Between-region genetic divergence reflects mode and tempo of tumor evolution

By Ruping Sun et al.

General Outline

- Problem Statement
- Methods
- Results

General Outline

- Problem Statement
  - Some Context
  - Key Biological Problem
- Methods
- Results

Problem Statement - Some Context

Problem Statement - Some Context

- In order to make precision (specialized medicine) work, there is a need to understand the specifics of how cells evolve across the spectrum of solid tumor types.

Problem Statement - Some Context

- In order to make precision (specialized medicine) work, there is a need to understand the specifics of how cells evolve across the spectrum of solid tumor types.
- Methods to infer the role of natural selection within established tumor types is lacking.

Problem Statement - Some Context

- In order to make precision (specialized medicine) work, there is a need to understand the specifics of how cells evolve across the spectrum of solid tumor types.
- Methods to infer the role of natural selection within established tumor types is lacking.
    - Tumor progression takes years,

Problem Statement - Some Context

- In order to make precision (specialized medicine) work, there is a need to understand the specifics of how cells evolve across the spectrum of solid tumor types.
- Methods to infer the role of natural selection within established tumor types is lacking.
    - Tumor progression takes years,
    - Often only detected in late stages

Problem Statement - Some Context

- In order to make precision (specialized medicine) work, there is a need to understand the specifics of how cells evolve across the spectrum of solid tumor types.
- Methods to infer the role of natural selection within established tumor types is lacking.
    - Tumor progression takes years,
    - Often only detected in late stages
    - This contributes to poor understanding of tumorigenesis and growth immediately after transformation

Clonal Evolution Model

Clonal Evolution Model

- Evolution is the product of 3 major underlying processes

Clonal Evolution Model

- Evolution is the product of 3 major underlying processes
  - Mutation
    - Readily measured in human tumours - SSNV calling, CNA estimation etc.

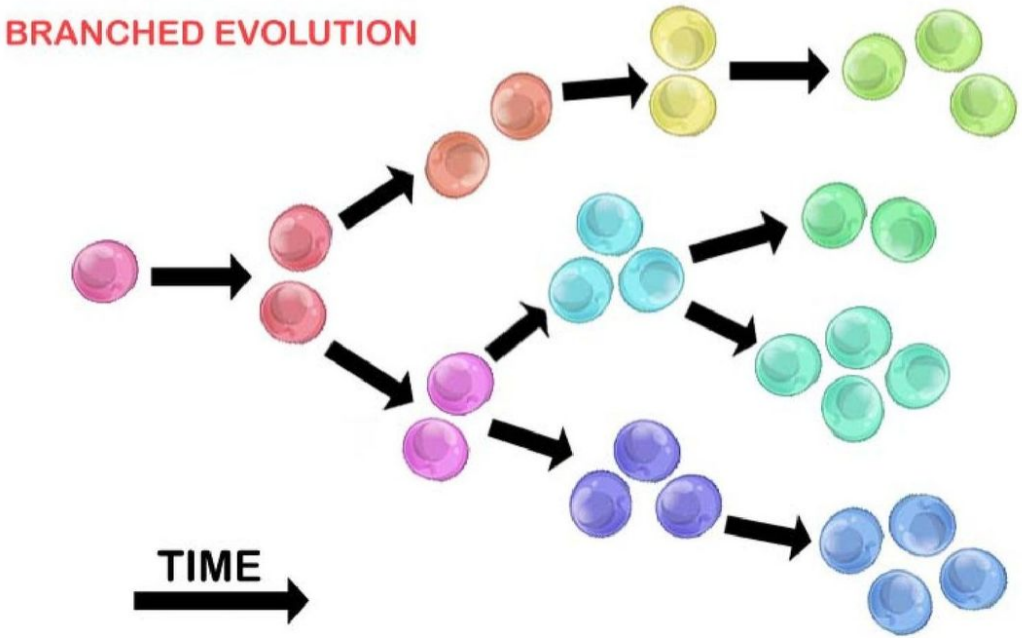
Clonal Evolution Model

- Evolution is the product of 3 major underlying processes
  - Mutation
    - Readily measured in human tumours
  - Selection
    - Assumed to govern the growth of an established tumor after tumorigenesis (transformation)
    - Acquisition of additional “driver” mutations results in multiple selective sweeps.
    - Due to the selection of driver mutations, hitchhiking passenger mutations can also attain high frequency and manifest as subclonal clusters in bulk sequencing data

Clonal Evolution Model

- Evolution is the product of 3 major underlying processes
  - Mutation
    - Readily measured in human tumours
  - Selection
    - Assumed to govern the growth of an established tumor after tumorigenesis (transformation)
    - Acquisition of additional “driver” mutations results in multiple selective sweeps.
    - Due to the selection of driver mutations, hitchhiking passenger mutations can also attan high frequency and manifest as subclonal clusters in bulk sequencing data
  - Drift
    - Difficult to distinguish from selection but may also cause extensive ITH.

Clonal Evolution Model



Goal

Goal

- Test hypotheses about underlying evolutionary processes using spatiotemporal patterns of genetic variation among cell populations.

Goal

- Test hypotheses about underlying evolutionary processes using spatiotemporal patterns of genetic variation among cell populations.
- Use simulations to make up lack of high quality early stage data

Goal

- Test hypotheses about underlying evolutionary processes using spatiotemporal patterns of genetic variation among cell populations.
- Use simulations to make up lack of high quality early stage data
- Leverage powerful statistical methods to reveal something about

Key Questions

- What kind of role does selection play in the life of a tumor?

Key Questions

- What kind of role does selection play in the life of a tumor?
- Can you distinguish strong positive selection from weak selection or neutral evolution during tumor progression?

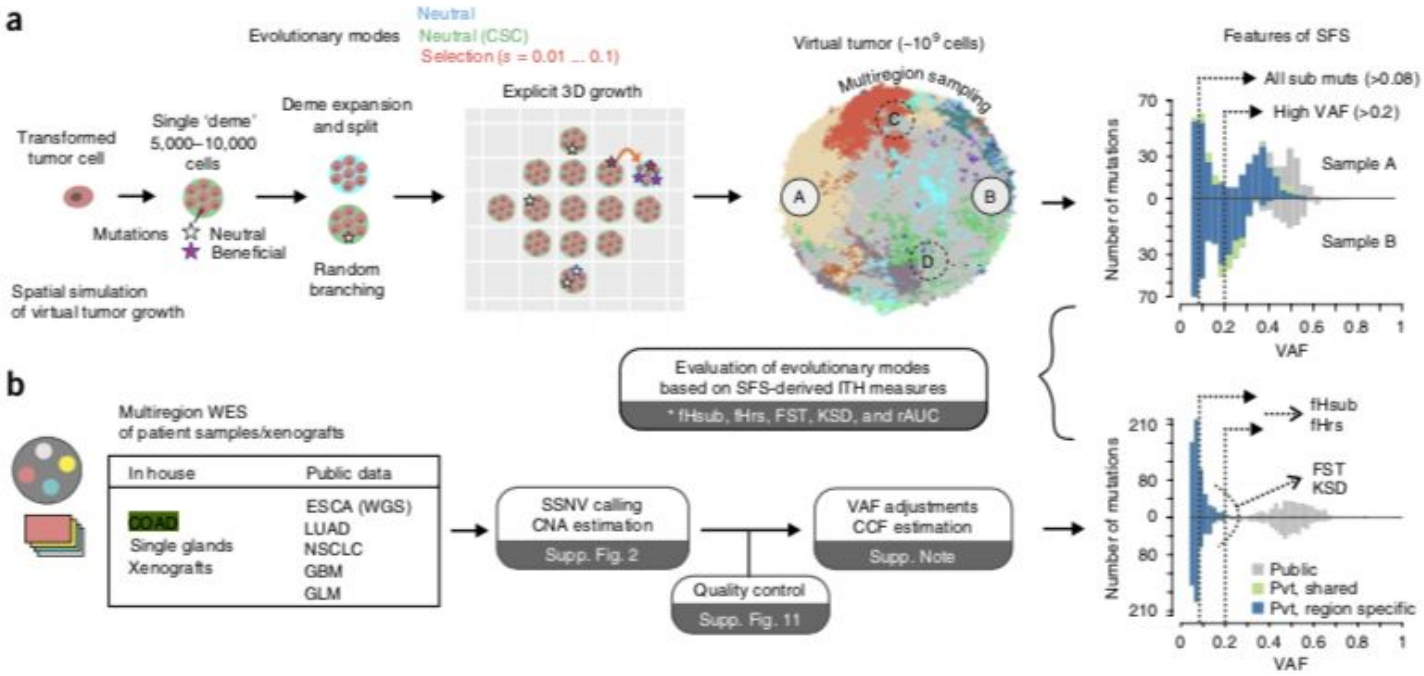
General Outline

- Problem Statement
- Methods
- Results

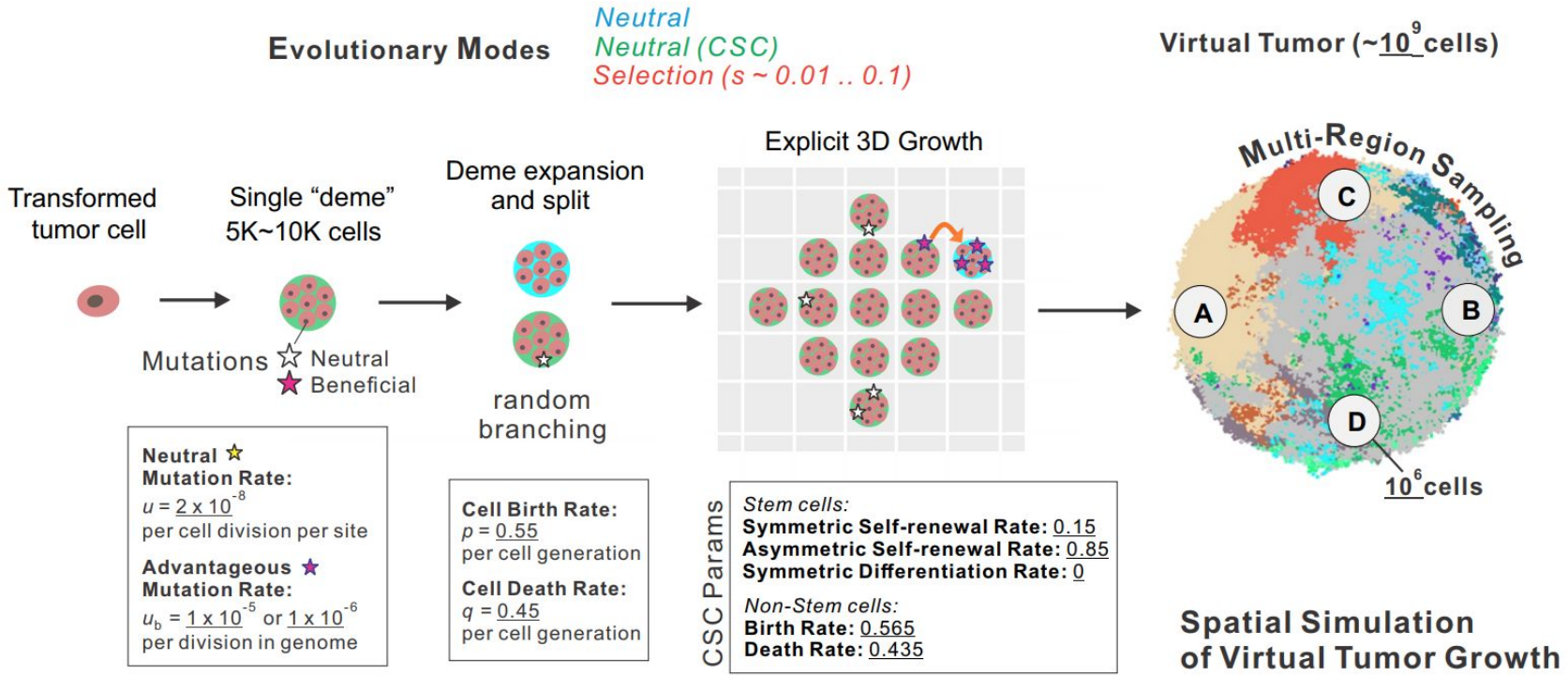
General Outline

- Problem Statement
- Methods
- Results

Methods


a
Evolutionary modes
Neutral
Neutral (CSC)
Selection (s = 0.01 ... 0.1)
Transformed tumor cell
Single 'deme' 5,000–10,000 cells
Deme expansion and split
Explicit 3D growth
Virtual tumor (~10^9 cells)
Multiregion sampling
Features of SFS
Mutations
Neutral
Beneficial
Random branching
Spatial simulation of virtual tumor growth
All sub muts (>0.08)
High VAF (>0.2)
Sample A
Sample B
Number of mutations
VAF
b
Multiregion WES of patient samples/xenografts
In house
Public data
COAD
Single glands
Xenografts
ESCA (WGS)
LUAD
NSCLC
GBM
GLM
SSNV calling
CNA estimation
Supp. Fig. 2
Quality control
Supp. Fig. 11
VAF adjustments
CCF estimation
Supp. Note
Evaluation of evolutionary modes based on SFS-derived ITH measures
* fHsub, fHrs, FST, KSD, and rAUC
fHsub
fHrs
FST
KSD
Public
Pvt, shared
Pvt, region specific

Methods - Spatial Computational Modelling



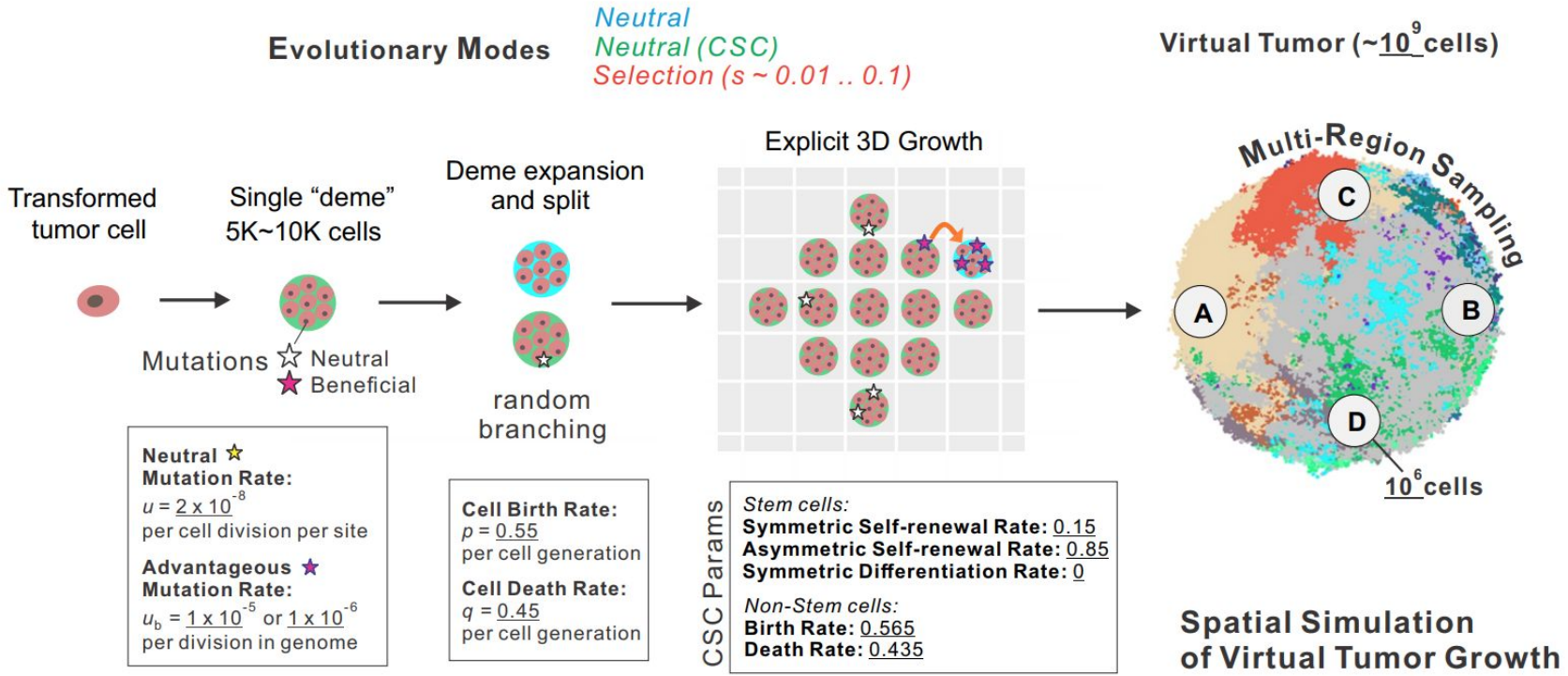
Methods - Spatial Computational Modelling

- Simulate various modes of evolution

Methods - Spatial Computational Modelling

- Simulate various modes of evolution
  - Neutral
  - Neutral CSC
  - Positive Selection (s = 0.01 -> 0.1 etc.) Acquisition of advantageous mutations adjusts birth-death rate according to s.

Methods - Spatial Computational Modelling



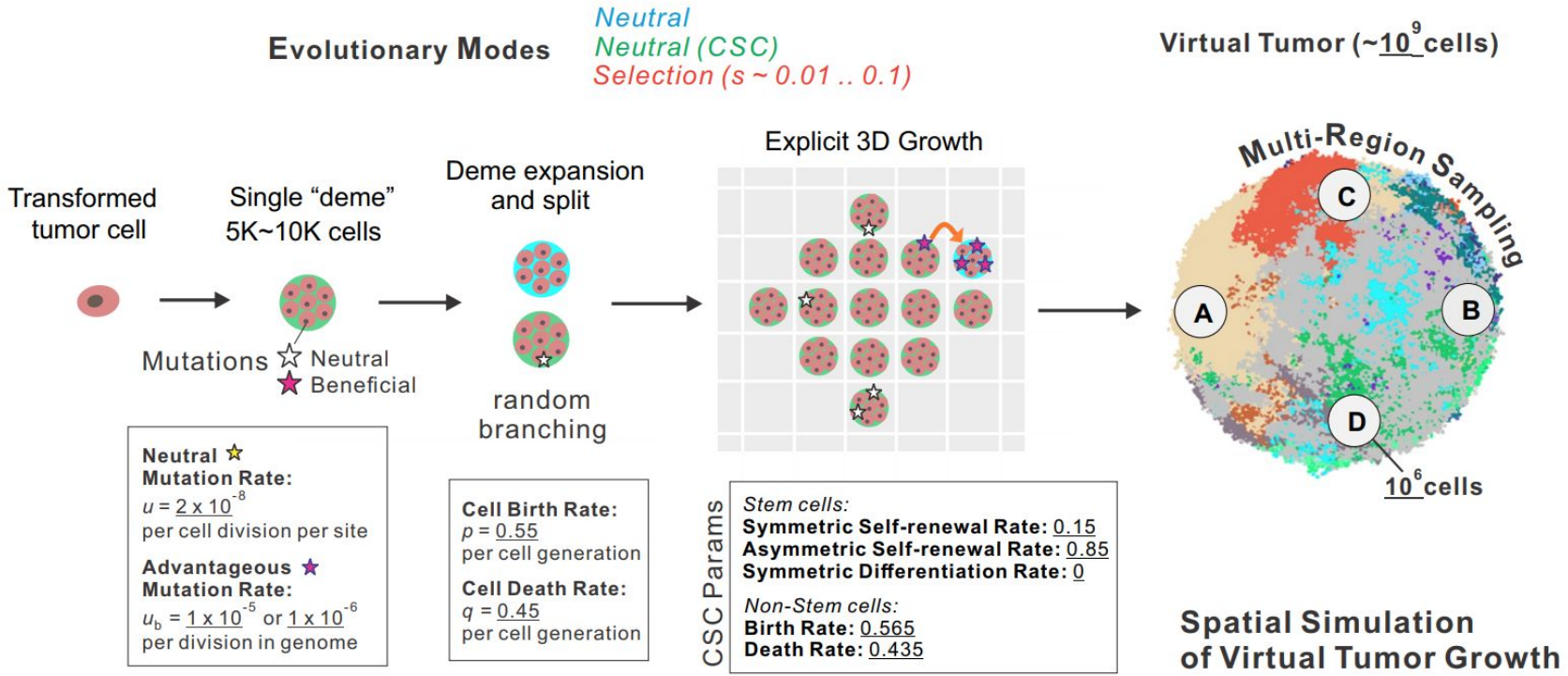
Methods - Spatial Computational Modelling

- Simulate various modes of evolution

Methods - Spatial Computational Modelling

- Simulate various modes of evolution.
- At each time step, a random deme near the periphery of the tumor is chosen for division.
    - This peripheral growth model is consistent with recent studies that show that cells at the periphery are more proliferative than those at the core of a tumor.
    - A random empty lattice site is chosen for the replicate deme.
    - Deme size is chosen conservatively as a maximum of 10k, as large demes hinder selection structurally
    - When a deme reaches maximum size it splits into two offspring demes. The split is modelled by a binomial distribution **(N_c, p=0.5)** where N_c is the current size of the deme

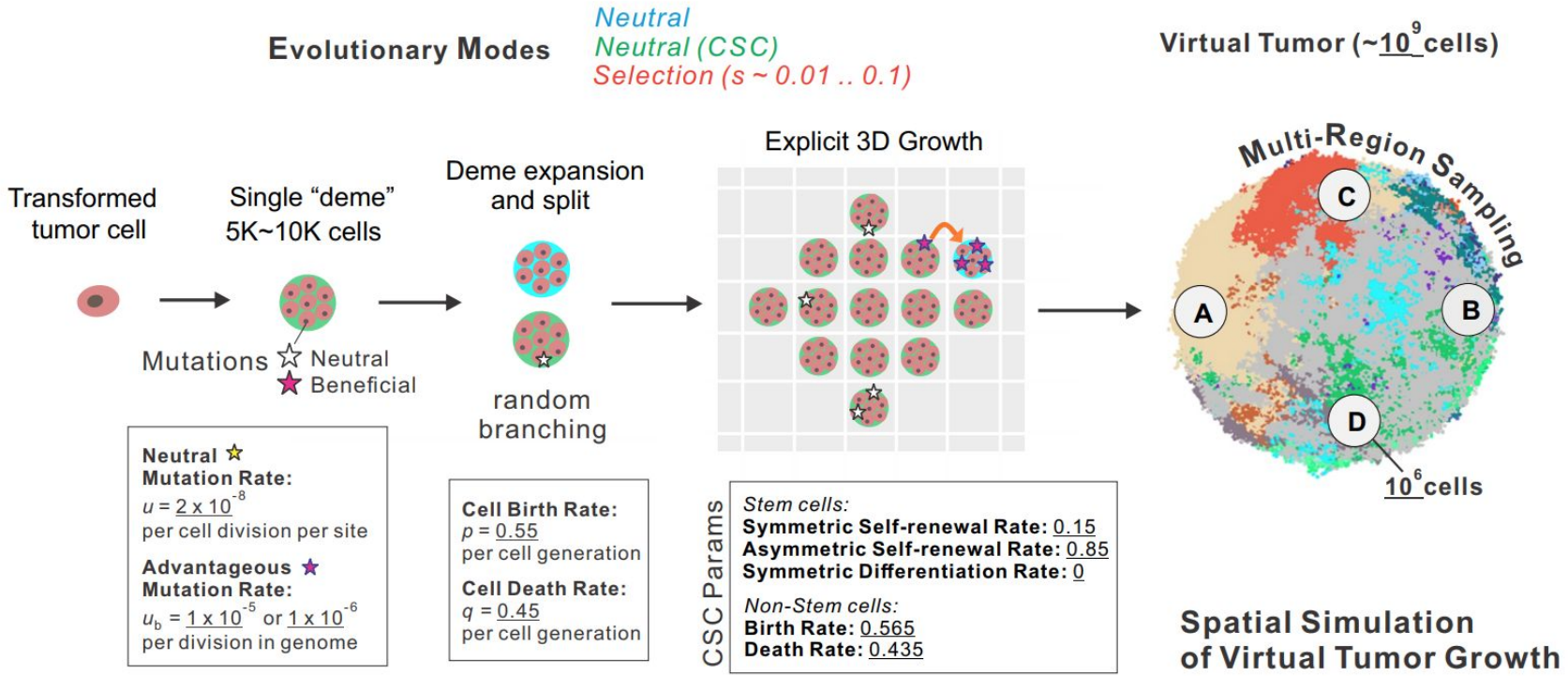
Methods - Spatial Computational Modelling



Methods - Spatial Computational Modelling

- Simulate various modes of evolution.
- At each time step, a random deme near the periphery of the tumor is chosen for division

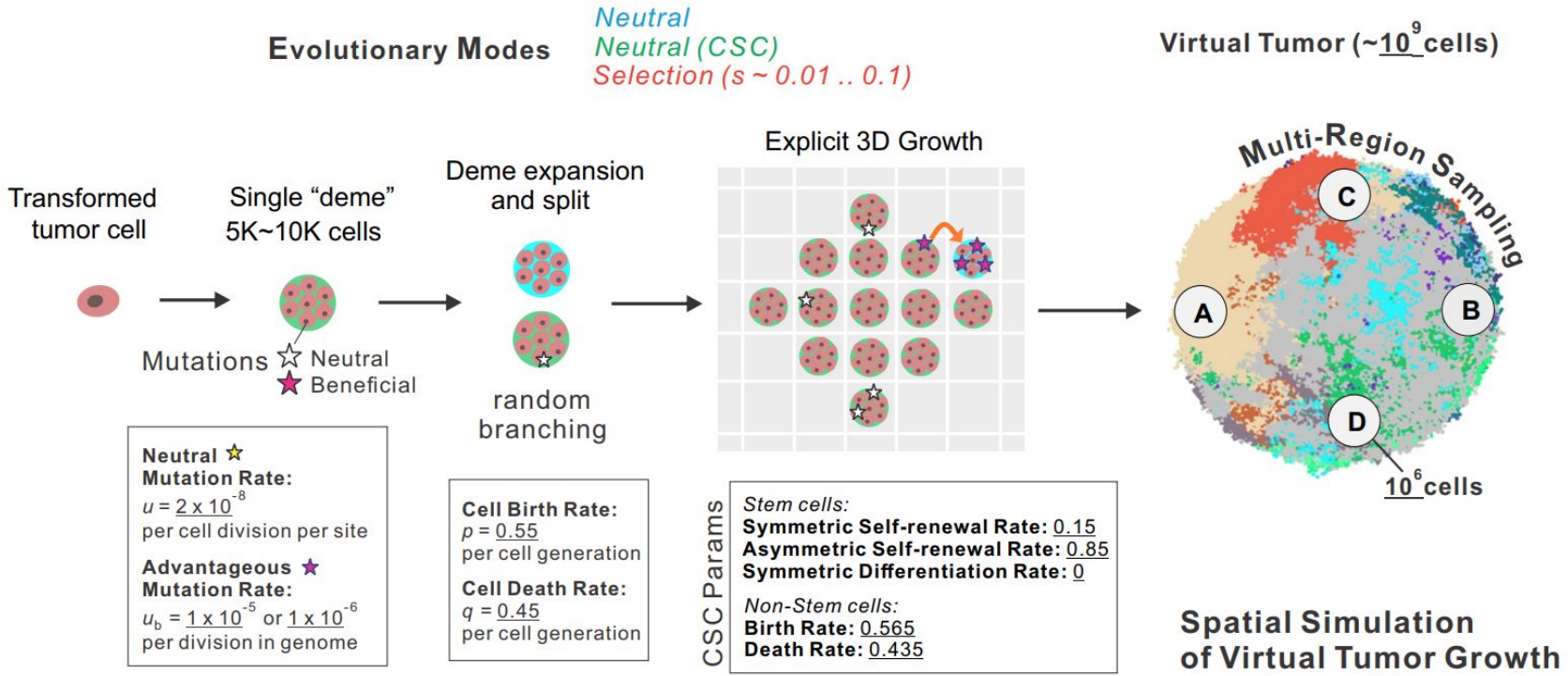
Methods - Spatial Computational Modelling



Methods - Spatial Computational Modelling

- Simulate various modes of evolution.
- At each time step, a random deme near the periphery of the tumor is chosen for division
- Cell division is a continuous time Markov Process known as a *Birth-Death* process
    - At any time step there a ***p*** probability that a cell will divide and a ***q = 1-p*** probability that it will die.
    - Death/Birth rate ratio is important and is characteristic of cancers. E.g. **h = q/p = 0.99 in early tumors but 0.72 in metastatic colorectal cancer.**

Methods - Spatial Computational Modelling



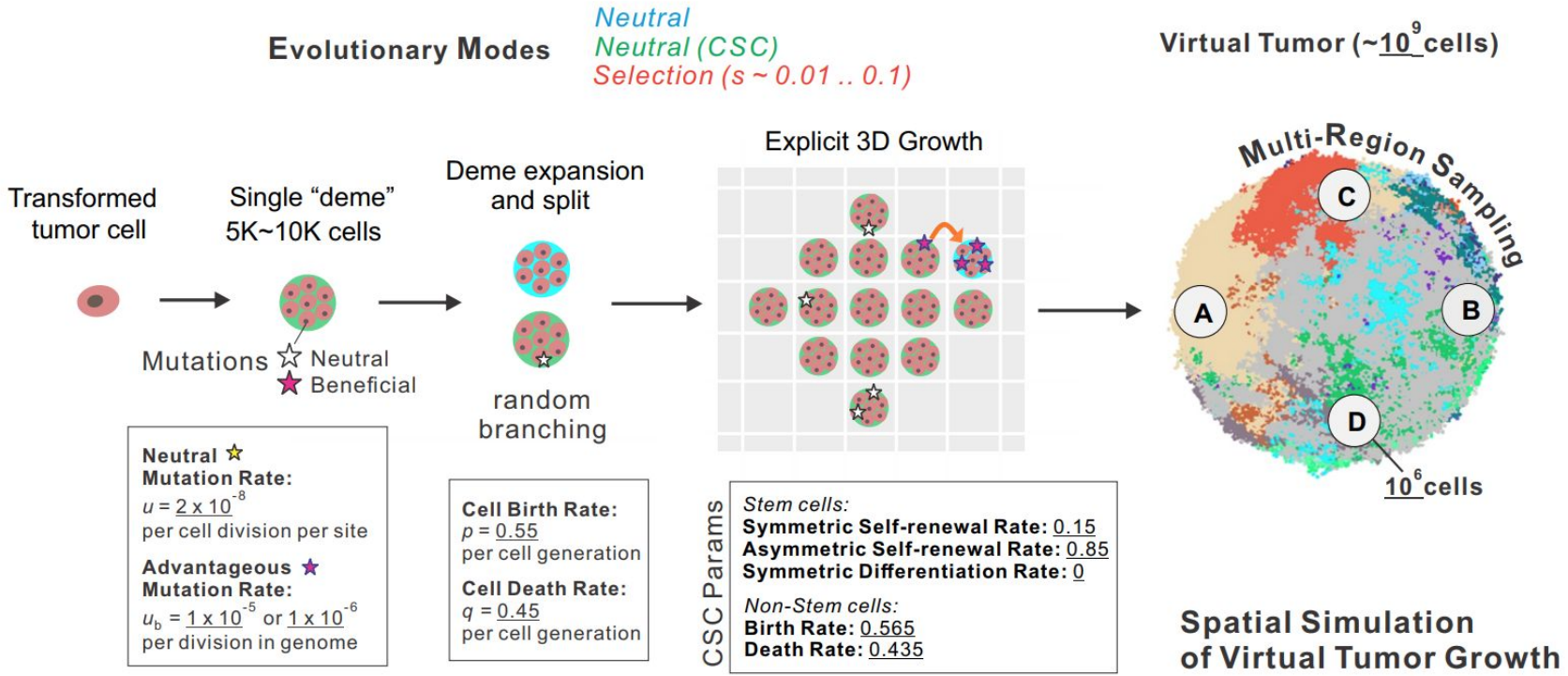
Methods - Spatial Computational Modelling

- Simulate various modes of evolution.
- At each time step, a random deme near the periphery of the tumor is chosen for division
- Cell division is a continuous time Markov Process known as a *Birth-Death* process

Methods - Spatial Computational Modelling

- Simulate various modes of evolution.
- At each time step, a random deme near the periphery of the tumor is chosen for division
- Cell division is a continuous time Markov Process known as a *Birth-Death* process
- Poisson process random point mutations ***assuming infinite sites model*****,** at deme division time.
    - Under null models, all mutations are neutral and don't confer fitness advantage
    - Under selection models, beneficial mutations occur as a poisson process with a different mean, and increase the birth rate of the mutated cell by ***s*****.**
    - **All beneficial mutations are given a unique index, and other statistics like host cells are recorded.**

Methods - Spatial Computational Modelling



Methods - Spatial Computational Modelling

- Simulate various modes of evolution.
- At each time step, a random deme near the periphery of the tumor is chosen for division
- Cell division is a continuous time Markov Process known as a *Birth-Death* process
- Poisson process random point mutations ***assuming infinite sites model*****,** at deme division time.

Methods - SSNV Calling, SCNA detection and VAF adjustment

- For each raw SNV call from MuTect, read alignmet features from all samples were reinspected in an automated fashion to assess confidence.
- TitanCNA was used to estimate somatic CNAs, and the observed VAF for each detected SSNV was adjusted on the basis of cancer cell fraction, tumor purities, as well as local copy number estimates.

$$VAF_a = \frac{CCF_{est}}{2}, \qquad CCF_{est} = \begin{cases} N_c \times \frac{VAF_o}{p'} - P_{CNA} \times (N_t - N_b - 1) & Early\ Major \\ N_c \times \frac{VAF_o}{p'} - P_{CNA} \times (N_b - 1) & Early\ Minor \\ N_c \times \frac{VAF_o}{p'} & Late/Independent \end{cases}$$

Methods - Subclonal SSNV identification in MRS

- An SSNV m is *subclonal* if all of the following criteria are met
  - Total Probability $P_m < 0.05$

Methods - Subclonal SSNV identification in MRS

$$P_m = \prod_{i=1}^{k} P_{mi}\left(X_{mi} \le S_{mi}, N_{mi}, f.pub_{mi}\right) < 0.05$$

$$f.pub_{mi} = \begin{cases} \mathrm{pu}_i \times \dfrac{\mathrm{nb}_{mi}}{\mathrm{nc}_{mi}} & \text{if } \mathrm{nb}_{mi} \ge 1,\ \mathrm{nt}_{mi} \ge 2 \\ \left(\mathrm{pu}_i \times \left(\mathrm{nt}_{mi} - \mathrm{nb}_{mi}\right)\right) / \mathrm{nc}_{mi} & \text{otherwise} \end{cases}$$

- P_{mi} is the probability for region *i* of observing less than S_{mi} reads having the mutant allele out of N_{mi} reads.
- This is provided the lower bound of *f* described above which is the expected frequency if the SSNV is public.

Methods - Subclonal SSNV identification in MRS

- An SSNV m is *subclonal* if all of the following criteria are met
  - Total Probability $P_m < 0.05$

Methods - Subclonal SSNV identification in MRS

- An SSNV m is *subclonal* if all of the following criteria are met
  - Total Probability $P_m < 0.05$
  - At least one region *i* with $CCF_m \pm 95\% CI_{mi} < 1$

Methods - Subclonal SSNV identification in MRS

- An SSNV m is *subclonal* if all of the following criteria are met
  - Total Probability $P_m < 0.05$
  - At least one region *i* with $CCF_m \pm 95\% CI_{mi} < 1$
  - At least one region *i* with adjusted $VAF_{mi} < 0.25$, which was chosen experimentally.

Methods - Subclonal SSNV identification in MRS

- An SSNV m is *subclonal* if all of the following criteria are met
  - Total Probability $P_m < 0.05$
  - At least one region *i* with $CCF_m \pm 95\% CI_{mi} < 1$
  - At least one region *i* with adjusted $VAF_{mi} < 0.25$, which was chosen experimentally.
- Any SSNV that fails to meet one of the above criteria is considered public!

Methods - ITH metrics

Methods - ITH metrics

- The first is fHrs, the fraction of high frequency region-specific subclonal SSNVs out of all region-specific subclonal SSNVs

Methods - ITH metrics

2)
$$\mathrm{fHrs} = \frac{1}{r} \times \sum_{j=1}^{r} \left(\frac{\mathrm{RSM}_{ja}^{\mathrm{high}}}{2 \times \mathrm{RSM}_{ja}^{\mathrm{all}}} + \frac{\mathrm{RSM}_{jb}^{\mathrm{high}}}{2 \times \mathrm{RSM}_{jb}^{\mathrm{all}}} \right)$$

where $\mathrm{RSM}_{ja}^{\mathrm{high}}$ and $\mathrm{RSM}_{ja}^{\mathrm{all}}$ represent the number of high-frequency (VAF > 0.2) region-specific SSNVs and the number of all region-specific SSNVs with VAF >0.08 for region a, in a pairwise comparison j between regions a and b.

Methods - ITH metrics

- The first is **fHrs**, the fraction of high frequency region-specific subclonal SSNVs out of all region-specific subclonal SSNVs.
- Next is **fHsub,** which is the fraction of subclonal SSNVs with high frequency (VAF > 0.2).

Methods - ITH metrics

1)
$$\mathrm{fHsub} = \frac{1}{k} \times \sum_{i=1}^{k} \frac{\mathrm{SM}_i^{\mathrm{high}}}{\mathrm{SM}_i^{\mathrm{all}}}$$

where $\mathrm{SM}_i^{\mathrm{high}}$ and $\mathrm{SM}_i^{\mathrm{all}}$ are the number of high-frequency subclonal SSNVs (adjusted VAF > 0.2, hereafter referred to as VAF) and the number of all subclonal SSNVs with VAF >0.08 for region *i*. The cutoff was set to 0.2 because above this value fHsub tends to plateau in its sensitivity to distinguishing the neutral and selection models (**Supplementary Fig. 36**). A lower cutoff of 0.08 was chosen empirically to satisfy the tradeoff between the number of subclonal SSNVs and variant calling errors.

Methods - ITH metrics

- The first is **fHrs**, the fraction of high frequency region-specific subclonal SSNVs out of all region-specific subclonal SSNVs.
- Next is **fHsub,** which is the fraction of subclonal SSNVs with high frequency (VAF > 0.2).
- **KSD** (Kolmogorov-Smirnov distance) - estimates the dissimilarity of the SFSs between the two regions

Methods - ITH metrics

4)

$$\mathrm{KSD} = \frac{1}{r} \times \sum_{j=1}^{r} \mathrm{KSD}_j$$

and

$$\mathrm{KSD}_j = \max \left| F_a - F_b \right|$$

where F_a is the cumulative SFS of region a, in a pairwise comparison j between regions a and b.

Methods - ITH metrics

- The first is **fHrs**, the fraction of high frequency region-specific subclonal SSNVs out of all region-specific subclonal SSNVs.
- Next is **fHsub,** which is the fraction of subclonal SSNVs with high frequency (VAF > 0.2).
- **KSD** (Kolmogorov-Smirnov distance) - estimates the dissimilarity of the SFSs between the two regions.
- **FST** (fixation index) - a measure of genetic divergence between the two regions.

Methods - ITH metrics

3)

$$\mathrm{FST} = \frac{1}{r} \times \sum_{j=1}^{r} \mathrm{FST}_j^{\mathrm{Hudson}}$$

and

$$\mathrm{FST}_j^{\mathrm{Hudson}} = \frac{\sum_{m=1}^{m_t} \left(f_a^m - f_b^m\right)^2 - \frac{f_a^m \times \left(1 - f_a^m\right)}{d_a^m - 1} - \frac{f_b^m \times \left(1 - f_b^m\right)}{d_b^m - 1}}{\sum_{m=1}^{m_t} f_a^m \times \left(1 - f_b^m\right) + f_b^m \times \left(1 - f_a^m\right)}$$

where f_a^m is the VAF for SSNV m and d_a^m is the sequencing depth for SSNV m in region a. The genetic variance components (numerator and denominator) are averaged separately to obtain a ratio combining the Hudson FST estimates across all m_t SSNVs[54].

Methods - ITH metrics

- The first is **fHrs**, the fraction of high frequency region-specific subclonal SSNVs out of all region-specific subclonal SSNVs.
- Next is **fHsub,** which is the fraction of subclonal SSNVs with high frequency (VAF > 0.2).
- **KSD** (Kolmogorov-Smirnov distance) - estimates the dissimilarity of the SFSs between the two regions.
- **FST** (fixation index) - a measure of genetic divergence between the two regions.
- **rAUC -** The ratio of the area under the pooled cumulative SFS to the area under the theoretical cumulative SFS assuming neutral growth in a well-mixed population. **For MRS, *pooled VAF = #alternative alleles/total read depth*.**

Methods - ITH metrics

5) $$\mathrm{rAUC} = {}^{\mathrm{AUC}_{\mathrm{merged}}}\!\!\diagup\!\!_{\mathrm{AUC}_{\mathrm{theoretical}}}$$

Methods - Evolutionary Mode Classifier

Methods - Evolutionary Mode Classifier

- A support vector machine classifier with a radial basis function (RBF) kernel was trained on 1400 simulated tumors derived from 7 different growth models first 3 models labelled as “effectively neutral” and the other 4 as “selection”, using 10-fold cross-validation with the 5 ITH metrics as features.

Methods - Evolutionary Mode Classifier

- A support vector machine classifier with a radial basis function (RBF) kernel was trained on 1400 simulated tumors derived from 7 different growth models first 3 models labelled as “effectively neutral” and the other 4 as “selection”, using 10-fold cross-validation with the 5 ITH metrics as features.
- Hyperparameter optimization/Grid Search for optimizing RBF parameters.

Methods - Evolutionary Mode Classifier

- A support vector machine classifier with a radial basis function (RBF) kernel was trained on 1400 simulated tumors derived from 7 different growth models first 3 models labelled as “effectively neutral” and the other 4 as “selection”, using 10-fold cross-validation with the 5 ITH metrics as features.
- Hyperparameter optimization/Grid Search for optimizing RBF parameters.
- Another SVM was trained on the ICs derived from ICA of the 5 ITH metrics above.

General Outline

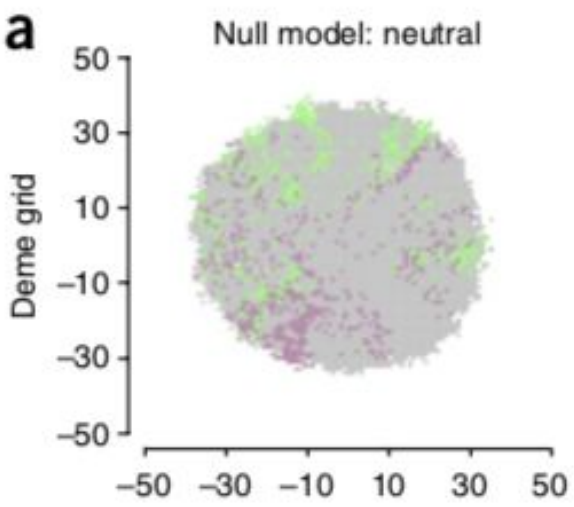
- Problem Statement
- Solution and Methods
- **Results**

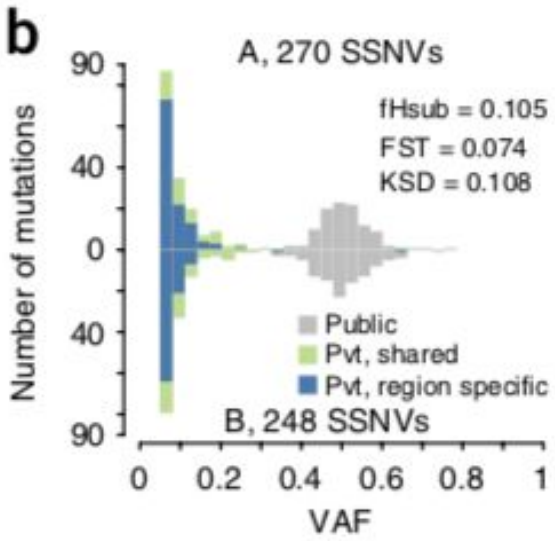
Results

Results

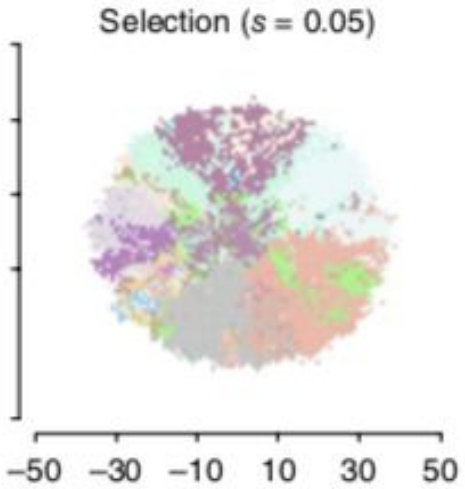
- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.

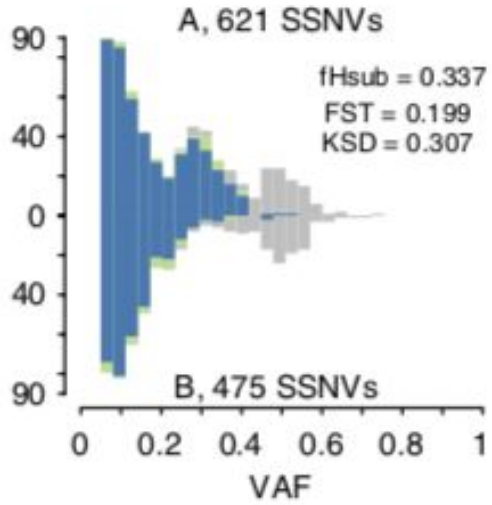
Results of Spatial Simulation



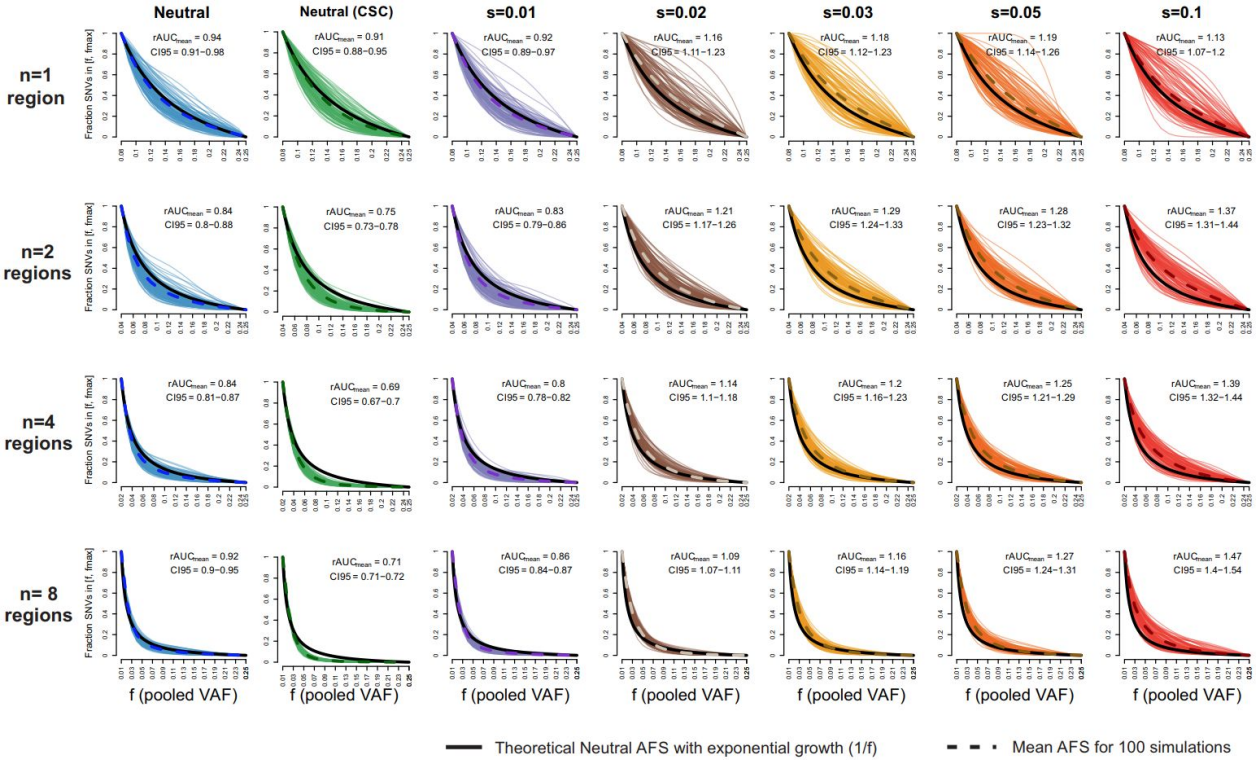


Results of Spatial Simulation





Results of Spatial Simulation



Results

- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.

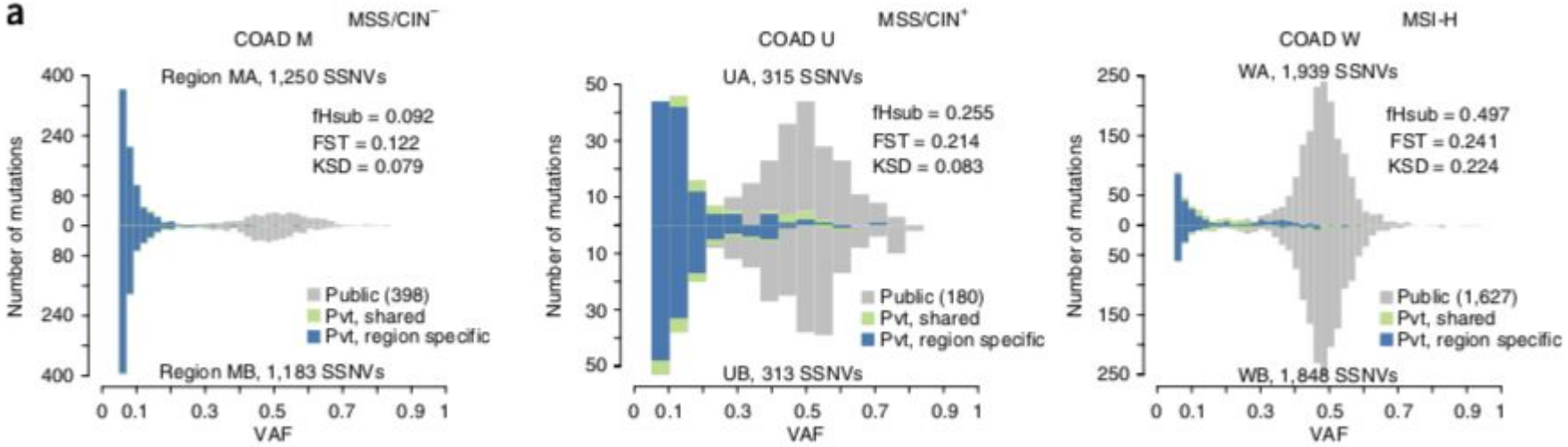
Results

- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.

Results

- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.
    - COAD - colorectal adenocarcinoma dataset (taken > 3cm apart)
        - Consists of 8 tumors W,M,G,N,U,O, and S.
        - These samples were passed through the VAP :
            - M, O, U showed ITH metrics comparable to virtual tumors under neutral growth
            - G,N,W, and S exhibited only slightly higher values, but consistent with weak selection.

Results


a
MSS/CIN⁻
COAD M
Region MA, 1,250 SSNVs
fHsub = 0.092
FST = 0.122
KSD = 0.079
Public (398)
Pvt, shared
Pvt, region specific
Region MB, 1,183 SSNVs
Number of mutations
VAF
MSS/CIN⁺
COAD U
UA, 315 SSNVs
fHsub = 0.255
FST = 0.214
KSD = 0.083
Public (180)
Pvt, shared
Pvt, region specific
UB, 313 SSNVs
MSI-H
COAD W
WA, 1,939 SSNVs
fHsub = 0.497
FST = 0.241
KSD = 0.224
Public (1,627)
Pvt, shared
Pvt, region specific
WB, 1,848 SSNVs

Results

- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.
    - COAD - colorectal adenocarcinoma dataset (taken > 3cm apart)
        - Consists of 8 tumors W,M,G,N,U,O, and S.
        - These samples were passed through the VAP :
            - M, O, U showed ITH metrics comparable to virtual tumors under neutral growth
            - G,N,W, and S exhibited only slightly higher values, but consistent with weak selection.

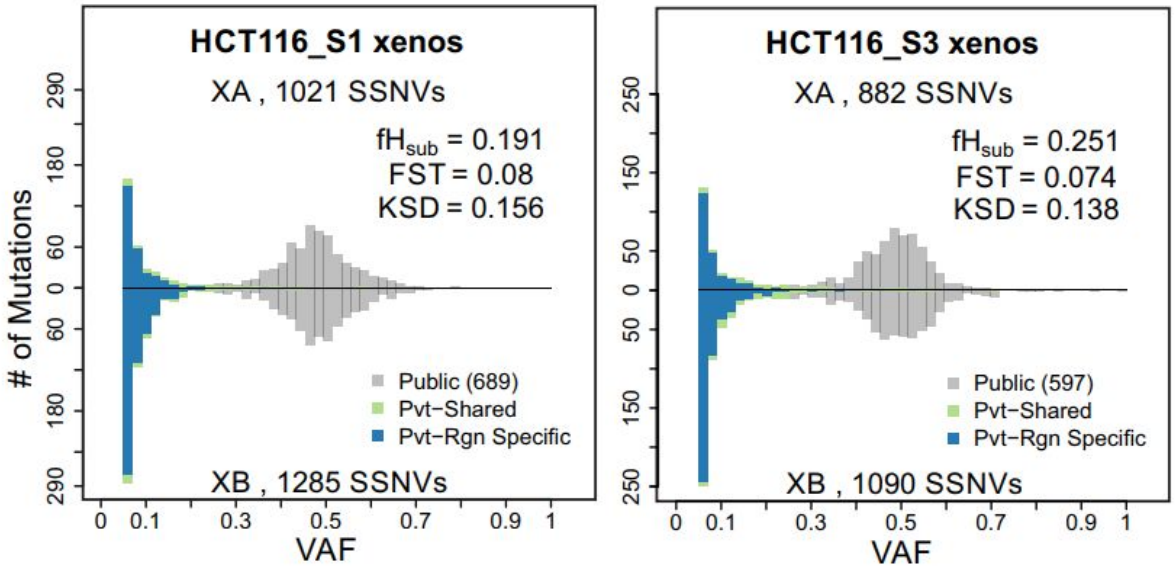
Results

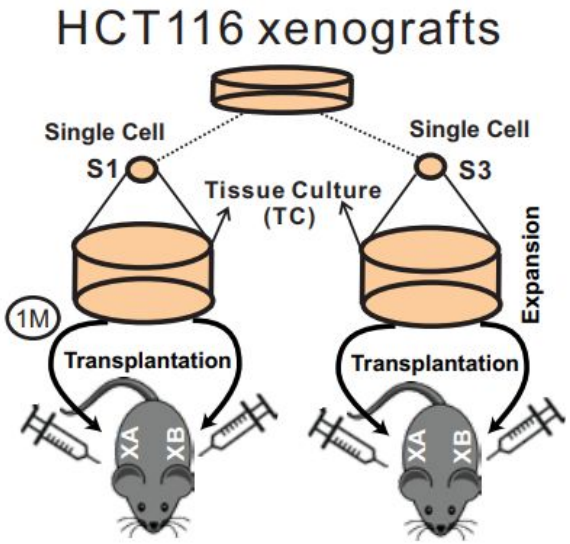
- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.
    - COAD - colorectal adenocarcinoma dataset (taken > 3cm apart)

Results

- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.
    - COAD - colorectal adenocarcinoma dataset (taken > 3cm apart)
    - In *vivo* COAD xenograft
        - Generated a type of cancerous COAD cell line and xenotransplanted it into immune-compromised mice.
        - The resulting tumors were sequenced.

Results


HCT116_S1 xenos
XA , 1021 SSNVs
fH$_{sub}$ = 0.191
FST = 0.08
KSD = 0.156
Public (689)
Pvt−Shared
Pvt−Rgn Specific
XB , 1285 SSNVs
# of Mutations
VAF
HCT116_S3 xenos
XA , 882 SSNVs
fH$_{sub}$ = 0.251
FST = 0.074
KSD = 0.138
Public (597)
Pvt−Shared
Pvt−Rgn Specific
XB , 1090 SSNVs
VAF


HCT116 xenografts
Single Cell
S1
Single Cell
S3
Tissue Culture
(TC)
Expansion
1M
Transplantation
Transplantation
XA
XB
XA
XB

Results

- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.
    - COAD - colorectal adenocarcinoma dataset (taken > 3cm apart)
    - In *vivo* COAD xenograft
        - Generated a type of cancerous COAD cell line and xenotransplanted it into immune-compromised mice.
        - The resulting tumors were sequenced.

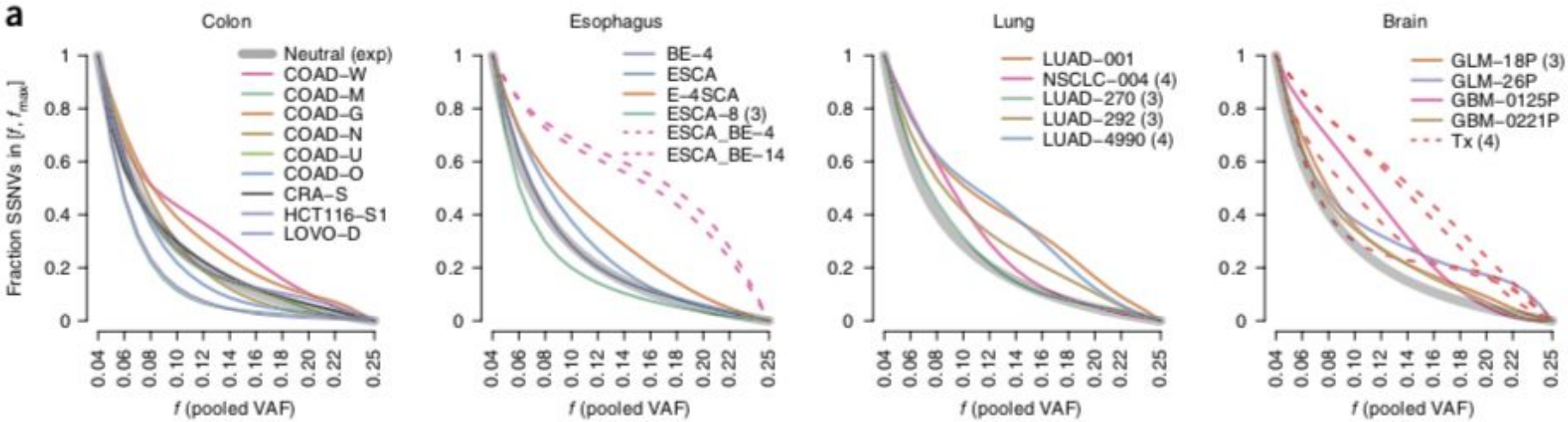
Results

- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.
    - COAD - colorectal adenocarcinoma dataset (taken > 3cm apart)
    - In *vivo* COAD xenograft

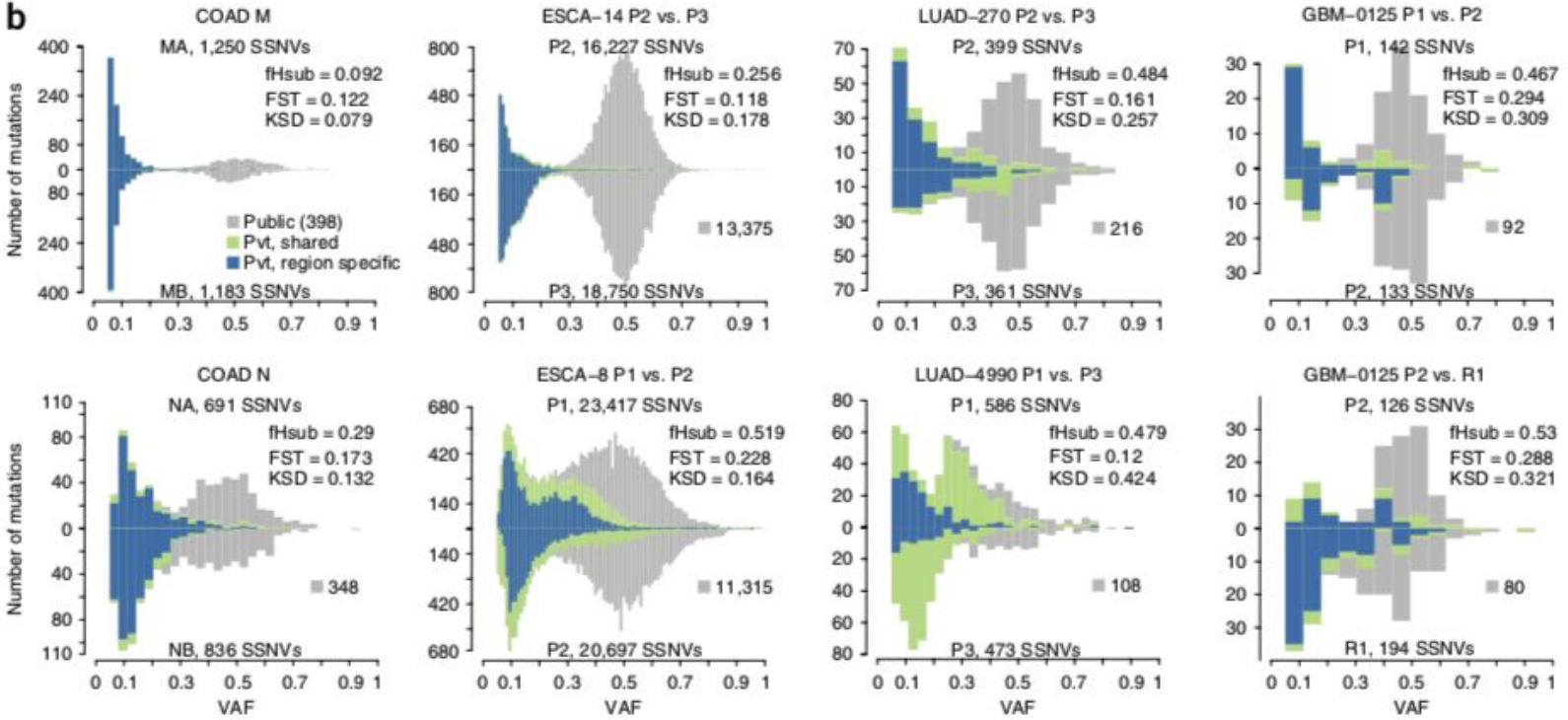
Results

- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.
    - COAD - colorectal adenocarcinoma dataset (taken > 3cm apart)
    - In *vivo* COAD xenograft
    - Samples from 4 types of solid tumors.
        - Additionally, for use as positive controls:
            - 2 tumors treated with a mutating agent known to impose positive selective pressure
            - Premalignant lesions (Barrett's esophagus lesions) were also used.

Results


a
Colon
Neutral (exp)
COAD–W
COAD–M
COAD–G
COAD–N
COAD–U
COAD–O
CRA–S
HCT116–S1
LOVO–D
Esophagus
BE–4
ESCA
E–4SCA
ESCA–8 (3)
ESCA_BE–4
ESCA_BE–14
Lung
LUAD–001
NSCLC–004 (4)
LUAD–270 (3)
LUAD–292 (3)
LUAD–4990 (4)
Brain
GLM–18P (3)
GLM–26P
GBM–0125P
GBM–0221P
Tx (4)
Fraction SSNVs in [f, f_max]
f (pooled VAF)
0
0.2
0.4
0.6
0.8
1
0.04
0.06
0.08
0.10
0.12
0.14
0.16
0.18
0.20
0.22
0.25

Results


b
COAD M
MA, 1,250 SSNVs
fHsub = 0.092
FST = 0.122
KSD = 0.079
Public (398)
Pvt, shared
Pvt, region specific
MB, 1,183 SSNVs
Number of mutations
ESCA–14 P2 vs. P3
P2, 16,227 SSNVs
fHsub = 0.256
FST = 0.118
KSD = 0.178
13,375
P3, 18,750 SSNVs
LUAD–270 P2 vs. P3
P2, 399 SSNVs
fHsub = 0.484
FST = 0.161
KSD = 0.257
216
P3, 361 SSNVs
GBM–0125 P1 vs. P2
P1, 142 SSNVs
fHsub = 0.467
FST = 0.294
KSD = 0.309
92
P2, 133 SSNVs
COAD N
NA, 691 SSNVs
fHsub = 0.29
FST = 0.173
KSD = 0.132
348
NB, 836 SSNVs
ESCA–8 P1 vs. P2
P1, 23,417 SSNVs
fHsub = 0.519
FST = 0.228
KSD = 0.164
11,315
P2, 20,697 SSNVs
LUAD–4990 P1 vs. P3
P1, 586 SSNVs
fHsub = 0.479
FST = 0.12
KSD = 0.424
108
P3, 473 SSNVs
GBM–0125 P2 vs. R1
P2, 126 SSNVs
fHsub = 0.53
FST = 0.288
KSD = 0.321
80
R1, 194 SSNVs
0 0.1 0.3 0.5 0.7 0.9 1
VAF

Results

- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.
    - COAD - colorectal adenocarcinoma dataset (taken > 3cm apart)
    - In *vivo* COAD xenograft
    - Samples from 4 types of solid tumors + 4 positive controls

Results

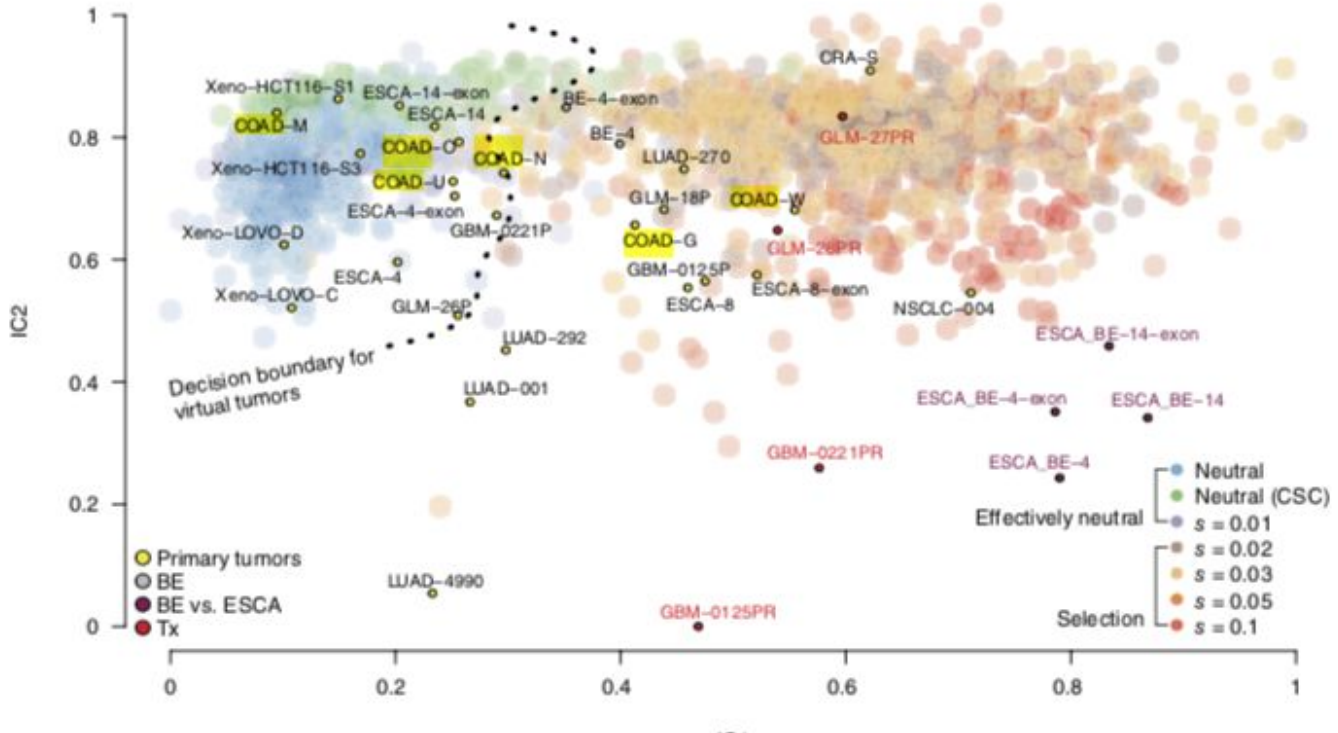
- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.

Results

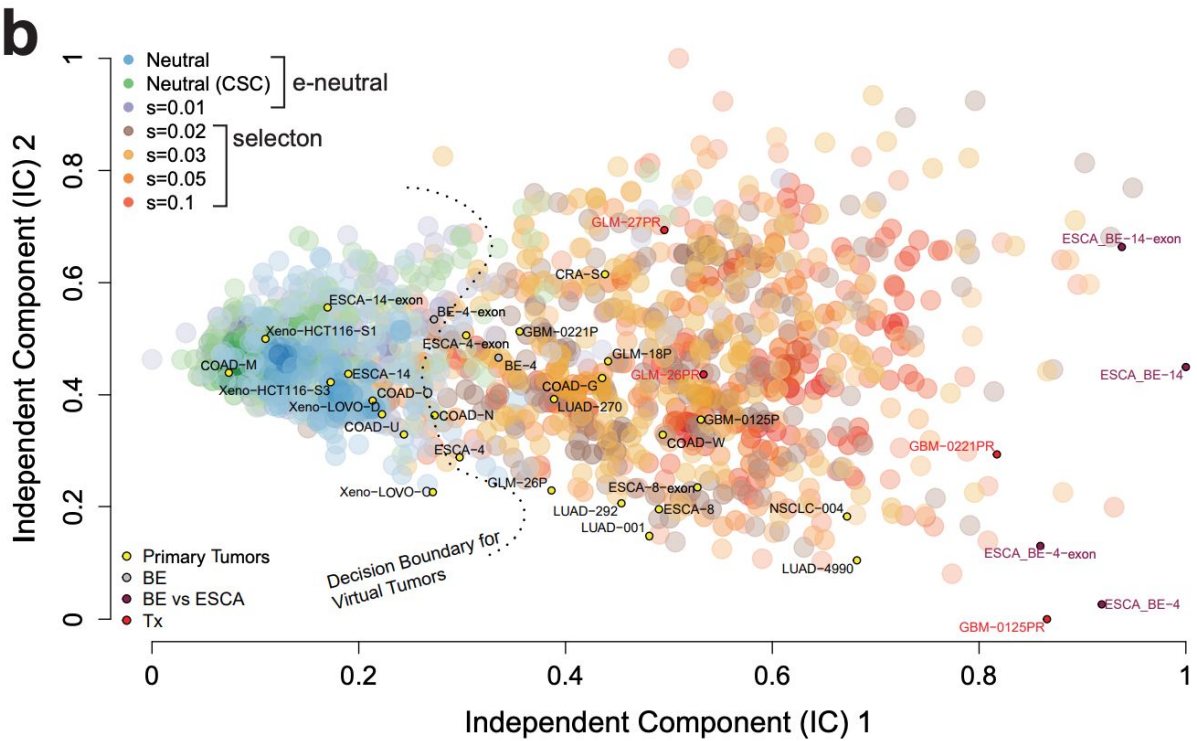
- Tumors evolving neutrally and through strong positive selection show fundamentally different patterns of intra-tumor heterogeneity, and these can be distinguished through multi-region sequencing.
- Patterns of ITH evaluated in several publicly available MRS datasets spanning multiple tumor types.
- ITH metrics were calculated for all datasets, classified, and projected into model space.
    - ICA (independent component analysis) using the five ITH metrics identified two distinct clusters, corresponding to selection with s>0.02.

Results

b


IC2
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
Xeno-HCT116-S1
ESCA-14-exon
ESCA-14
BE-4-exon
BE-4
CRA-S
GLM-27PR
COAD-M
COAD-O
COAD-N
LUAD-270
Xeno-HCT116-S3
COAD-U
GLM-18P
COAD-W
ESCA-4-exon
Xeno-LOVO-D
GBM-0221P
COAD-G
GLM-26PR
ESCA-4
GBM-0125P
Xeno-LOVO-C
ESCA-8-exon
ESCA-8
NSCLC-004
GLM-26P
LUAD-292
ESCA_BE-14-exon
Decision boundary for virtual tumors
LUAD-001
ESCA_BE-4-exon
ESCA_BE-14
GBM-0221PR
ESCA_BE-4
Neutral
Neutral (CSC)
Effectively neutral
s = 0.01
s = 0.02
s = 0.03
s = 0.05
s = 0.1
Selection
Primary tumors
BE
BE vs. ESCA
Tx
LUAD-4990
GBM-0125PR

Results


b
Neutral
Neutral (CSC)
e-neutral
s=0.01
s=0.02
s=0.03
s=0.05
s=0.1
selecton
Primary Tumors
BE
BE vs ESCA
Tx
Decision Boundary for Virtual Tumors
Independent Component (IC) 2
Independent Component (IC) 1
0
0.2
0.4
0.6
0.8
1
GLM-27PR
CRA-S
ESCA-14-exon
BE-4-exon
Xeno-HCT116-S1
GBM-0221P
ESCA-4-exon
GLM-18P
COAD-M
BE-4
GLM-26PR
ESCA-14
Xeno-HCT116-S3
COAD-G
COAD-D
Xeno-LOVO-D
LUAD-270
COAD-N
GBM-0125P
COAD-U
COAD-W
ESCA-4
GLM-26P
Xeno-LOVO-C
ESCA-8-exon
LUAD-292
ESCA-8
NSCLC-004
LUAD-001
LUAD-4990
ESCA_BE-14-exon
ESCA_BE-14
GBM-0221PR
ESCA_BE-4-exon
ESCA_BE-4
GBM-0125PR