Analysis of HapMap Lymphoblastic Cell Line Surface Protein Expression

Whitney Green

Summer Research Program in Genomics

PI: Philip De Jager, MD PhD

Supervisor: Elizabeth Rossin

BROAD INSTITUTE

Project Overview

- Approach: use the HapMap lymphoblastic cell lines as a model system to study gene expression
- Goals: Correlate genetic variation captured by the HapMap with:
  1. cell surface protein expression
  2. poly-marker immunological expression

Data Collection: The Samples

- HapMap Cell Lines (270 Total Samples)
  - 30 Yoruban Trios (YRI)
  - 45 Japanese Individuals (JPT)
  - 45 Han Chinese Individuals (CHB)
  - 30 U.S. Trios (CEU)

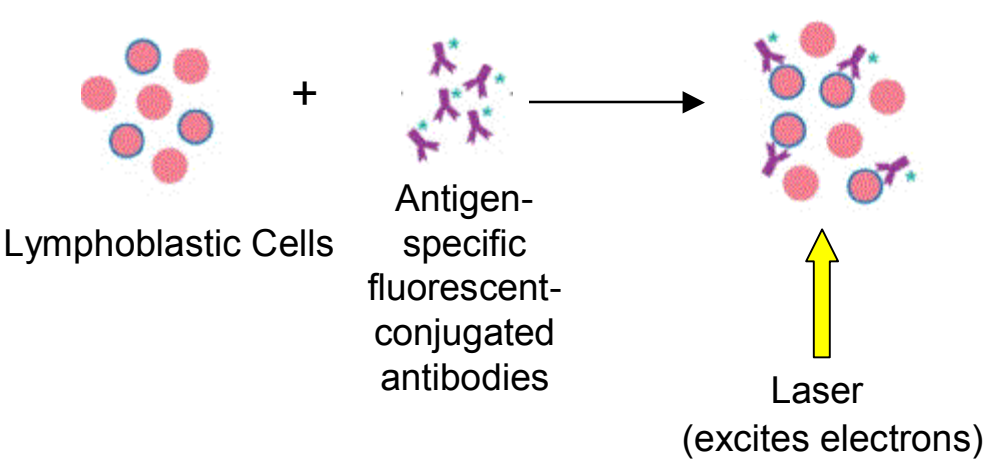
Data Collection: The Markers

1. B-cell specific markers
2. Markers expressed on B-cells that showed evidence of heritability in previous RNA expression studies

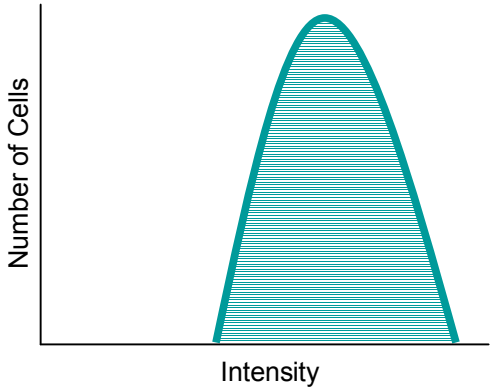
CD19	CD20
IL6R	IgD
HLA-DQ	HLA-DR
CD227	CD21
CD40	IgG
IgM	CD58
CD95	CD86
CD80	

BROAD INSTITUTE

Data Collection: The Method



Cell #	FSC-H	SSC-H	FL1-H	FL2-H	FL3-H	FL2-A	FL4-H
1	484	282	117	195	88	0	192
2	538	308	106	36	49	0	15
3	509	374	227	253	205	0	191
4	595	386	182	188	129	0	237

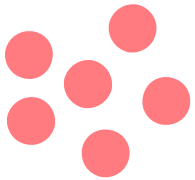


Fluorescent-activated cell sorting

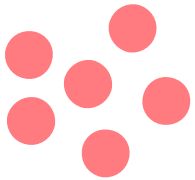
BROAD INSTITUTE

Images courtesy of http://www2.hawaii.edu/~johnb/micro/micro161/antigen-antibody_reactions/Chap5_Ag_Ab_reactions.htm

Data Collection: The Method



Lymphoblastic cells



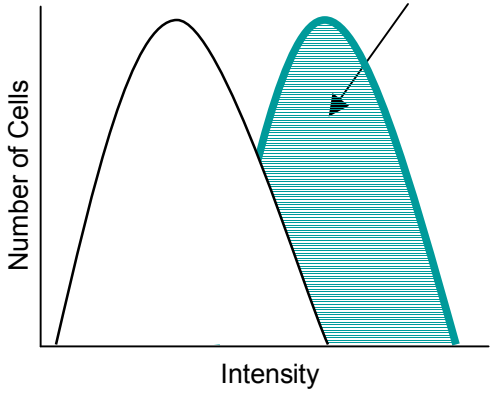
Laser
(excites electrons)

light
emitted

Detector

Cell #	FSC-H	SSC-H	FL1-H	FL2-H	FL3-H	FL2-A	FL4-H
1	484	282	117	195	88	0	192
2	538	308	106	36	49	0	15
3	509	374	227	253	205	0	191
4	595	386	182	188	129	0	237

phenotype



Fluorescent-activated cell sorting

BROAD INSTITUTE

Images courtesy of http://www2.hawaii.edu/~johnb/micro/micro161/antigen-antibody_reactions/Chap5_Ag_Ab_reactions.htm

Data Processing and Analysis

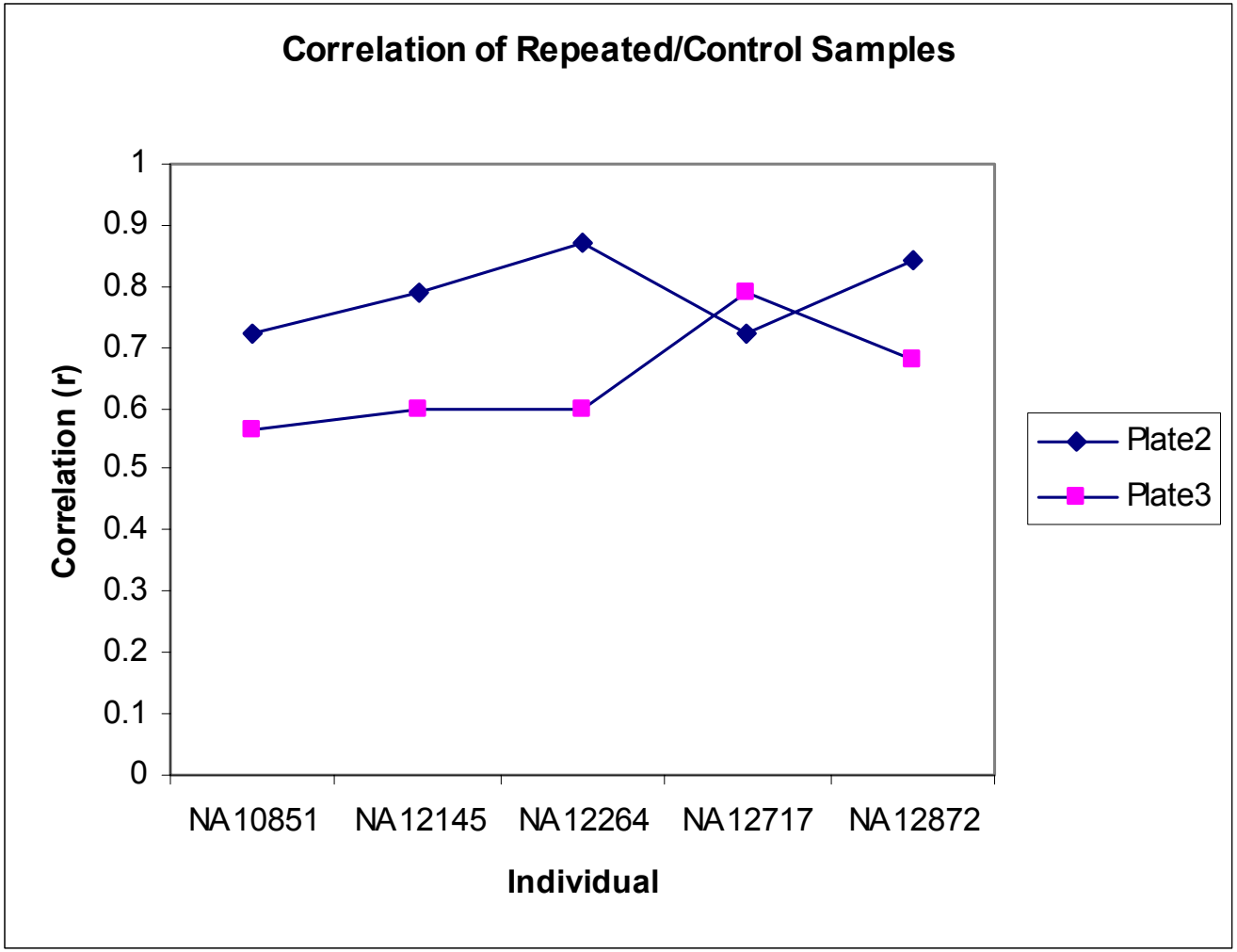
1. Perform quality control analysis
2. Correlate marker expression with genetic variants
3. Define poly-marker immunological expression profiles
4. Correlate immunological expression profiles with genetic variants

BROAD INSTITUTE

Sources of Variation

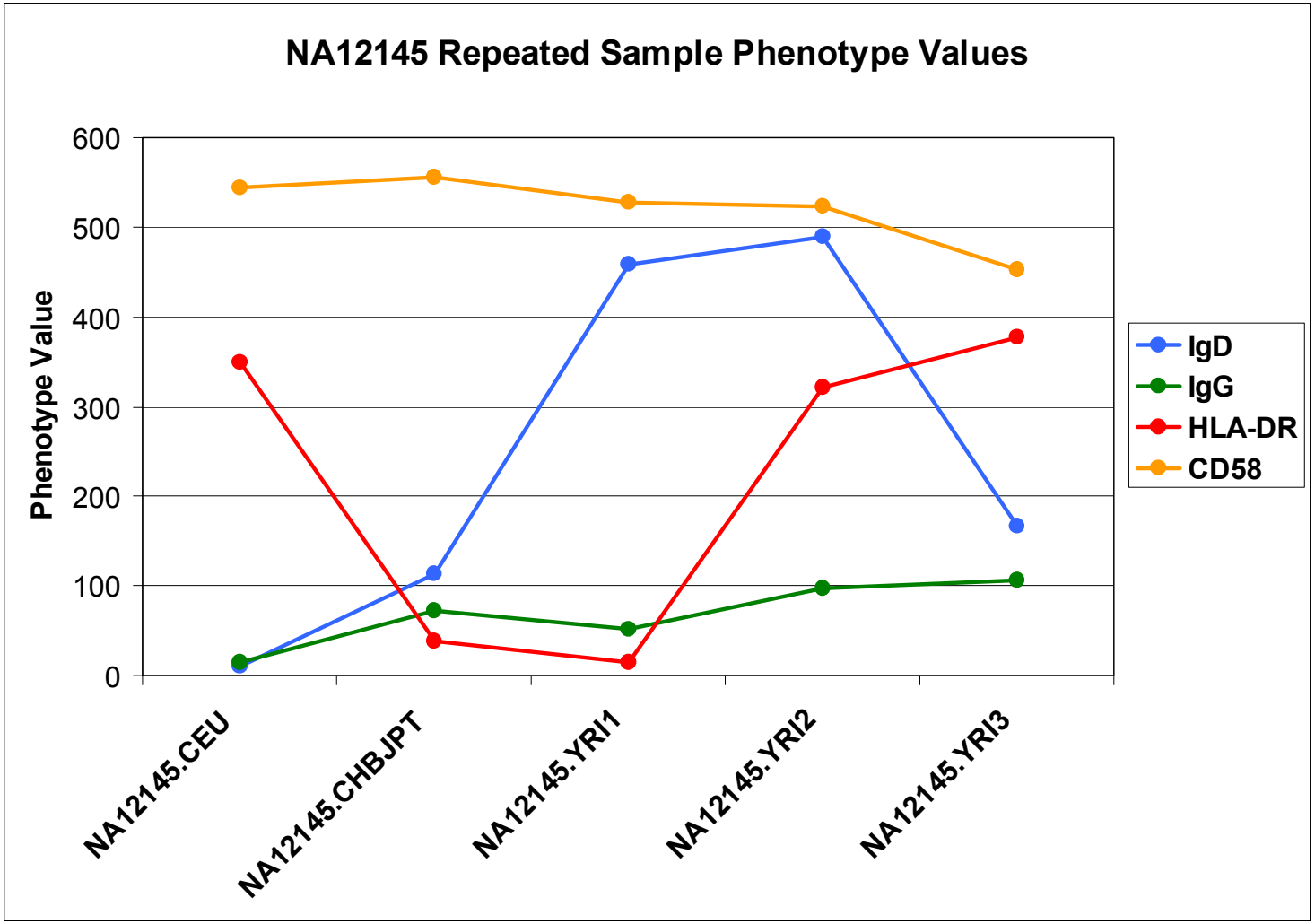
- Manual handling of samples (staining protocol/loading FACS machine)
- Cell count (sampling)
- Other (cell culture, user manipulation required during detection)

Correlation of Repeated/Control Samples



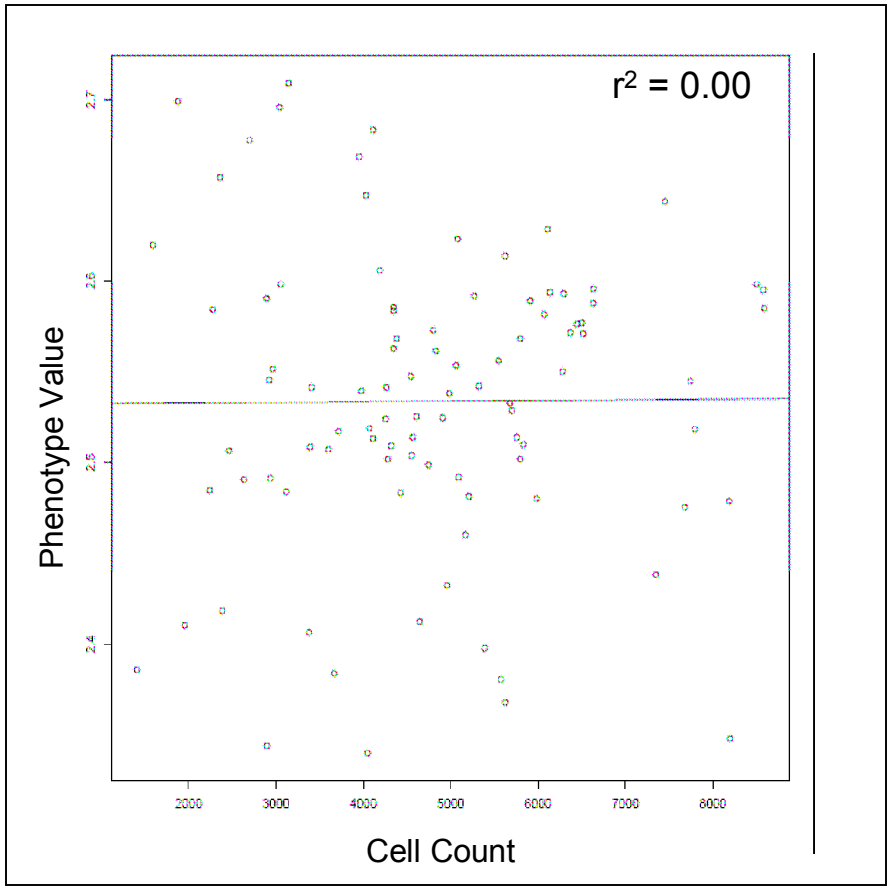
BROAD INSTITUTE

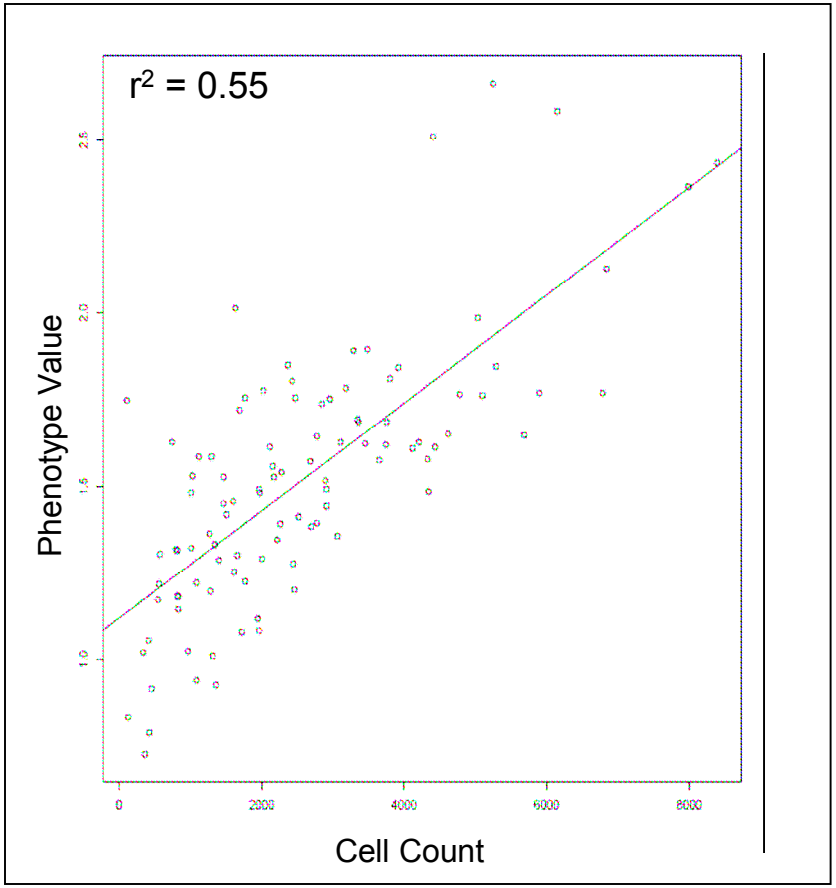
Correlation of Repeated/Control Samples



BROAD INSTITUTE

Phenotype vs. Cell Count





BROAD INSTITUTE

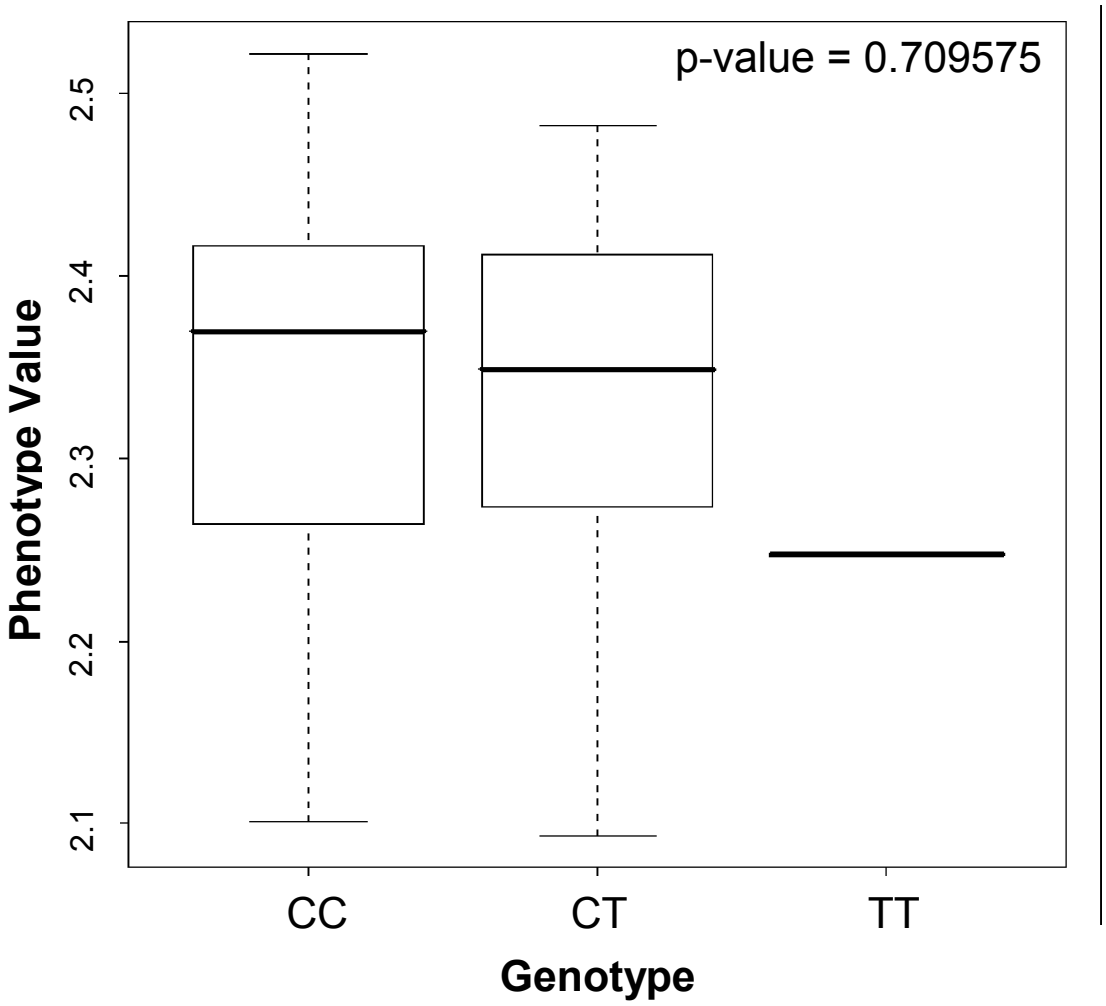
Courtesy of Roman Yelensky

Markers that Passed QC in CEU

CD19	**CD20**
IL6R	IgD
HLA-DQ	**HLA-DR**
CD227	**CD21**
CD40	**IgG**
IgM	CD58
CD95	**CD86**
CD80	

BROAD INSTITUTE

Correlate Marker Expression with Genetic Variation



Attempt to validate a known correlation between CD40 protein expression and dose of the rs1883832^{C} allele (in Kozak sequence)

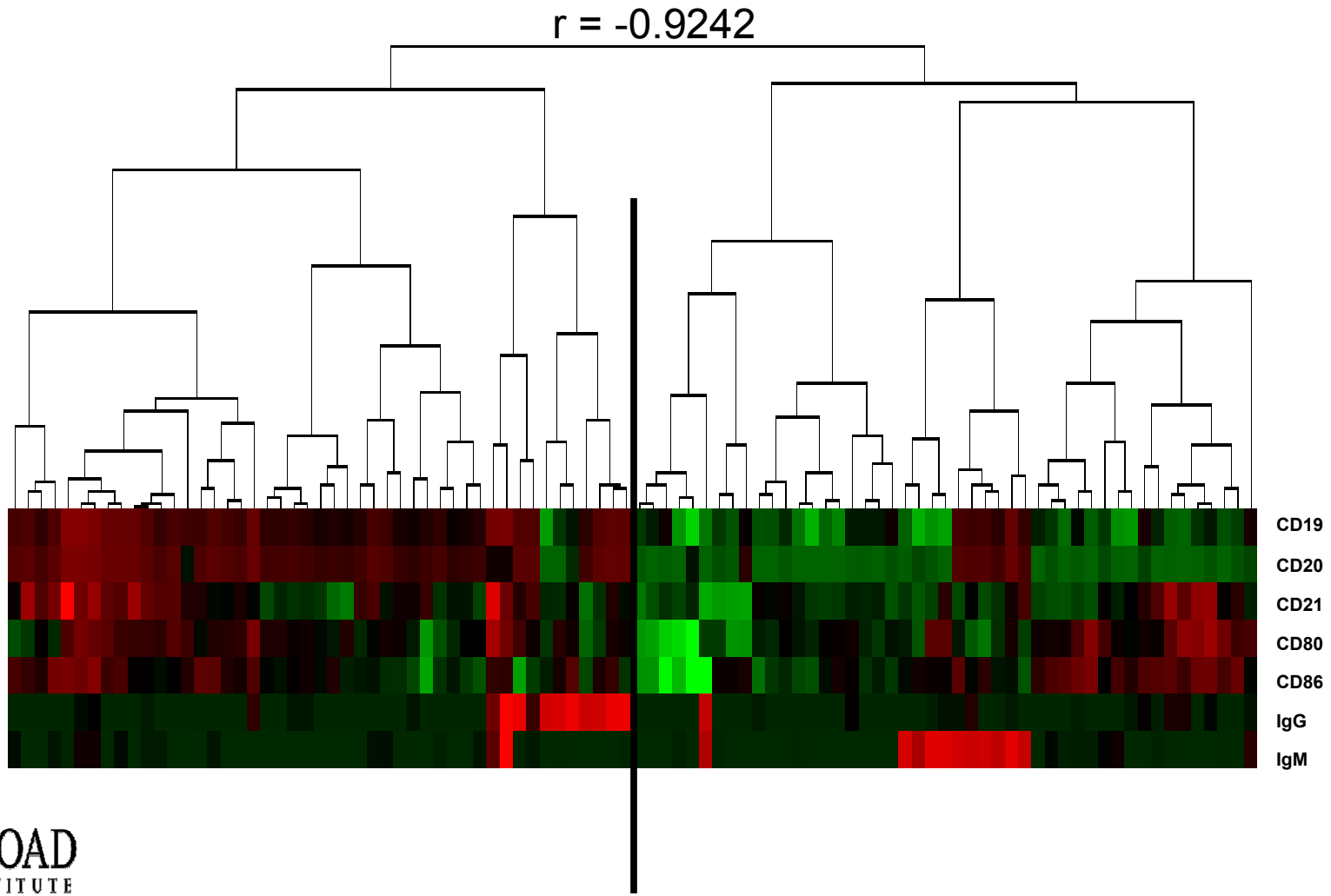
C allele = susceptibility to Grave's disease, high protein expression (*in vitro* & B cells *ex vivo* – 23 individuals)

T allele = low expression

BROAD INSTITUTE

Courtesy of Roman Yelensky
Jacobson *et al., Endocrinology* 2005, 146:2684

Define poly-marker immunological expression profiles



BROAD INSTITUTE

Conclusion

- Completed quality control analysis
  - Defined a curated marker set
- Sub-populations of cell lines that share particular immunological profiles may exist

Future Work

- Further define poly-marker immunological expression profiles
- Further correlate marker expression phenotype with genetic variation
- Correlate marker expression phenotype with other relevant phenotypes

Acknowledgements

Philip De Jager
Elizabeth Rossin
Roman Yelensky

Angela Brunache
Bruce Birren
Sarah Calvo

Edwin Choy
Jackie Slavik
Chavie Wolfish
Sasha Bonakdar
David Altshuler
David Hafler

Eric Lander
Broad Institute
Program in Medical and Population Genetics
Inflammatory Disease Research Group

BROAD INSTITUTE