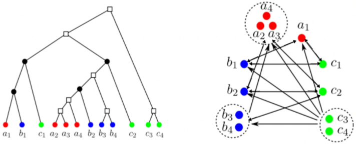
UNIVERSITÄT
LEIPZIG

Combinatorics 2022
Mantua, May 30 - June 3

Best match graphs and generalizations

Annachiara Korchmaros
Leipzig Bioinformatics Group

- Gene trees T
  - Rooted trees with leaf coloring $\sigma : L \to S$
  - Leaves set L = mutated copies of a gene
  - Leaves colors S = species

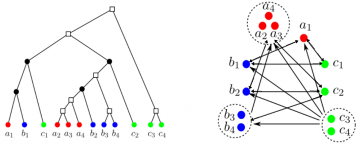


[1] D. Schaller *et al.*, *Journal of Mathematical Biology* **82**, 1–9 (2021).

Gene trees

- Gene trees T
  - Rooted trees with leaf coloring $\sigma : L \to S$
  - Leaves set L = mutated copies of a gene
  - Leaves colors S = species
- Best matches = most closely related leaves

  $y \in L$ is a **best match**[1] of $x \in L$ if $\sigma(x) \neq \sigma(y)$ and
  $\mathrm{lca}(x, y) \preceq \mathrm{lca}(x, z)$ for all $z \in L$ with $\sigma(z) = \sigma(y)$



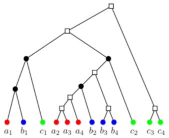
[1] D. Schaller *et al.*, *Journal of Mathematical Biology* **82**, 1–9 (2021).

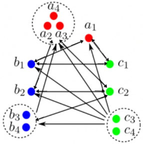
- Gene trees T
  - Rooted trees with leaf coloring $\sigma : L \to S$
  - Leaves set L = mutated copies of a gene
  - Leaves colors S = species
- Best matches = most closely related leaves

  $y \in L$ is a **best match**[1] of $x \in L$ if $\sigma(x) \neq \sigma(y)$ and

  $\mathrm{lca}(x, y) \preceq \mathrm{lca}(x, z)$ for all $z \in L$ with $\sigma(z) = \sigma(y)$
- $G = (V, E)$ vertex colored digraph is **BMG** of T if

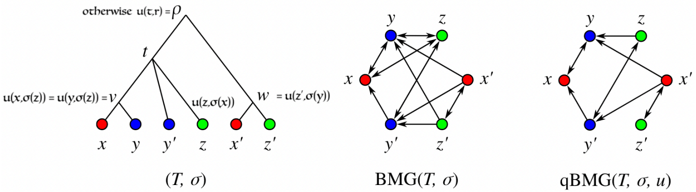
  $V = L$ and $x \to y$ iff y is a best match of x





[1] D. Schaller *et al.*, *Journal of Mathematical Biology* **82**, 1–9 (2021).

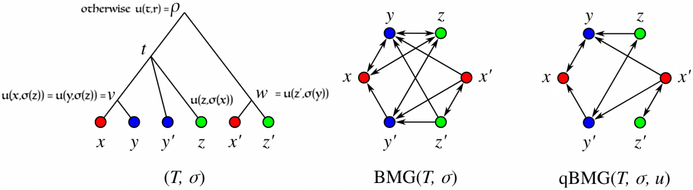
Gene trees



[2] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

Truncation map of T is $u : L \times S \to V$; V (vertex set of T) st

- $u(x, \sigma(x)) = x$
- $u(x, s), s \neq \sigma(x)$ is a node on the path from x to ρ



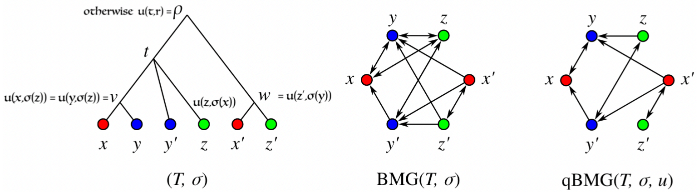
[2] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

Truncation map of T is $u : L \times S \to V$; V (vertex set of T) st

- $u(x, \sigma(x)) = x$
- $u(x, s), s \neq \sigma(x)$ is a node on the path from x to ρ

$y \in L$ is a **quasi-best match**[2] of $x \in L$ if

- y is a best match of x
- $\text{lca}(x, y) \preceq \text{u}(x, \sigma(y))$



[2] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

Gene trees

Truncation map of T is $u : L \times S \to V$; V (vertex set of T) st

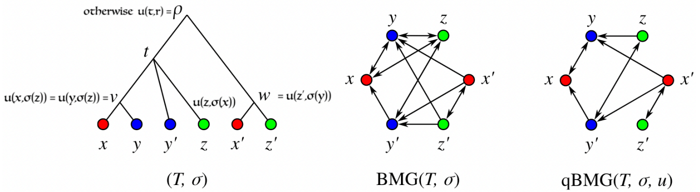
- $u(x, \sigma(x)) = x$
- $u(x, s), s \neq \sigma(x)$ is a node on the path from x to ρ

$y \in L$ is a **quasi-best match**[2] of $x \in L$ if

- y is a best match of x
- $\mathrm{lca}(x, y) \preceq \mathrm{u}(x, \sigma(y))$

$G = (V, E)$ vertex colored digraph is **qBMG** of T if

$V = L$ and $x \to y$ iff y is a quasi-best match of x in T



[2]A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

Truncation map of T is $u : L \times S \to V$; V (vertex set of T) st

- $u(x, \sigma(x)) = x$
- $u(x, s), s \neq \sigma(x)$ is a node on the path from x to ρ

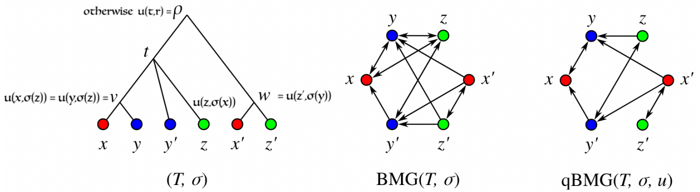
$y \in L$ is a **quasi-best match**[2] of $x \in L$ if

- y is a best match of x
- $\text{lca}(x, y) \preceq \text{u}(x, \sigma(y))$

$G = (V, E)$ vertex colored digraph is **qBMG** of T if

$V = L$ and $x \to y$ iff y is a quasi-best match of x in T

BMGs are qBMGs !! ($u(x, r) = \rho$)

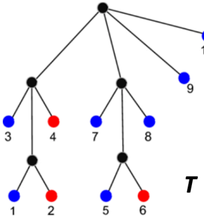


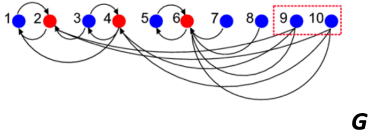
[2] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

Outline

1. Gene trees
2. 2-BMGs
3. 2-qBMGs

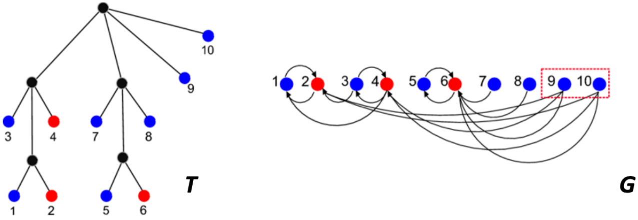
Theorem 1 (Schaller et al. 2021)





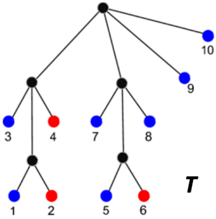
Theorem 1 (Schaller et al. 2021)

1. Let G be a vertex colored digraph. G is a BMG iff its connected components are BMGs on the same color sets.

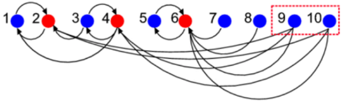


Theorem 1 (Schaller et al. 2021)

1. Let G be a vertex colored digraph. G is a BMG iff its connected components are BMGs on the same color sets.
2. If a connected vertex colored digraph is a BMG, then all its induced subgraphs on 2 colors are BMGs (**2-BMGs**).



T



G

Theorem 2 (Schaller et al. 2021)

Let $G = (V, E)$ be a connected bipartite vertex colored digraph and $N(x)$ the set of outneighbours of x. G is a 2-BMG iff

Theorem 2 (Schaller et al. 2021)

Let $G = (V, E)$ be a connected bipartite vertex colored digraph and $N(x)$ the set of outneighbours of x. G is a 2-BMG iff

N0 G is sink free.

Theorem 2 (Schaller et al. 2021)

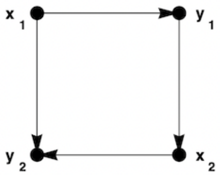
Let $G = (V, E)$ be a connected bipartite vertex colored digraph and $N(x)$ the set of outneighbours of x. G is a 2-BMG iff

- **N0** G is sink free.
- **N1** If u and v are independent vertices, then there exist no vertices w and t st $ut, vw, tw \in E$.

Theorem 2 (Schaller et al. 2021)

Let $G = (V, E)$ be a connected bipartite vertex colored digraph and $N(x)$ the set of outneighbours of x. G is a 2-BMG iff

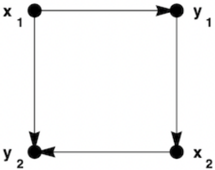
- **N0** G is sink free.
- **N1** If u and v are independent vertices, then there exist no vertices w and t st $ut, vw, tw \in E$.
- **N2** G is bitransitive, ie if $x_1y_1, y_1x_2, x_2y_2 \in E \Rightarrow x_1y_2 \in E$.



Theorem 2 (Schaller et al. 2021)

Let $G = (V, E)$ be a connected bipartite vertex colored digraph and $N(x)$ the set of outneighbours of x. G is a 2-BMG iff

- **N0** G is sink free.
- **N1** If u and v are independent vertices, then there exist no vertices w and t st $ut, vw, tw \in E$.
- **N2** G is bitransitive, ie if $x_1y_1, y_1x_2, x_2y_2 \in E \Rightarrow x_1y_2 \in E$.
- **N3** If $N(u) \cap N(v) \neq \emptyset \Rightarrow N(u) \subseteq N(v)$ or $N(v) \subseteq N(u)$.



$G = (V, E)$ digraph
$N(x)$ out-neighborhood, $N^-(x)$ in-neighborhood

[3] A. Korchmaros, *Discrete Applied Mathematics* **304**, 397–416 (2021).

$G = (V, E)$ digraph
$N(x)$ out-neighborhood, $N^{-}(x)$ in-neighborhood

- **Orientation** of G is any underlying oriented digraph of G (select one edge for each symmetric edge)

[3] A. Korchmaros, *Discrete Applied Mathematics* **304**, 397–416 (2021).

$G = (V, E)$ digraph
$N(x)$ out-neighborhood, $N^-(x)$ in-neighborhood

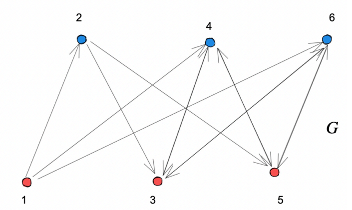
- **Orientation** of G is any underlying oriented digraph of G (select one edge for each symmetric edge)
- $x \sim y$ with $x, y \in V$ if $N(x) = N(y)$ and $N^-(x) = N^-(y)$

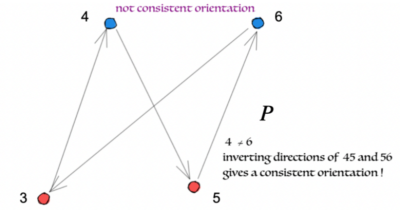
[3] A. Korchmaros, *Discrete Applied Mathematics* **304**, 397–416 (2021).

$G = (V, E)$ digraph
$N(x)$ out-neighborhood, $N^-(x)$ in-neighborhood

- **Orientation** of G is any underlying oriented digraph of G (select one edge for each symmetric edge)
- $x \sim y$ with $x, y \in V$ if $N(x) = N(y)$ and $N^-(x) = N^-(y)$
- **Consistent orientation**[3] of G if it preserves $\sim$

Not all orientations are consistent !!





[3] A. Korchmaros, *Discrete Applied Mathematics* **304**, 397–416 (2021).

Theorem 3 (A.K. 2021)

Let G be a connected 2-BMG.

1. Every consistent orientation of G is acyclic.

Theorem 3 (A.K. 2021)

Let G be a connected 2-BMG.

1. Every consistent orientation of G is acyclic.
2. G has a topological vertex ordering.

Topological vertex ordering of an oriented digraph is a relabeling of vertices st $x_i x_j \Rightarrow i < j$

Theorem 4 (Schaller, Stadler, Hellmuth 2021)

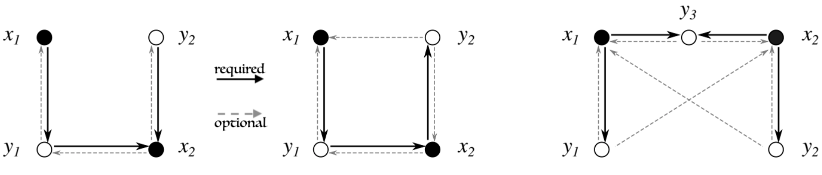
A bipartite (non necessary connected) vertex colored digraph G is 2-BMG iff

[4] D. Schaller *et al.*, *Theoretical Computer Science* **865**, 63–84 (2021).

Theorem 4 (Schaller, Stadler, Hellmuth 2021)

A bipartite (non necessary connected) vertex colored digraph *G* is 2-BMG iff

- *G* is sink-free and
- *G* does not contain an induced F1-, F2-, F3-graphs.



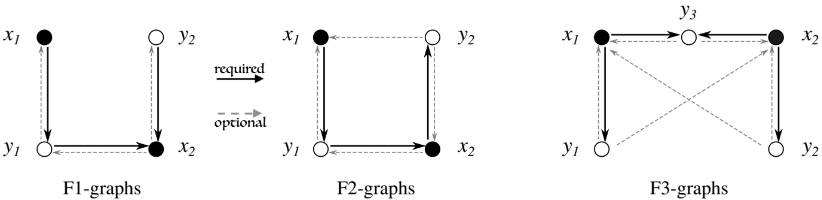
F1-graphs F2-graphs F3-graphs

[4] D. Schaller *et al.*, *Theoretical Computer Science* **865**, 63–84 (2021).

Theorem 4 (Schaller, Stadler, Hellmuth 2021)

A bipartite (non necessary connected) vertex colored digraph G is 2-BMG iff

- G is sink-free and
- G does not contain an induced F1-, F2-, F3-graphs.



There are 17 non-isomorphic forbidden induced subgraphs[4]

[4] D. Schaller *et al.*, *Theoretical Computer Science* **865**, 63–84 (2021).

- **2-qBMG**[5] is a bipartite qBMG

[5] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

- **2-qBMG**[5] is a bipartite qBMG

Theorem 5 (A.K., Schaller, Stadler, Hellmuth 2021)

A bipartite (connected) vertex-colored digraph *G* is 2-qBMG iff it satisfies N1, N2 and N3.

[5] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

- **2-qBMG**[5] is a bipartite qBMG

Theorem 5 (A.K., Schaller, Stadler, Hellmuth 2021)

A bipartite (connected) vertex-colored digraph G is 2-qBMG iff it satisfies N1, N2 and N3.

2-BMG = sink-free 2-qBMG

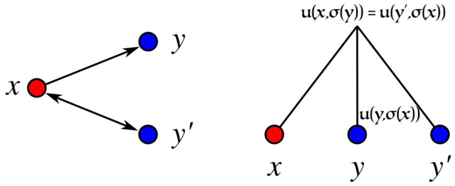
[5] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

- **2-qBMG**[5] is a bipartite qBMG

Theorem 5 (A.K., Schaller, Stadler, Hellmuth 2021)

A bipartite (connected) vertex-colored digraph *G* is 2-qBMG iff it satisfies N1, N2 and N3.

2-BMG = sink-free 2-qBMG



(G, σ) (T, σ, u)

[5] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

Proposition 1 (A.K., Schaller, Stadler, Hellmuth 2021)

2-qBMGs form a hereditary digraph class (every induced subgraph of a 2-qBMG is a 2-qBMG).

Proposition 1 (A.K., Schaller, Stadler, Hellmuth 2021)

2-qBMGs form a hereditary digraph class (every induced subgraph of a 2-qBMG is a 2-qBMG).

2-BMGs do not form an hereditary digraph class
(sink-freeness is not hereditary) !!

Theorem 6 (A.K., Schaller, Stadler, Hellmuth 2021)

Let G be a bipartite (non necessary connected) vertex colored digraph. If G does not contain an induced F1-, F2-, or F3-graph, then G is a 2-qBMG.

Theorem 6 (A.K., Schaller, Stadler, Hellmuth 2021)

Let G be a bipartite (non necessary connected) vertex colored digraph. If G does not contain an induced F1-, F2-, or F3-graph, then G is a 2-qBMG.

Proposition 2 (A.K., Schaller, Stadler, Hellmuth 2021)

Let G be digraph satisfying N1 and N2, then G is bipartite.

Open questions

[6] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

1. Which family of directed graphs G admits an L-coloring σ st (G, σ) is an L-qBMG? ($n = 2$ use N1,N2,N3)

[6] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

1. Which family of directed graphs G admits an L-coloring σ st (G, σ) is an L-qBMG? ($n = 2$ use N1,N2,N3)
2. **qBMG-coloring problem**:
   Given a digraph $G = (V, E)$ and integer $1 \leq L \leq |V|$.
   (a) Is there a L-coloring σ st (G, σ) is an L-qBMG?

[6] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

Open questions

1. Which family of directed graphs G admits an L-coloring σ st (G, σ) is an L-qBMG? ($n = 2$ use N1,N2,N3)
2. **qBMG-coloring problem**:
   Given a digraph $G = (V, E)$ and integer $1 \leq L \leq |V|$.
   (a) Is there a L-coloring σ st (G, σ) is an L-qBMG?
   Case $L = |V|$: Rooted start tree
   $u(x, \sigma(y)) = \rho$ if $xy \in E$; $u(x, \sigma(y)) = x$ if $xy \notin E$.

[6] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

1. Which family of directed graphs G admits an L-coloring σ st (G, σ) is an L-qBMG? ($n = 2$ use N1,N2,N3)
2. **qBMG-coloring problem**:
   Given a digraph $G = (V, E)$ and integer $1 \leq L \leq |V|$.
   (a) Is there a L-coloring σ st (G, σ) is an L-qBMG?
   Case $L = |V|$: Rooted start tree
   $u(x, \sigma(y)) = \rho$ if $xy \in E$; $u(x, \sigma(y)) = x$ if $xy \notin E$.
   (b) What is the qBMG-chromatic number of (G, σ)?
   (minimum number of colors for which G is a qBMG).
   If a L-coloring exists this can be transformed in an $L + 1$-coloring (by arbitrarily splitting a color class)[6].

[6] A. Korchmaros *et al.*, *arXiv preprint arXiv:2109.10235* (2021).

Thank you

- Peter Stadler
- David Schaller
- Marc Hellmut

Theorem 3 (A.K.)

Let G be a connected 2-BMG.

1. Every consistent orientation of G is acyclic.
2. G has a topological vertex ordering.

Topological vertex ordering of an oriented digraph is a relabeling of vertices st $x_i x_j \Rightarrow i < j$

Is every orientation of G acyclic? No, P has a 4-cycle.