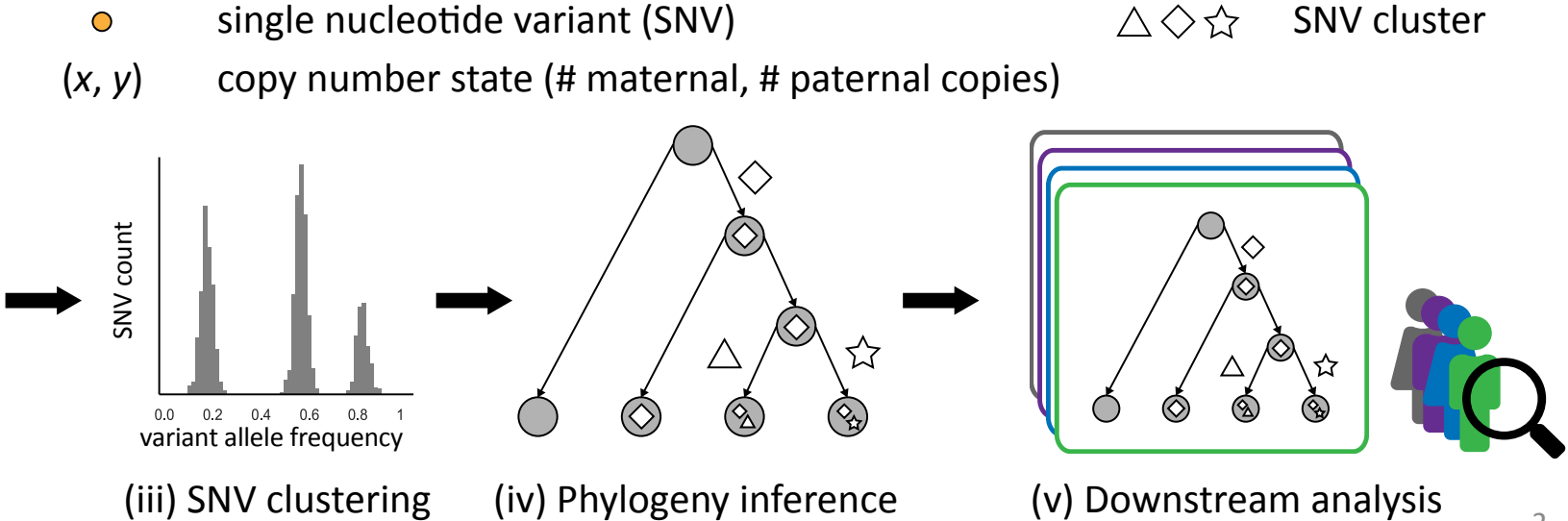# CS 598MEB
Computational Cancer Biology
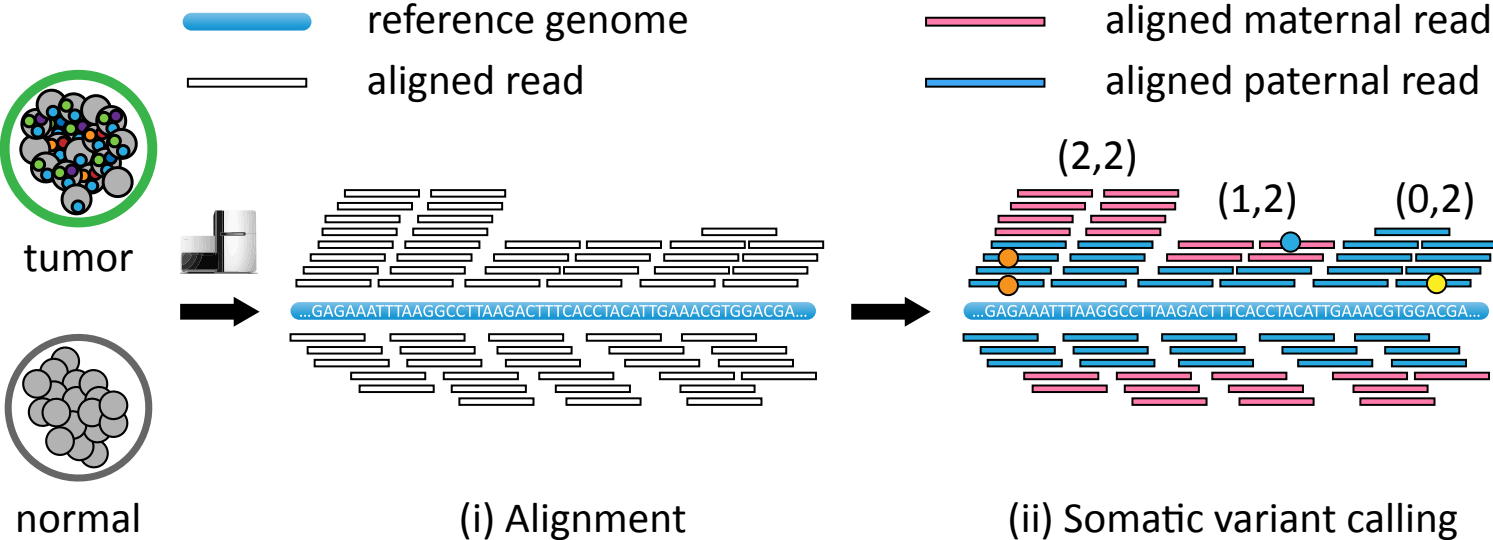## Lecture 11

Mohammed El-Kebir
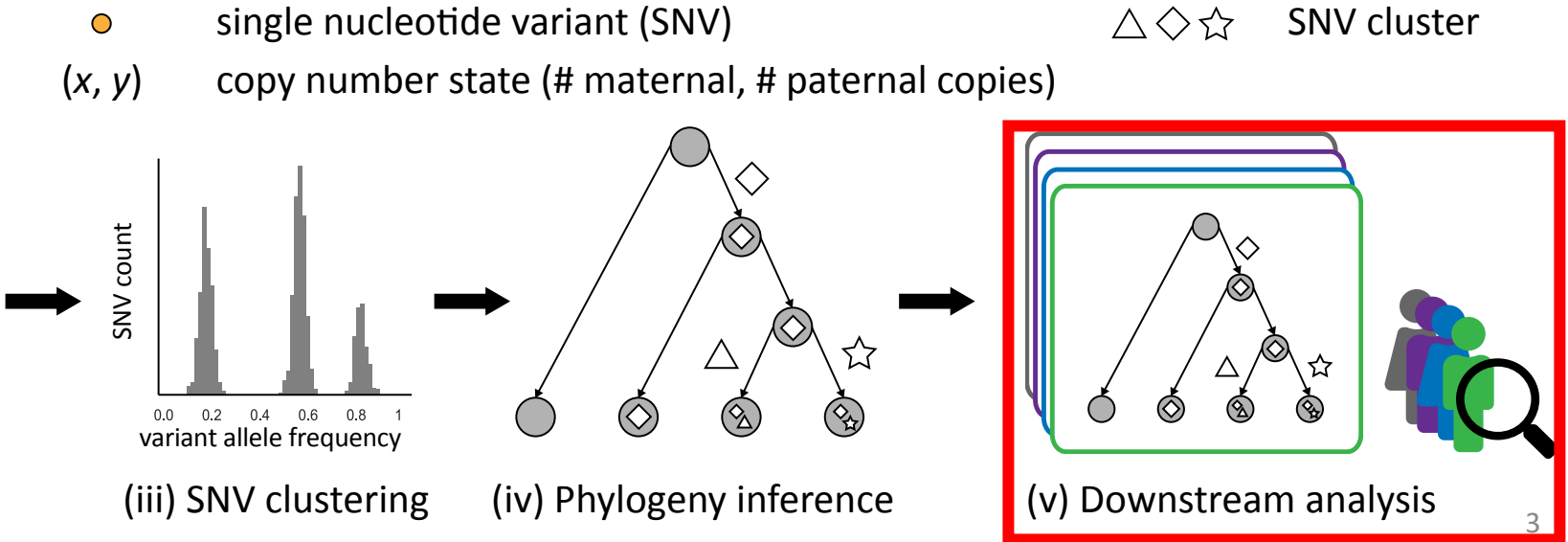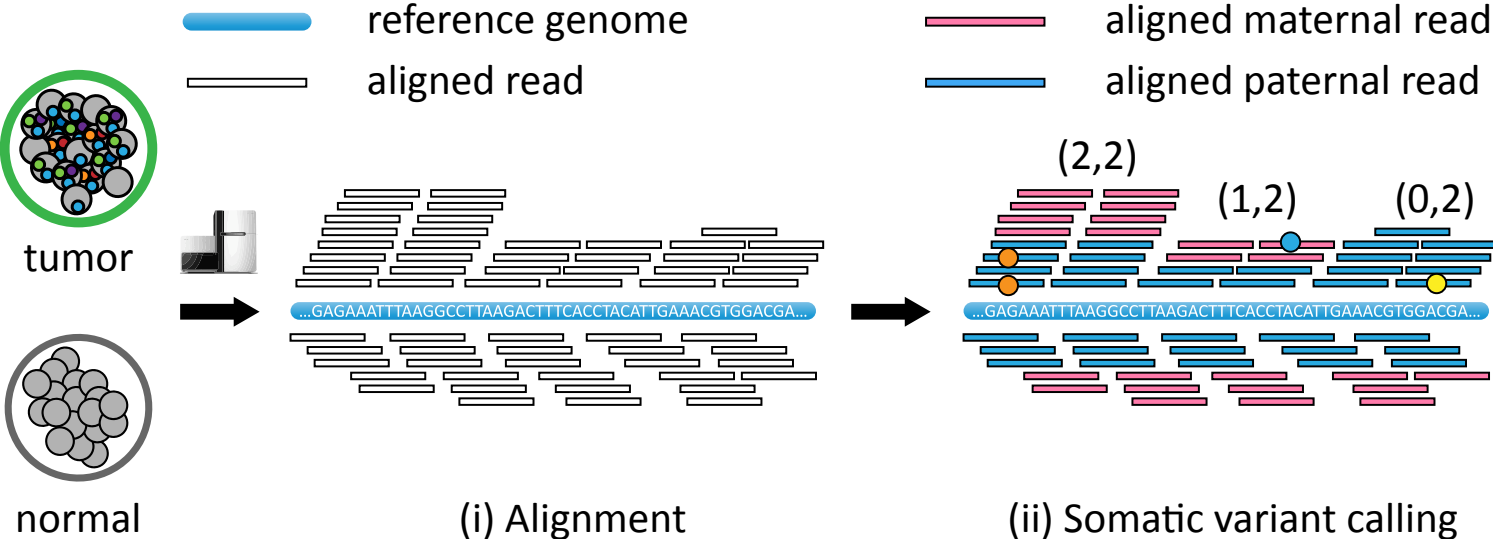
February 25, 2020

Cancer Phylogenetics Pipeline

reference genome
aligned read
aligned maternal read
aligned paternal read

(2,2) (1,2) (0,2)

...GAGAAATTTAAGGCCTTAAGACTTTCACCTACATTGAAACGTGGACGA...

...GAGAAATTTAAGGCCTTAAGACTTTCACCTACATTGAAACGTGGACGA...

tumor

normal

(i) Alignment

(ii) Somatic variant calling

single nucleotide variant (SNV)

(x, y)   copy number state (# maternal, # paternal copies)

△ ◇ ☆   SNV cluster

SNV count

variant allele frequency

(iii) SNV clustering   (iv) Phylogeny inference   (v) Downstream analysis

2

Cancer Phylogenetics Pipeline

reference genome
aligned read
aligned maternal read
aligned paternal read

tumor

normal

(i) Alignment

(2,2) (1,2) (0,2)

...GAGAAATTTAAGGCCTTAAGACTTTCACCTACATTGAAACGTGGACGA...

(ii) Somatic variant calling

● single nucleotide variant (SNV)

(x, y) copy number state (# maternal, # paternal copies)

△ ◇ ☆ SNV cluster

(iii) SNV clustering

(iv) Phylogeny inference

(v) Downstream analysis

3

Outline

- Metastasis
- Maximum parsimony
- Problem statement
- Complexity
- Algorithm & results
- Problem variants

Reading:

- **M. El-Kebir**, G. Satas and B.J. Raphael. Inferring parsimonious migration histories for metastatic cancers. Nature Genetics, 50:718-726, 2018.

- **M. El-Kebir**[†]. Parsimonious Migration History Problem: Complexity and Algorithms. WABI 2018, Helsinki, Finland, August 20-22, 2018.
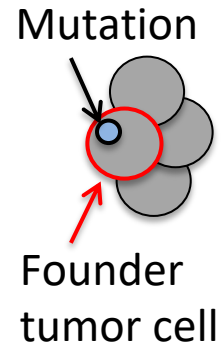
Formulating a combinatorial problem

Analyzing complexity & combinatorial structure

Designing an algorithm

Interpreting solutions and validating the algorithm

? **Biological question**

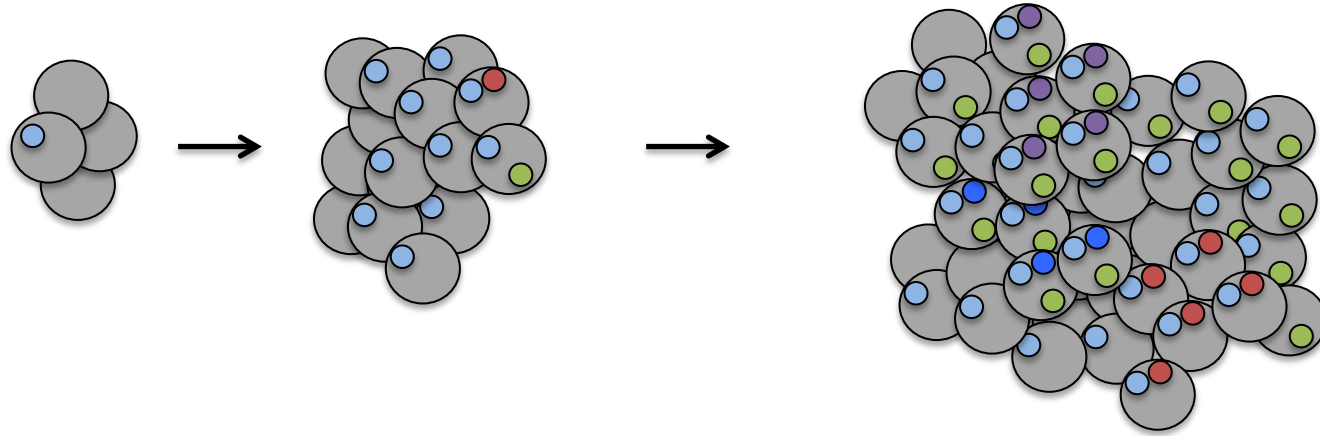# Tumorigenesis: (i) Cell Mutation

Clonal Theory of Cancer

[Nowell, 1976]

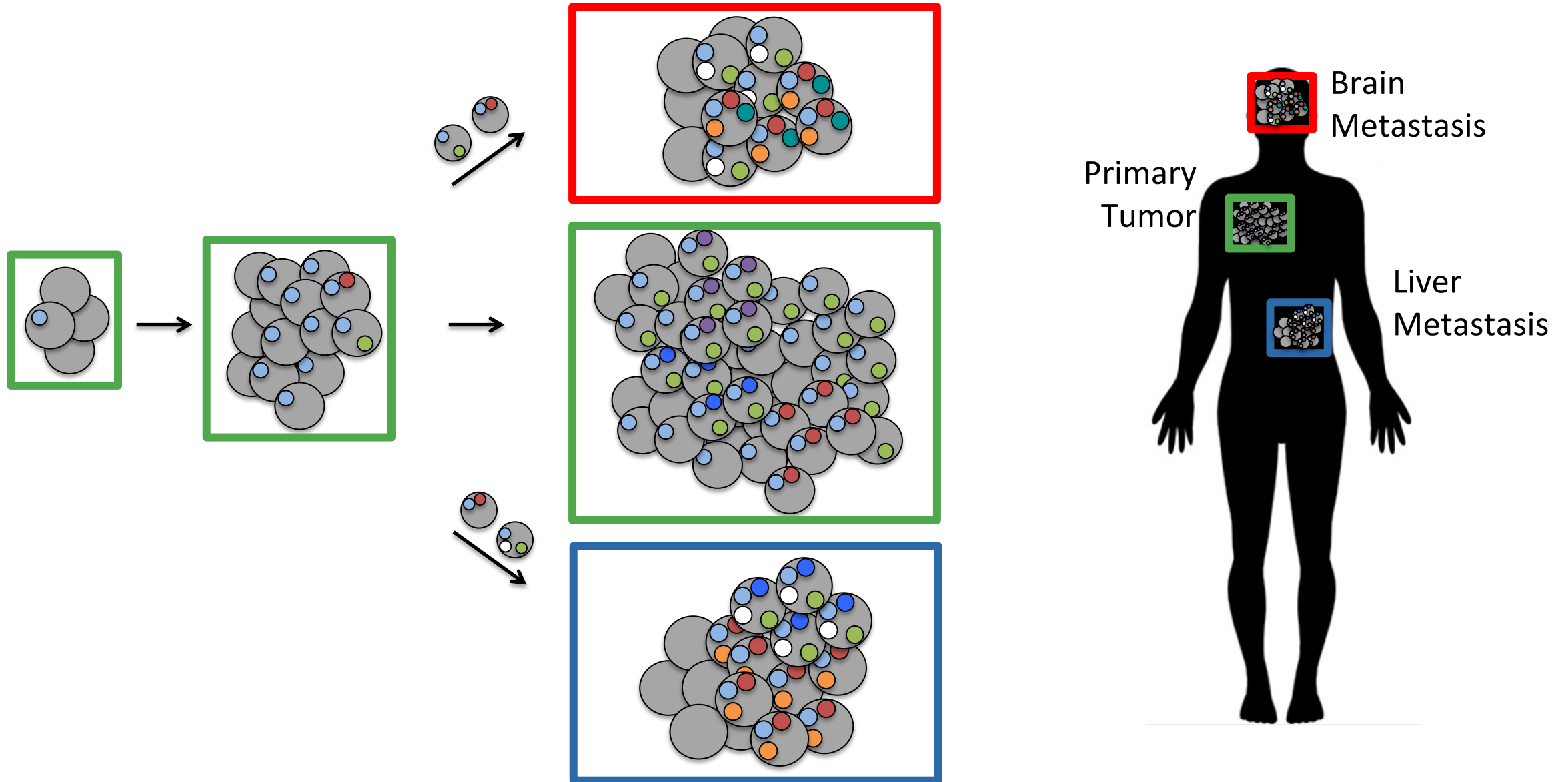# Tumorigenesis: (i) Cell Mutation, (ii) Cell Division
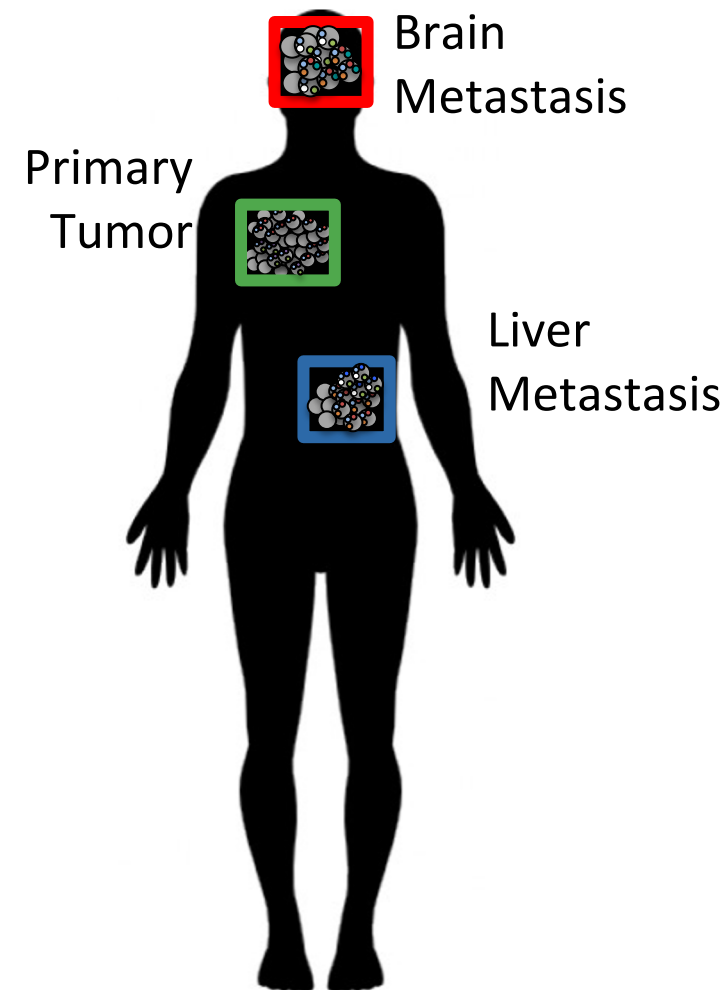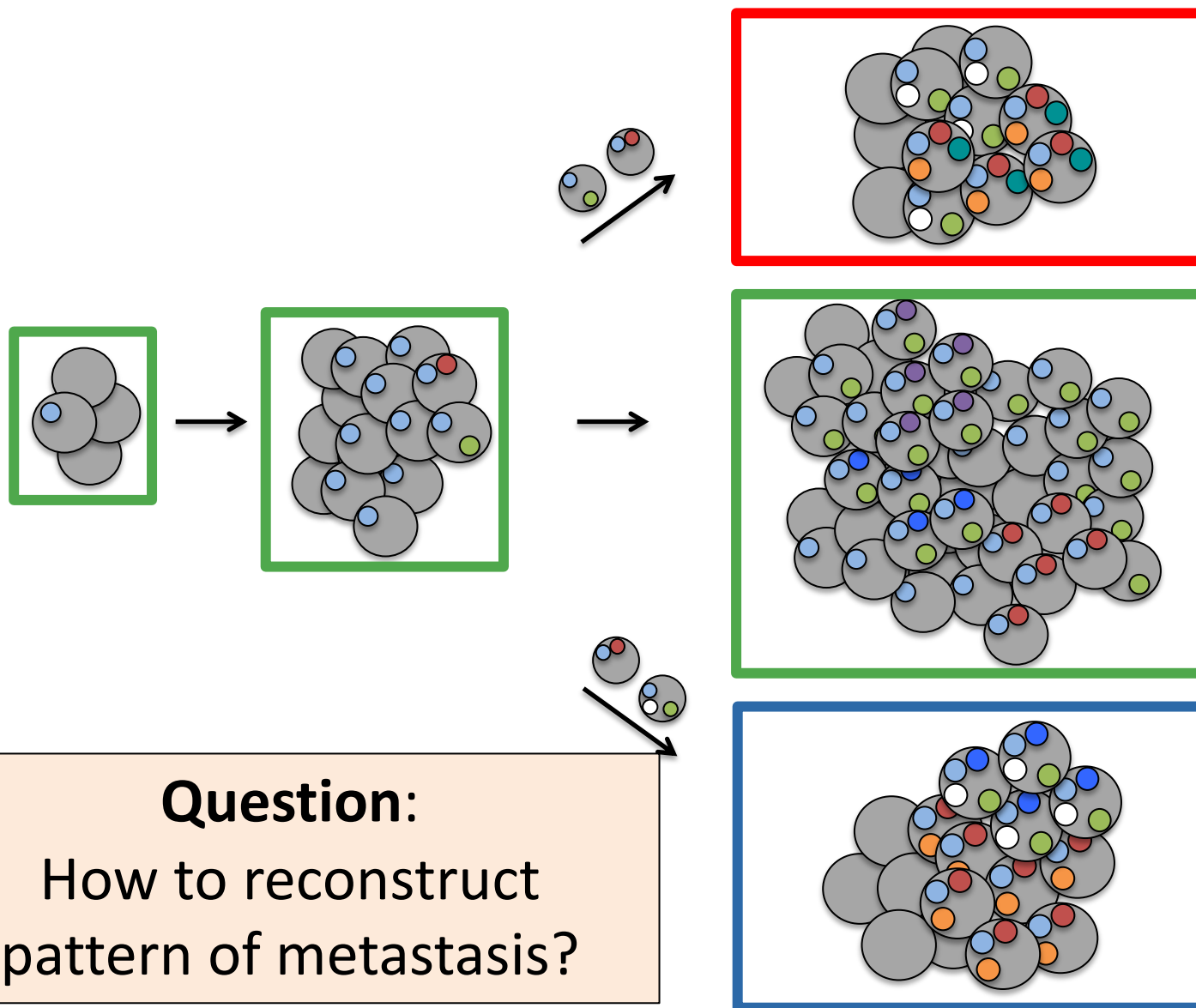
Clonal Theory of Cancer
[Nowell, 1976]

Heterogeneous Tumor

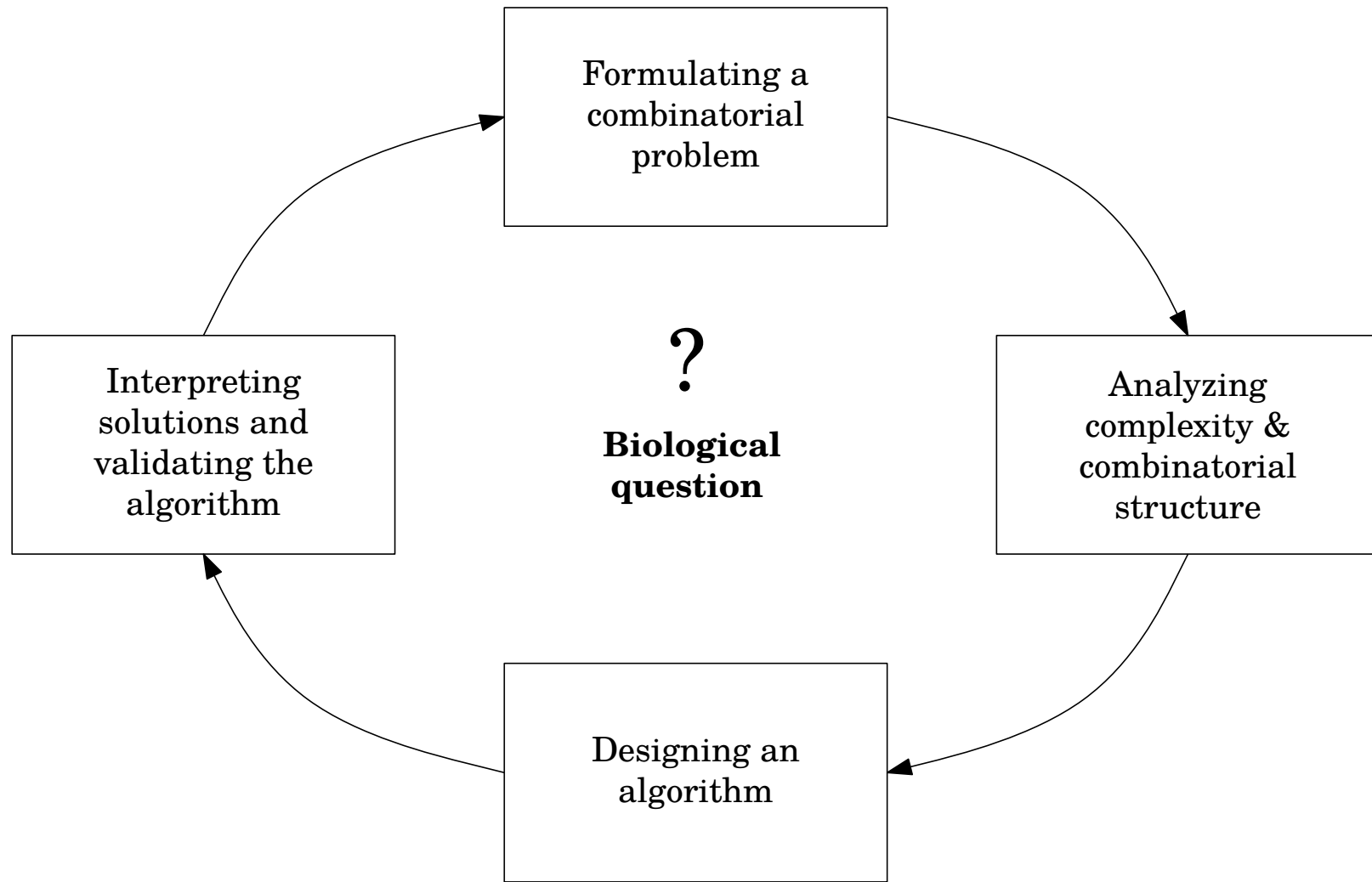# Tumorigenesis: (i) Cell Division, (ii) Mutation & (iii) Migration

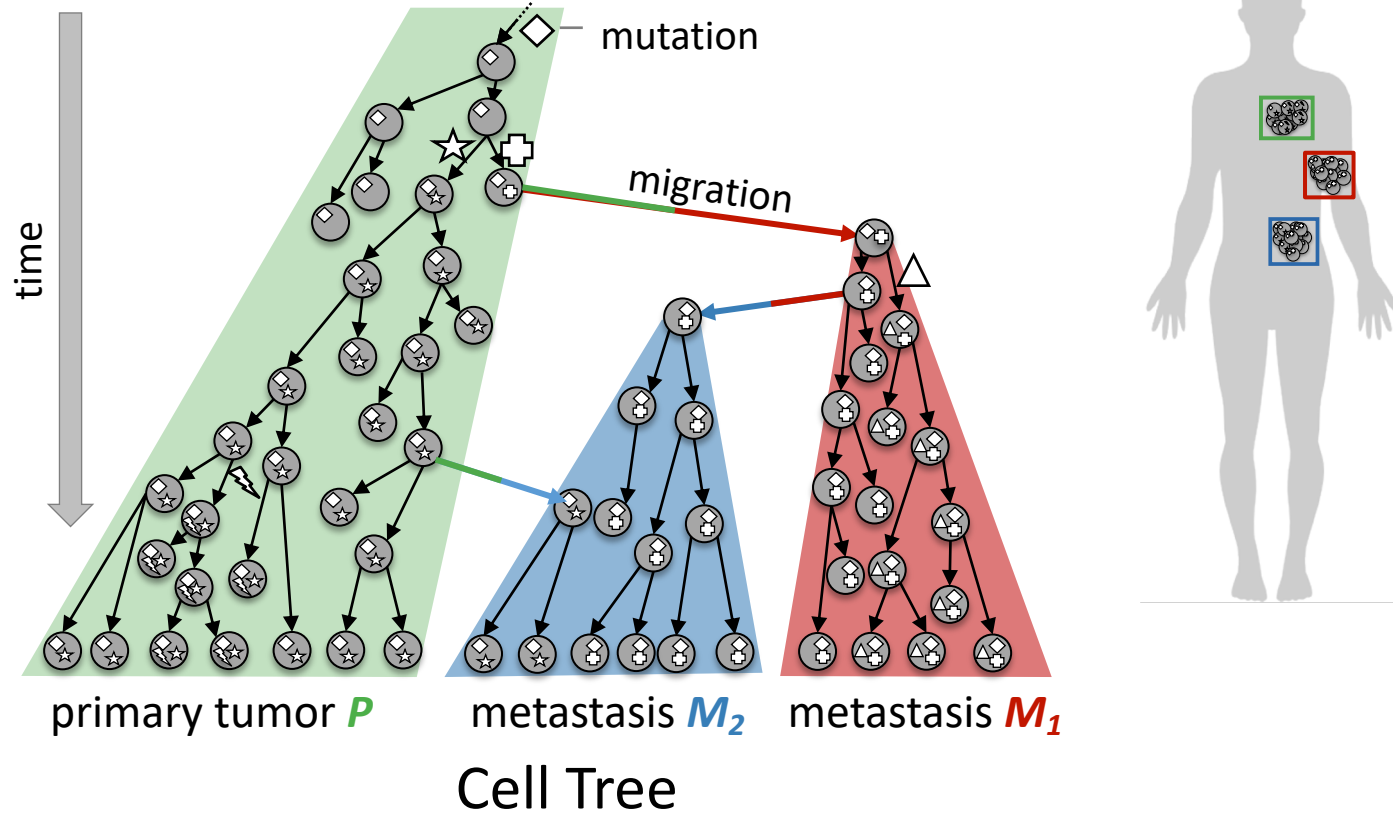# Tumorigenesis: (i) Cell Division, (ii) Mutation & (iii) Migration

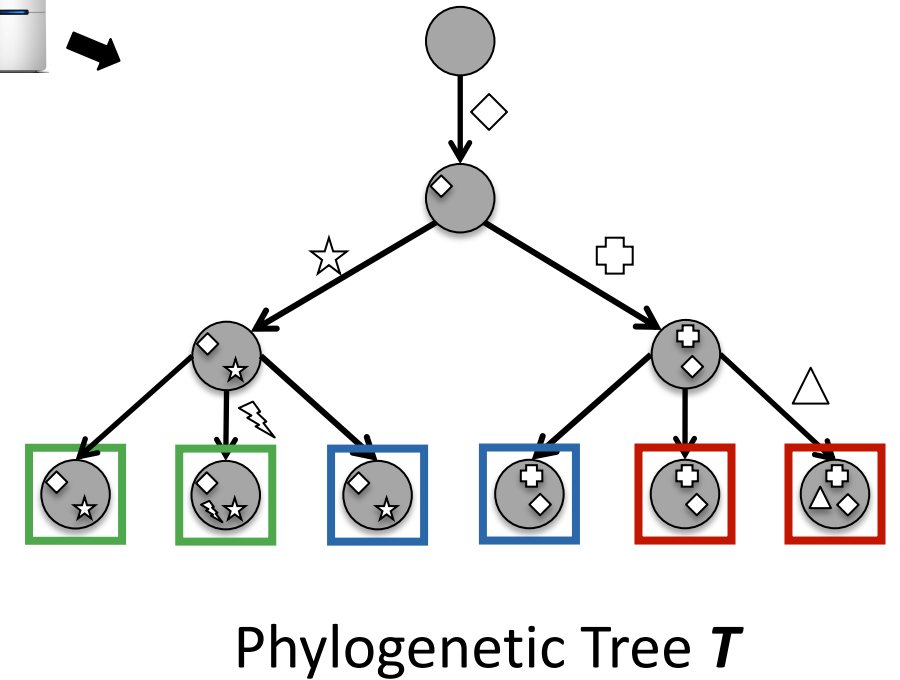**Question**: How to reconstruct pattern of metastasis?
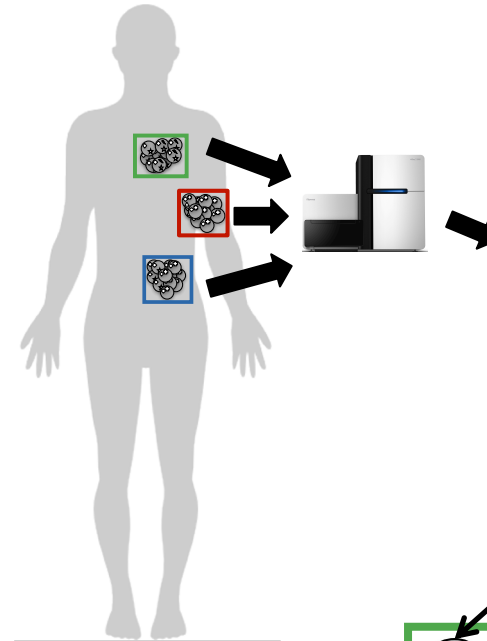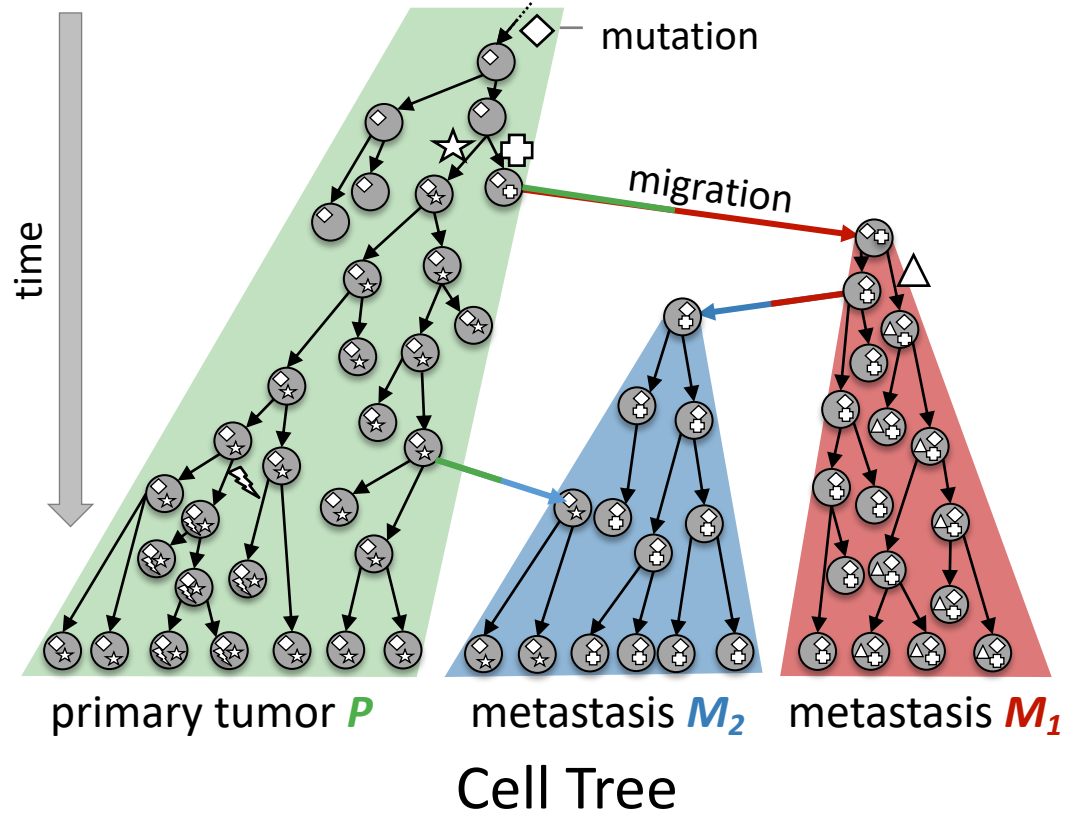
Key Challenge in Computational Biology

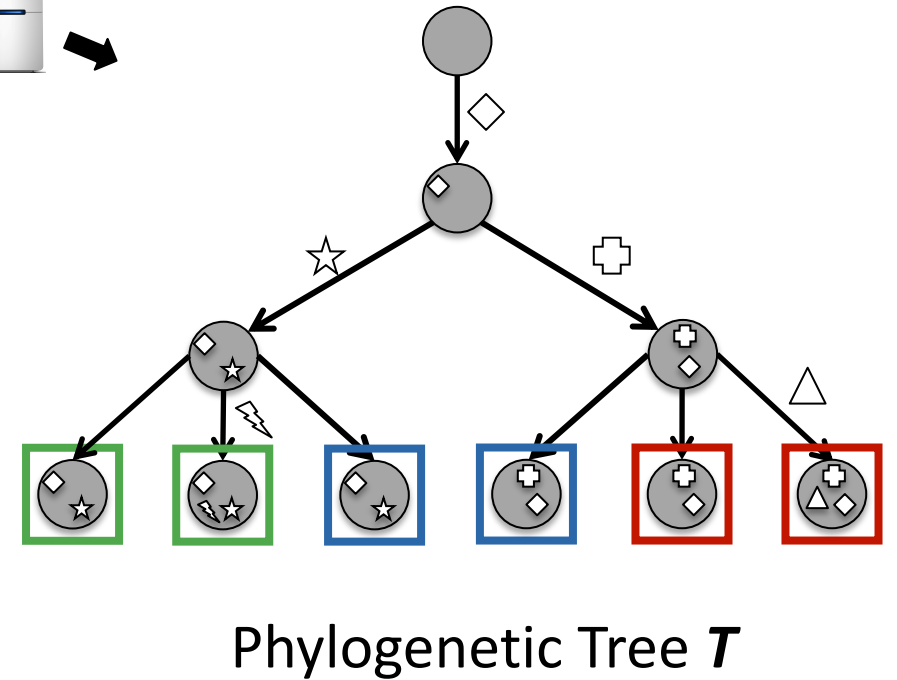Translating a biological problem into computer science
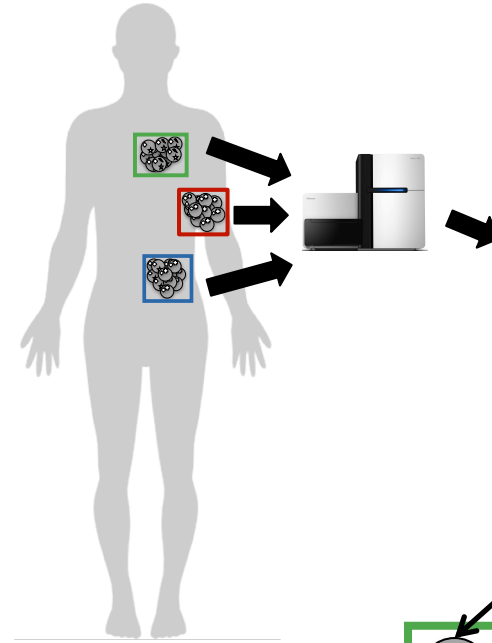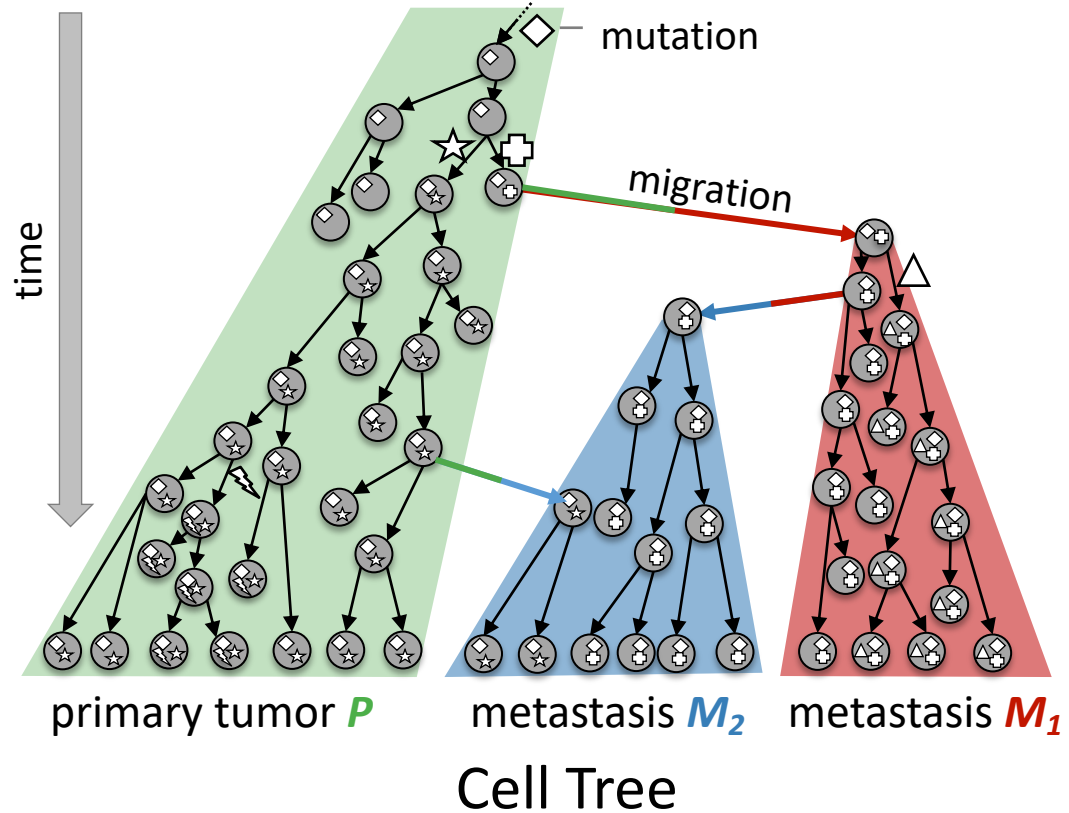
Tumorigenesis: (i) Cell Division, (ii) Mutation & (iii) Migration

Tumorigenesis: (i) Cell Division, (ii) Mutation & (iii) Migration

Cell Tree

Phylogenetic Tree **T**

Tumorigenesis: (i) Cell Division, (ii) Mutation & (iii) Migration

Cell Tree

Phylogenetic Tree **T**

Goal: Given phylogenetic tree **T**, find *parsimonious* vertex labeling **ℓ** with fewest migrations

Tumorigenesis: (i) Cell Division, (ii) Mutation & (iii) Migration

Question: Find parsimonious vertex labeling

Cell Tree

Phylogenetic Tree **T**

Goal: Given phylogenetic tree **T**, find *parsimonious* vertex labeling **ℓ** with fewest migrations

Tumorigenesis: (i) Cell Division, (ii) Mutation & (iii) Migration

Cell Tree
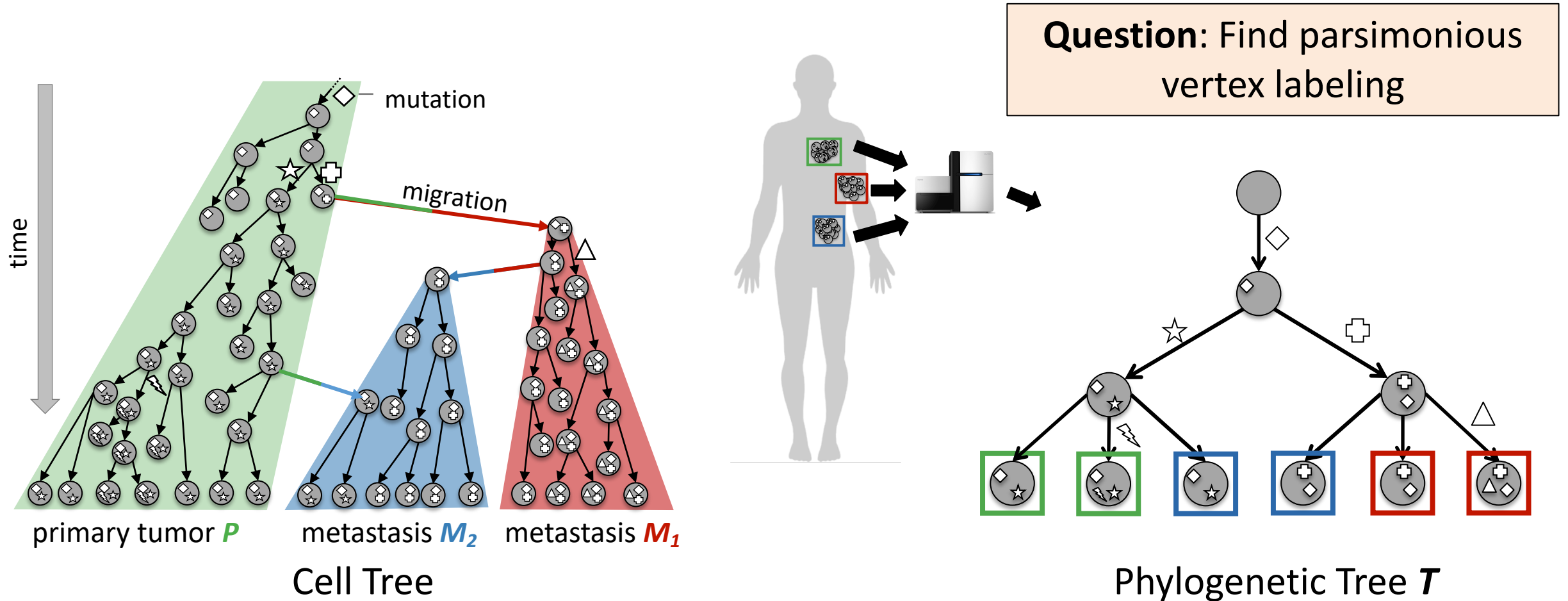
Phylogenetic Tree **T**

Goal: Given phylogenetic tree **T**, find ***parsimonious*** vertex labeling **ℓ** with fewest migrations

Outline

- Metastasis
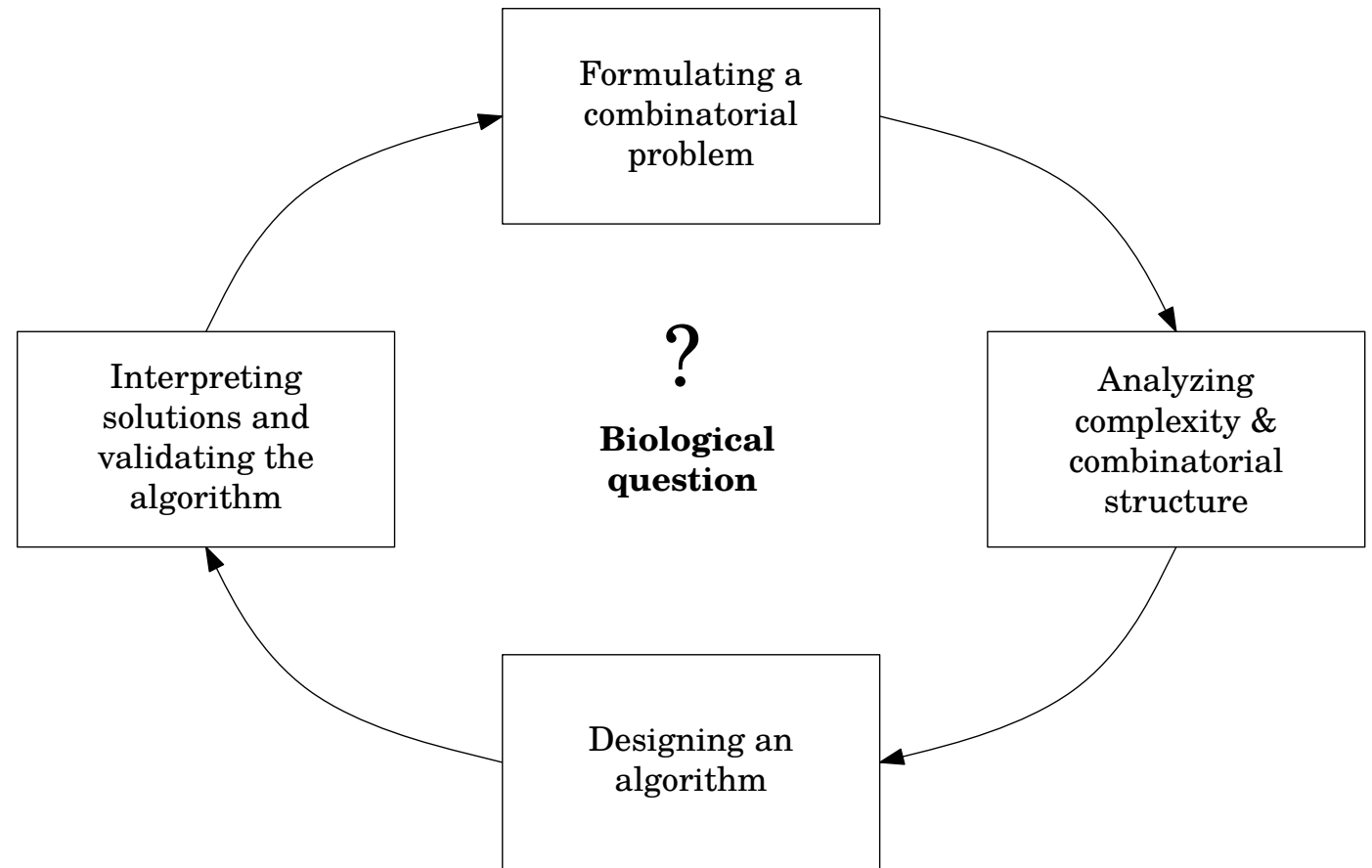- Maximum parsimony
- Problem statement
- Complexity
- Algorithm & results
- Problem variants

Reading:

- **M. El-Kebir**, G. Satas and B.J. Raphael. Inferring parsimonious migration histories for metastatic cancers. Nature Genetics, 50:718-726, 2018.

- **M. El-Kebir**[†]. Parsimonious Migration History Problem: Complexity and Algorithms. WABI 2018, Helsinki, Finland, August 20-22, 2018.

Formulating a combinatorial problem → Analyzing complexity & combinatorial structure → Designing an algorithm → Interpreting solutions and validating the algorithm

Biological question ?

Character-Based Phylogeny Reconstruction: Criterion
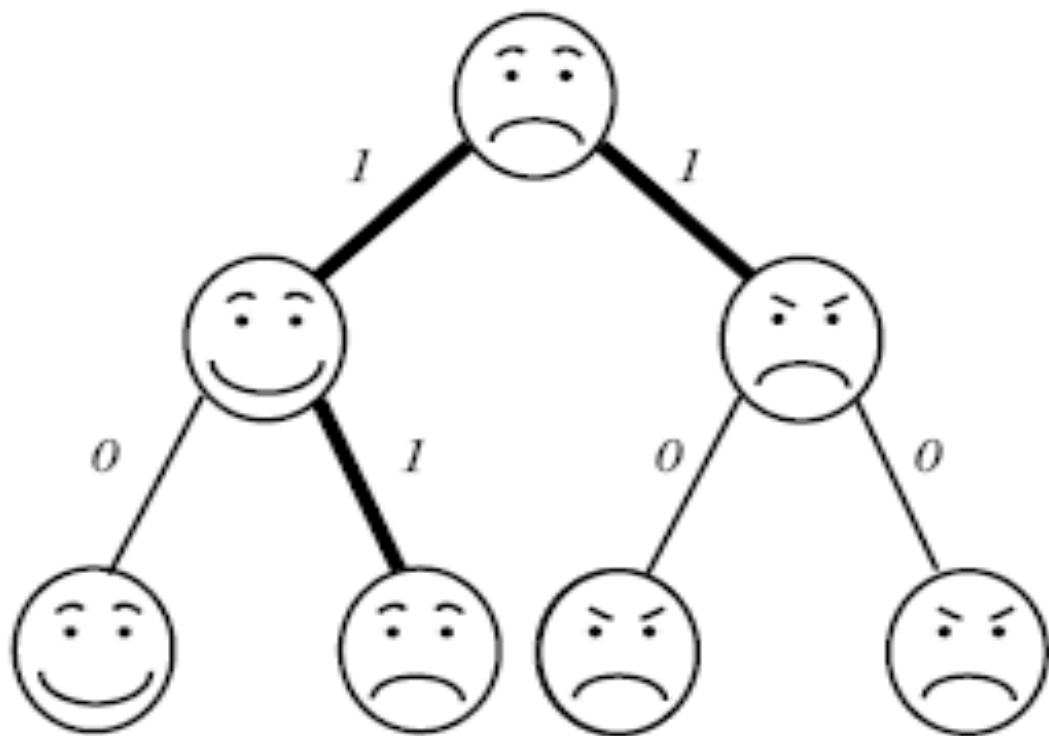
(a) *Parsimony Score=3*
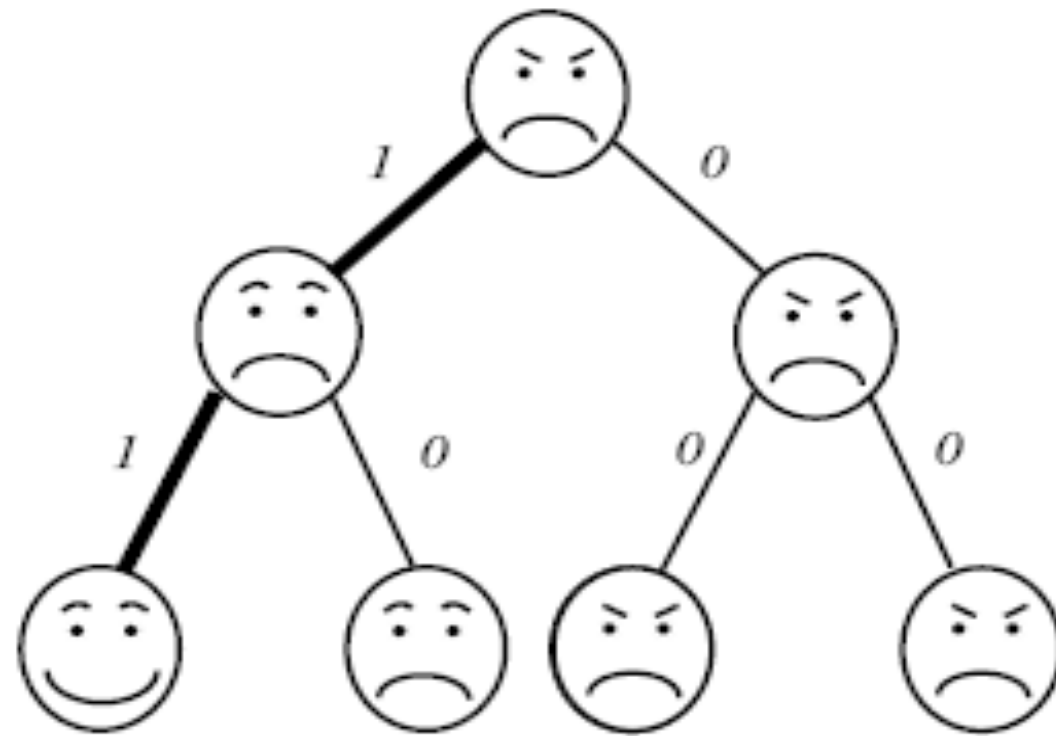
(b) *Parsimony Score=2*

Parsimony: minimize number of changes on edges of tree

A Small and a Large Problem

Small Maximum Parsimony Phylogeny Problem:
Given $m \times n$ matrix $A = [a_{i,j}]$ and tree T with m leaves, find assignment of character states to each internal vertex of T with minimum parsimony score.

Large Maximum Parsimony Phylogeny Problem:
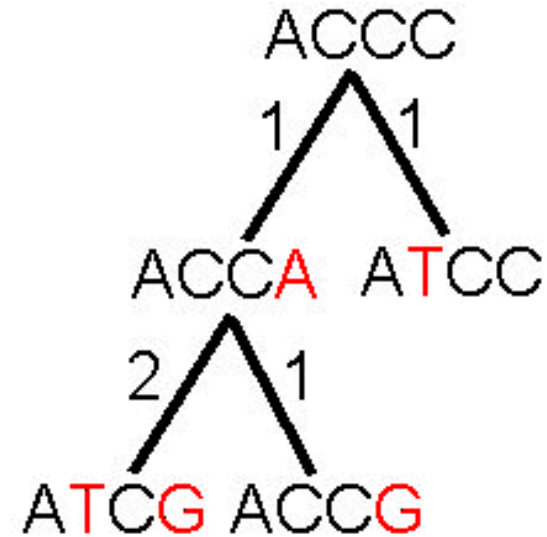Given $m \times n$ matrix $A = [a_{i,j}]$, find a tree T with m leaves labeled according to A and an assignment of character states to each internal vertex of T with minimum parsimony score.

Question: Are both problems easy (i.e. in P)?

Small Maximum Parsimony Phylogeny Problem
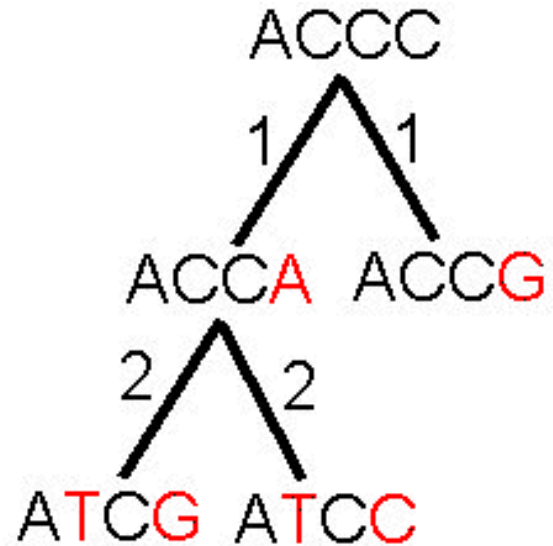
Key observations: (1) Characters can be solved independently.
(2) Optimal substructure in subtrees.

Recurrence for Small Maximum Parsimony Problem

Small Maximum Parsimony Phylogeny Problem:
Given rooted tree T whose leaves are labeled by $\sigma : L(T) \rightarrow \Sigma$, find assignment of states to each internal vertex of T with minimum parsimony score.

Recurrence for Small Maximum Parsimony Problem

Small Maximum Parsimony Phylogeny Problem:
Given rooted tree T whose leaves are labeled by $\sigma : L(T) \to \Sigma$, find assignment of states to each internal vertex of T with minimum parsimony score.

Let $\mu(v, s)$ be the minimum number of mutations in the subtree rooted at v when assigning state s to v.

Let $\delta(v)$ be the set of children of v.

Recurrence for Small Maximum Parsimony Problem

Small Maximum Parsimony Phylogeny Problem:
Given rooted tree T whose leaves are labeled by $\sigma : L(T) \to \Sigma$, find assignment of states to each internal vertex of T with minimum parsimony score.

Let $\mu(v, s)$ be the minimum number of mutations in the subtree rooted at v when assigning state s to v.

$$c(s, t) = \begin{cases} 0, & \text{if } s = t \\ 1, & \text{if } s \neq t, \end{cases}$$

Let $\delta(v)$ be the set of children of v.

$$\mu(v, s) = \min \begin{cases} \infty, & \text{if } v \in L(T) \text{ and } s \neq \sigma(v), \\ 0, & \text{if } v \in L(T) \text{ and } s = \sigma(v), \\ \sum_{w \in \delta(v)} \min_{t \in \Sigma} \{c(s, t) + \mu(w, t)\}, & \text{if } v \notin L(T). \end{cases}$$

Filling out DP Table and Traceback

Filling out μ Fill$(T, r(T), \sigma, \Sigma)$

Fill(T, v, σ, Σ) $O(m|\Sigma|^2)$

if $v \in L(T)$ then

For $s \in \Sigma$

if $s = \sigma(v)$ then

$\mu(v, s) = 0$

else

$\mu(v, s) = \infty$

else

For $w \in \delta(v)$ // children

Fill(T, w, σ, Σ)

$\mu(v, s) = 0$

For $w \in \delta(v)$

$\mu(v, s) += \min_{t \in \Sigma} \{ c(s, t) + \mu(w, t) \}$

Backtrace(T, v, μ)

if $v = r(T)$

$\sigma(r(T)) = \arg\min_{s \in \Sigma} \{ \mu(r(T), s) \}$

$= \sigma(u)$

else Let u be the parent of v and let s be the state of u

$\sigma(v) = \arg\min_{t \in \Sigma} \{ c(s, t) + \mu(v, t) \}$
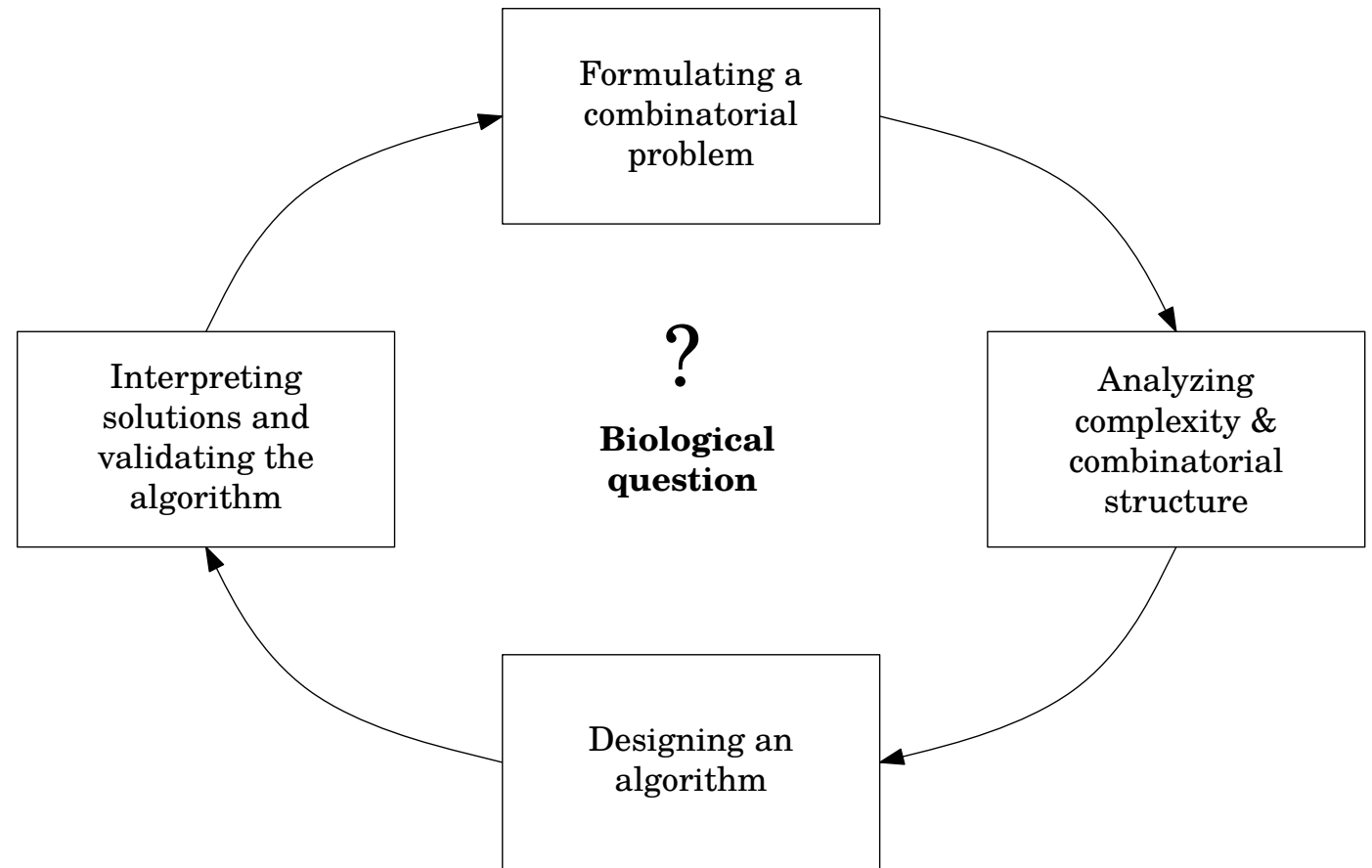
For $w \in \delta(v)$

Backtrace(T, w, μ)
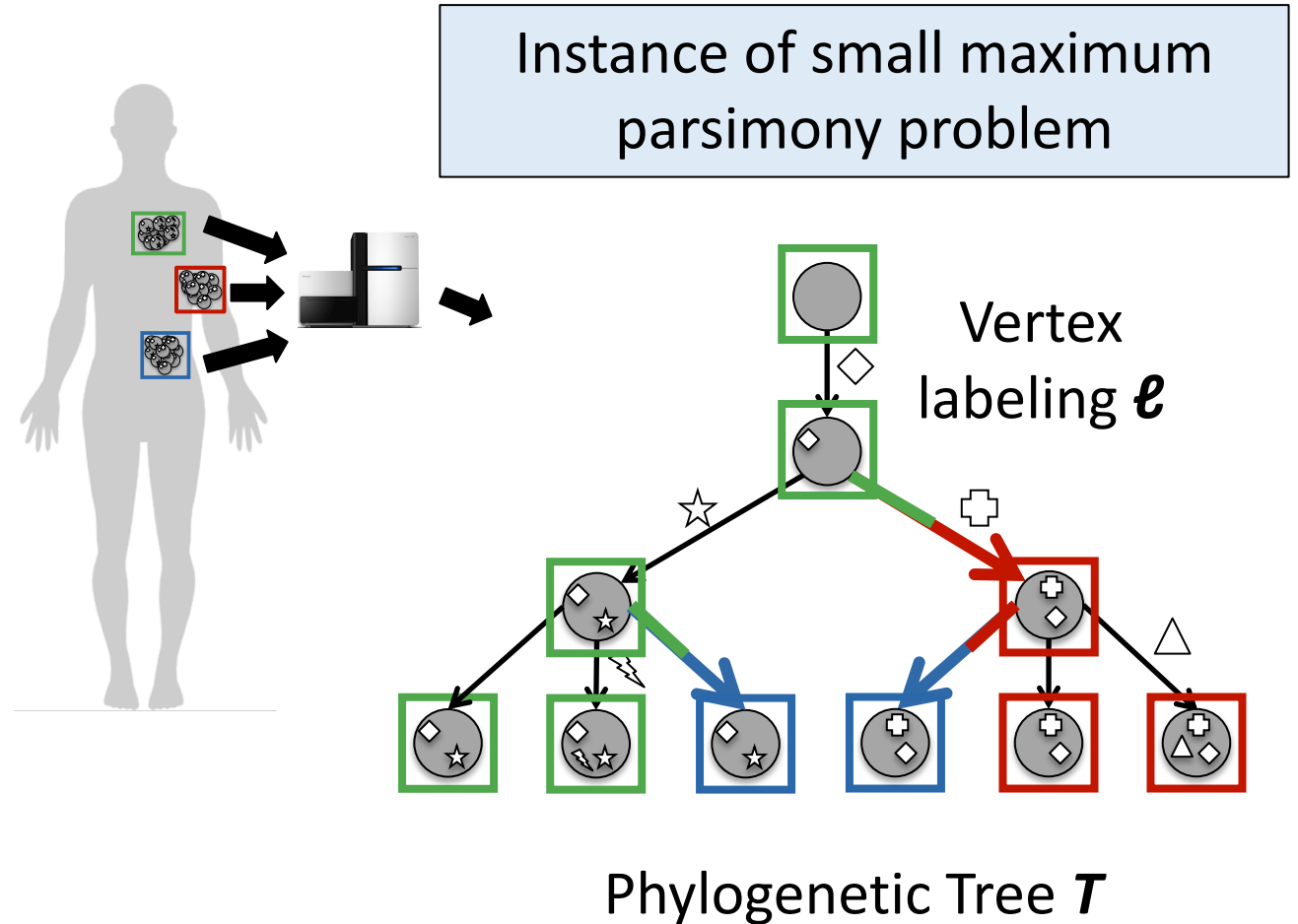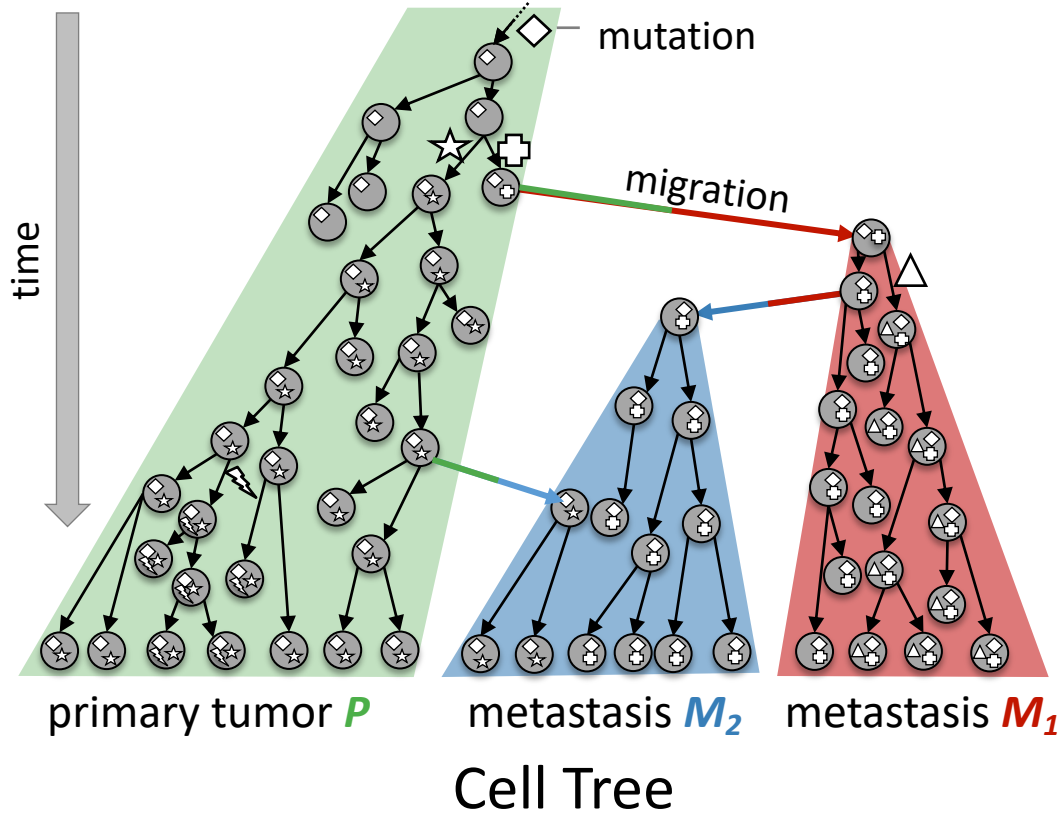
Let $r(T)$ be the root vertex

Outline

- Metastasis
- Maximum parsimony
- Problem statement
- Complexity
- Algorithm & results
- Problem variants

Reading:

- **M. El-Kebir**, G. Satas and B.J. Raphael. Inferring parsimonious migration histories for metastatic cancers. Nature Genetics, 50:718-726, 2018.

- **M. El-Kebir**[†]. Parsimonious Migration History Problem: Complexity and Algorithms. WABI 2018, Helsinki, Finland, August 20-22, 2018.

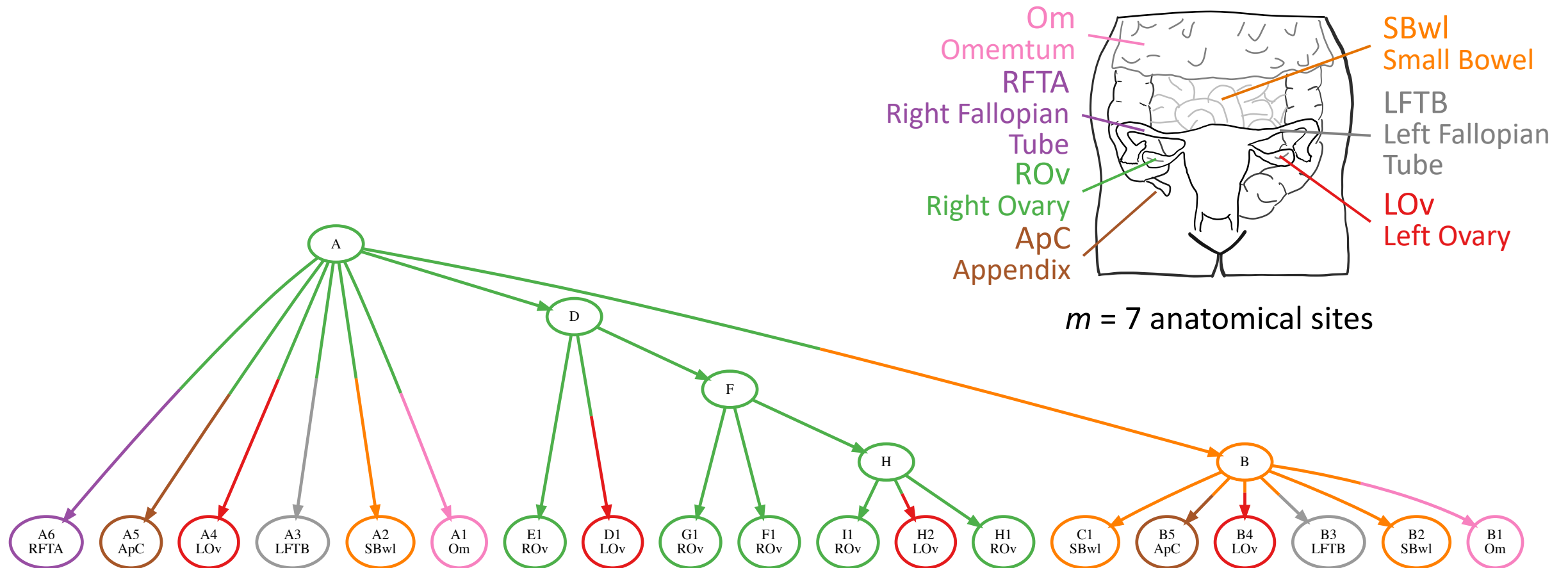# Tumorigenesis: (i) Cell Division, (ii) Mutation & (iii) Migration

Instance of small maximum parsimony problem

Vertex labeling ℓ

primary tumor **P**   metastasis **M₂**   metastasis **M₁**

Cell Tree

Phylogenetic Tree **T**

Goal: Given phylogenetic tree **T**, find *parsimonious* vertex labeling ℓ with fewest migrations

Slatkin, M. and Maddison, W. P. (1989). A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics, 123*(3), 603–613.

Minimum Migration Analysis in Ovarian Cancer

McPherson et al. (2016). Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer. *Nature Genetics*.

- Instance of the maximum parsimony small phylogeny problem [Fitch, 1971; Sankoff, 1975]
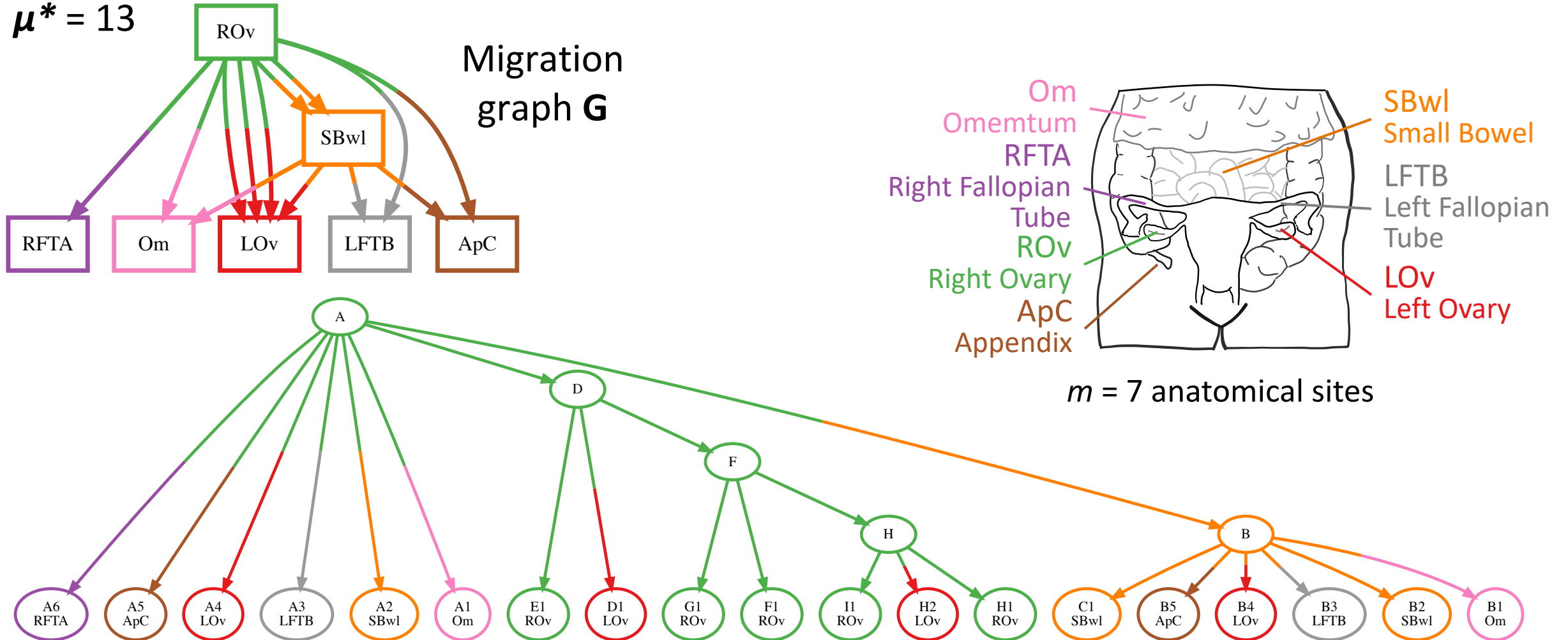

m = 7 anatomical sites

Minimum Migration Analysis in Ovarian Cancer

McPherson et al. (2016). Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer. *Nature Genetics*.
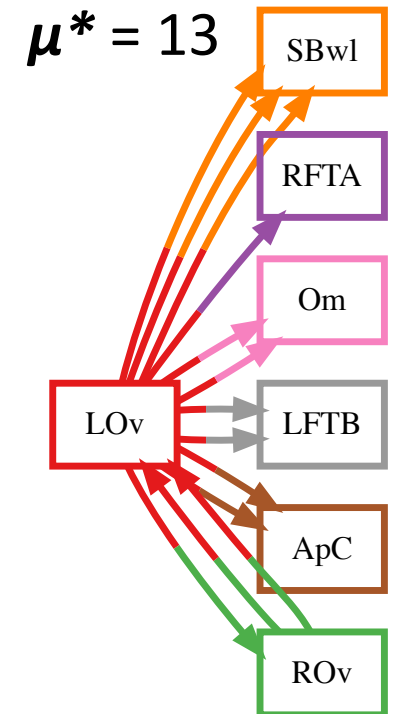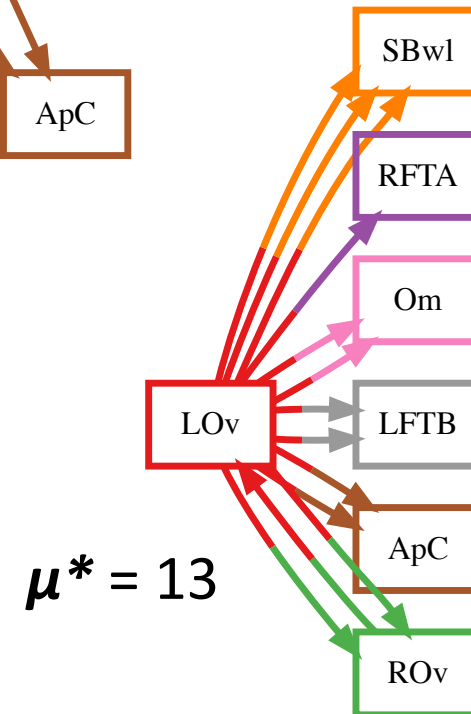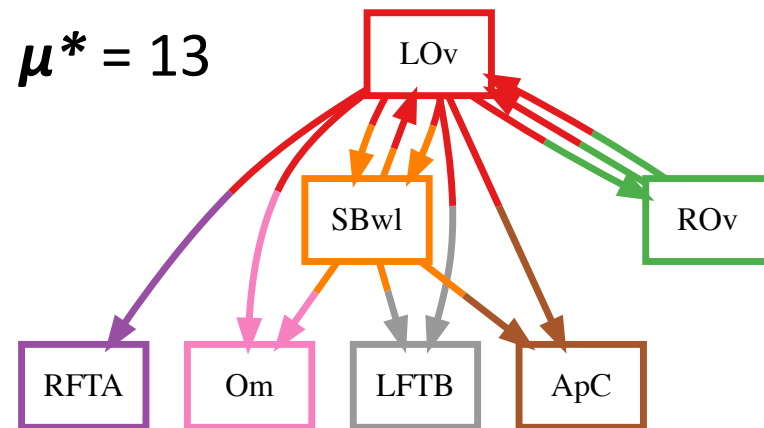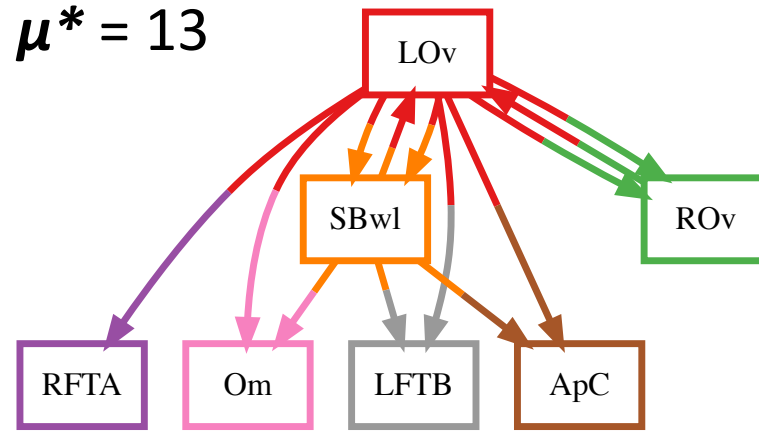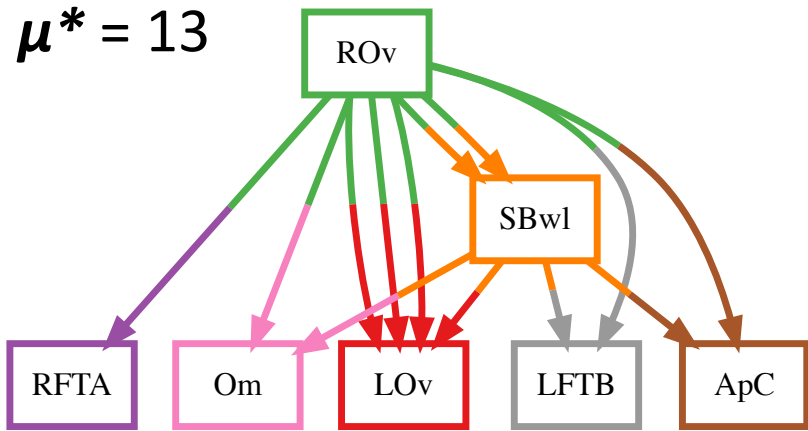
- Instance of the maximum parsimony small phylogeny problem [Fitch, 1971; Sankoff, 1975]

$\mu^* = 13$

Migration graph **G**

$m = 7$ anatomical sites

Minimum Migration History is *Not* Unique

- Enumerate all minimum-migration vertex labelings in the backtrace step

$\boldsymbol{\mu}^* = 13$

$\boldsymbol{\mu}^* = 13$

$\boldsymbol{\mu}^* = 13$

$\boldsymbol{\mu}^* = 13$

$\boldsymbol{\mu}^* = 13$

ApC	Appendix
LFTB	Left Fallopian Tube
LOv	Left Ovary
RFTA	Right Fallopian Tube
ROv	Right Ovary
SBwl	Small Bowel
Om	Omentum

5

Comigrations: Simultaneous Migrations of Multiple Clones

- Multiple tumor cells migrate simultaneously through the blood stream [Cheung et al., 2016]

- Second objective: number γ of comigrations is the number of multi-edges in migration graph G [†]

[†] Not necessarily true in the case of directed cycles

$\mu^* = 13$
$\gamma = 10$

Migration Graph G
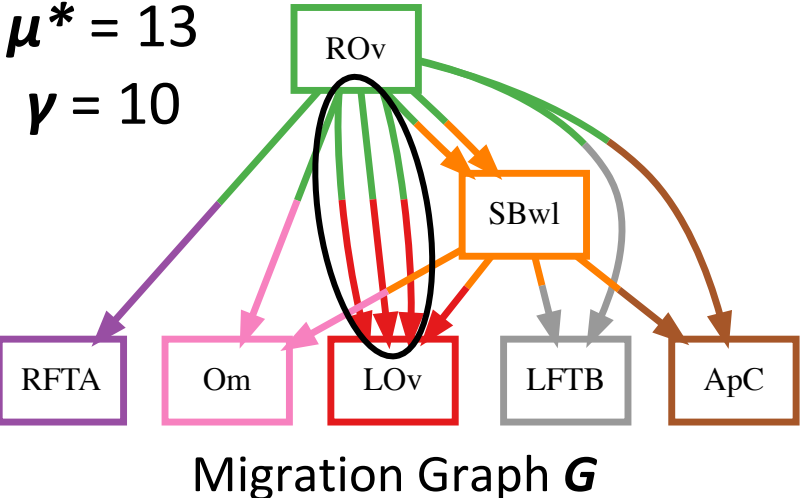
Clone Tree T

F Circulating tumor cell clusters

| Multicolored | Red Only | Cyan Only |

mTomato
CFP

DAPI
K14

10μm

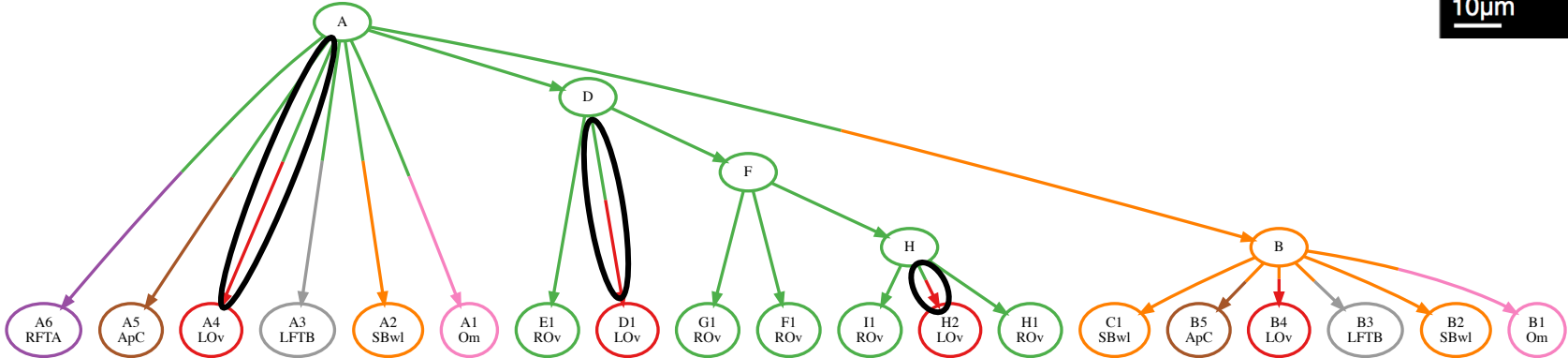| ApC | Appendix |
LFTB	Left Fallopian Tube
LOv	Left Ovary
RFTA	Right Fallopian Tube
ROv	Right Ovary
SBwl	Small Bowel
Om	Omentum

6

Comigrations: Simultaneous Migrations of Multiple Clones

- Multiple tumor cells migrate simultaneously through the blood stream [Cheung et al., 2016]

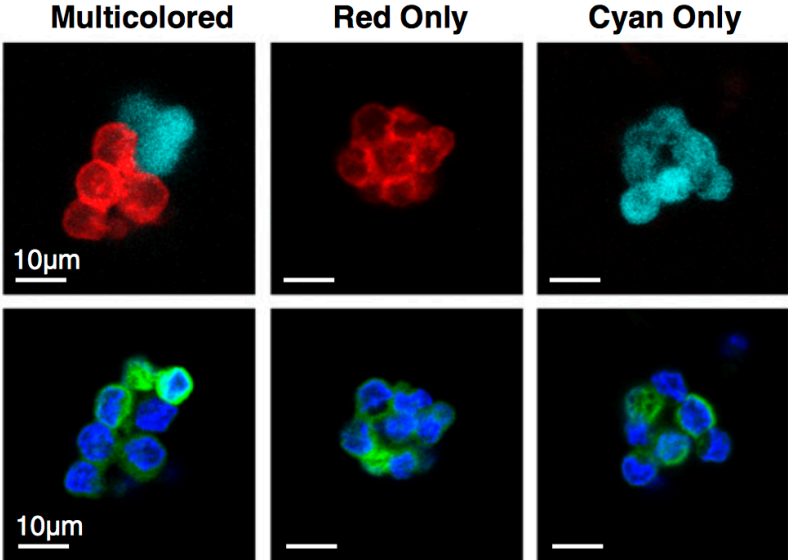- Second objective: number γ of comigrations is the number of multi-edges in migration graph G [†]

[†] Not necessarily true in the case of directed cycles

$\mu* = 13$
$\gamma = 10$

$\mu* = 13$
$\gamma = 11$

$\mu* = 13$
$\gamma = 11$

$\mu* = 13$
$\gamma = 7$

$\mu* = 13$
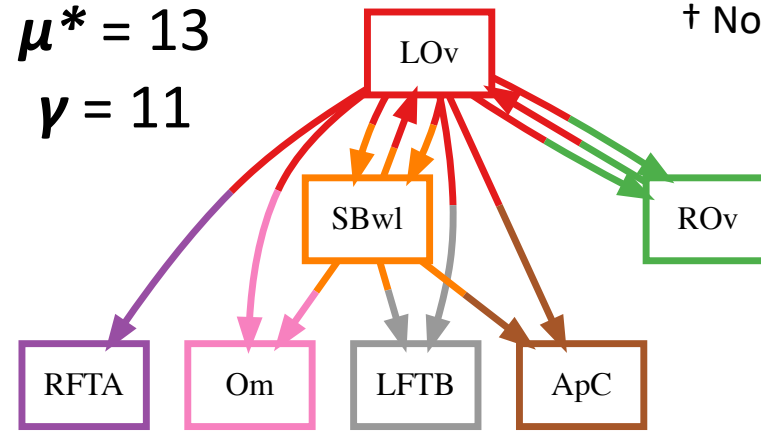$\gamma = 7$
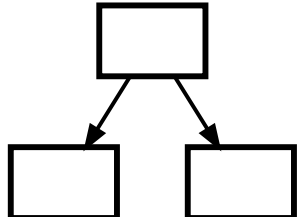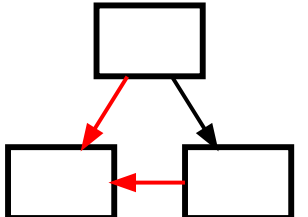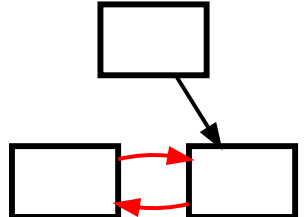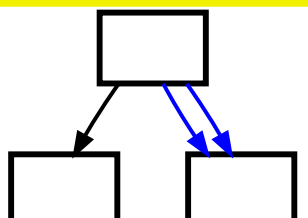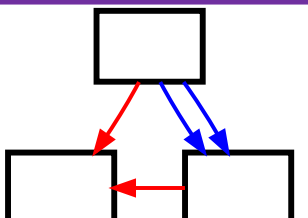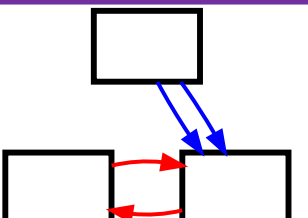
ApC	Appendix
LFTB	Left Fallopian Tube
LOv	Left Ovary
RFTA	Right Fallopian Tube
ROv	Right Ovary
SBwl	Small Bowel
Om	Omentum

Tradeoffs between Migrations, Comigrations and Migration Pattern

Constrained Multi-objective Optimization Problem

Parsimonious Migration History (PMH): Given a phylogenetic tree T and a set $\mathcal{P} \subseteq \{S, M, R\}$ of allowed migration patterns, find vertex labeling ℓ with minimum migration number $\mu^*(T)$ and smallest comigration number $\hat{\gamma}(T)$.

single-source seeding (S)

multi-source seeding (M)

reseeding (R)

a $\mathcal{P} = \{S\}$ $(\mu^*, \hat{\gamma}) = (6, 2)$

Vertex labeling ℓ_A Migration graph G_A (S)

Phylogenetic tree T Leaf labeling $\hat{\ell}$

b $\mathcal{P} = \{S, M\}$ $(\mu^*, \hat{\gamma}) = (5, 3)$ ℓ_B G_B (M) $(T, \hat{\ell})$

c $\mathcal{P} = \{S, M, R\}$ $(\mu^*, \hat{\gamma}) = (4, 4)$ ℓ_C G_C (R) $(T, \hat{\ell})$

El-Kebir, M., Satas, G., & Raphael, B. J. (2018). Inferring parsimonious migration histories for metastatic cancers. *Nature Genetics, 50*(5), 718–726.

8
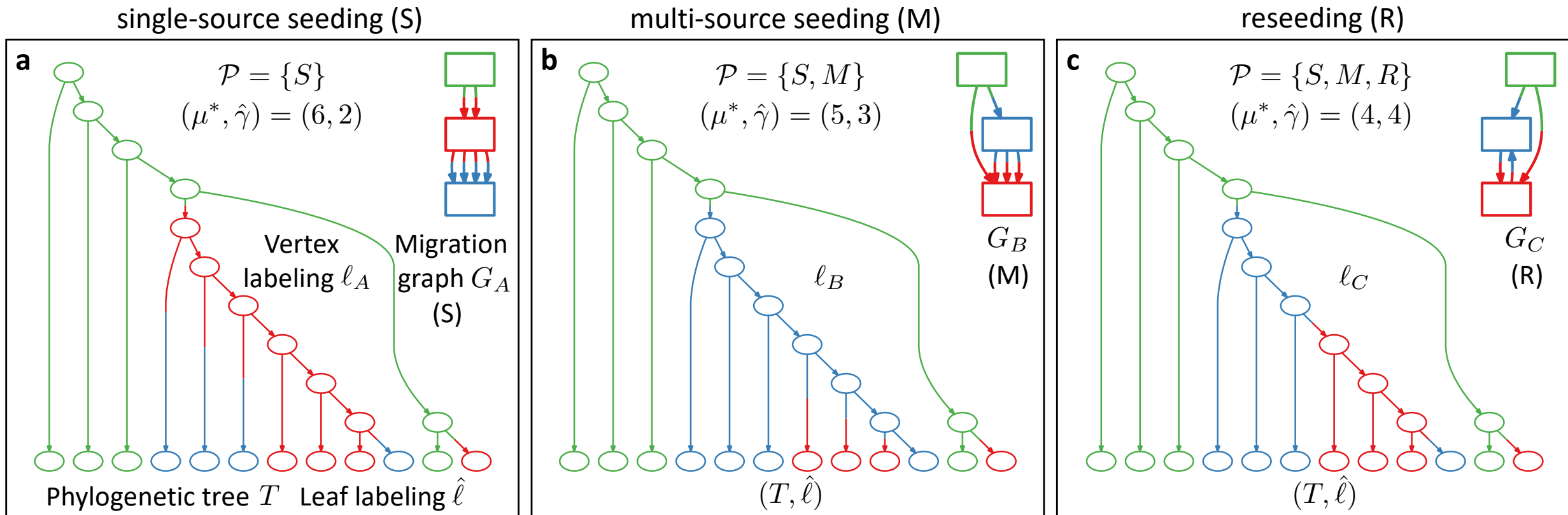
Outline

- Metastasis
- Maximum parsimony
- Problem statement
- Complexity
- Algorithm & results
- Problem variants

Reading:

- **M. El-Kebir**, G. Satas and B.J. Raphael. Inferring parsimonious migration histories for metastatic cancers. Nature Genetics, 50:718-726, 2018.

- **M. El-Kebir**[†]. Parsimonious Migration History Problem: Complexity and Algorithms. WABI 2018, Helsinki, Finland, August 20-22, 2018.

Results [El-Kebir, WABI 2018]

Parsimonious Migration History (PMH): Given a phylogenetic tree T and a set $\mathcal{P} \subseteq \{S, M, R\}$ of allowed migration patterns, find vertex labeling ℓ with minimum migration number $\mu^*(T)$ and smallest comigration number $\hat{\gamma}(T)$.
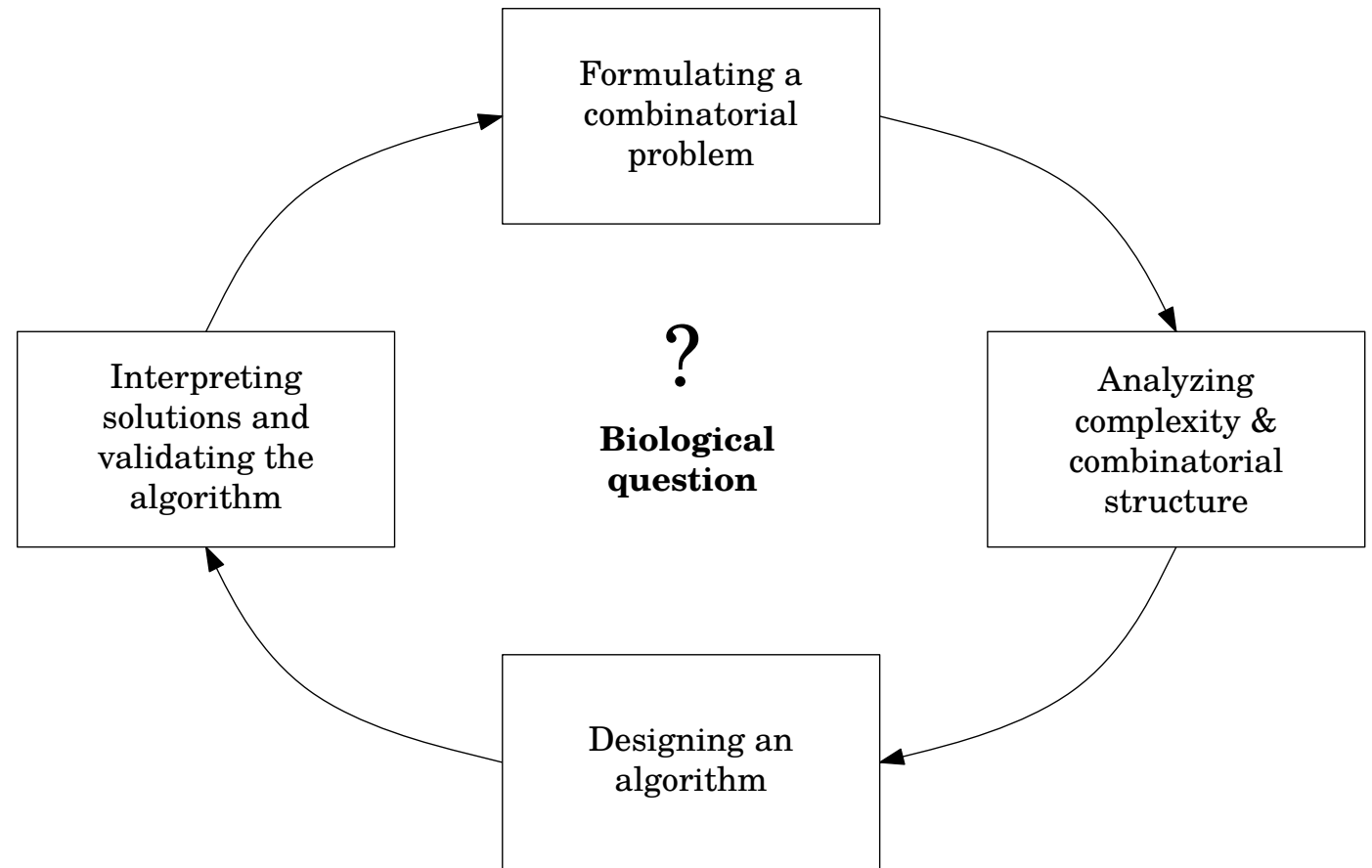
single-source seeding (S)

a

$\mathcal{P} = \{S\}$
$(\mu^*, \hat{\gamma}) = (6, 2)$

Vertex labeling ℓ_A Migration graph G_A (S)

Phylogenetic tree T Leaf labeling $\hat{\ell}$
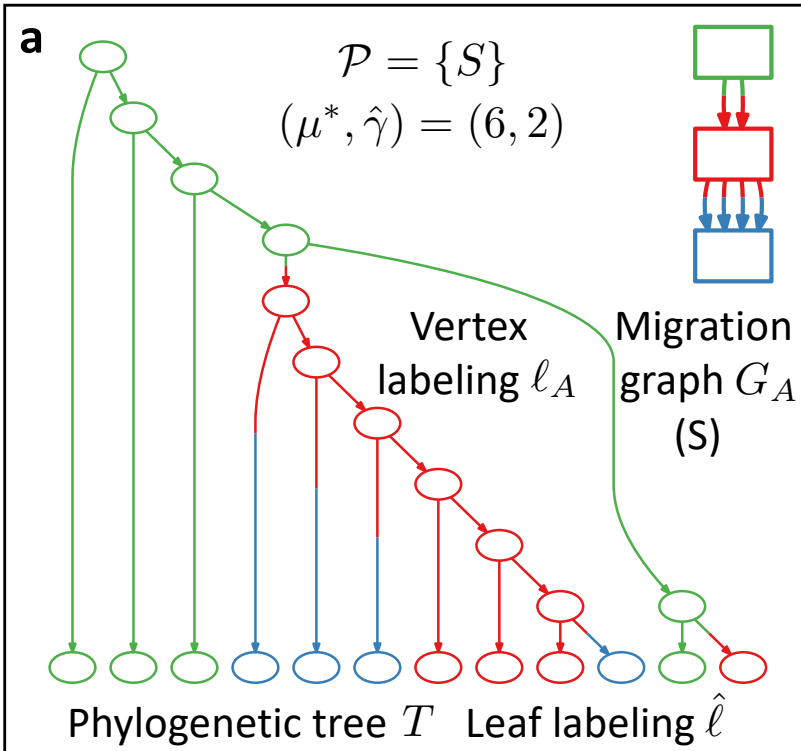
Theorem 1: PMH is NP-hard when $\mathcal{P} = \{S\}$

Theorem 2: PMH is fixed parameter tractable in the number m of locations when $\mathcal{P} = \{S\}$

PMH is NP-hard when $\mathcal{P} = \{S\}$

3-SAT: Given $\varphi = \bigwedge_{i=1}^{k}(y_{i,1} \lor y_{i,2} \lor y_{i,3})$ with variables $\{x_1, \dots, x_n\}$ and k clauses, find $\phi : [n] \to \{0,1\}$ satisfying φ
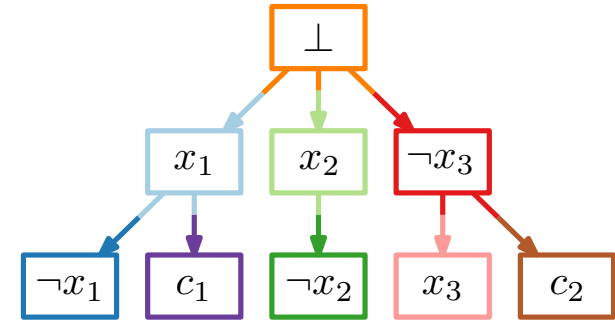
$$\Sigma = \{x_1, \dots, x_n, \neg x_1, \dots, \neg x_n, c_1, \dots c_k, \bot\}$$

PMH is NP-hard when $\mathcal{P} = \{S\}$

$\Sigma = \{x_1, \dots, x_n, \neg x_1, \dots, \neg x_n, c_1, \dots c_k, \bot\}$

3-SAT: Given $\varphi = \bigwedge_{i=1}^{k}(y_{i,1} \vee y_{i,2} \vee y_{i,3})$ with variables $\{x_1, \dots, x_n\}$ and k clauses, find $\phi : [n] \to \{0,1\}$ satisfying φ

Three ideas:

1. Ensure that $(x, \neg x) \in E(G)$ or $(\neg x, x) \in E(G)$

2. Ensure that $\ell^*(r(T)) = \bot$

3. Ensure that φ is satisfiable if and only if ℓ^* encodes a satisfying truth assignment

a Phylogenetic tree T

b Variable gadget $T[x_j]$ $(2B$ leaves$)$

c Root gadget $T[\bot]$ $(A$ leaves$)$

d Clause literal gadget $T[y_{i,j}]$ $(9$ leaves$)$
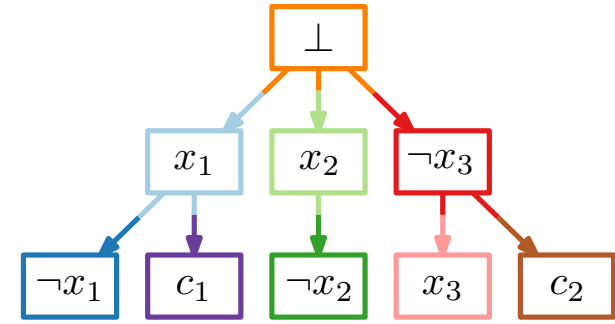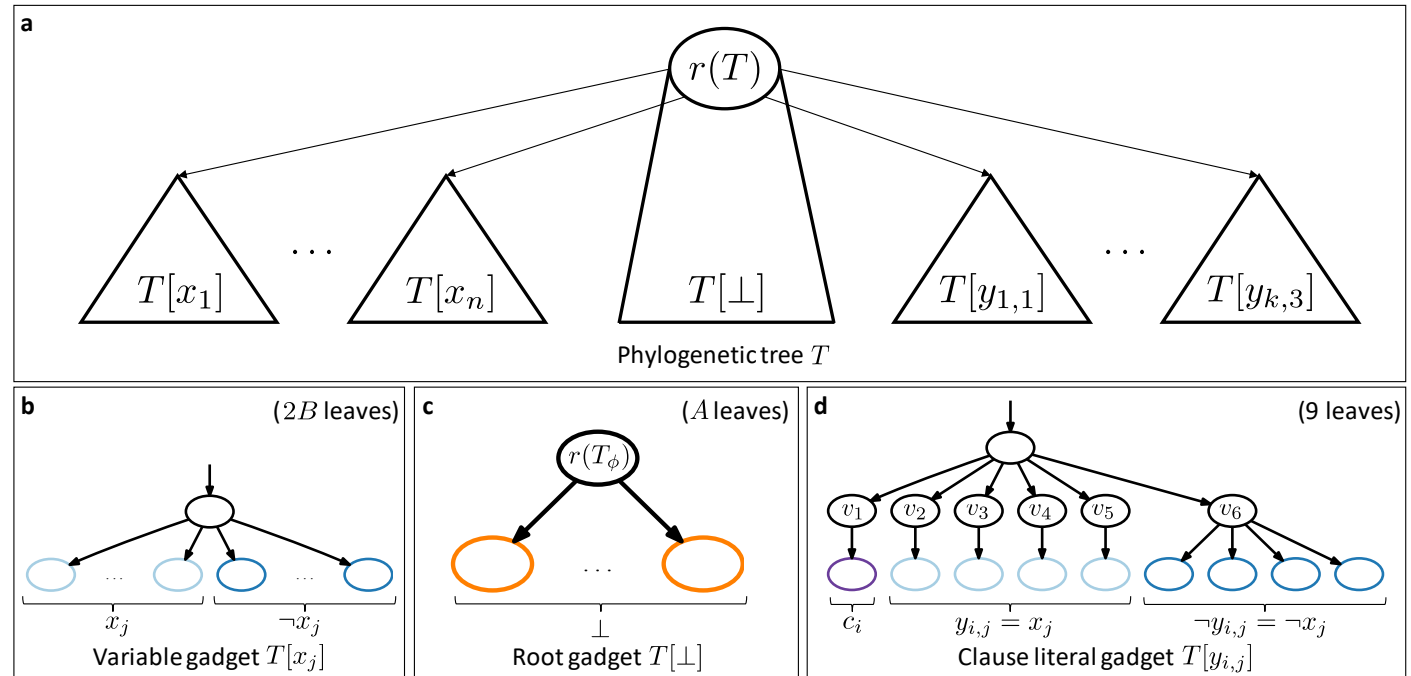
35

PMH is NP-hard when $\mathcal{P} = \{S\}$

3-SAT: Given $\varphi = \bigwedge_{i=1}^{k}(y_{i,1} \vee y_{i,2} \vee y_{i,3})$ with variables $\{x_1, \ldots, x_n\}$ and k clauses, find $\phi : [n] \to \{0,1\}$ satisfying φ

$\Sigma = \{x_1, \ldots, x_n, \neg x_1, \ldots, \neg x_n, c_1, \ldots c_k, \bot\}$

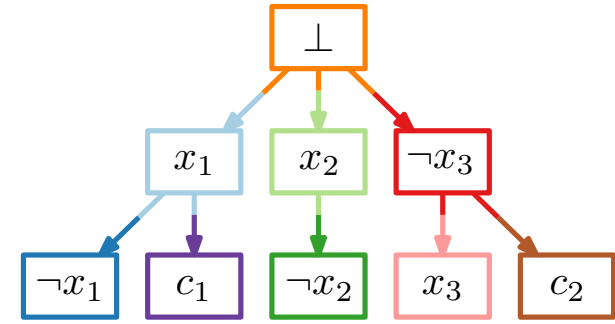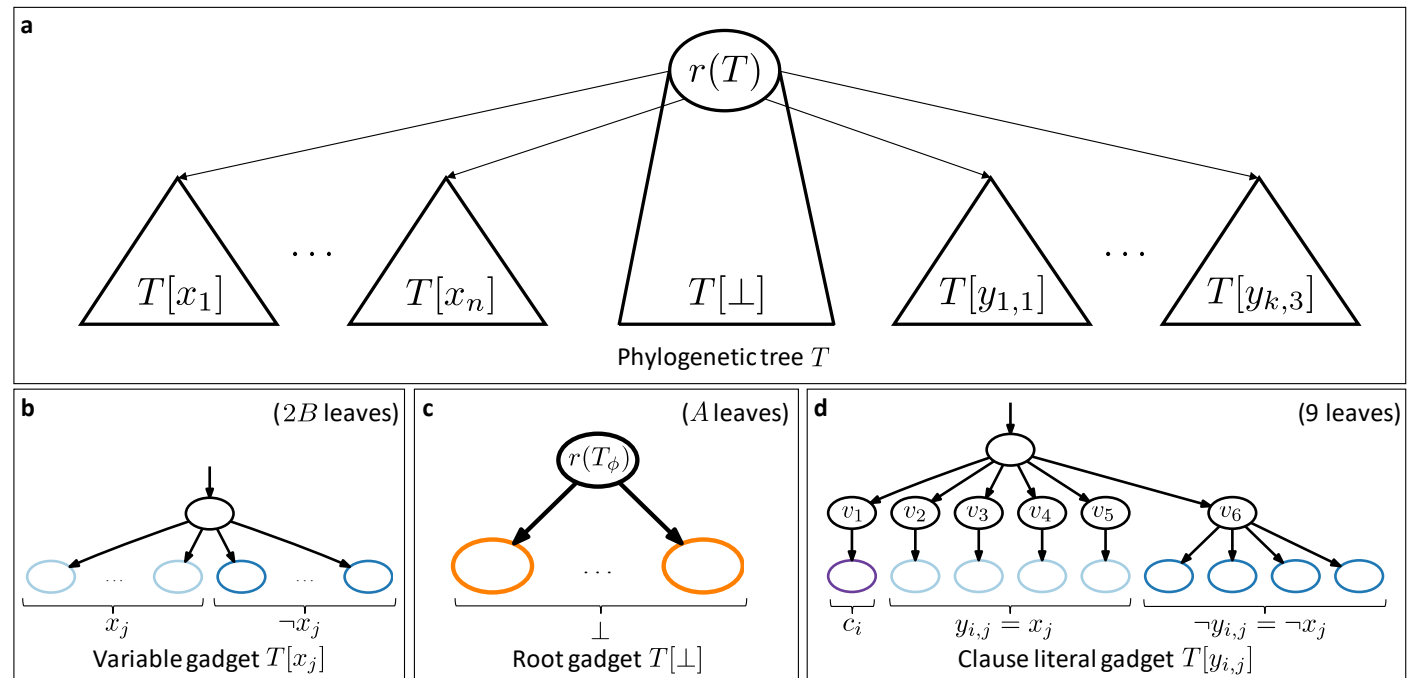Three ideas:

1. Ensure that $(x, \neg x) \in E(G)$ or $(\neg x, x) \in E(G)$

2. Ensure that $\ell^*(r(T)) = \bot$

3. Ensure that φ is satisfiable if and only if ℓ^* encodes a satisfying truth assignment

a

$r(T)$

$T[x_1]$ ... $T[x_n]$ $T[\bot]$ $T[y_{1,1}]$ ... $T[y_{k,3}]$

Phylogenetic tree T

b (2B leaves)

x_j $\neg x_j$
Variable gadget $T[x_j]$

c (A leaves)

$r(T_\phi)$

$\bot$
Root gadget $T[\bot]$

d (9 leaves)

v_1 v_2 v_3 v_4 v_5 v_6

c_i $y_{i,j} = x_j$ $\neg y_{i,j} = \neg x_j$
Clause literal gadget $T[y_{i,j}]$

Lemma: Let $B > 10k + 1$ and $A > 2Bn + 27k$.
Then, φ is satisfiable if and only if $\mu^*(T) = (B + 1)n + 25k$

PMH is NP-hard when $\mathcal{P} = \{S\}$

$$\mu^*(T) = (B+1)n + 25k$$
$$= 23 * 3 + 50 * 2 = 119$$

$$\varphi = (x_1 \lor x_2 \lor \neg x_3) \land (\neg x_1, \neg x_2, \neg x_3)$$

$$k = 2, n = 3$$
$$B = 10k + 2 = 22$$
$$A = 2Bn + 27k + 1 = 187$$

$$\Sigma = \{x_1, x_2, x_3, \neg x_1, \neg x_2, \neg x_3, c_1, c_2, \perp\}$$

Lemma: Let $B > 10k + 1$ and $A > 2Bn + 27k$.
Then, φ is satisfiable if and only if $\mu^*(T) = (B+1)n + 25k$

Leaf labeling $\hat{\ell}$
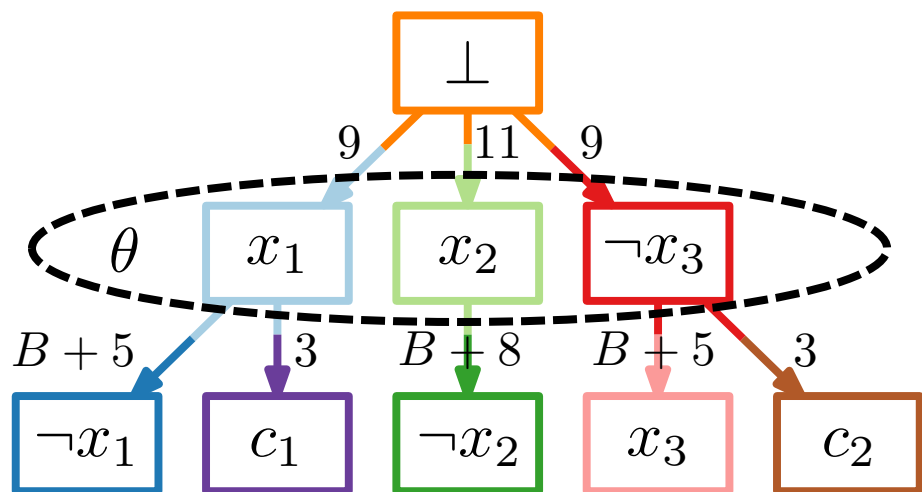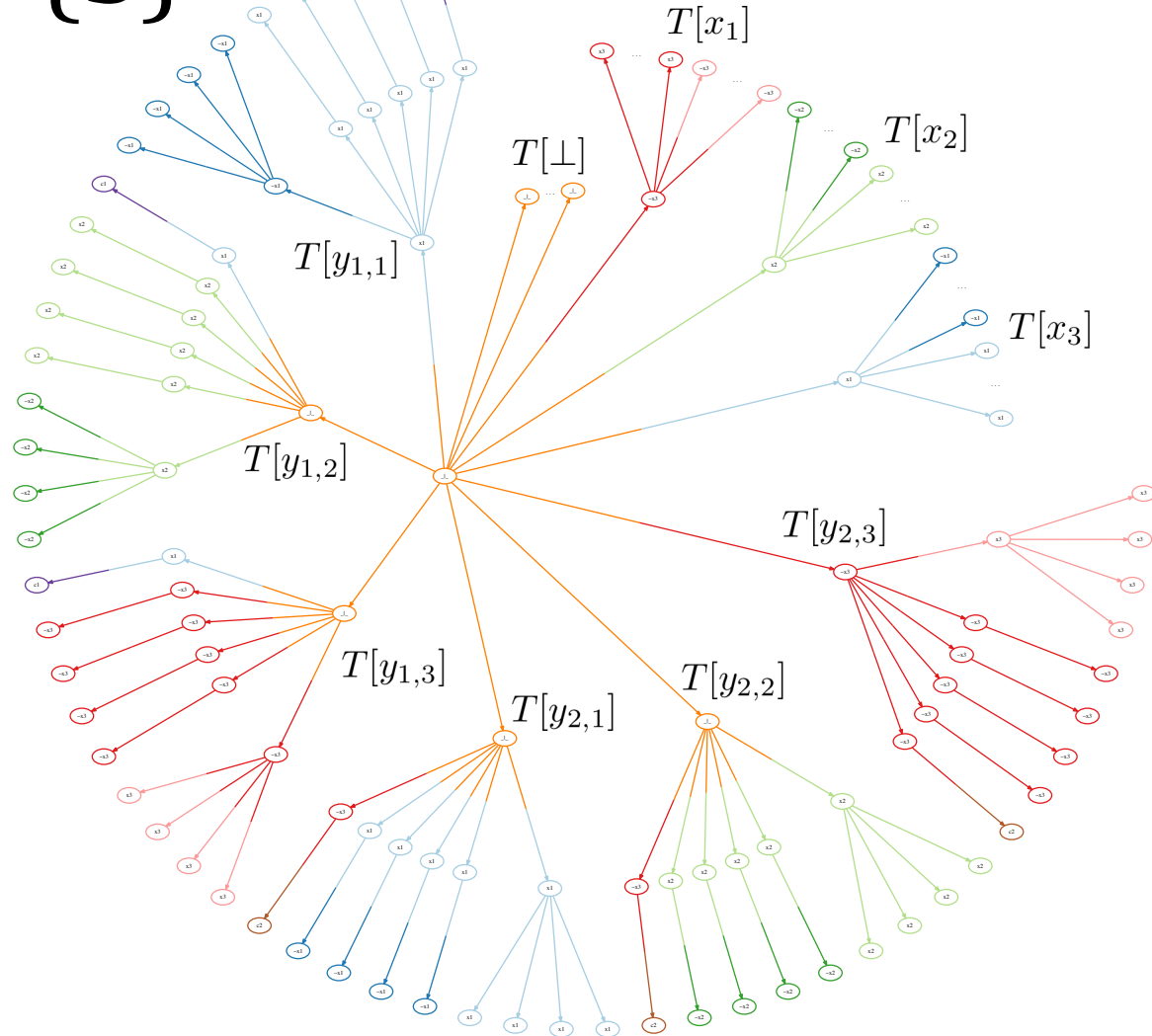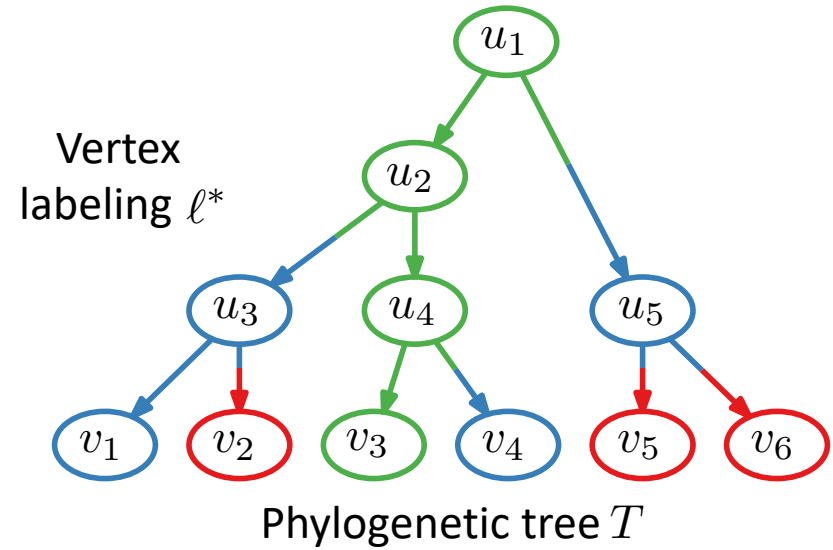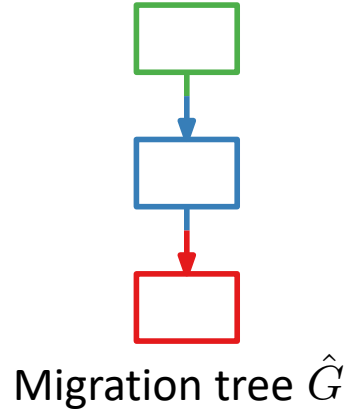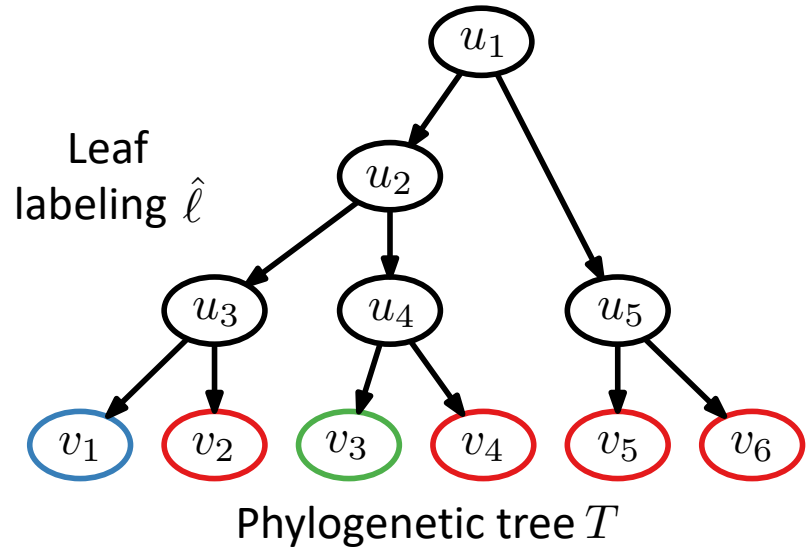
Phylogenetic tree T

Migration tree $\hat{G}$

Vertex labeling ℓ^*

Phylogenetic tree T

Lemma: If there exists labeling ℓ consistent with $\hat{G}$ then
$$d_T(u,v) \geq d_{\hat{G}}(\mathrm{lca}_{\hat{G}}(u), \hat{\ell}(v)) \qquad \forall u,v \in V(T) \text{ such that } u \preceq_T v. \quad (1)$$

$$\ell^*(v) = \begin{cases} \mathrm{LCA}_{\hat{G}}(r(T)), & \text{if } v = r(T), \\ \sigma(\ell^*(\pi(v)), \mathrm{LCA}_{\hat{G}}(v)), & \text{if } v \neq r(T), \end{cases}$$

where $\sigma(s,t) = s$ if $s = t$ and otherwise $\sigma(s,t)$ is the unique child of s that lies on the path from s to t in $\hat{G}$.

Lemma: If (1) holds then ℓ^* is a minimum migration labeling consistent with $\hat{G}$.

PMH is FPT in number m of locations when $\mathcal{P} = \{S\}$

Leaf labeling $\hat{\ell}$

Phylogenetic tree T

Migration tree $\hat{G}$

Vertex labeling ℓ^*

Phylogenetic tree T

$O(nm^m)$ time

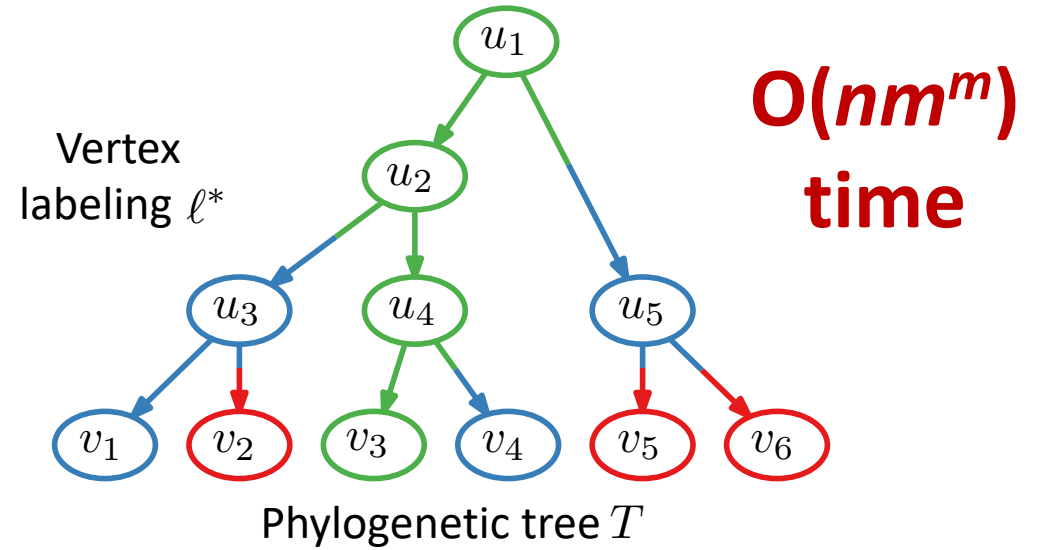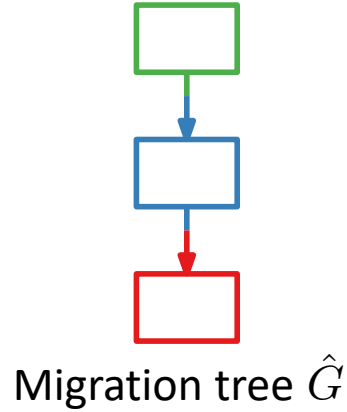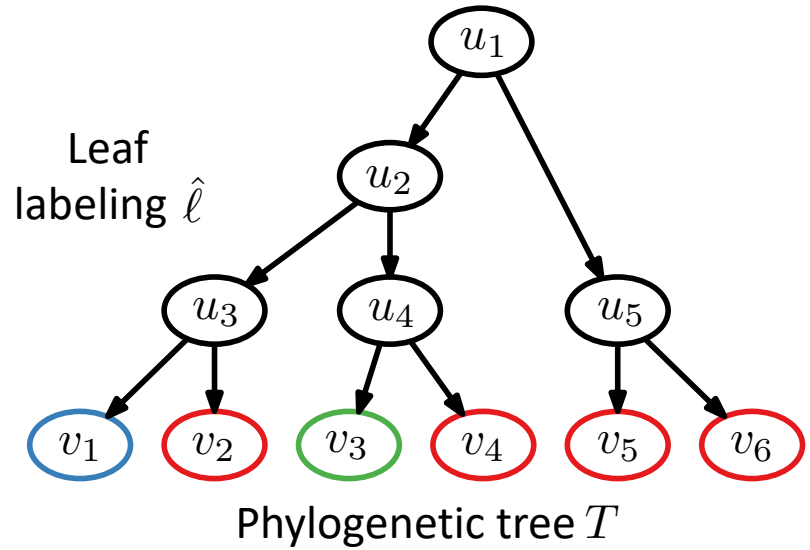**Lemma:** If there exists labeling ℓ consistent with $\hat{G}$ then
$$d_T(u,v) \geq d_{\hat{G}}(\mathrm{lca}_{\hat{G}}(u), \hat{\ell}(v)) \qquad \forall u,v \in V(T) \text{ such that } u \preceq_T v. \quad (1)$$

$$\ell^*(v) = \begin{cases} \mathrm{LCA}_{\hat{G}}(r(T)), & \text{if } v = r(T), \\ \sigma(\ell^*(\pi(v)), \mathrm{LCA}_{\hat{G}}(v)), & \text{if } v \neq r(T), \end{cases}$$

where $\sigma(s,t) = s$ if $s = t$ and otherwise $\sigma(s,t)$ is the unique child of s that lies on the path from s to t in $\hat{G}$.
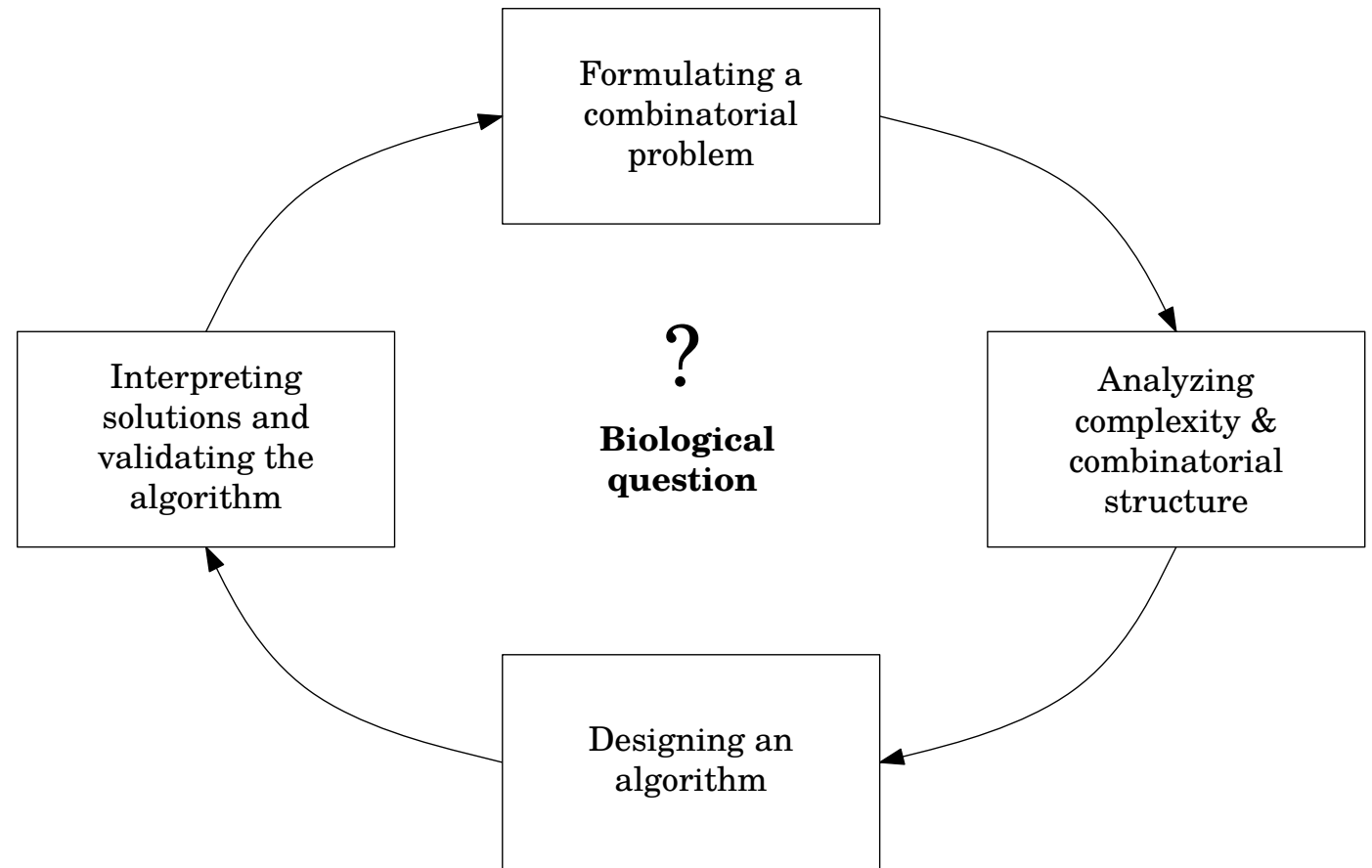
Lemma: If (1) holds then ℓ^* is a minimum migration labeling consistent with $\hat{G}$.

Outline

- Metastasis
- Maximum parsimony
- Problem statement
- Complexity
- Algorithm & results
- Problem variants

Reading:

- **M. El-Kebir**, G. Satas and B.J. Raphael. Inferring parsimonious migration histories for metastatic cancers. Nature Genetics, 50:718-726, 2018.

- **M. El-Kebir**[†]. Parsimonious Migration History Problem: Complexity and Algorithms. WABI 2018, Helsinki, Finland, August 20-22, 2018.

Simulations

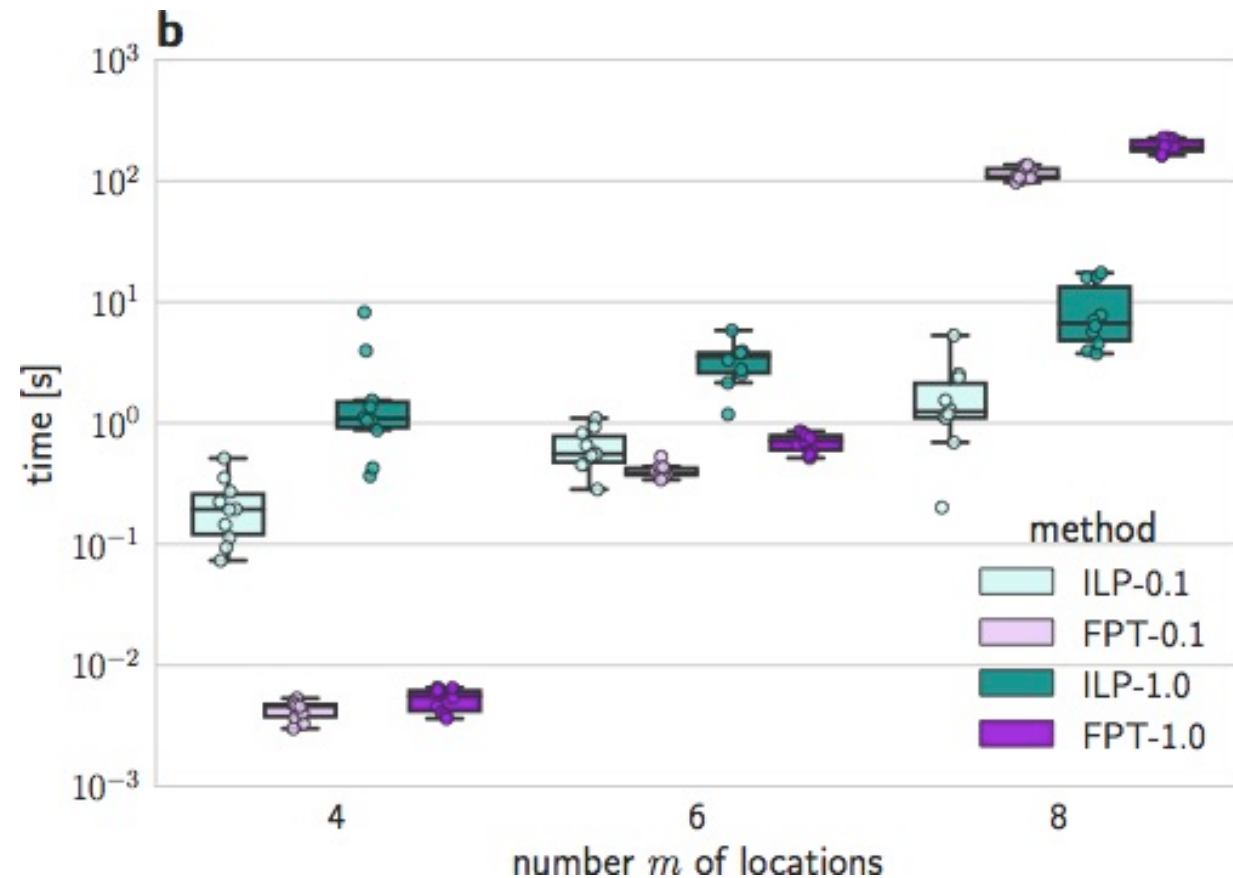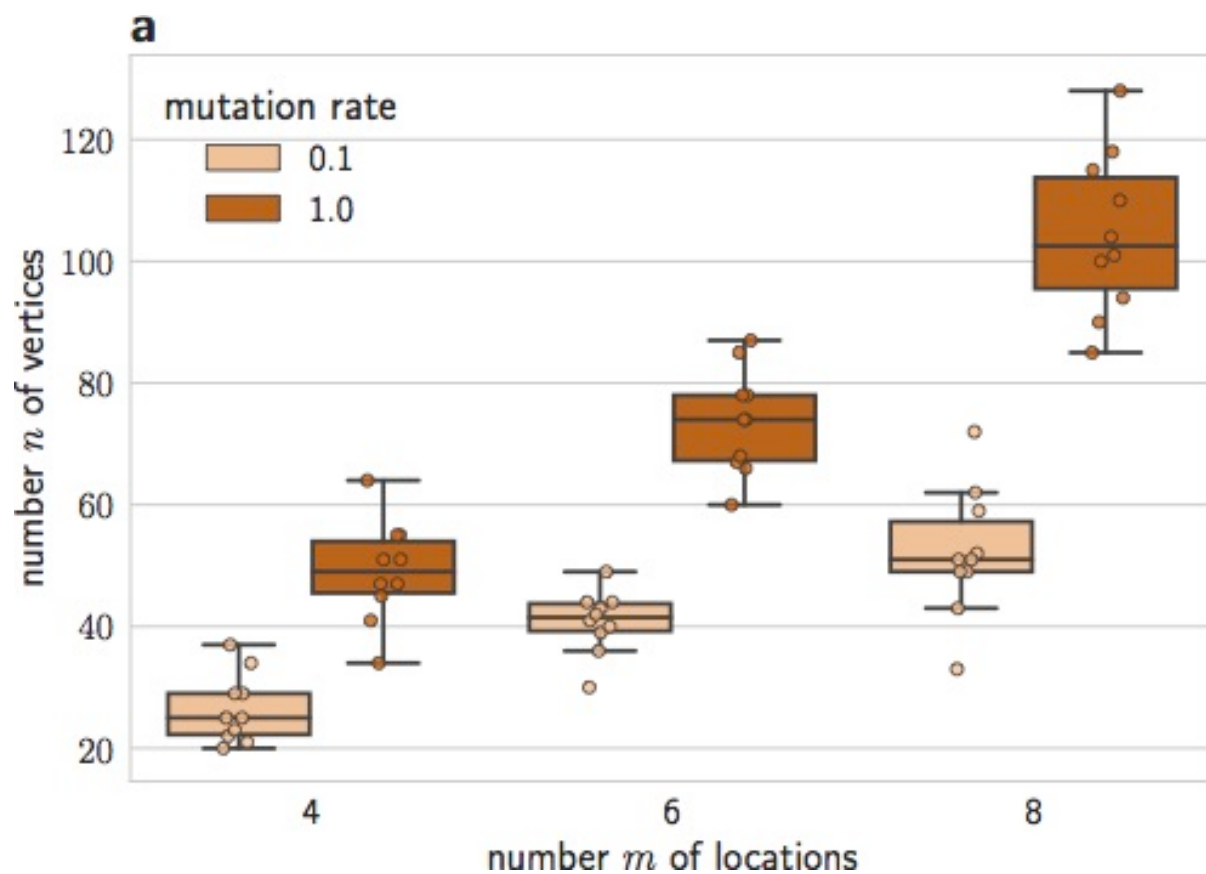Available on: https://github.com/elkebir-group/PMH-S

Outline

- Metastasis
- Maximum parsimony
- Problem statement
- Complexity
- Algorithm & results
- Problem variants

Reading:

- **M. El-Kebir**, G. Satas and B.J. Raphael. Inferring parsimonious migration histories for metastatic cancers. Nature Genetics, 50:718-726, 2018.

- **M. El-Kebir**[†]. Parsimonious Migration History Problem: Complexity and Algorithms. WABI 2018, Helsinki, Finland, August 20-22, 2018.

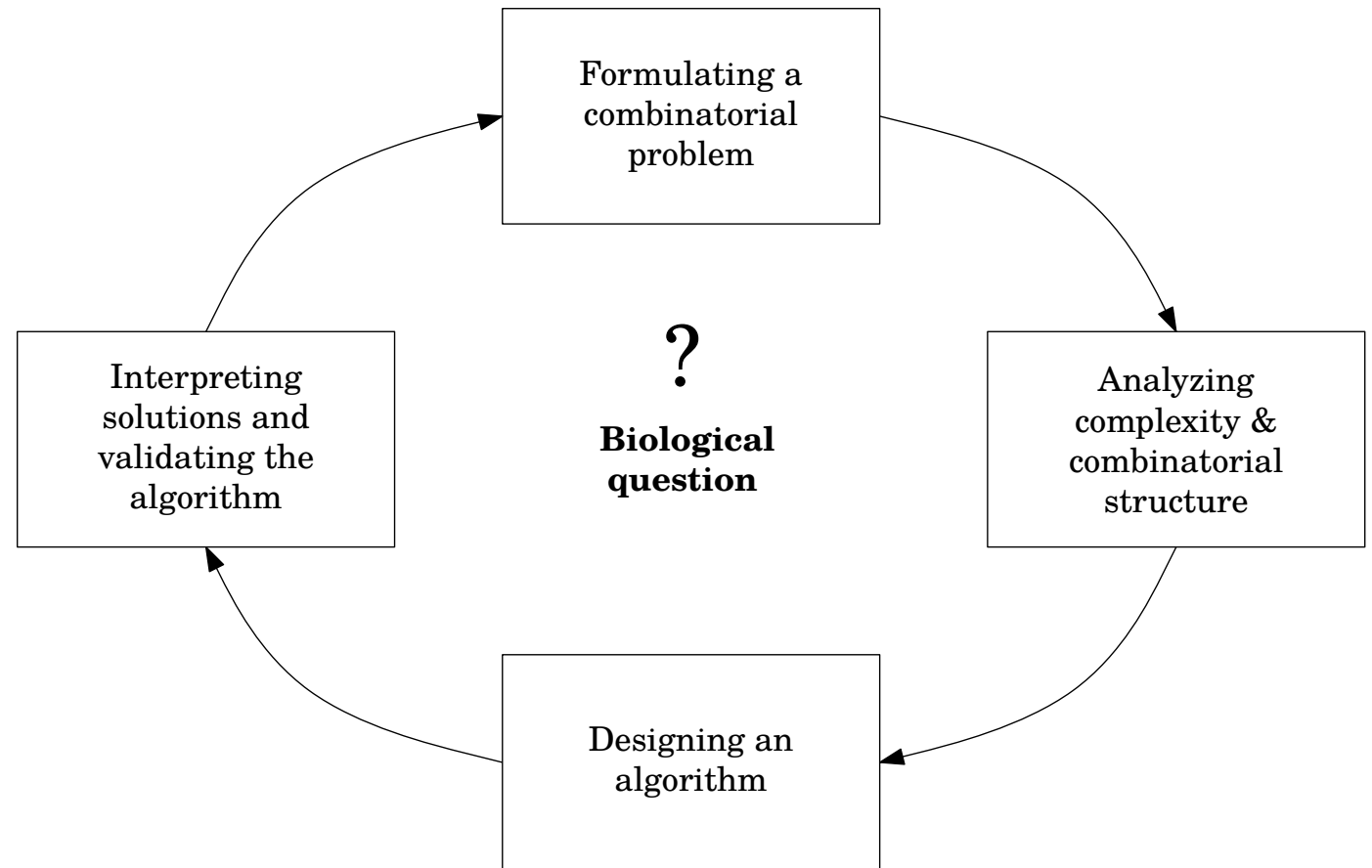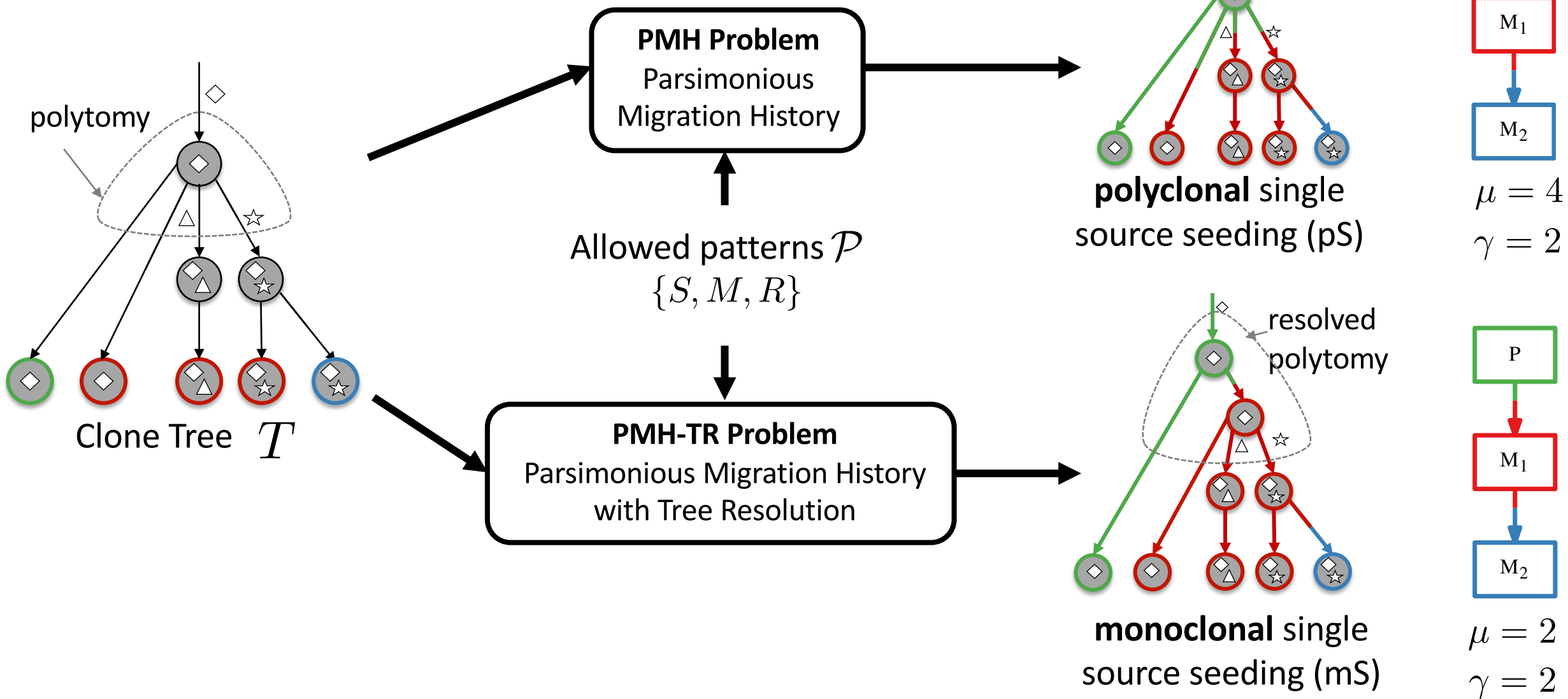Formulating a combinatorial problem → Analyzing complexity & combinatorial structure → Designing an algorithm → Interpreting solutions and validating the algorithm → (cycle)

? **Biological question**

Resolving Clone Tree Ambiguities

polytomy

Clone Tree T

PMH Problem
Parsimonious Migration History

Allowed patterns $\mathcal{P}$
$\{S, M, R\}$

PMH-TR Problem
Parsimonious Migration History with Tree Resolution

polyclonal single source seeding (pS)

resolved polytomy

monoclonal single source seeding (mS)

$\mu = 4$
$\gamma = 2$

$\mu = 2$
$\gamma = 2$

Resolving Clone Tree Ambiguities

variant allele frequency

confidence intervals

$$F^- = \begin{array}{c} \\ P \\ M_1 \\ M_2 \end{array} \begin{array}{ccc} \diamond & \triangle & \star \\ \left[\begin{array}{ccc} 0.75 & 0 & 0 \\ 0.82 & 0.15 & 0.56 \\ 0.66 & 0 & 0.68 \end{array} \right] \end{array}$$

$$F^+ = \begin{array}{c} \\ P \\ M_1 \\ M_2 \end{array} \begin{array}{ccc} \diamond & \triangle & \star \\ \left[\begin{array}{ccc} 0.86 & 0 & 0 \\ 0.92 & 0.30 & 0.68 \\ 0.78 & 0 & 0.77 \end{array} \right] \end{array}$$

Clone Tree T

resolved polytomy

$$\hat{F} = \begin{array}{c} \\ P \\ M_1 \\ M_2 \end{array} \begin{array}{ccc} \diamond & \triangle & \star \\ \left[\begin{array}{ccc} 0.82 & 0 & 0 \\ 0.88 & 0.24 & 0.64 \\ 0.73 & 0 & 0.73 \end{array} \right] \end{array} = \begin{array}{c} \\ P \\ M_1 \\ M_2 \end{array} \begin{array}{ccc} \left[\begin{array}{ccc} 0.82 & 0 & 0 \\ 0 & 0.24 & 0.64 \\ 0 & 0 & 0.73 \end{array} \right] \end{array} \underbrace{\begin{array}{ccc} \diamond & \triangle & \star \\ \left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{array} \right] \end{array}}_{B}$$

$\underbrace{\qquad}_{\hat{F}}$ $\underbrace{\qquad}_{U}$

PMH-TI Problem
Parsimonious Migration History
with Tree Inference

$\mathcal{P}$

$\mu = 2$

$\gamma = 2$

monoclonal single
source seeding (mS)

44

Mutation Matrix

mutations

$$P \begin{pmatrix} \diamond & \clubsuit & \star & \triangle & \lightning \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

samples M_1 M_2

unobserved clones

Standard Phylogenetic Techniques

Sample Tree

$\diamond$

$\star\lightning$* $\clubsuit$

$\triangle$ *$\star$

P M_1 M_2

*homoplasy
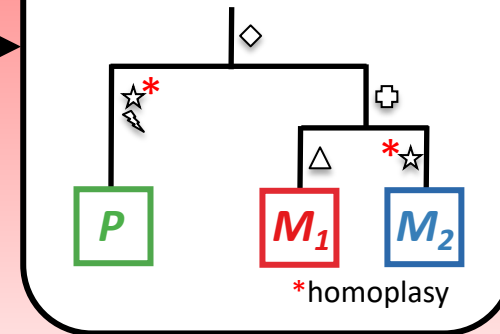
Cellular History of Metastatic Cancer

$\diamond$ — mutation

migration

time

primary P metastasis M_2 metastasis M_1
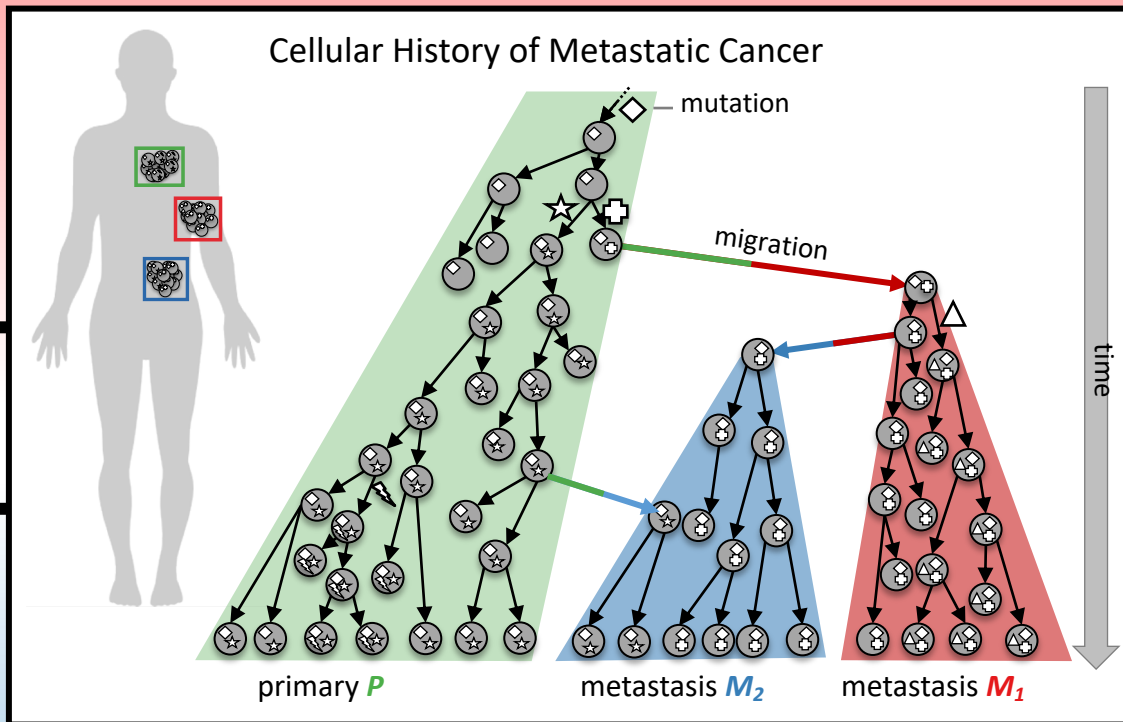
Sequencing and Mutation Calling

Tumor Phylogenetic Techniques
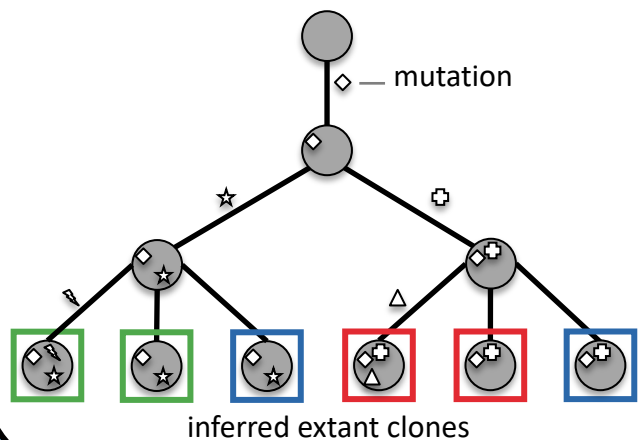
Cell Division and Mutation History

Cell Migration History

MACHINA

Clone Tree

$\diamond$ — mutation

$\star$ $\clubsuit$

$\triangle$

inferred extant clones
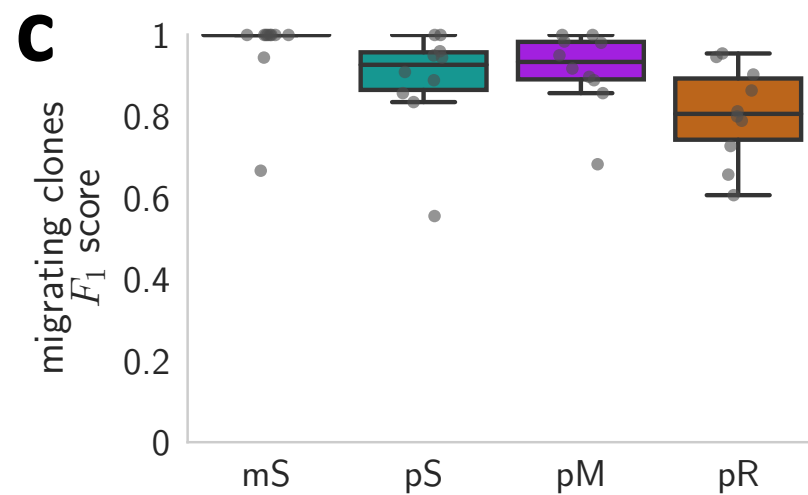
Migration Graph

P

M_1

— migration

M_2
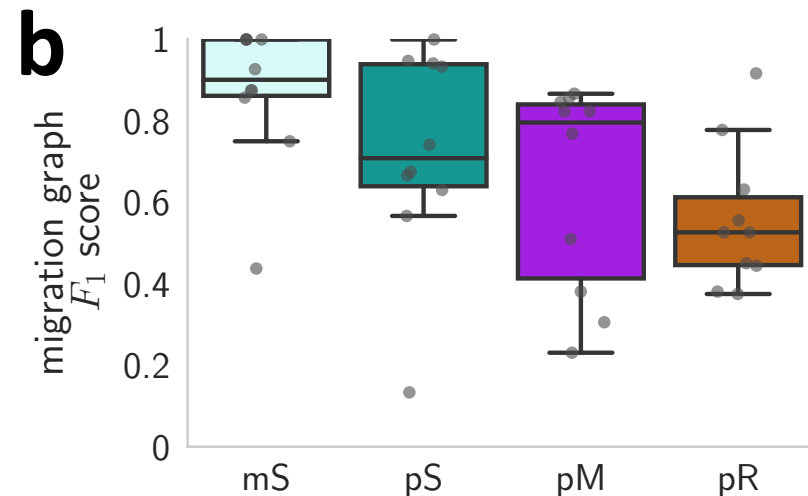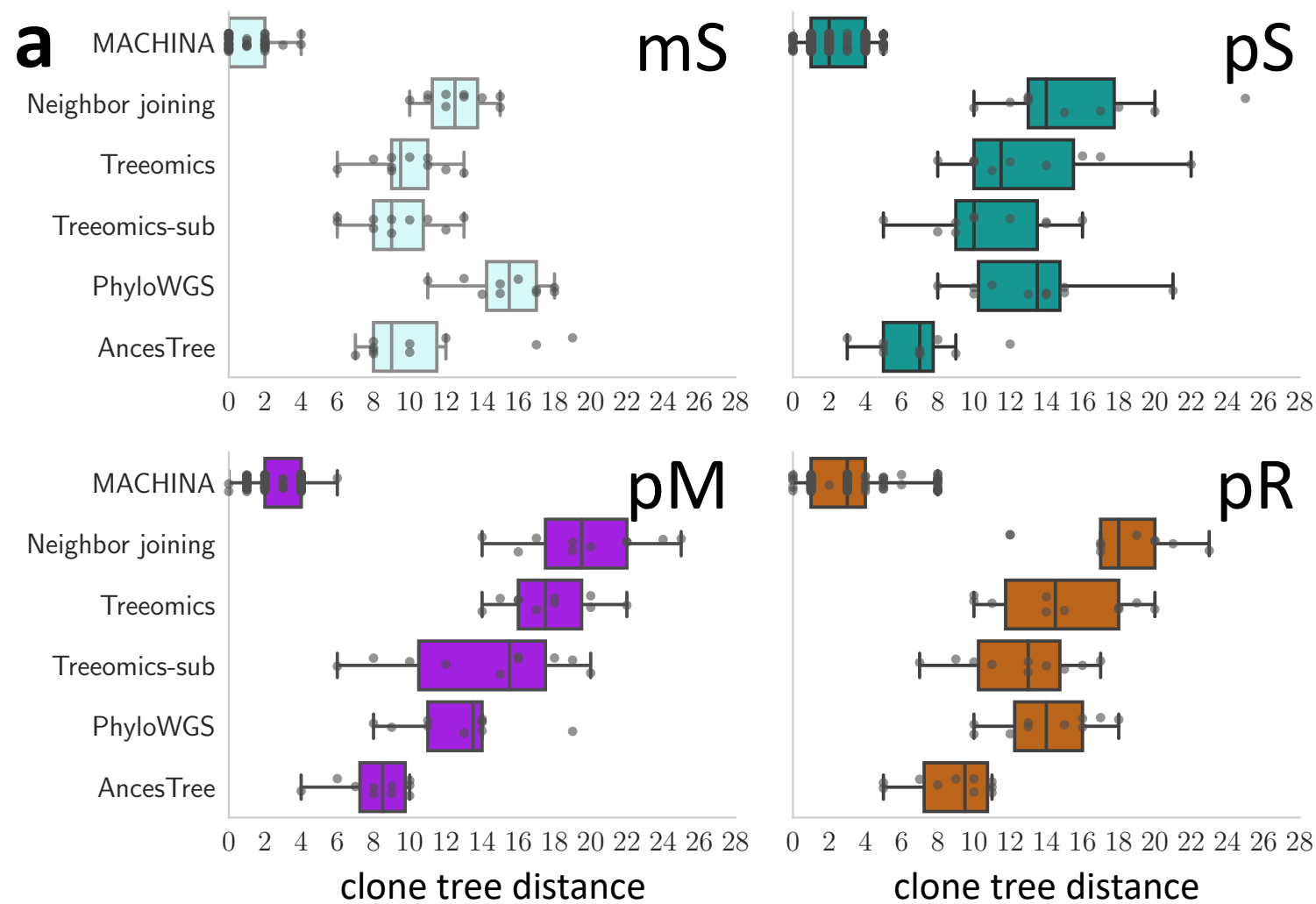
Label ancestral vertices by anatomical sites

Resolve clone tree ambiguities
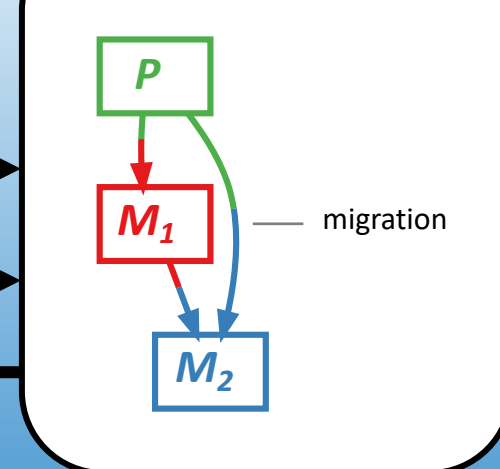
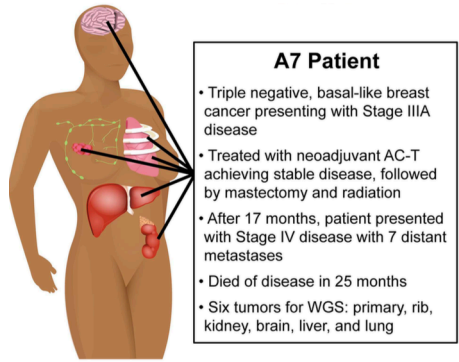# MACHINA accurately infers clone trees and migration histories on simulated data

Applying MACHINA to Metastatic Breast Cancer

Hoadley et al. Tumor Evolution in Two Patients with Basal-like Breast Cancer: A Retrospective Genomics Study of Multiple Metastases. *PLOS Med, 13*(12) 2016

A7 Patient
- Triple negative, basal-like breast cancer presenting with Stage IIIA disease
- Treated with neoadjuvant AC-T achieving stable disease, followed by mastectomy and radiation
- After 17 months, patient presented with Stage IV disease with 7 distant metastases
- Died of disease in 25 months
- Six tumors for WGS: primary, rib, kidney, brain, liver, and lung

(B) **Reported** $(\mu, \gamma) = (12, 6)$

polyclonal multi-source seeding (pM)

(C) **Inferred** $(\mu, \gamma) = (5, 5)$

resolved polytomy

resolved polytomy

monoclonal single-source seeding (mS)