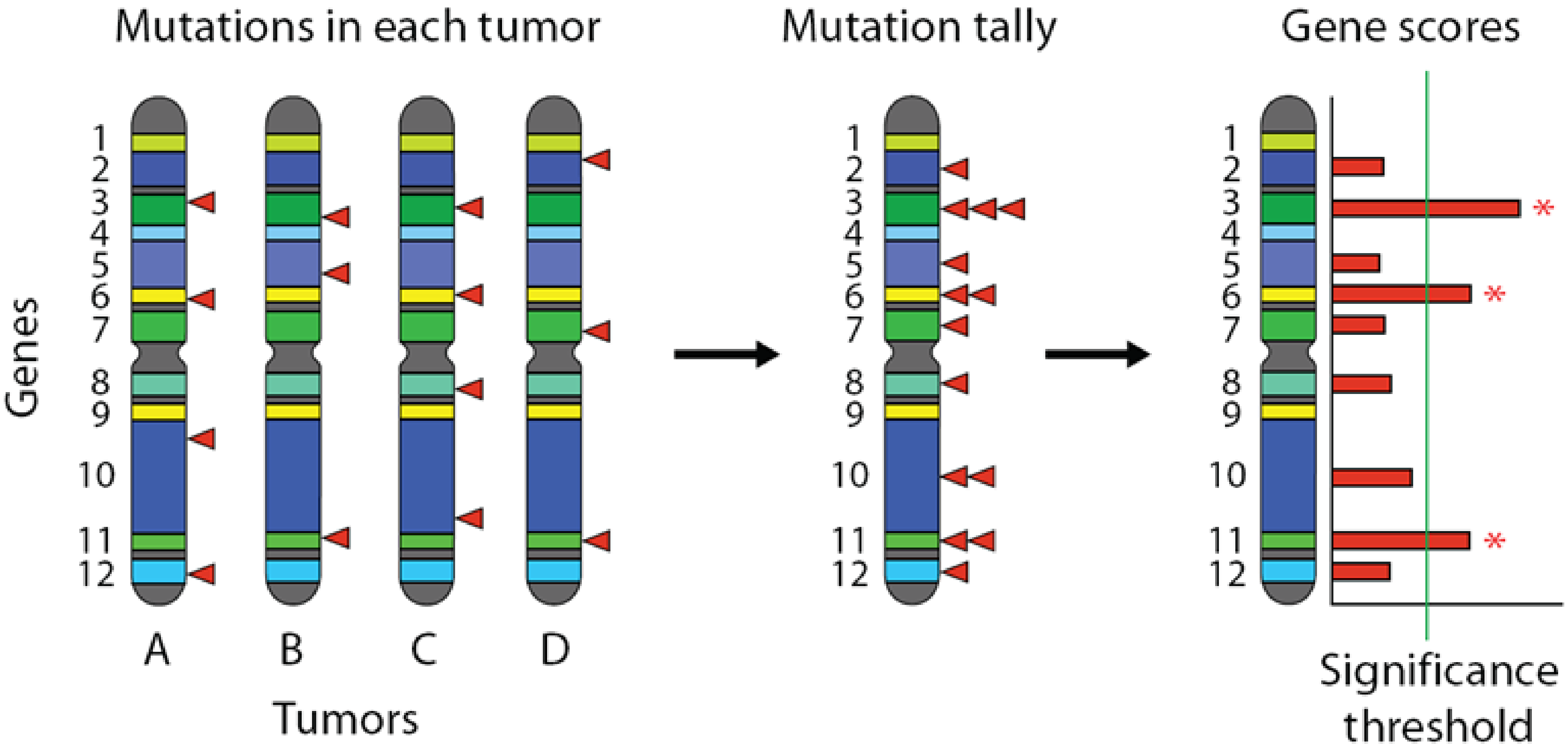
MutSig



http://www.broadinstitute.org/cancer/cga/mutsig

Covered Bases/Total Possible Mutations for that Gene & Category

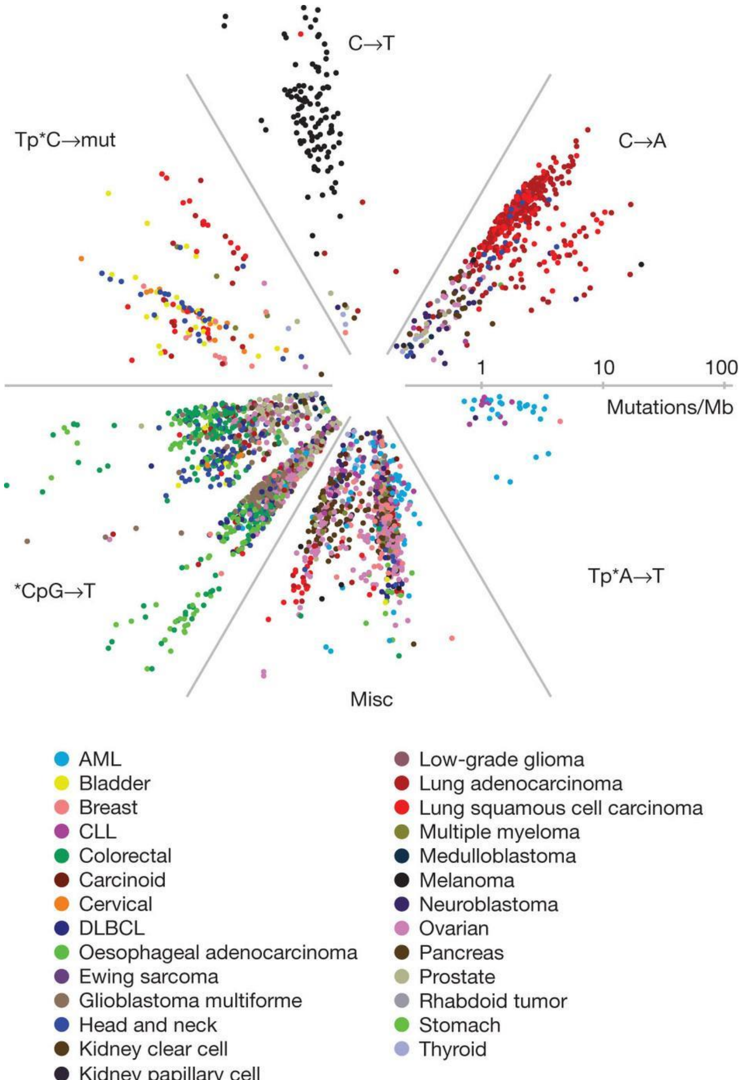
$$s_g = \sum_{c} -log_{10}(binomial(n_c, N_c, \mu_c))$$

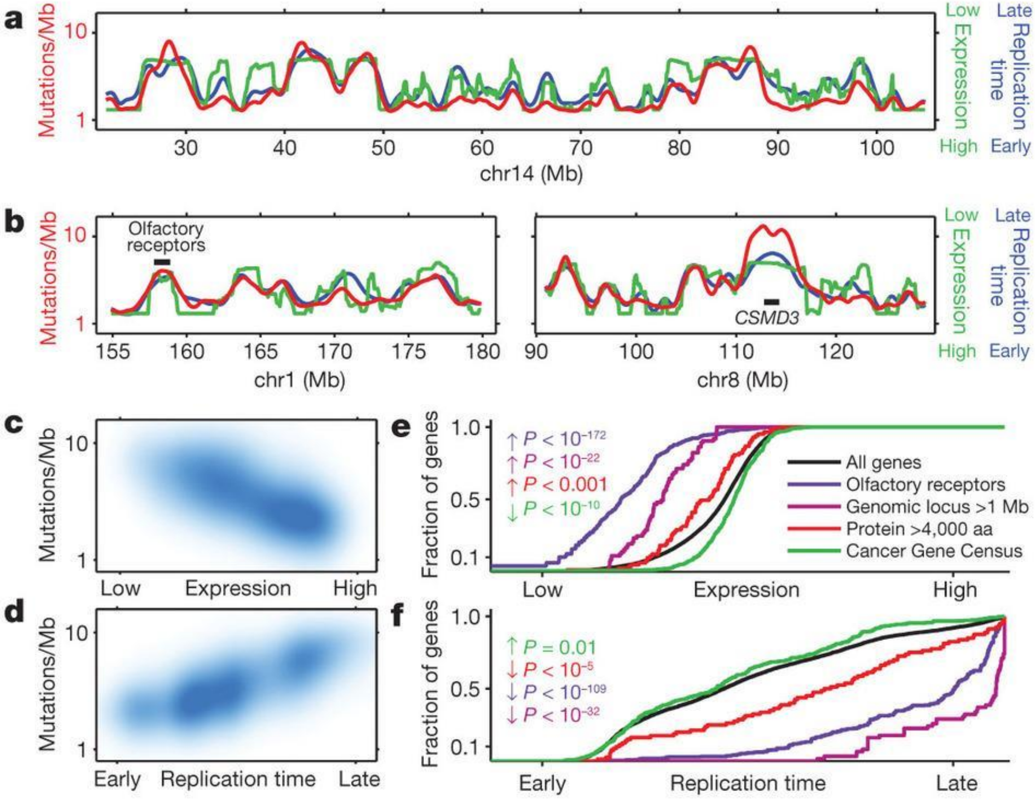
Score for that gene

Sum over each mutation category

Observed mutations in that gene & category

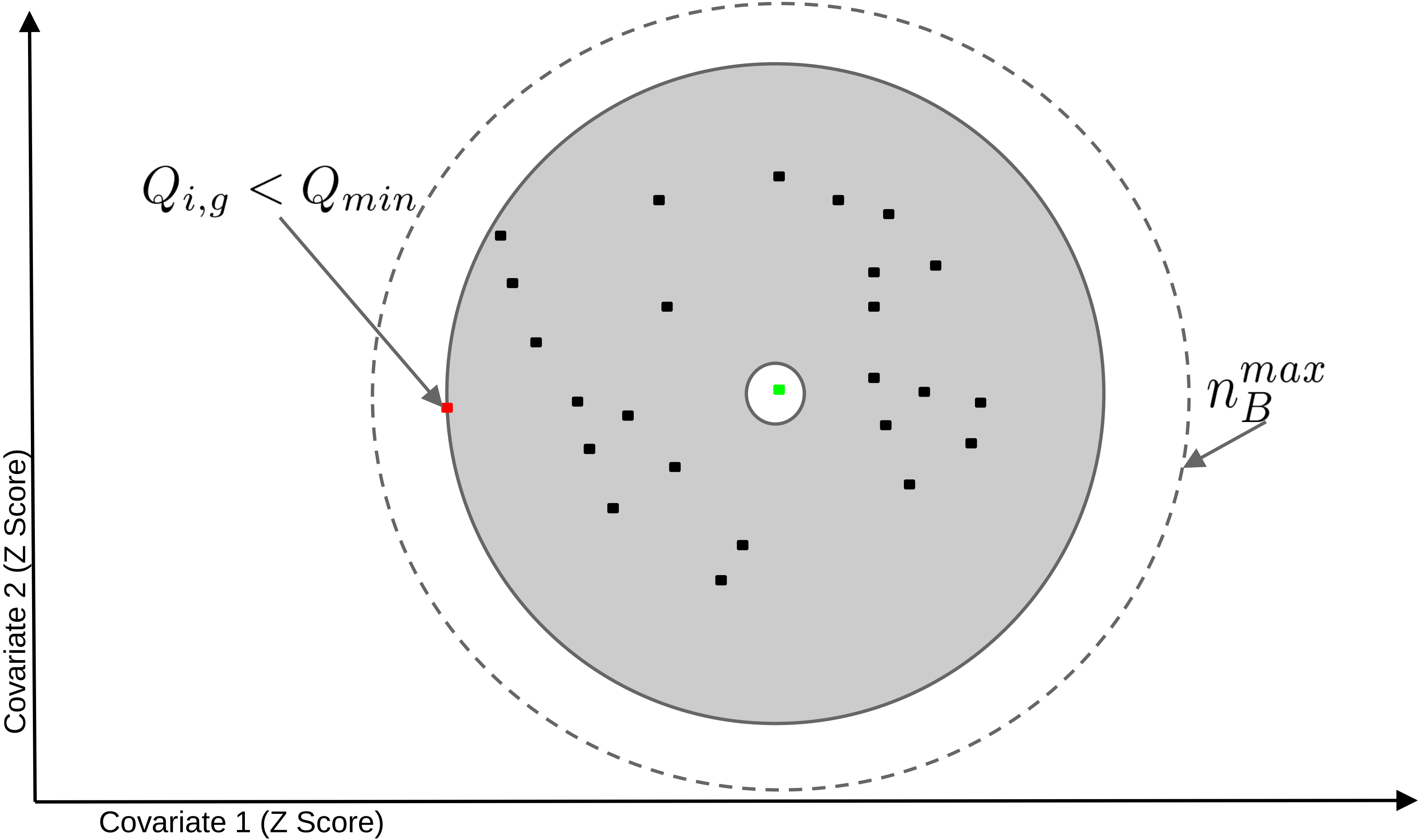
Background Mutation Rate


C→T
Tp*C→mut
C→A
1
10
100
Mutations/Mb
Tp*A→T
*CpG→T
Misc
AML
Bladder
Breast
CLL
Colorectal
Carcinoid
Cervical
DLBCL
Oesophageal adenocarcinoma
Ewing sarcoma
Glioblastoma multiforme
Head and neck
Kidney clear cell
Kidney papillary cell
Low-grade glioma
Lung adenocarcinoma
Lung squamous cell carcinoma
Multiple myeloma
Medulloblastoma
Melanoma
Neuroblastoma
Ovarian
Pancreas
Prostate
Rhabdoid tumor
Stomach
Thyroid


a
Mutations/Mb
10
1
30
40
50
60
70
80
90
100
chr14 (Mb)
Low
High
Expression
Late
Early
Replication time
b
Olfactory receptors
155
160
165
170
175
180
chr1 (Mb)
CSMD3
90
100
110
120
chr8 (Mb)
c
Low
Expression
High
d
Early
Replication time
Late
e
Fraction of genes
1.0
0.5
0.1
↑ P < 10⁻¹⁷²
↑ P < 10⁻²²
↑ P < 0.001
↓ P < 10⁻¹⁰
All genes
Olfactory receptors
Genomic locus >1 Mb
Protein >4,000 aa
Cancer Gene Census
f
↑ P = 0.01
↓ P < 10⁻⁵
↓ P < 10⁻¹⁰⁹
↓ P < 10⁻³²

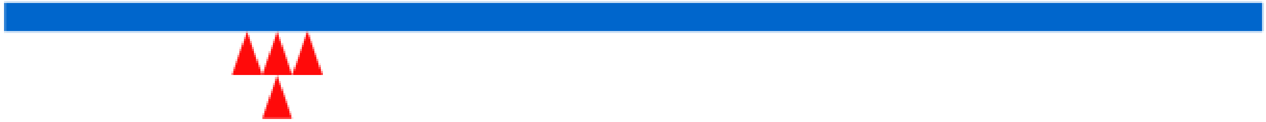
+

ITSTOCK FREE

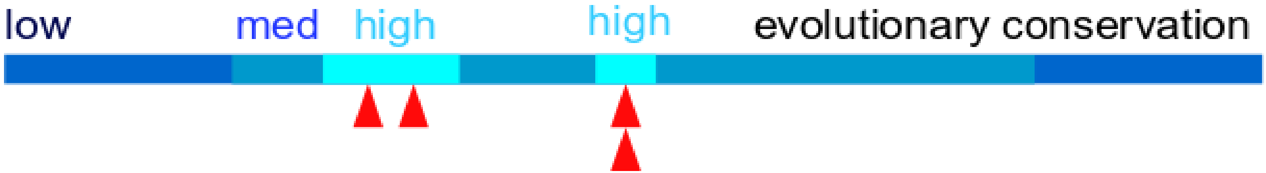

$Q_{i,g} < Q_{min}$
n_B^{max}
Covariate 2 (Z Score)
Covariate 1 (Z Score)

MutSigCV **Burden** of protein-altering mutations (compared to background)

MutSigCL **Clustering** of mutations in hotspots



MutSigFN **Functional** impact of the mutations



Lawrence et al. Nature, Jan 2014

Do you agree with their chosen inputs?

Could a simpler method get the same result?

Are there any improvements that could be made?