simple as well as complex, and a flexible approach to carefully match the tools to the scientific questions.

Acknowledgements

We thank Richard Raubertas, Ting Wang, Shubhankar Ray and Vladimir Svetnik for their contribution.

Thank You