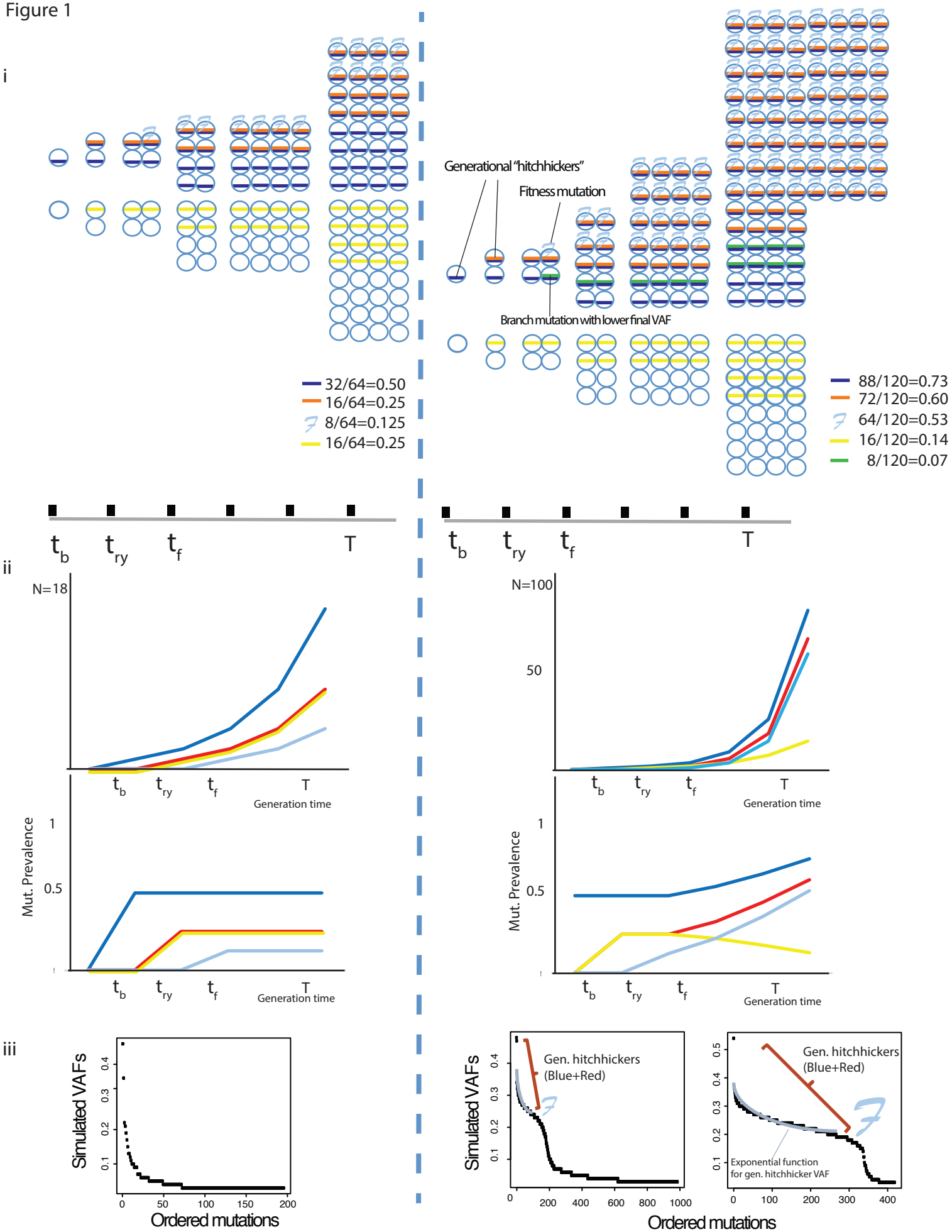
Figure 1


i
Generational "hitchhickers"
Fitness mutation
Branch mutation with lower final VAF
32/64=0.50
16/64=0.25
8/64=0.125
16/64=0.25
88/120=0.73
72/120=0.60
64/120=0.53
16/120=0.14
8/120=0.07
ii
N=18
N=100
50
Generation time
Mut. Prevalence
1
0.5
iii
Simulated VAFs
Ordered mutations
Gen. hitchhickers
(Blue+Red)
Exponential function
for gen. hitchhicker VAF

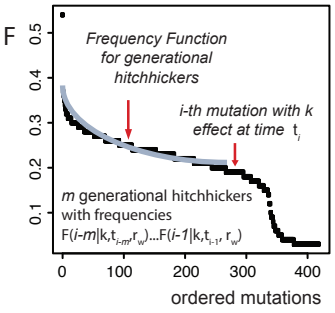
Using a single tumor

VCF file

Normalize VAF (based on ploidy, CNV and purity) to assign mutational frequency F

Use subclonal Information for each mutation to select non-overlapping subclones

Order mutations based on frequency F



For each mutation i > m using m generational hitchickers

A) estimate of growth 'r'

B) NLS optimization for effect 'k' as:

$$\mathbb{Q}(k, t_g \mid r, F, m) = \sum_m (F_m - (A * e^{-r(m+t_g)} + B))^2$$

C) Identify peaks 'K_p'

Output

$$\vec{r} = \{r_1, r_2, .. r_n\} \in \mathcal{R}^N_{+/-}$$
$$\vec{k} = \{k_1, k_2, .. k_n\} \in \mathcal{R}^N_{+}$$
$$\overrightarrow{Kp} = \{k_1, k_2, .. k_{p<n}\} \in \mathcal{R}^N_{+},$$
$$where\ \overrightarrow{Kp} \subset \vec{k}$$

Frequency function for m "hitchhickers" (with $F_m > F_i$)

$$F(T, t_{i-m}) = \frac{e^{-r_w * t_{i-m}} * [\, N_{tot} - F(t_i) * N_{tot} + \sqrt[k_i]{F(t_i) * N_{tot}}\,] + F(t_i) * N_{tot} - \sqrt[k_i]{F(t_i) * N_{tot}}}{N_{tot}}$$

r_w: growth r corresponding to window [*i-m*, *i*-1]

t_{i-m}: the time when the (*i-m*)-th mutation occured

k_i: the effect K of the i-th mutation

$F(t_i)$: the frequency of the hypothetical fitness mutation *i*

N_{tot}: the total number of mutations

Frequency function for m "hitchhickers" with local reoptimization

$$F(T, \mathbf{t_g}, t_{i-m}) = \frac{e^{-r(\mathbf{t_g} + t_{i-m})} * (N_{tot} - F(t_i) * N_{tot} + \sqrt[k_i]{F(t_i) * N_{tot}}) + F(t_i) * N_{tot} - \sqrt[k_i]{F(t_i) * N_{tot}}}{N_{tot}}$$

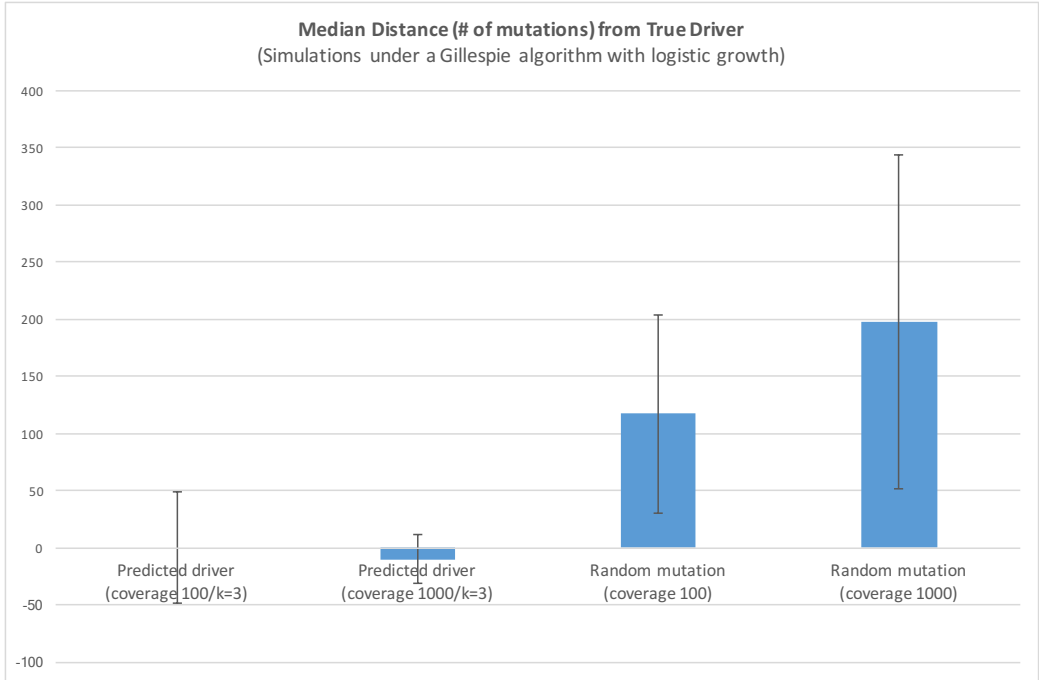
t_g: locally optimized generational time to adjust for local hitchhickers

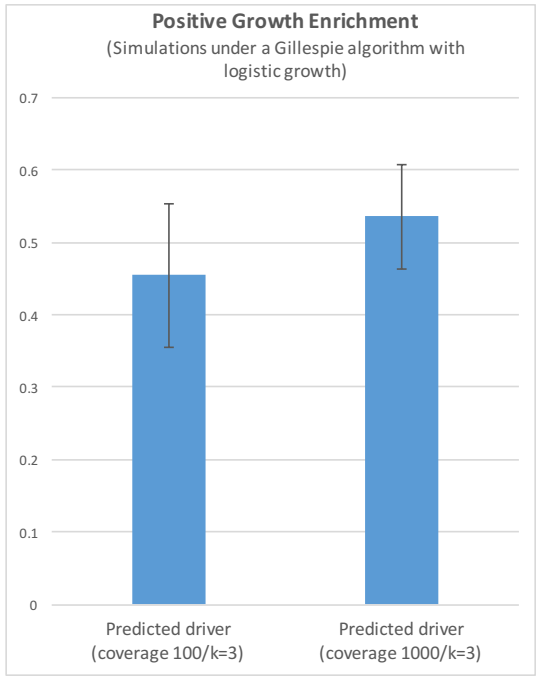
Using M multiple tumors

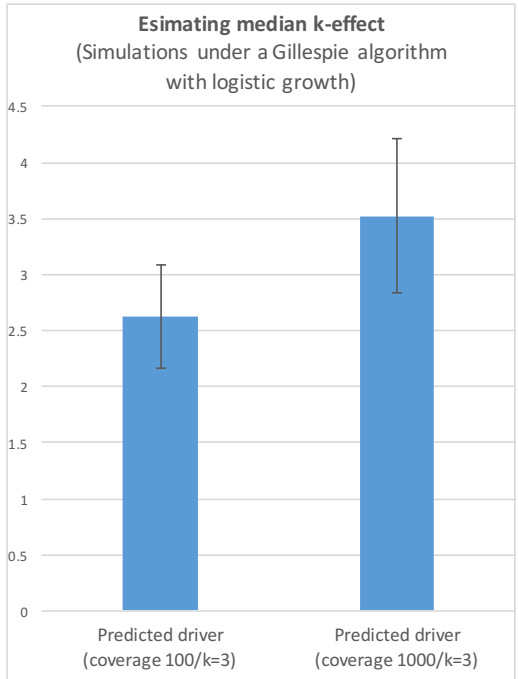
A) Estimate **Positive Growth Enrichment (PGE):**

Across all M tumors,
A **type** of mutation (eg. all *missense TP53* mutations, or *all premature-stop* mutations *in Tumor Suppressor Genes*), found w times in M tumors,
is **enriched during positive growth** if:
mutational growth r_{mut} > 0 more often than random

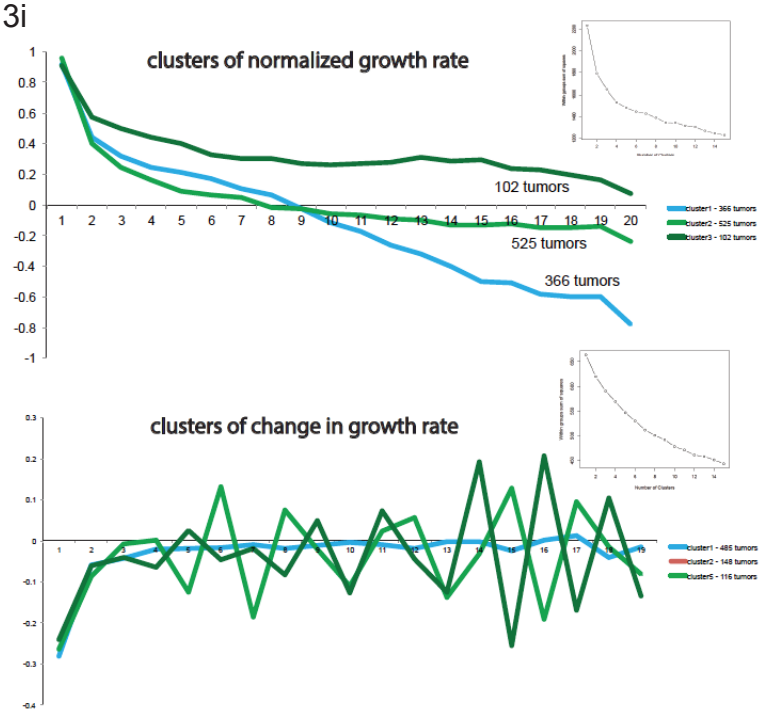
B) Estimate the range of effect k (eg. [1.2-1.4]) within a type of mutation is more enriched than random


Median Distance (# of mutations) from True Driver
(Simulations under a Gillespie algorithm with logistic growth)
400
350
300
250
200
150
100
50
0
-50
-100
Predicted driver (coverage 100/k=3)
Predicted driver (coverage 1000/k=3)
Random mutation (coverage 100)
Random mutation (coverage 1000)

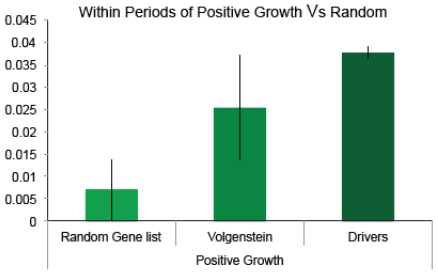

Positive Growth Enrichment
(Simulations under a Gillespie algorithm with logistic growth)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Predicted driver (coverage 100/k=3)
Predicted driver (coverage 1000/k=3)

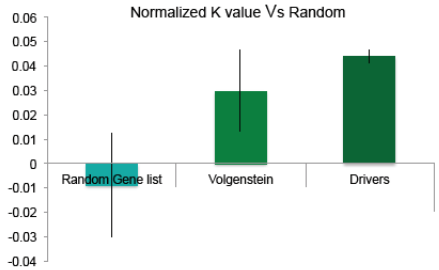

Esimating median k-effect
(Simulations under a Gillespie algorithm with logistic growth)
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Predicted driver (coverage 100/k=3)
Predicted driver (coverage 1000/k=3)

3i

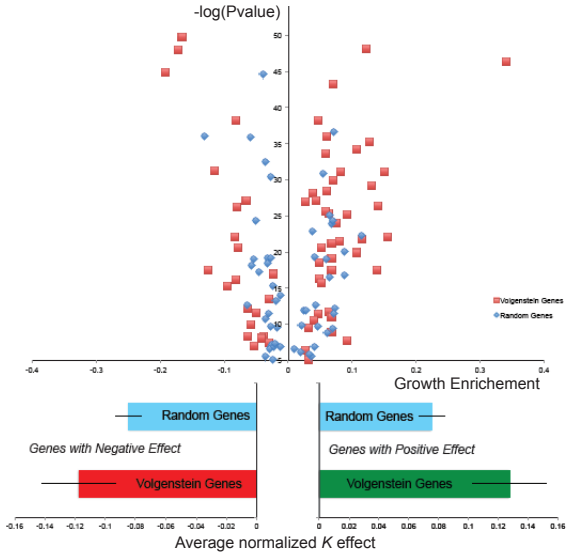


3ii



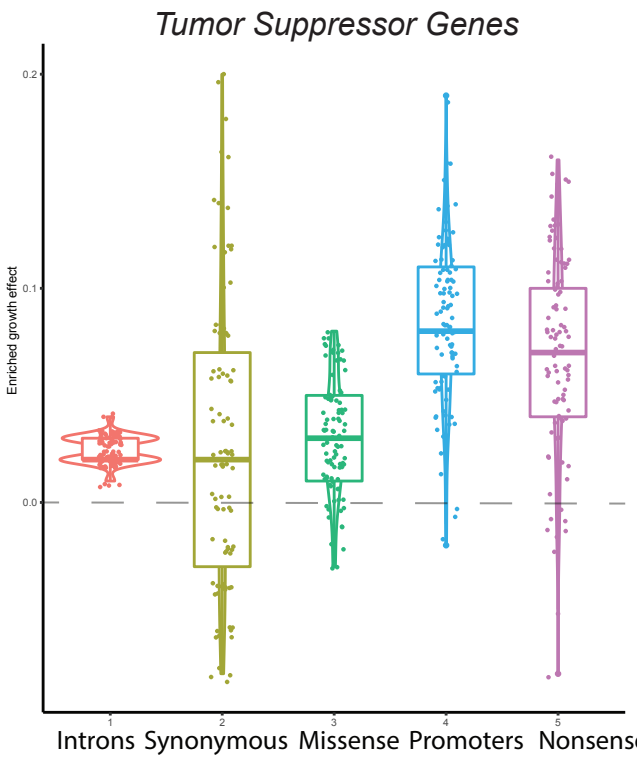


3iii

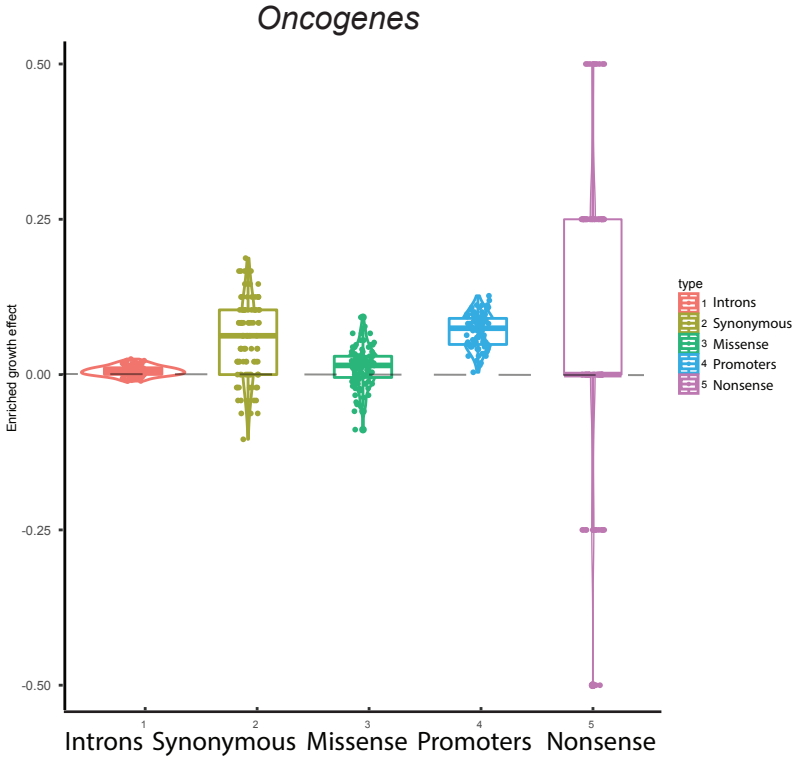


3iv

4i


Tumor Suppressor Genes
Enriched growth effect
0.2
0.1
0.0
1
2
3
4
5
Introns
Synonymous
Missense
Promoters
Nonsense

4ii


Oncogenes
Enriched growth effect
0.50
0.25
0.00
-0.25
-0.50
1
2
3
4
5
Introns
Synonymous
Missense
Promoters
Nonsense
type
1 Introns
2 Synonymous
3 Missense
4 Promoters
5 Nonsense

4iii

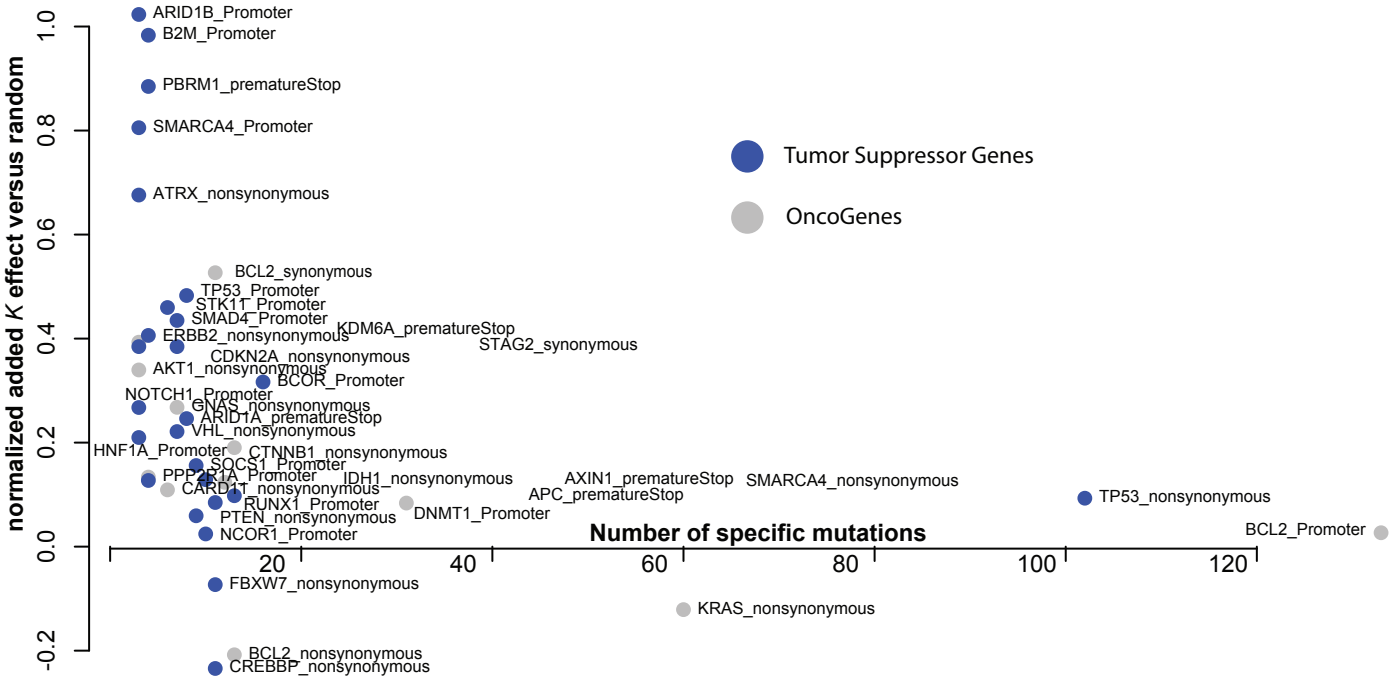
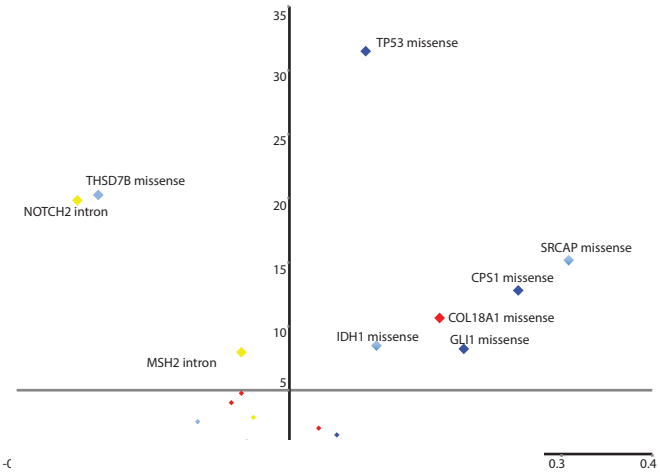
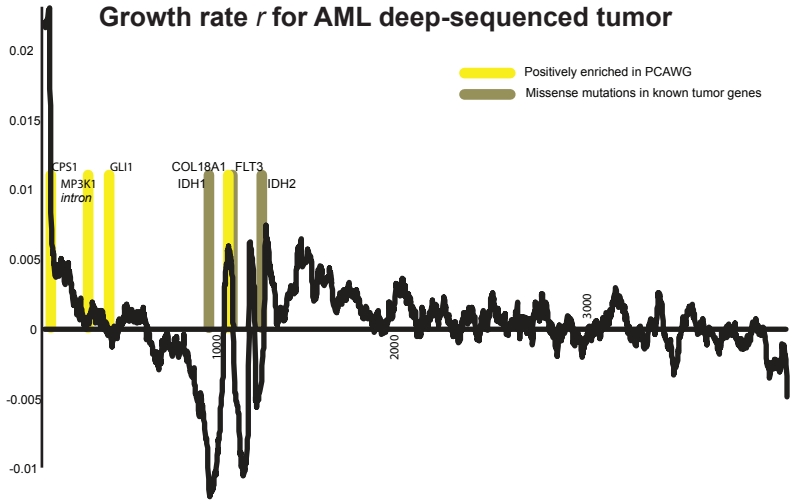

normalized added K effect versus random
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
Number of specific mutations
20
40
60
80
100
120
Tumor Suppressor Genes
OncoGenes
ARID1B_Promoter
B2M_Promoter
PBRM1_prematureStop
SMARCA4_Promoter
ATRX_nonsynonymous
BCL2_synonymous
TP53_Promoter
STK11_Promoter
SMAD4_Promoter
ERBB2_nonsynonymous
KDM6A_prematureStop
STAG2_synonymous
CDKN2A_nonsynonymous
AKT1_nonsynonymous
BCOR_Promoter
NOTCH1_Promoter
GNAS_nonsynonymous
ARID1A_prematureStop
VHL_nonsynonymous
HNF1A_Promoter
CTNNB1_nonsynonymous
SOCS1_Promoter
PPP2R1A_Promoter
IDH1_nonsynonymous
AXIN1_prematureStop
SMARCA4_nonsynonymous
CARD11_nonsynonymous
APC_prematureStop
TP53_nonsynonymous
RUNX1_Promoter
DNMT1_Promoter
PTEN_nonsynonymous
NCOR1_Promoter
BCL2_Promoter
FBXW7_nonsynonymous
KRAS_nonsynonymous
BCL2_nonsynonymous
CREBBP_nonsynonymous

Figure 5

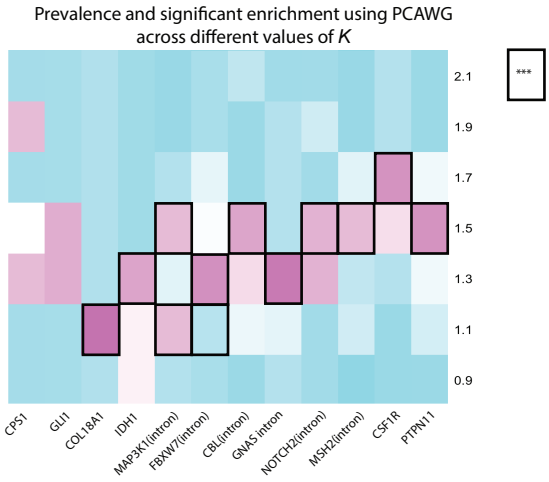
5i



5ii



5iii



5iv