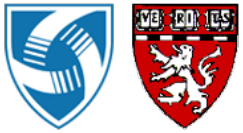
Epigenetic control of genetics: the impact of epigenome on mutation

Shamil Sunyaev



Division of Genetics
Department of Medicine
Brigham and Women's Hospital / Harvard Medical School

Broad Institute of M.I.T. and Harvard

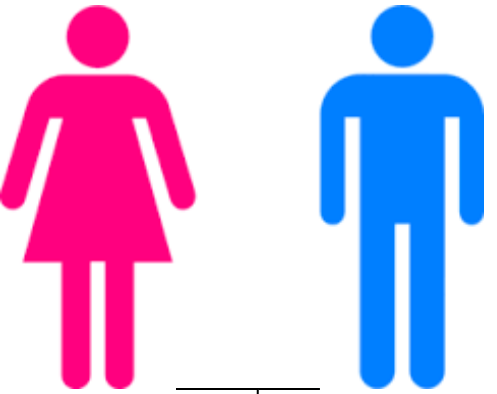
Complex relationship between epigenetics and genetics

The field is interested in the effect of genetic variation on epigenetic features: (favorite_feature)QTL studies.

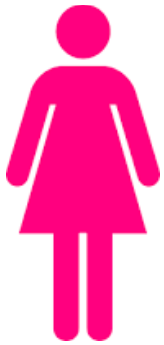
We are interested in the effect of epigenetic features on genetic variation via control of mutation rate.

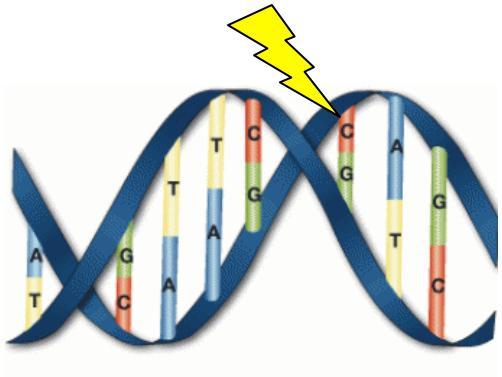
Data on *de novo* mutations

Germ-line mutations

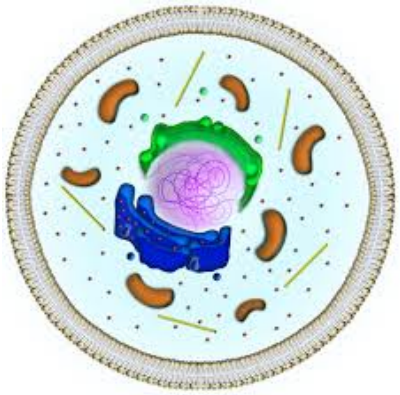


Change in DNA sequence

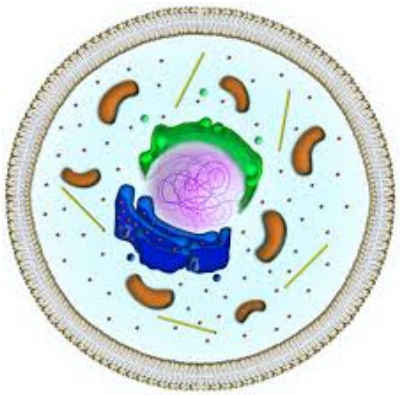


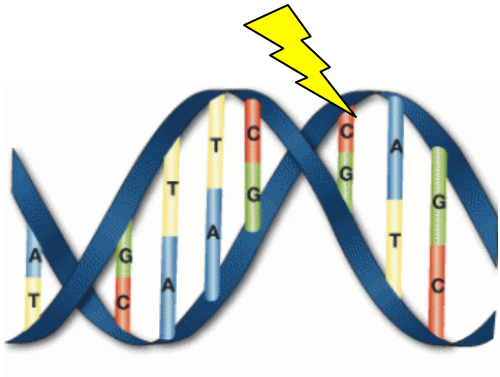


Somatic mutations



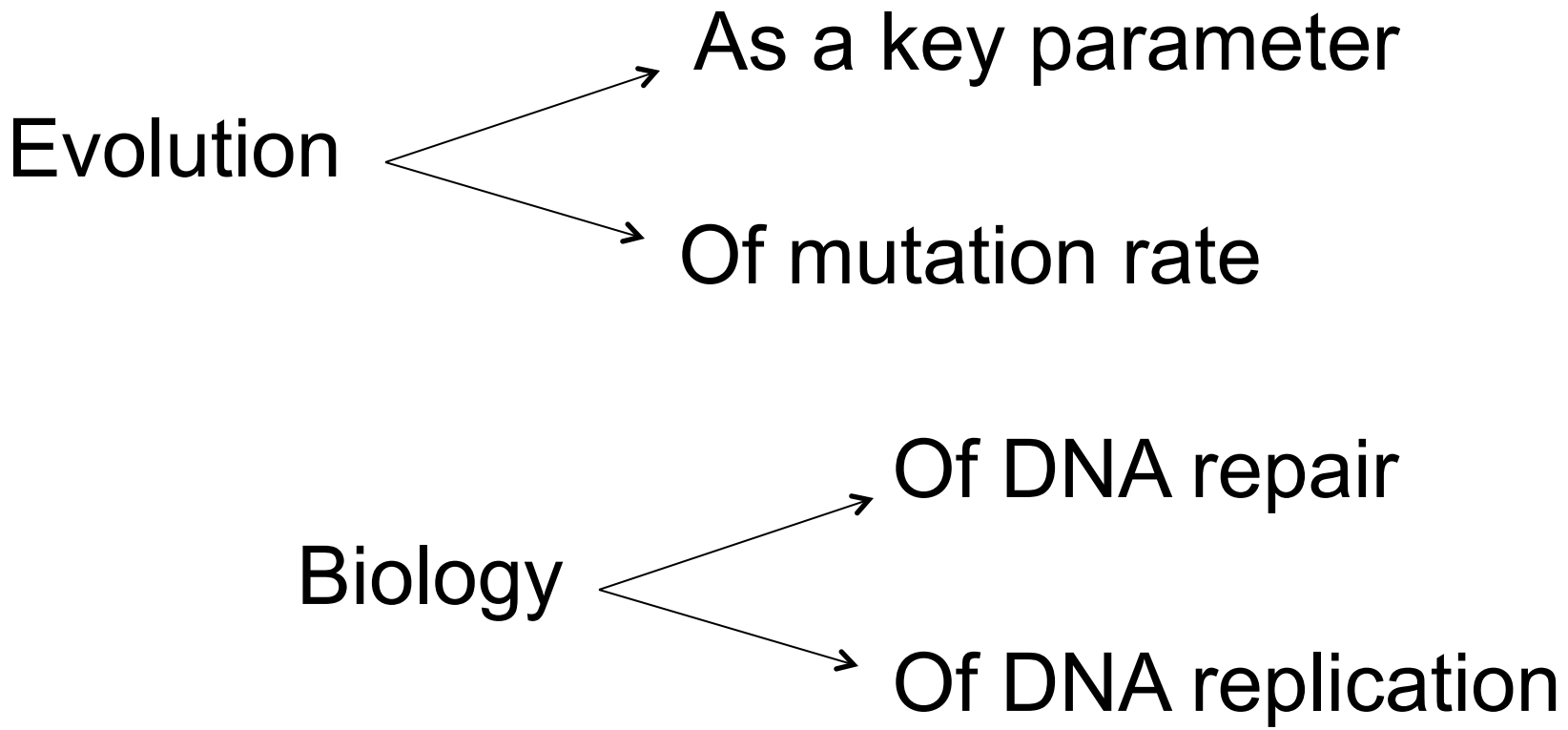
Change in DNA sequence



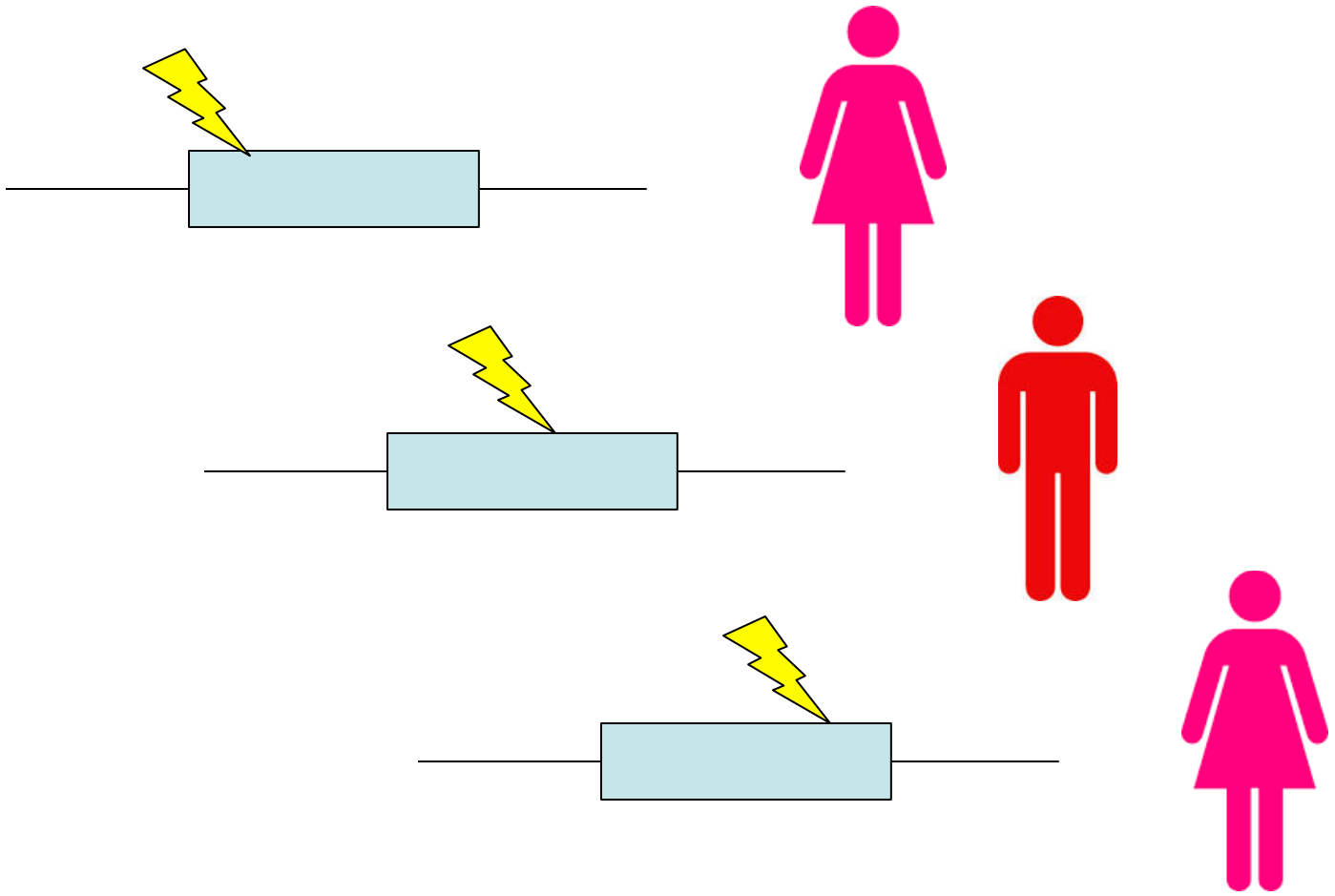


Why is this of any interest?

Statistical / medical genetics ——— Methods



Gene mapping by recurrence



Possible approaches that do not involve controls

Estimate genomic mutation rate using independent samples

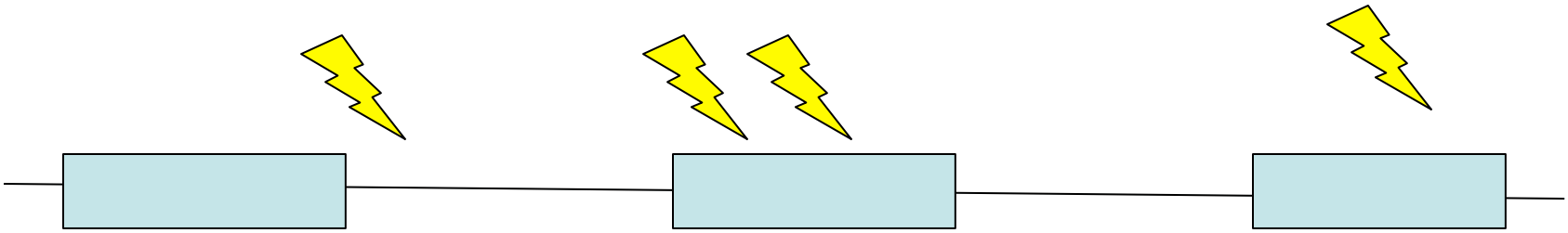
Evaluate probability to observe recurrent events in a given gene

Correct for multiple testing

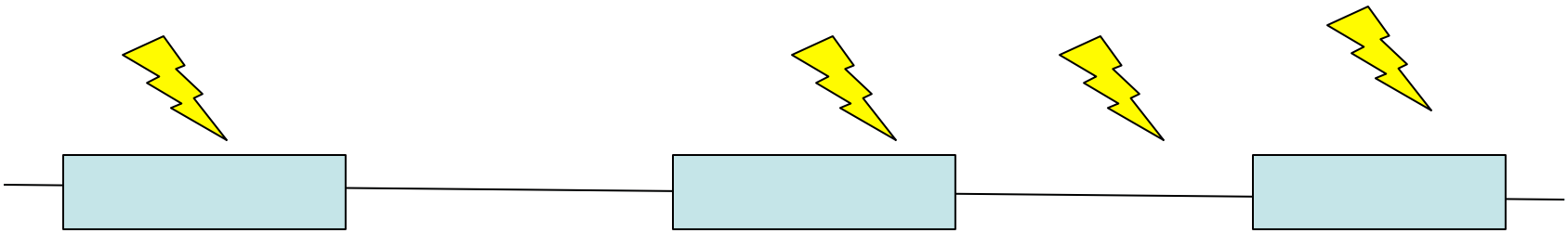
PROBLEM: heterogeneity among samples

Possible approaches that do not involve controls

Real data

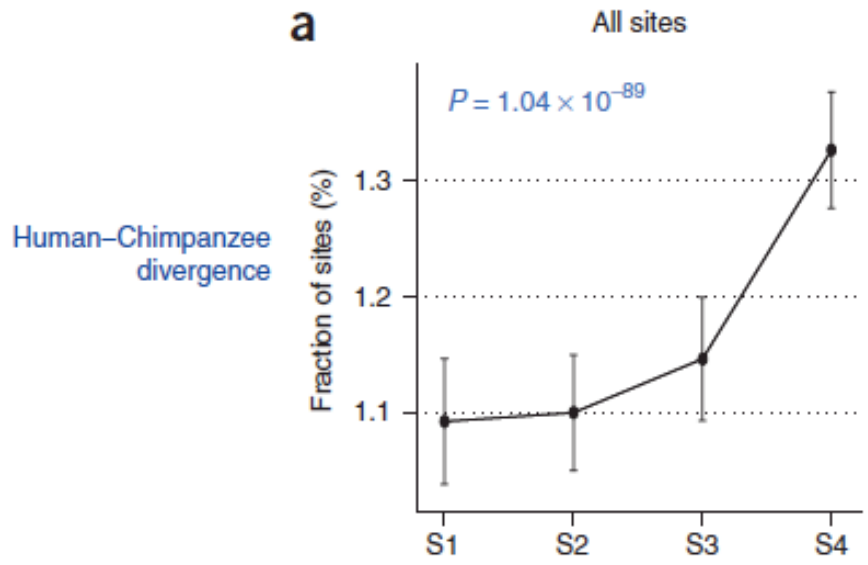


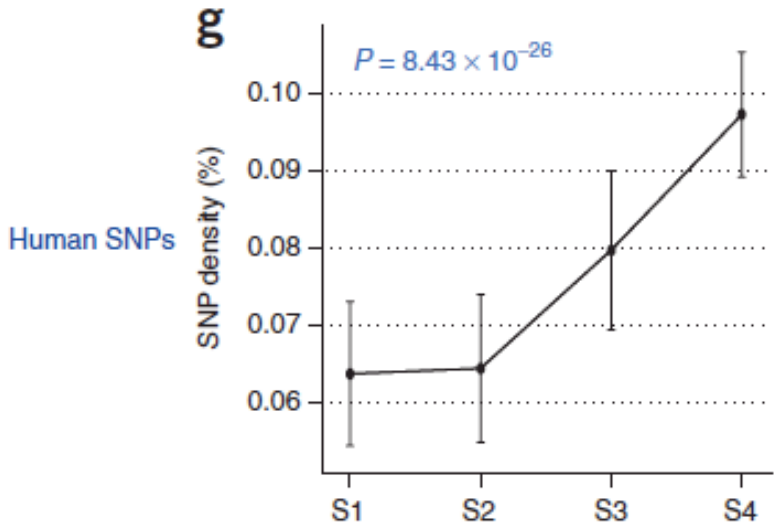
Random permutations



PROBLEM: heterogeneity of mutation rate along the genome

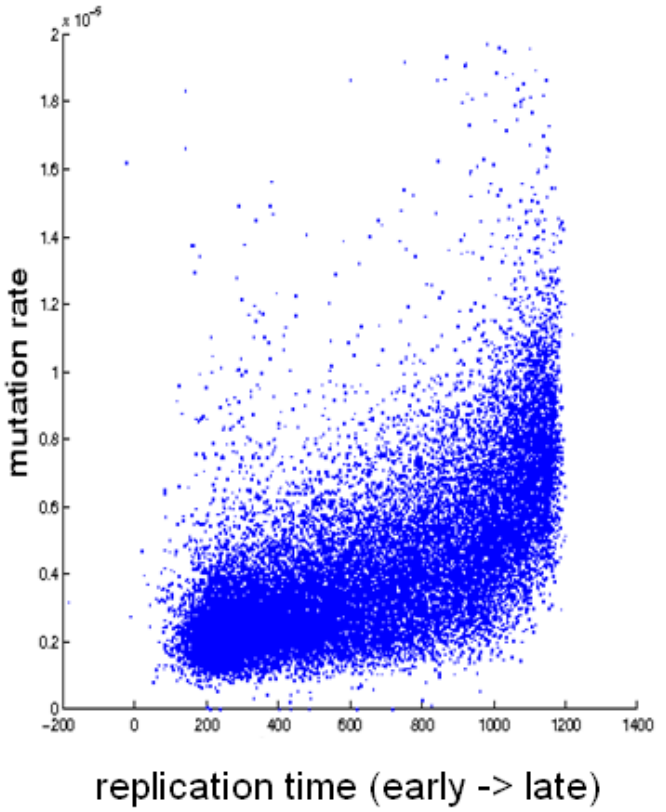
Germ line mutation rates are associated with replication timing

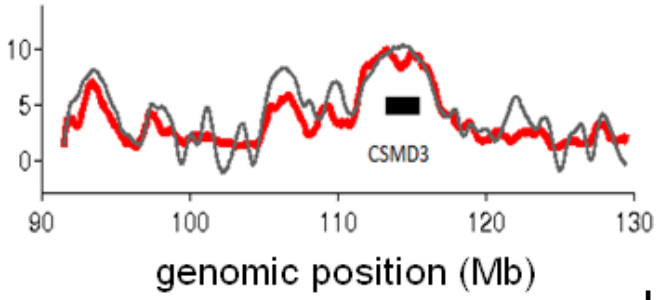




Stamatoyannopoulos, Adzhubei et al., Nature Genetics 2009

Somatic cancer mutation density is associated with replication timing





Lawrence, *et al.*, *Nature* 2013

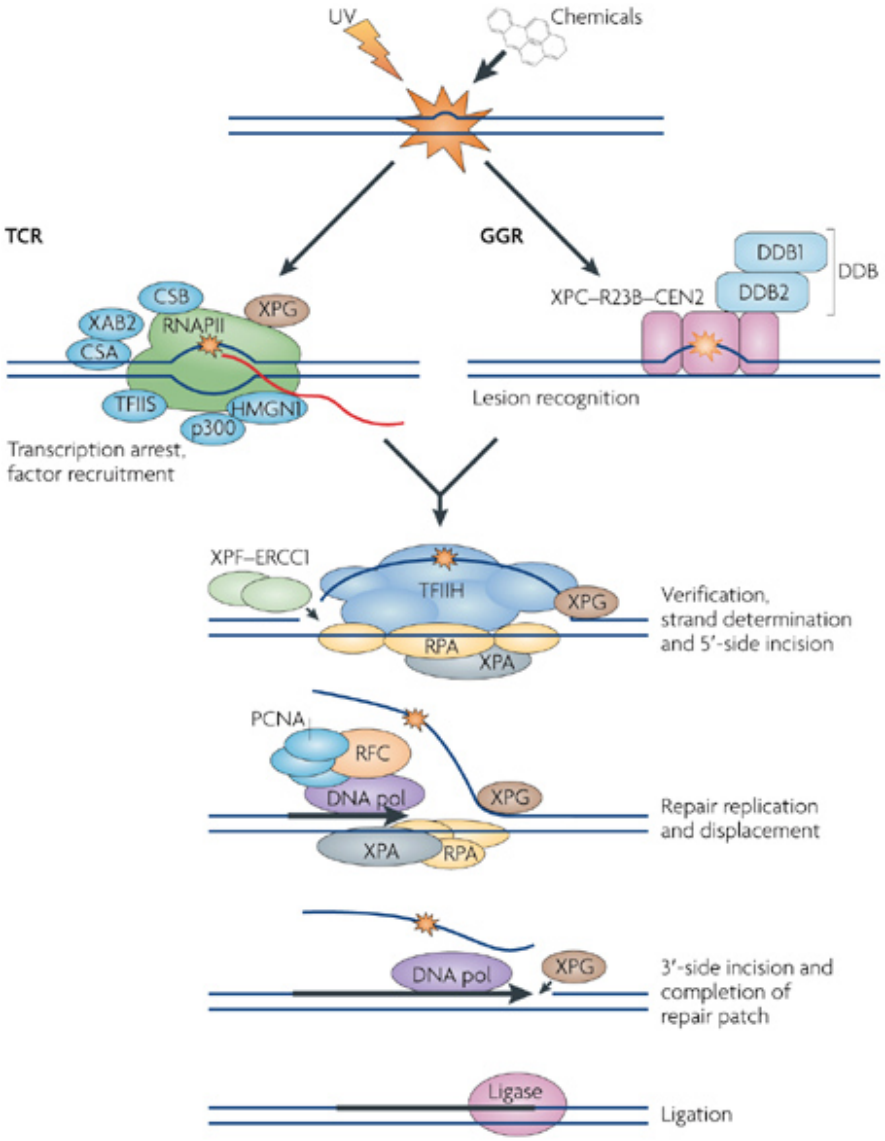
Somatic mutation rate depends on expression

Mutation rate is reduced in transcribed regions compared to intergenic regions

The reduction of mutation rate is proportional to expression level

The effect is attributed to transcription coupled repair (TCR), which is supported by the strand bias

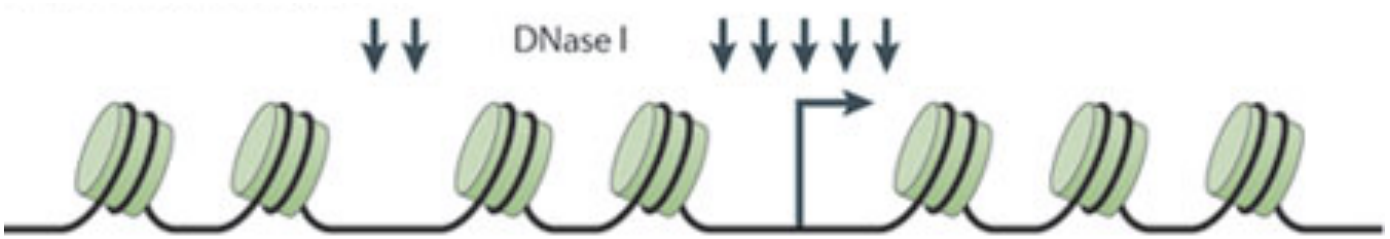
Hanawalt & Spivak, *Nat Rev Mol Cell Biol* 2008



Nature Reviews | Molecular Cell Biology

Regulatory regions and chromatin accessibility

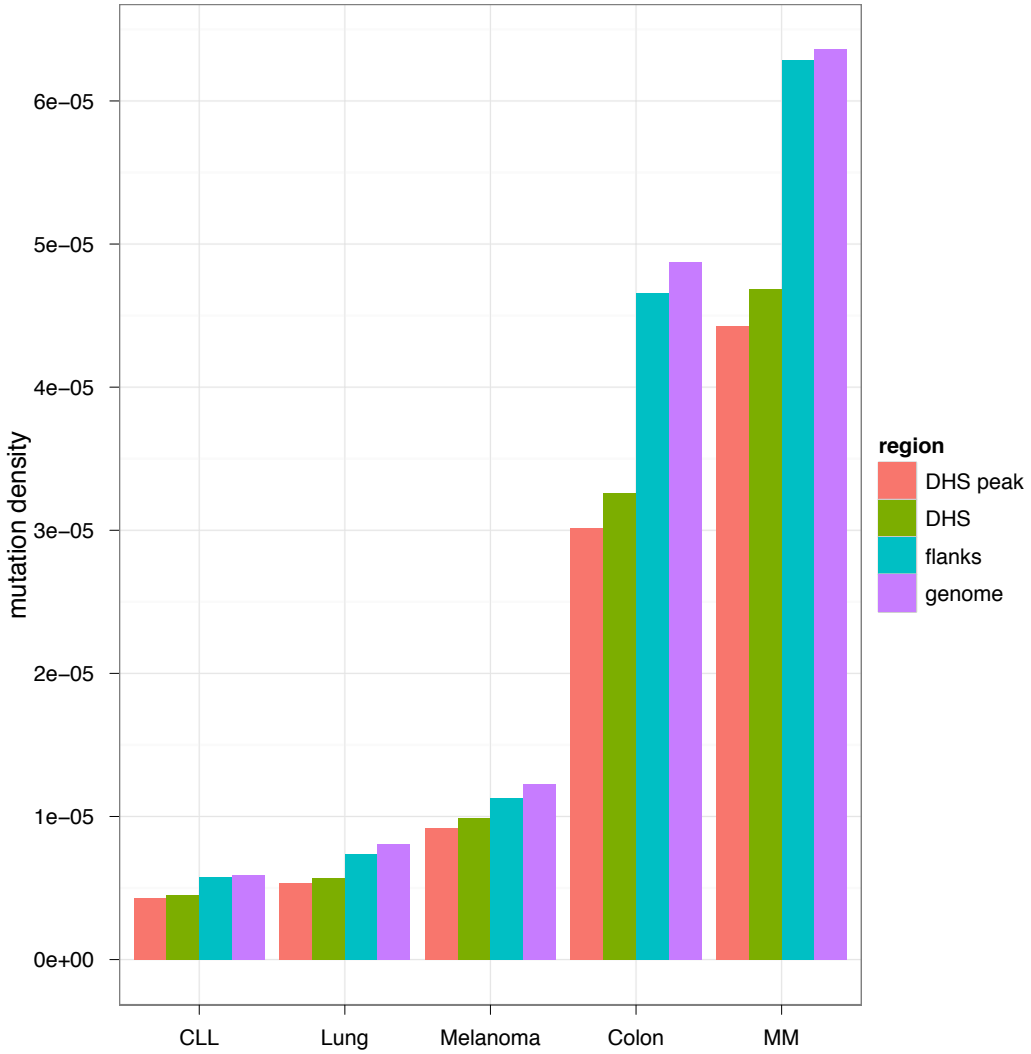
Hypersensitivity to DNase I is a hallmark of regulatory regions



DNase seq is used to map regulatory regions by assessing chromatin accessibility

adopted from Bell et al., NRG 2011

Mutation rate is reduced in regulatory regions marked by accessible chromatin



Polak *et al.*, *Nature Biotech. 2014*

Analysis of melanoma genome sequences

LETTER

doi:10.1038/nature11071

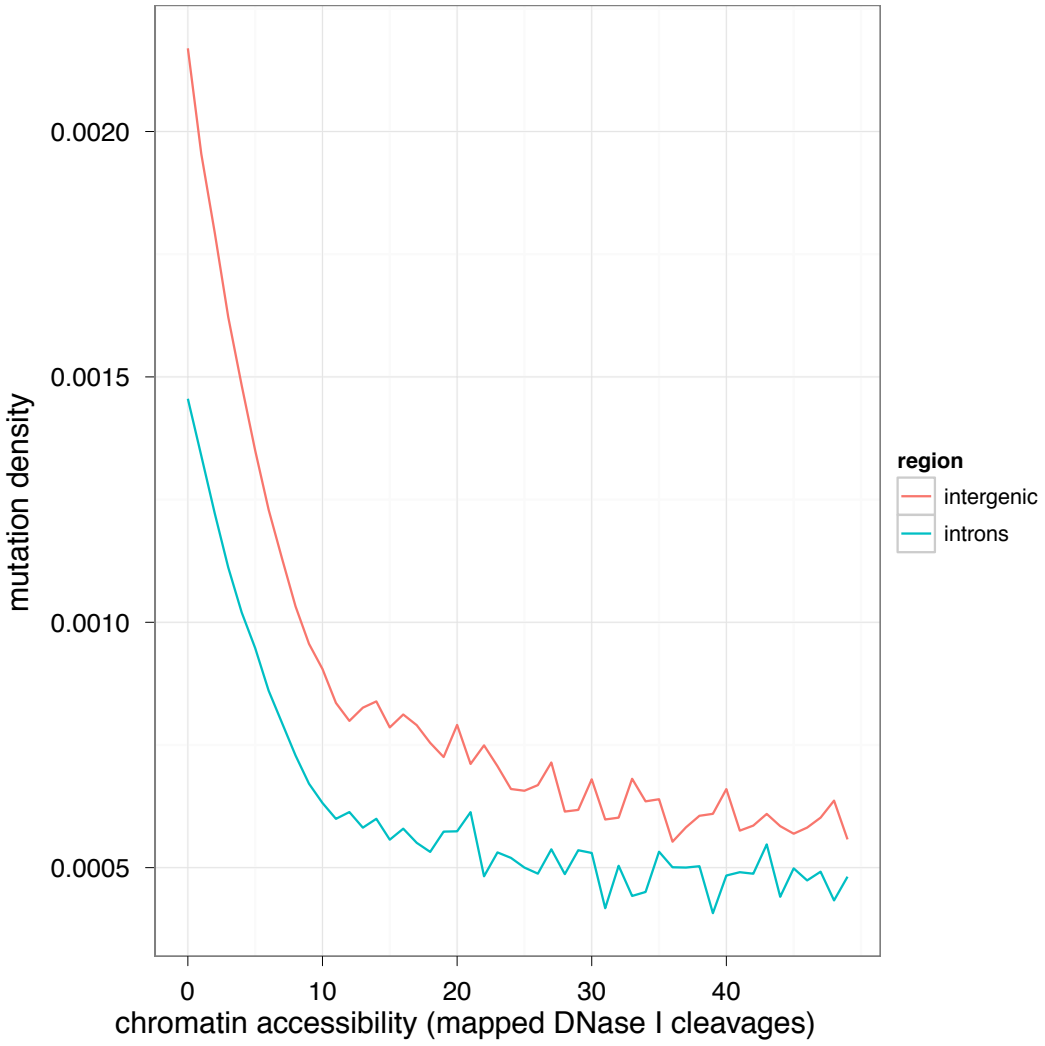
Melanoma genome sequencing reveals frequent *PREX2* mutations

Michael F. Berger[1]†*, Eran Hodis[1]*, Timothy P. Heffernan[2]†*, Yonathan Lissanu Deribe[2]†*, Michael S. Lawrence[1], Alexei Protopopov[2]†, Elena Ivanova[2], Ian R. Watson[2]†, Elizabeth Nickerson[1], Papia Ghosh[2], Hailei Zhang[2], Rhamy Zeid[2], Xiaojia Ren[2], Kristian Cibulskis[1], Andrey Y. Sivachenko[1], Nikhil Wagle[2,3], Antje Sucker[4], Carrie Sougnez[1], Robert Onofrio[1], Lauren Ambrogio[1], Daniel Auclair[1], Timothy Fennell[1], Scott L. Carter[1], Yotam Drier[5], Petar Stojanov[1], Meredith A. Singer[2]†, Douglas Voet[1], Rui Jing[1], Gordon Saksena[1], Jordi Barretina[1], Alex H. Ramos[1,3], Trevor J. Pugh[1,2,3], Nicolas Stransky[1], Melissa Parkin[1], Wendy Winckler[1], Scott Mahan[1], Kristin Ardlie[1], Jennifer Baldwin[1], Jennifer Wargo[6], Dirk Schadendorf[4], Matthew Meyerson[1,2,3,7], Stacey B. Gabriel[1], Todd R. Golub[1,7,8,9], Stephan N. Wagner[10], Eric S. Lander[1,11]*, Gad Getz[1]*, Lynda Chin[1,2,3]†* & Levi A. Garraway[1,2,3,7]*

Reasons:

1) Multiple samples
2) Abundance of mutations
3) Most mutations originate from UV lesions repaired by NER

Continuous dependency on number of DNase I cleavages in melanoma samples



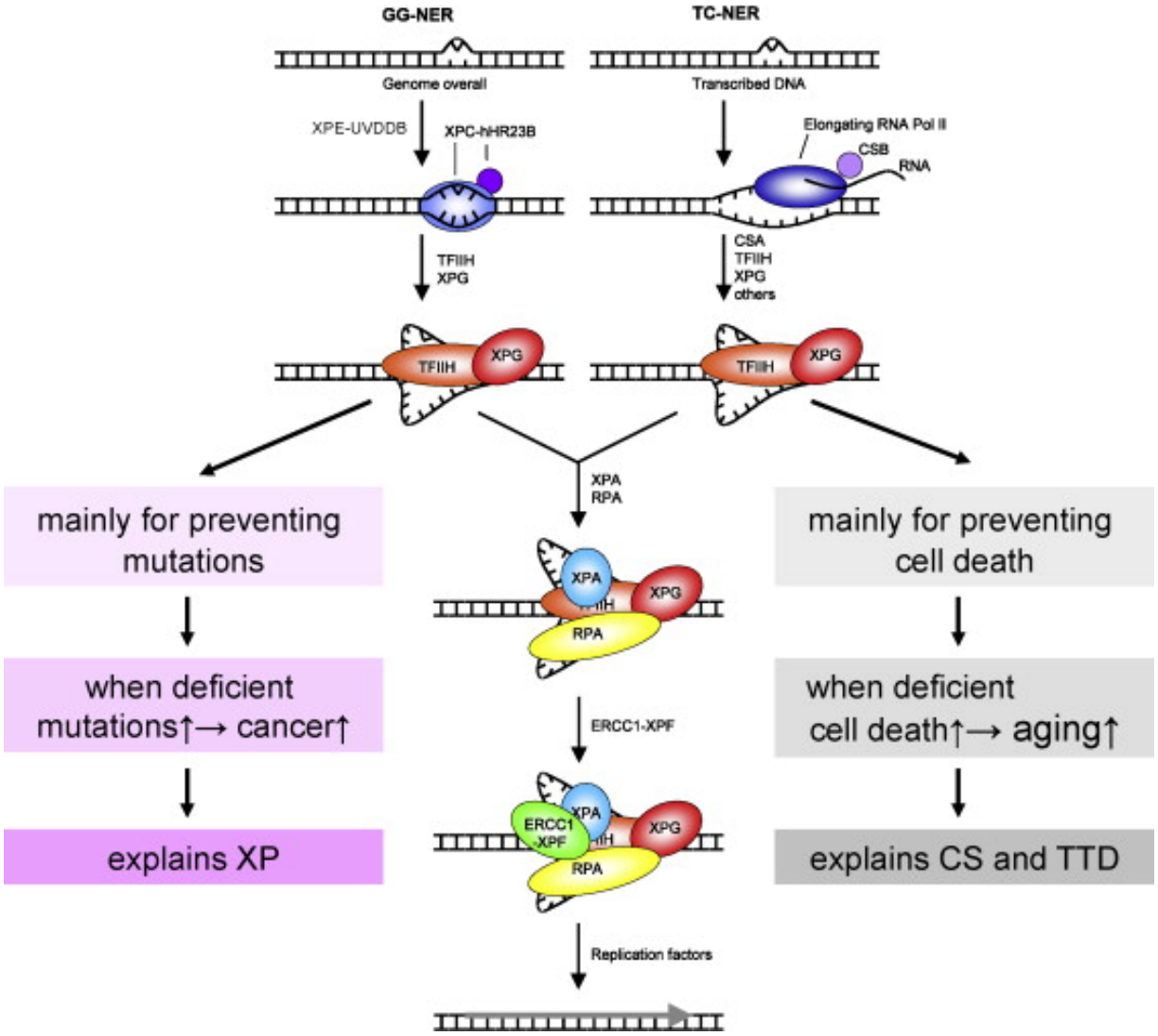
Potential mechanisms

1) Purifying selection in regulatory elements: *seems unlikely because the selection must be assumed much stronger than in coding regions.*

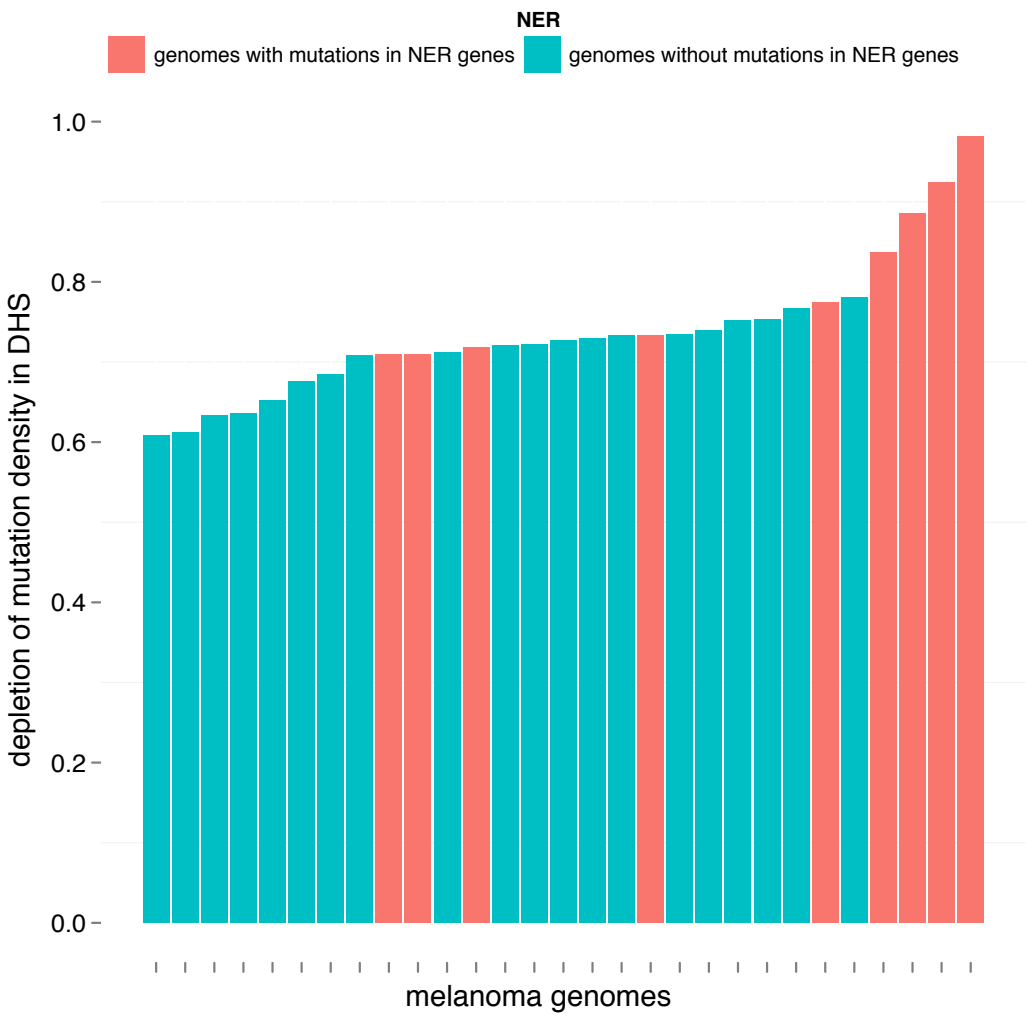
2) Association with replication timing: *seems unlikely due to the scale of the effect and is not supported by multivariate regression analysis*

3) Accessibility to DNA repair: *is the effect associated with NER function?*

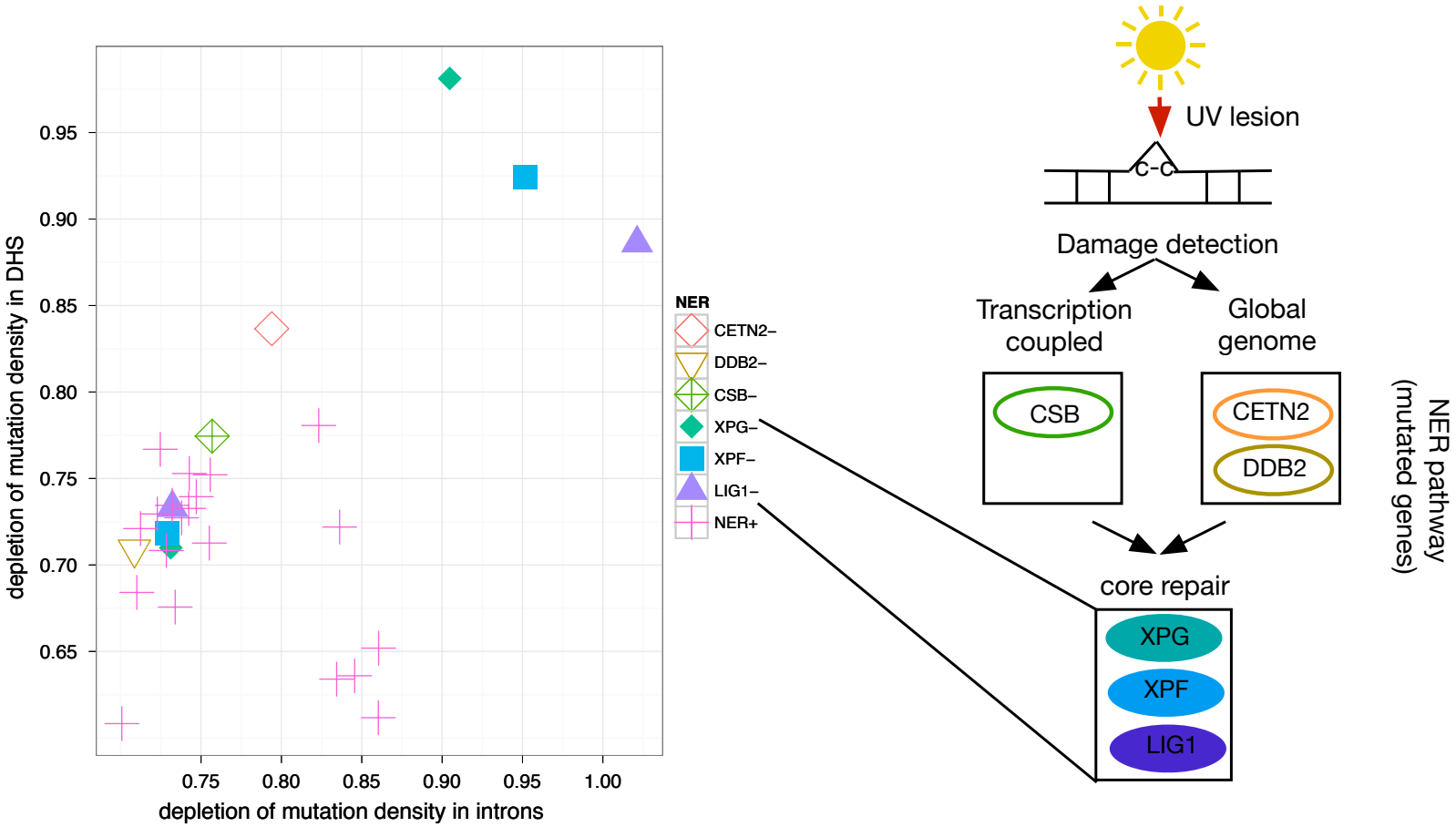
Nucleotide excision repair



Implicating nucleotide excision repair (NER)



Implicating nucleotide excision repair (NER)

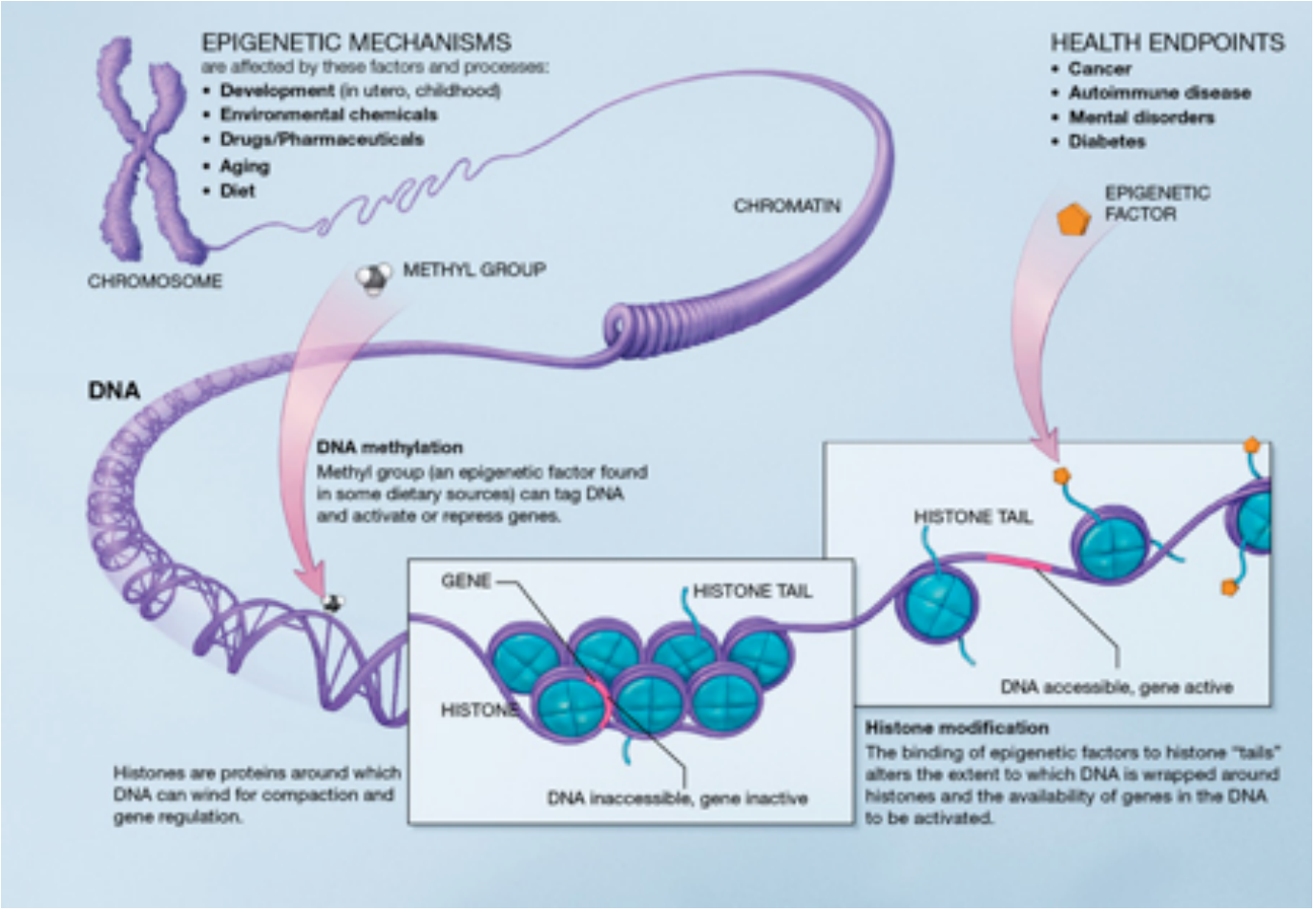


Conclusion

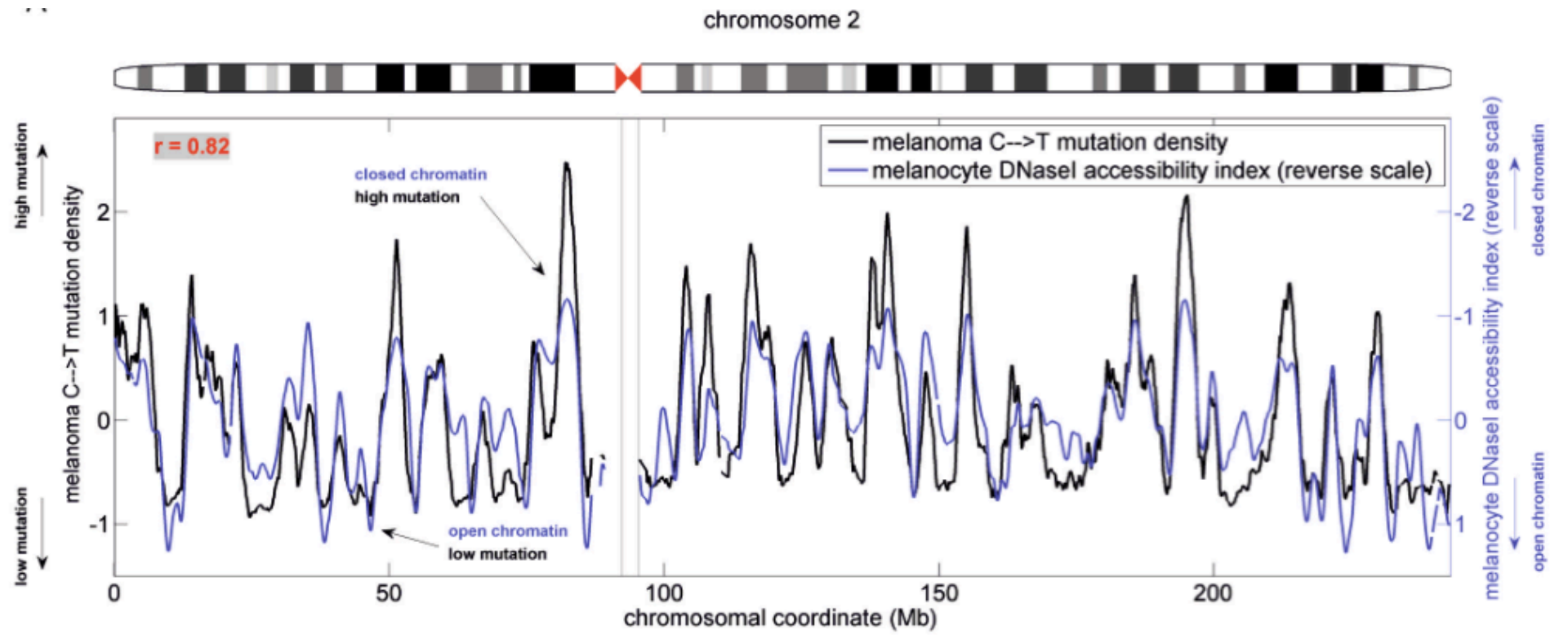
Mutation density is reduced in regulatory regions marked by DHS.

This effect is likely mediated by Global Genome Repair.

Epigenome Roadmap

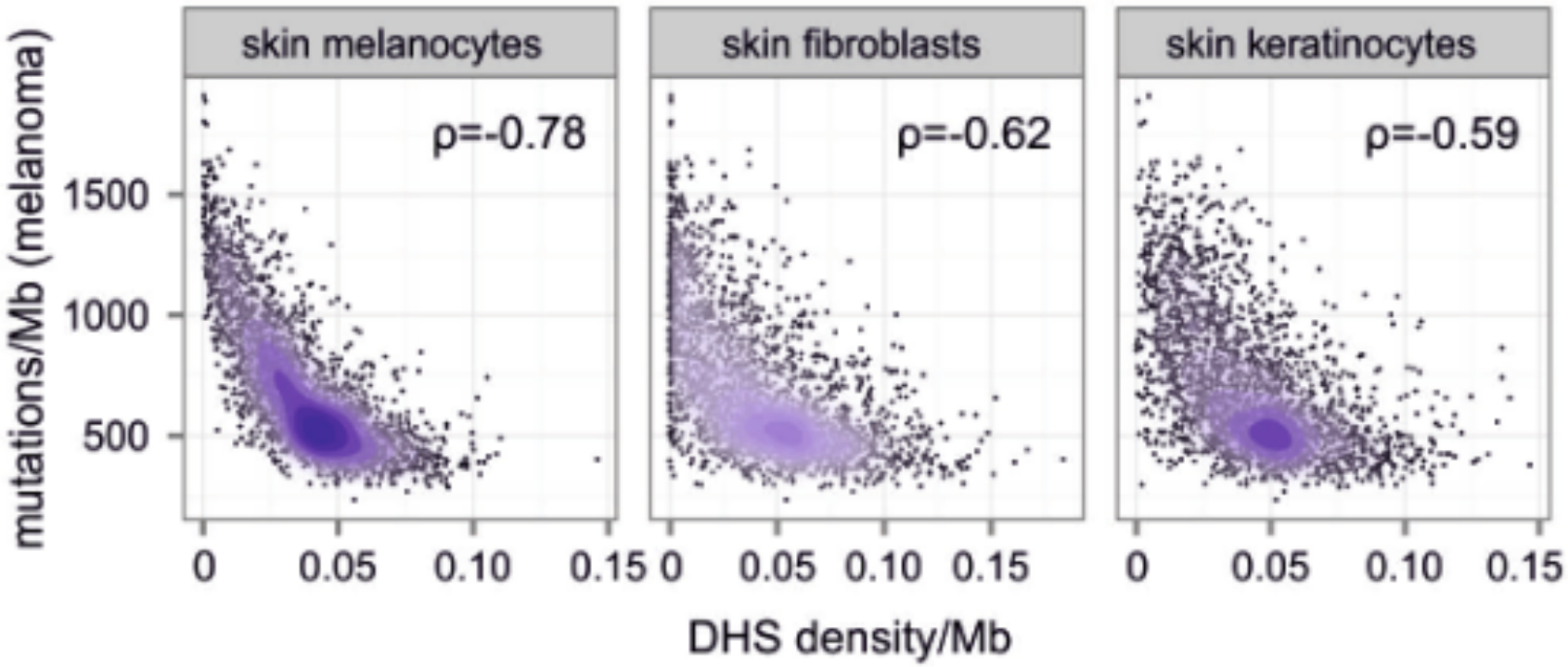


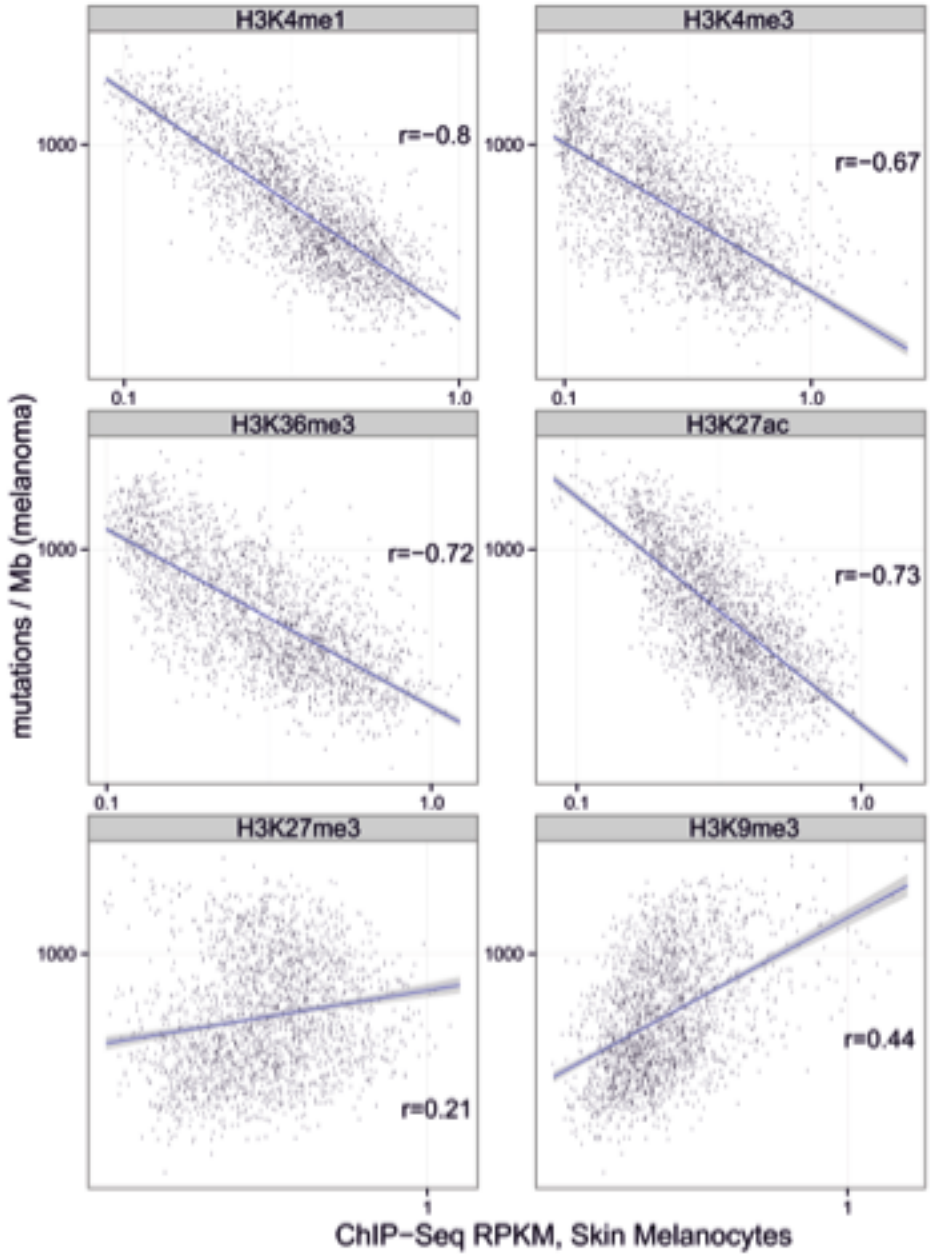
Predicting local mutation rate at 1Mb scale



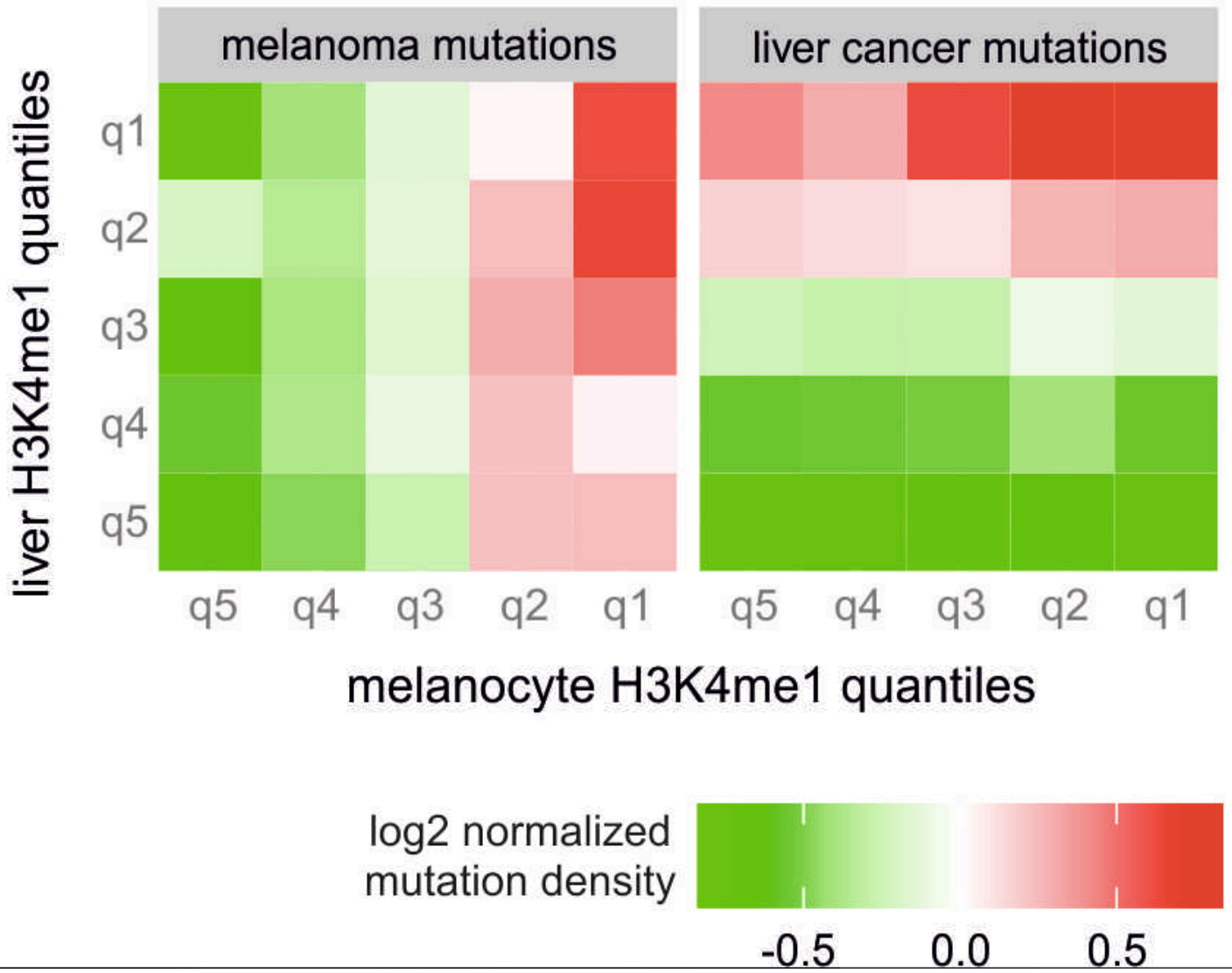
Polak, Karlic *et al.*, *Nature* 2015

Predicting local mutation rate at 1Mb scale

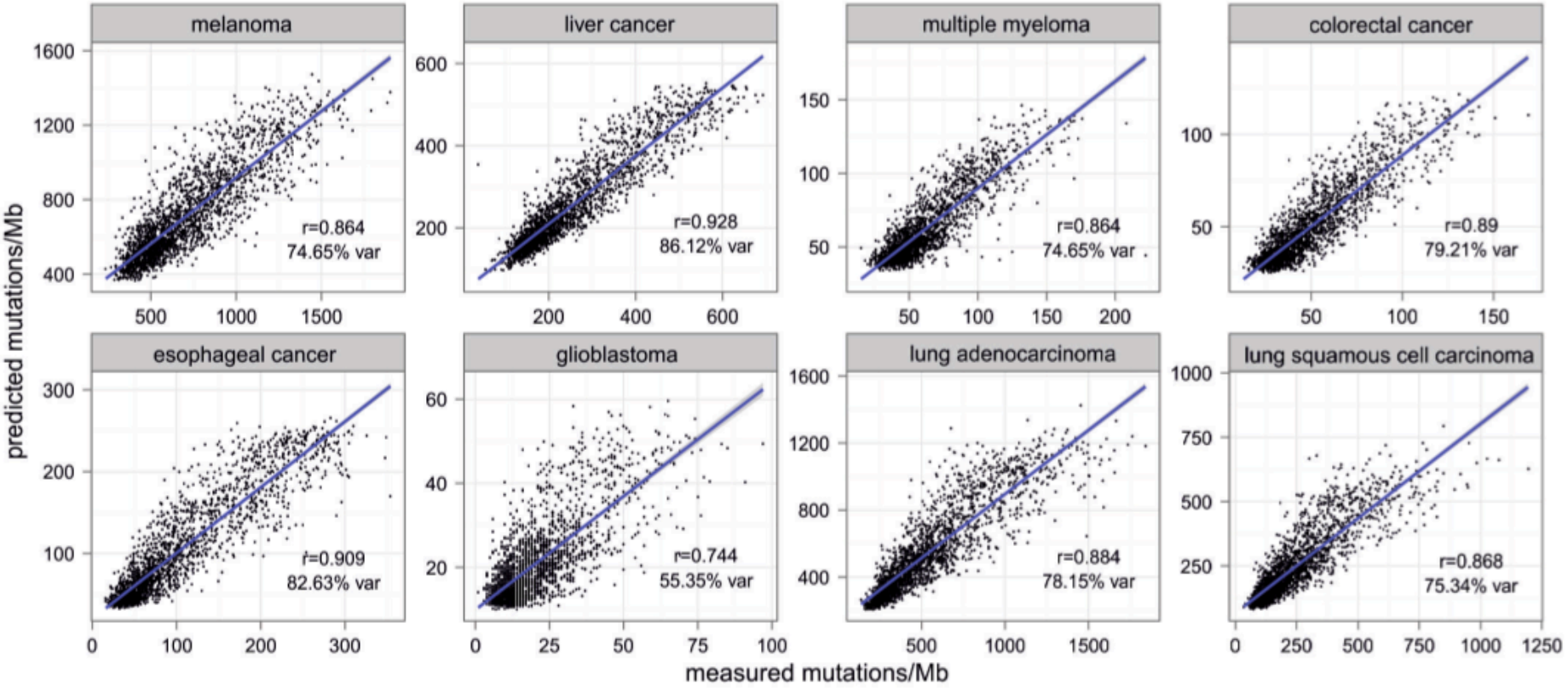


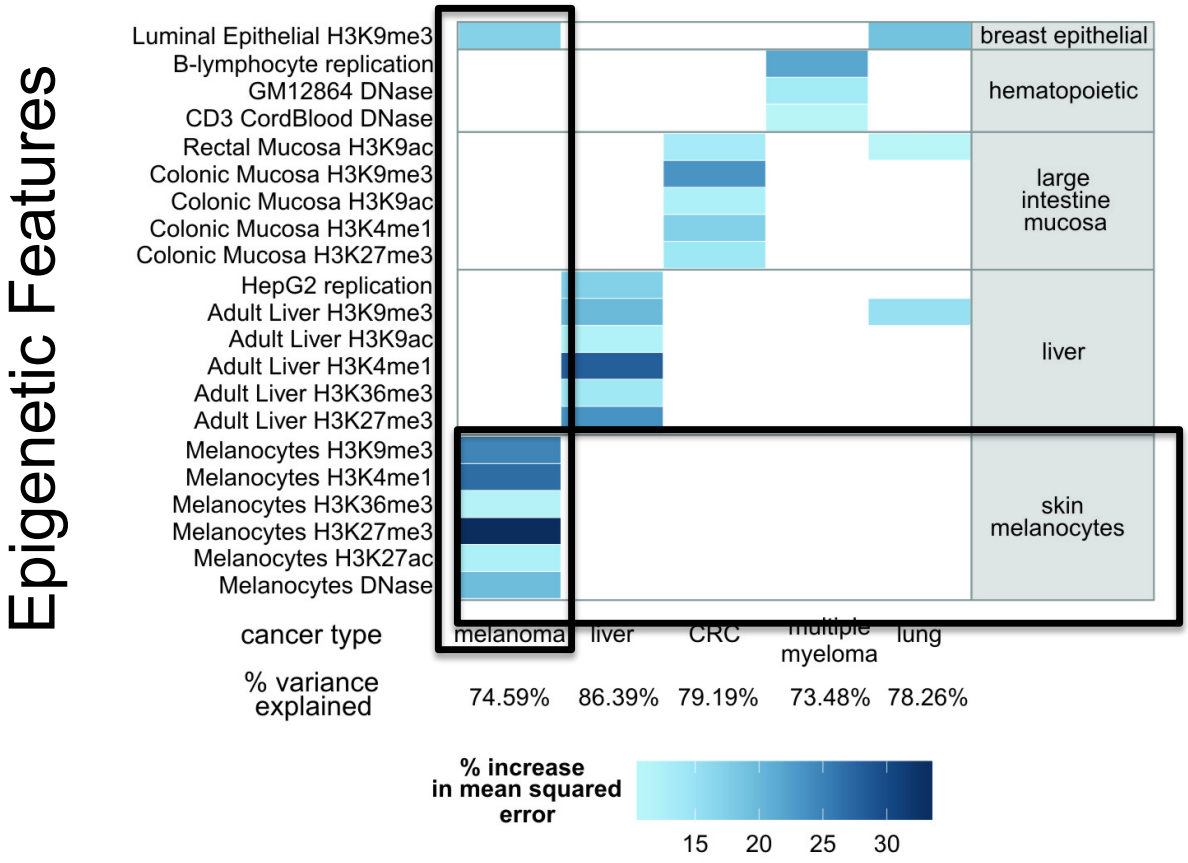

H3K4me1
r=−0.8
H3K4me3
r=−0.67
H3K36me3
r=−0.72
H3K27ac
r=−0.73
H3K27me3
r=0.21
H3K9me3
r=0.44
1000
0.1
1.0
1
mutations / Mb (melanoma)
ChIP−Seq RPKM, Skin Melanocytes

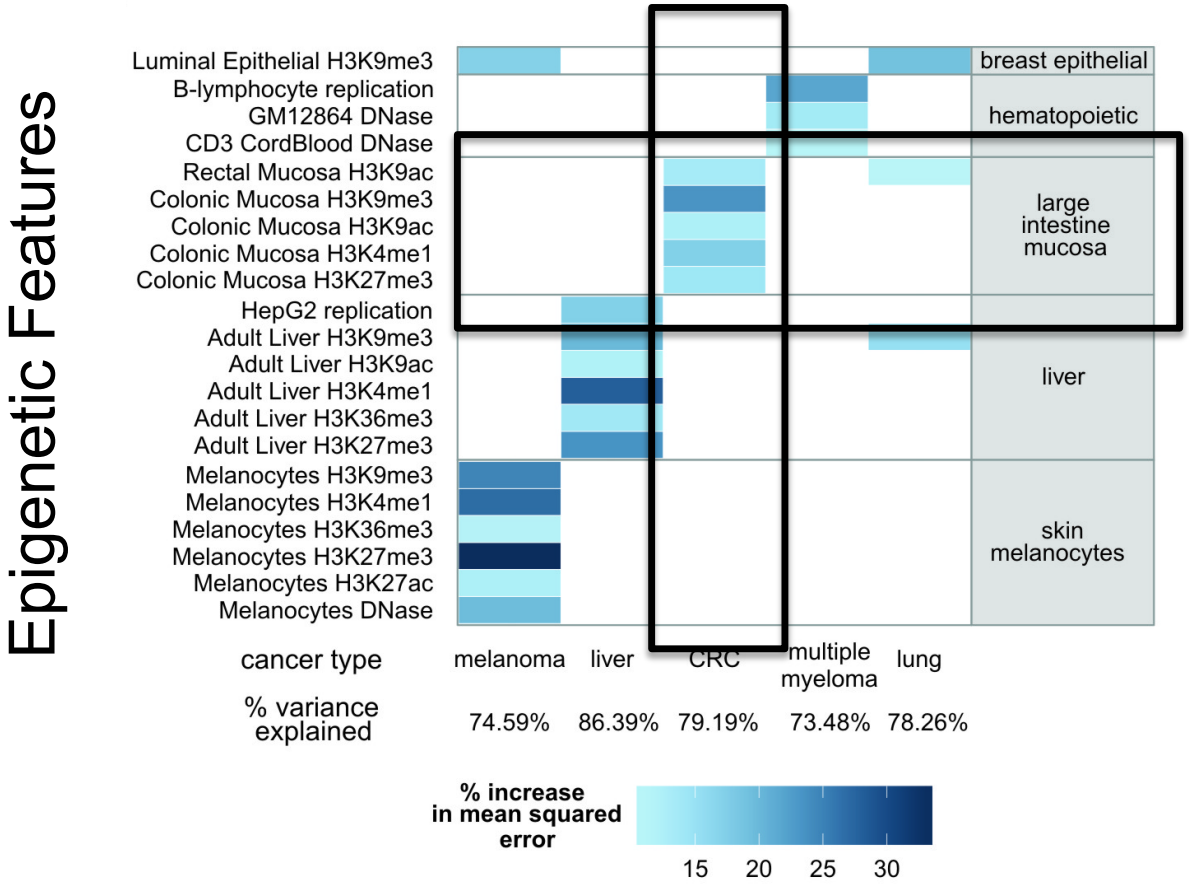
Cell type specificity



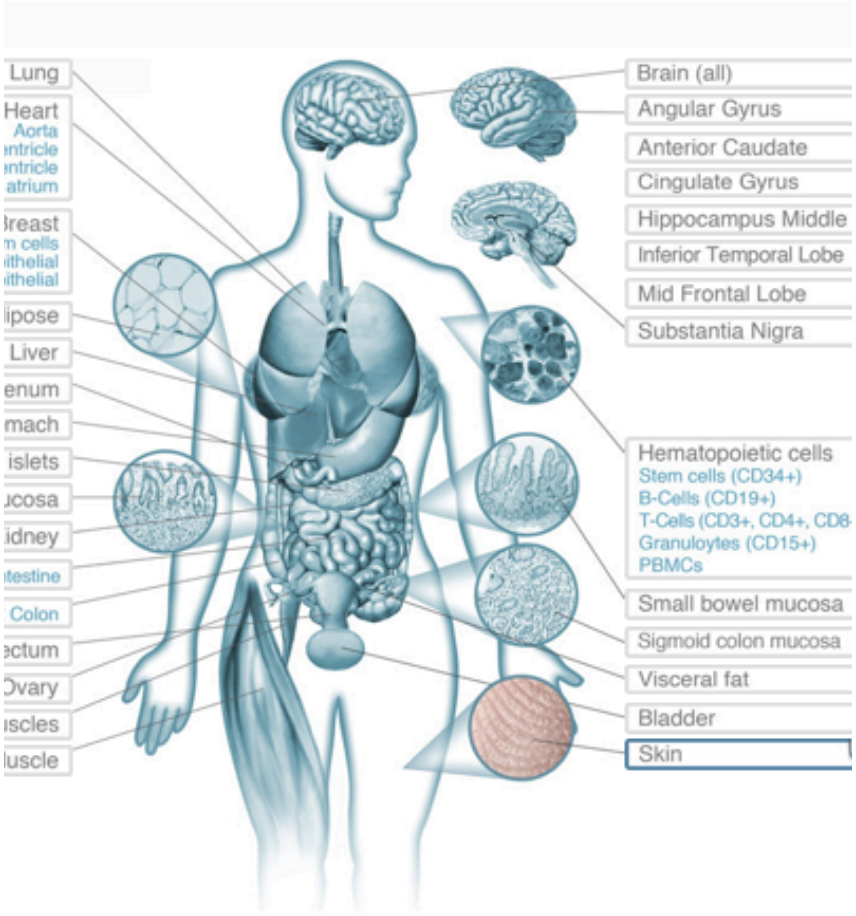
55-86% of regional variation is explained by 184 chromatin tracks from more than 80 tissues

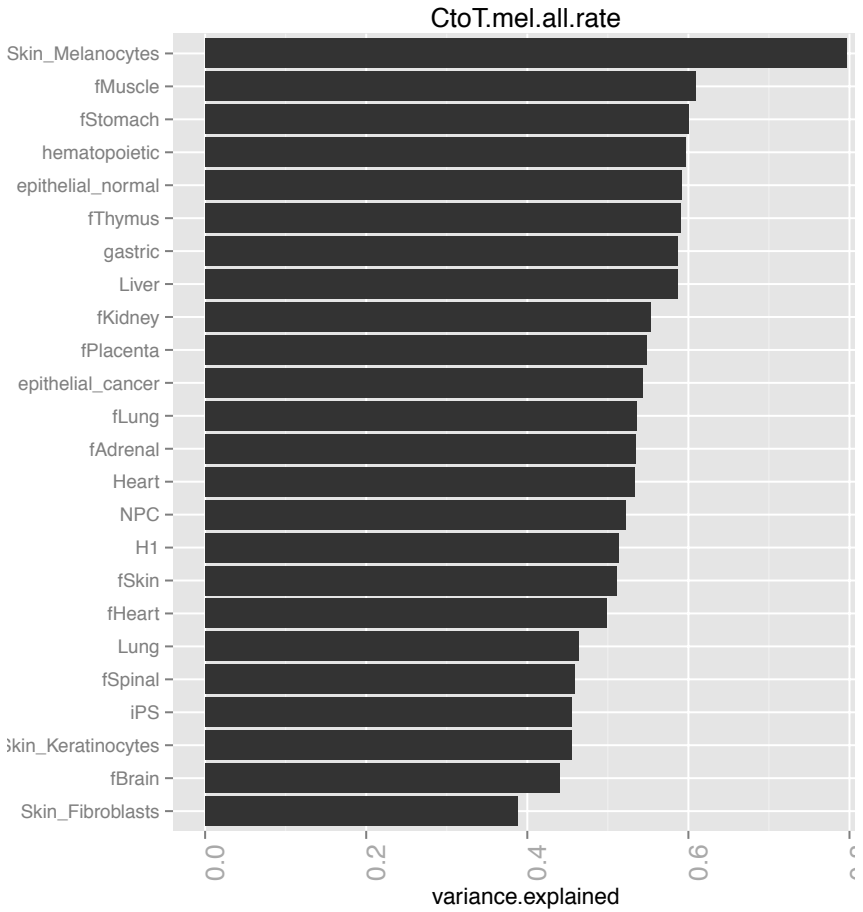



Epigenetic Features
Luminal Epithelial H3K9me3
B-lymphocyte replication
GM12864 DNase
CD3 CordBlood DNase
Rectal Mucosa H3K9ac
Colonic Mucosa H3K9me3
Colonic Mucosa H3K9ac
Colonic Mucosa H3K4me1
Colonic Mucosa H3K27me3
HepG2 replication
Adult Liver H3K9me3
Adult Liver H3K9ac
Adult Liver H3K4me1
Adult Liver H3K36me3
Adult Liver H3K27me3
Melanocytes H3K9me3
Melanocytes H3K4me1
Melanocytes H3K36me3
Melanocytes H3K27me3
Melanocytes H3K27ac
Melanocytes DNase
breast epithelial
hematopoietic
large intestine mucosa
liver
skin melanocytes
cancer type
melanoma
liver
CRC
multiple myeloma
lung
% variance explained
74.59%
86.39%
79.19%
73.48%
78.26%
% increase in mean squared error
15
20
25
30
Cancers

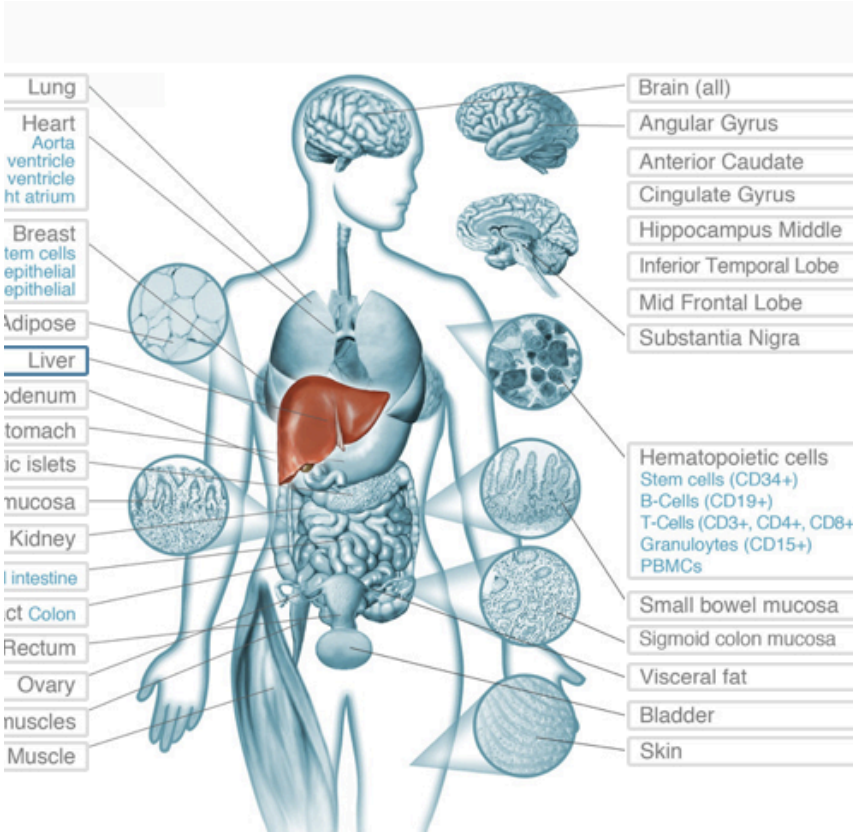

Epigenetic Features
Luminal Epithelial H3K9me3
B-lymphocyte replication
GM12864 DNase
CD3 CordBlood DNase
Rectal Mucosa H3K9ac
Colonic Mucosa H3K9me3
Colonic Mucosa H3K9ac
Colonic Mucosa H3K4me1
Colonic Mucosa H3K27me3
HepG2 replication
Adult Liver H3K9me3
Adult Liver H3K9ac
Adult Liver H3K4me1
Adult Liver H3K36me3
Adult Liver H3K27me3
Melanocytes H3K9me3
Melanocytes H3K4me1
Melanocytes H3K36me3
Melanocytes H3K27me3
Melanocytes H3K27ac
Melanocytes DNase
breast epithelial
hematopoietic
large intestine mucosa
liver
skin melanocytes
cancer type
melanoma
liver
CRC
multiple myeloma
lung
% variance explained
74.59%
86.39%
79.19%
73.48%
78.26%
% increase in mean squared error
15
20
25
30
Cancers

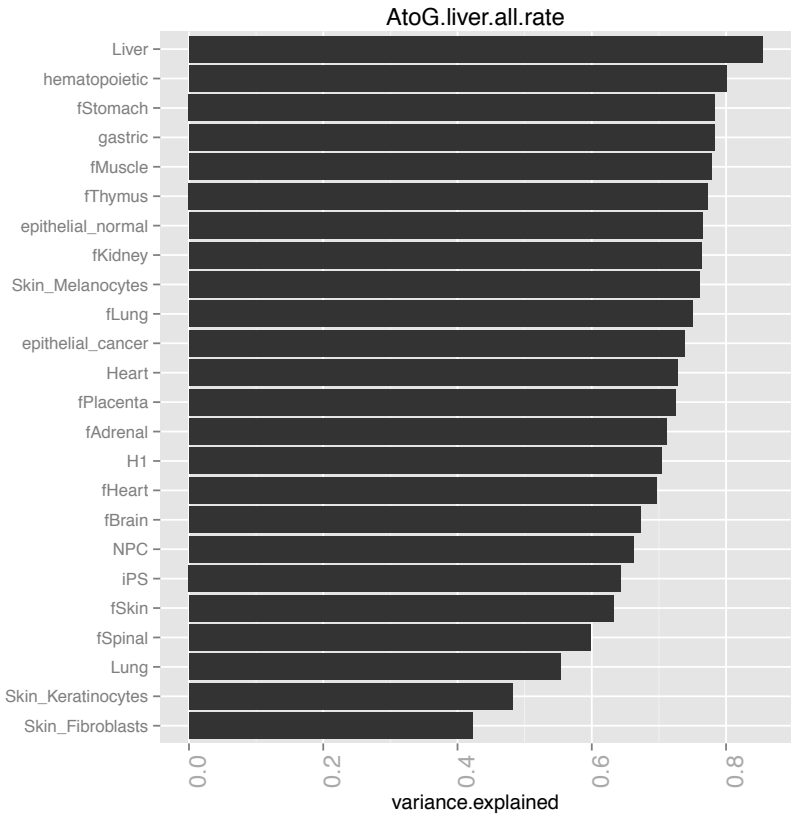
Cell type specificity



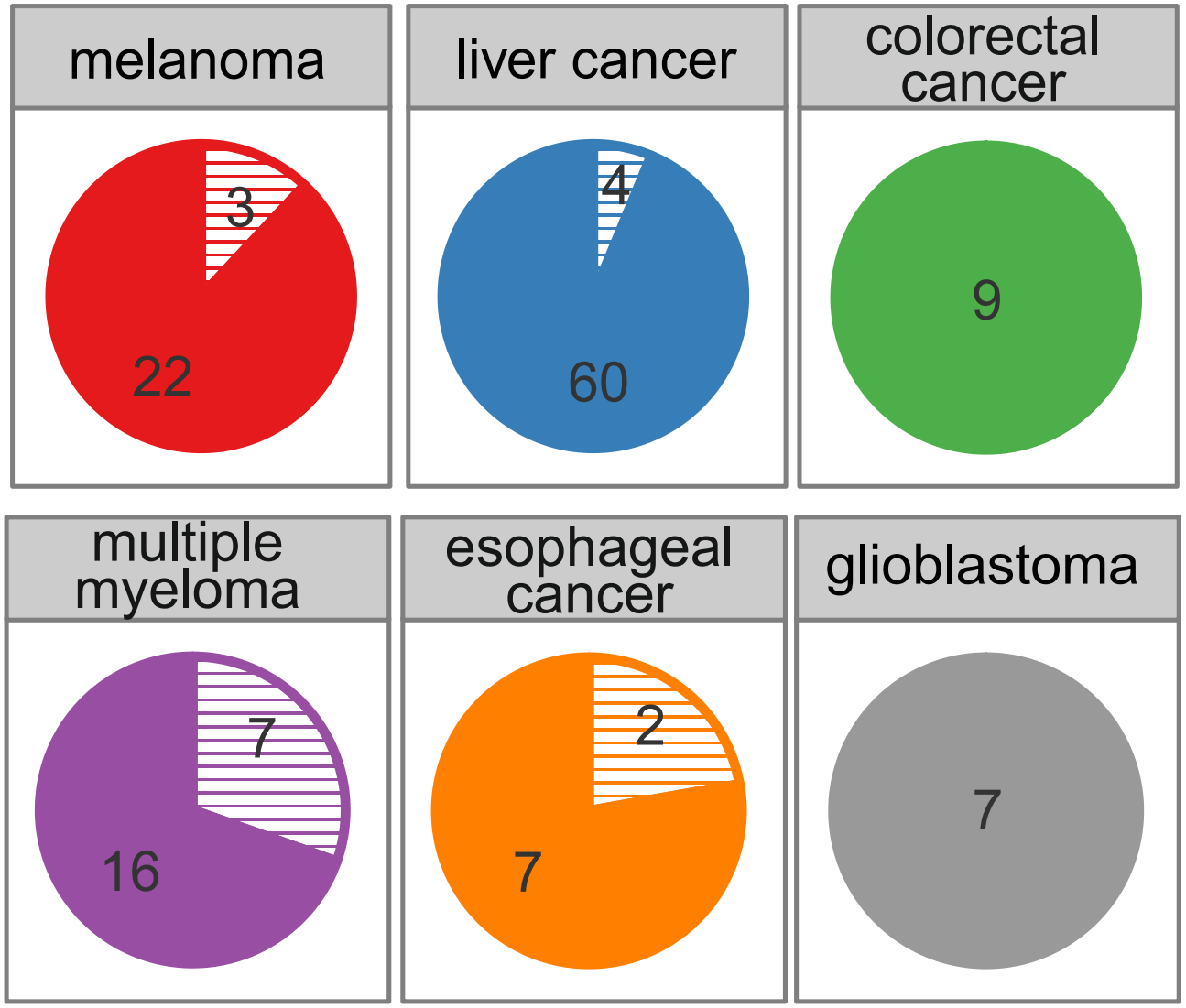


Cell type specificity

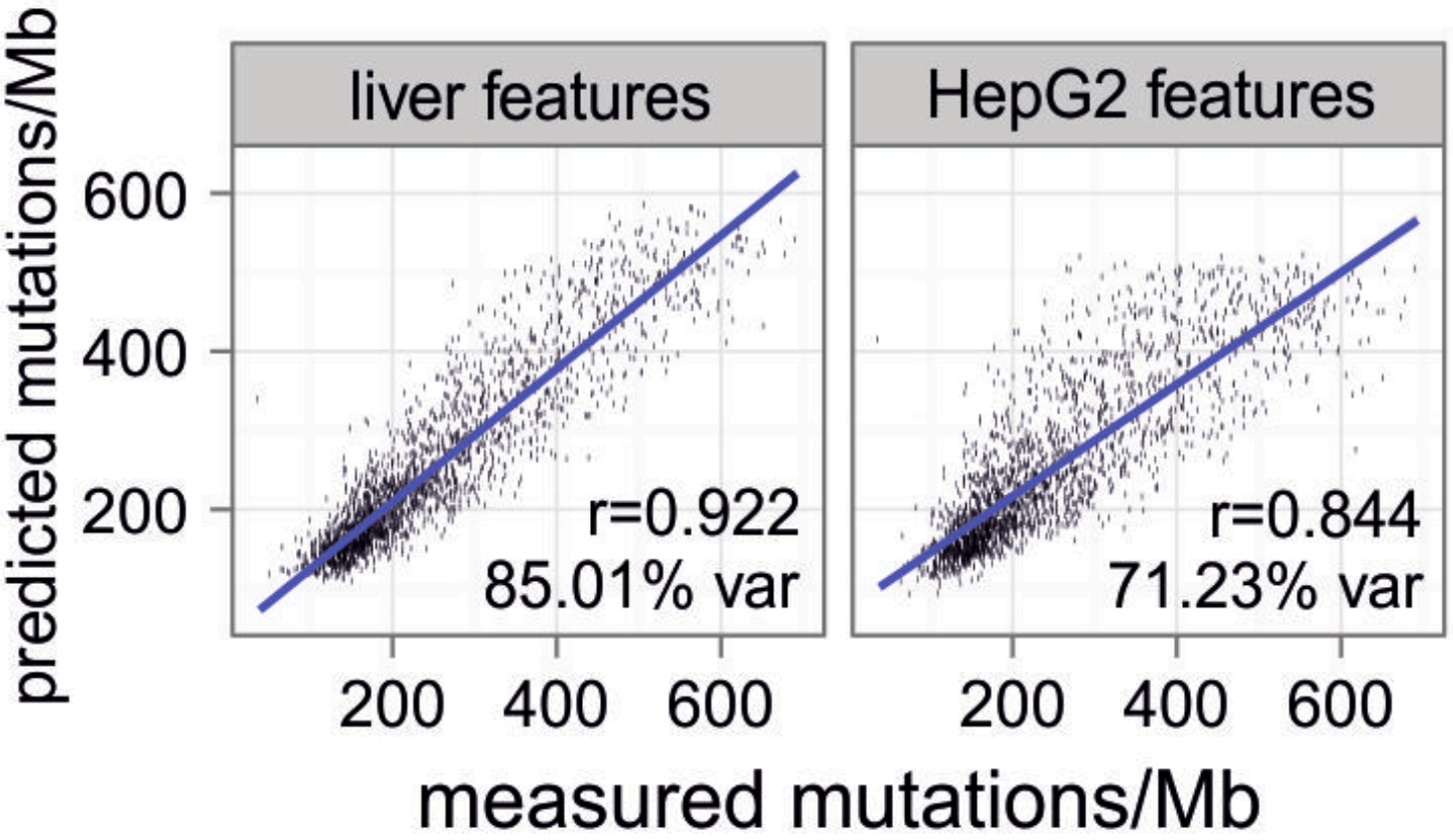




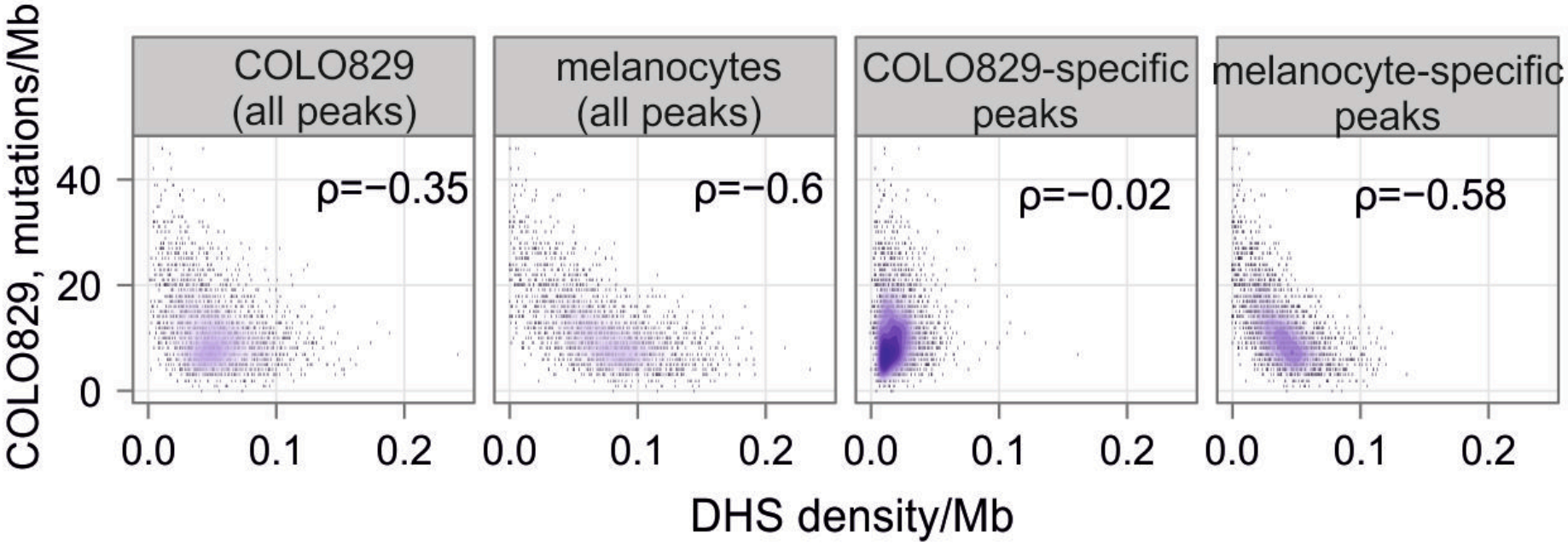
Predicting cell type of origin for 88% of samples



Is chromatin organization of cancer more informative?



Is chromatin organization of cancer more informative?



Conclusion

Mutation density at 1Mb scale is strongly associated with the chromatin organization.

This association is highly specific with respect to cell of origin.

Cancer genome sequence has enough information to predict cell of origin.

Acknowledgments

Acknowledgments

The lab: Daniel Jordan, Ivan Adzhubei, Paz Polak, Tobias Lenz, Mashaal Sohail, Dana Vuzman, Daniel Balick, Sung Chun, Jae-Hoon Sul, Chris Cassa, Sebastian Akle, David Radke

Collaborators:
John Stamatoyannopoulos, Bob Thurman, Rosa Karlic, Amnon Koren