On polyclonality of intestinal tumors

Michael A. Newton

University of Wisconsin

Chaos and Complex Systems April 2006

Thanks

Linda Clipson
W.F. Dove
Rich Halberg
Stephen Stanhope
Ruth Sullivan
Andrew Thliveris

Outline

- Bio
- Three statistical questions
  1. What is the polyclonal fraction?
     - estimation
  2. Is random collision plausible?
     - testing
  3. What is the spatial extent of interactions?
     - spatial modeling

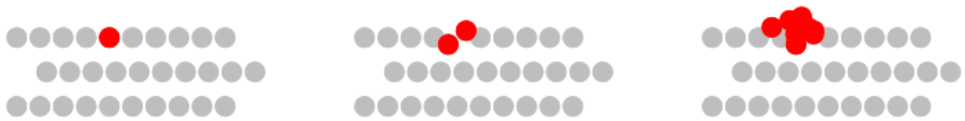
Multiple Intestinal Neoplasia (*Min*) mouse

- Inherits mutation in the tumor suppressor gene Apc (adenomatous polyposis coli)
- Presents X intestinal tumors (quantitative trait)
- Provides an animal model of intestinal cancer

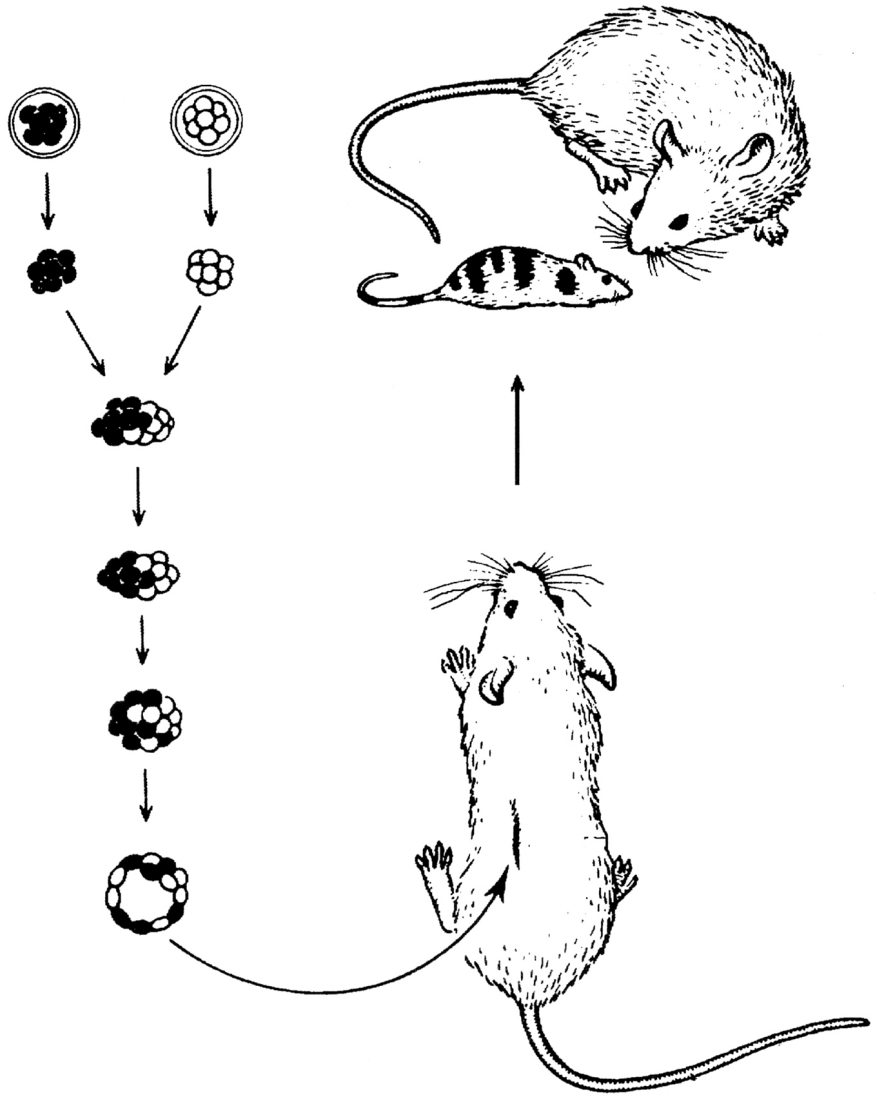
- Biology of tumor initiation not well understood
- Full Apc inactivation is an early event in tumor formation
- Distribution of X is affected by modifier genes

Clonal or polyclonal origin?

clonal	cells of a tumor descend from a single initiated aberrant cell
polyclonal	cells descend from multiple initiated cells

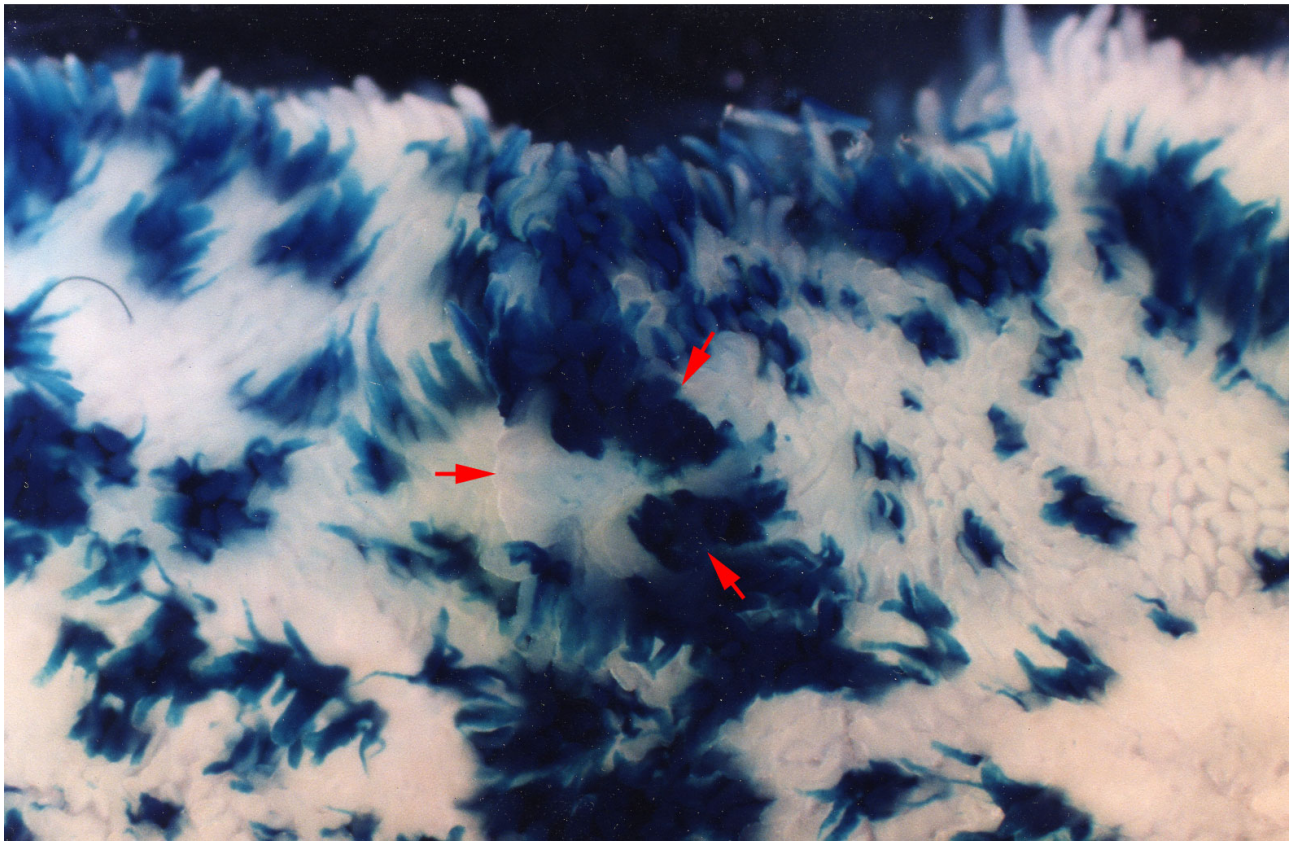


Aggregation chimeras enable detection of polyclonality

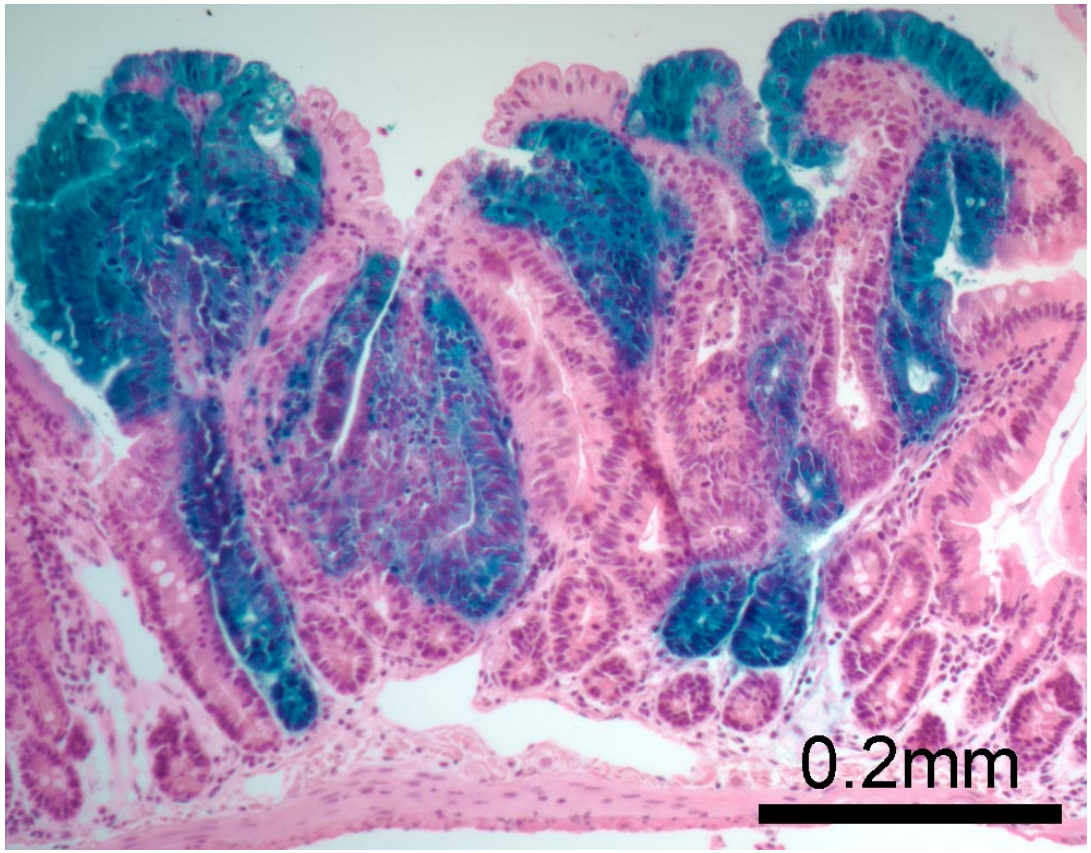


Intestinal epithelium of B6 chimera: patchwork, tumor

$$\mathrm{B6}\,Apc^{Min/+}\,Mom1^{R/R} \longleftrightarrow \mathrm{B6}\,Apc^{Min/+}\,Mom1^{R/R}\,ROSA26/+$$



Section shows tumor cells from both embryonic lineages



Heterotypic tumors not infrequent at low tumor multiplicity

		Counts of small intestinal tumors				
Mouse	%blue	Total	Heterotypic	Pure blue	Pure white	Ambiguous
1	20	19	5	5	6	3
2	85	24	3	13	6	2
3	20	9	2	2	5	0
4	60	19	3	2	10	4
5	30	24	2	0	21	1
6	50	9	2	2	3	2
7	40	8	5	0	3	0
Total		112	22	24	54	12

Heterotypic $\Rightarrow$ polyclonal, *but*, polyclonal $\not\Rightarrow$ heterotypic

	Clonality		
Phenotype	monoclonal ($C = 1$)	polyclonal ($C > 1$)	
blue (B)	$P\{B \cap (C = 1)\}$	$P\{B \cap (C > 1)\}$	$P(B)$
white (W)	$P\{W \cap (C = 1)\}$	$P\{W \cap (C > 1)\}$	$P(W)$
heterotypic (HET)	0	$P\{\mathrm{HET} \cap (C > 1)\}$	$P(\mathrm{HET})$
	$P(C = 1)$	$\theta = P(C > 1)$	100%

Q1: What fraction θ of tumors are polyclonal?

- $\mathrm{HET} \subset (C > 1) \Rightarrow P(\mathrm{HET}) \leq \theta$
  - $22/(22 + 24 + 54) = 22\%$
- Is there a better estimate or lower bound?
  - Novelli *et al.* 1996 proposed the lower bound

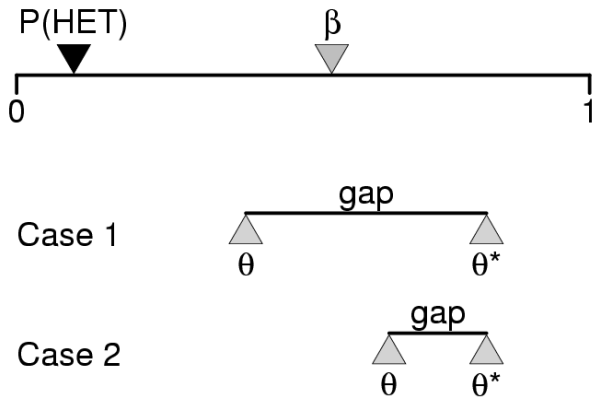
$$\beta = \frac{P(\mathrm{HET})}{P(\mathrm{HET} \cup \mathrm{HOM}_{\min})}$$

  - $22/(22 + 24) = 48\%$

Novelli's bound β is not valid

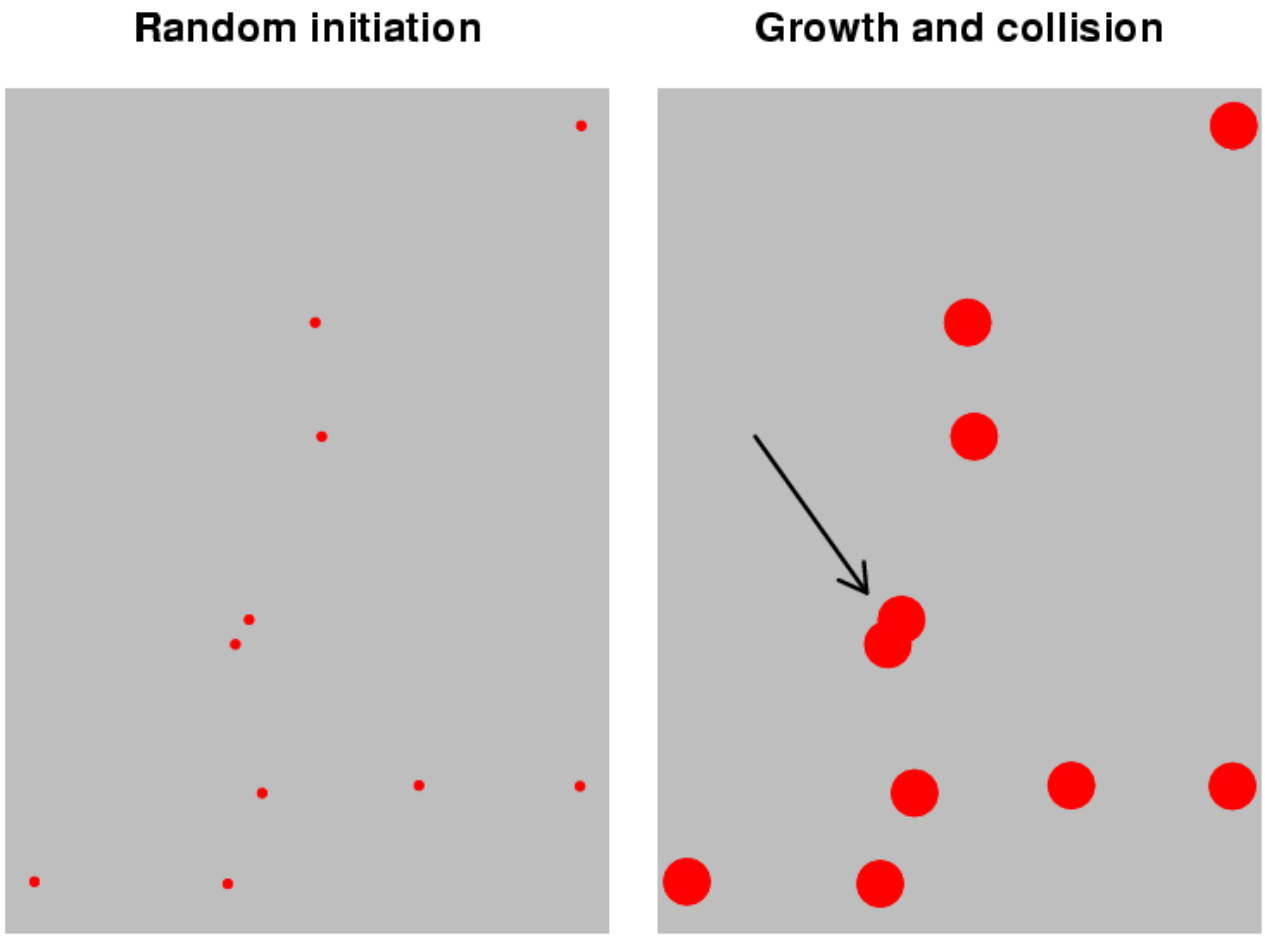
Theorem: Depending on system, either $\beta \geq \theta$ or $\beta \leq \theta$.

Sketch: Novelli's β is ok if $\theta = \theta^* = P(C > 1 | \mathrm{HET} \cup \mathrm{HOM}_{\min})$, but there is a gap $\theta < \theta^*$, under some regularity. $\square$



Problem: Estimation of θ is sensitive to assumptions about the mechanisms by which clones are bound into polyclonal tumors.

Random collision: a simple mechanism of polyclonality



Q2: Are the data consistent with random collision?

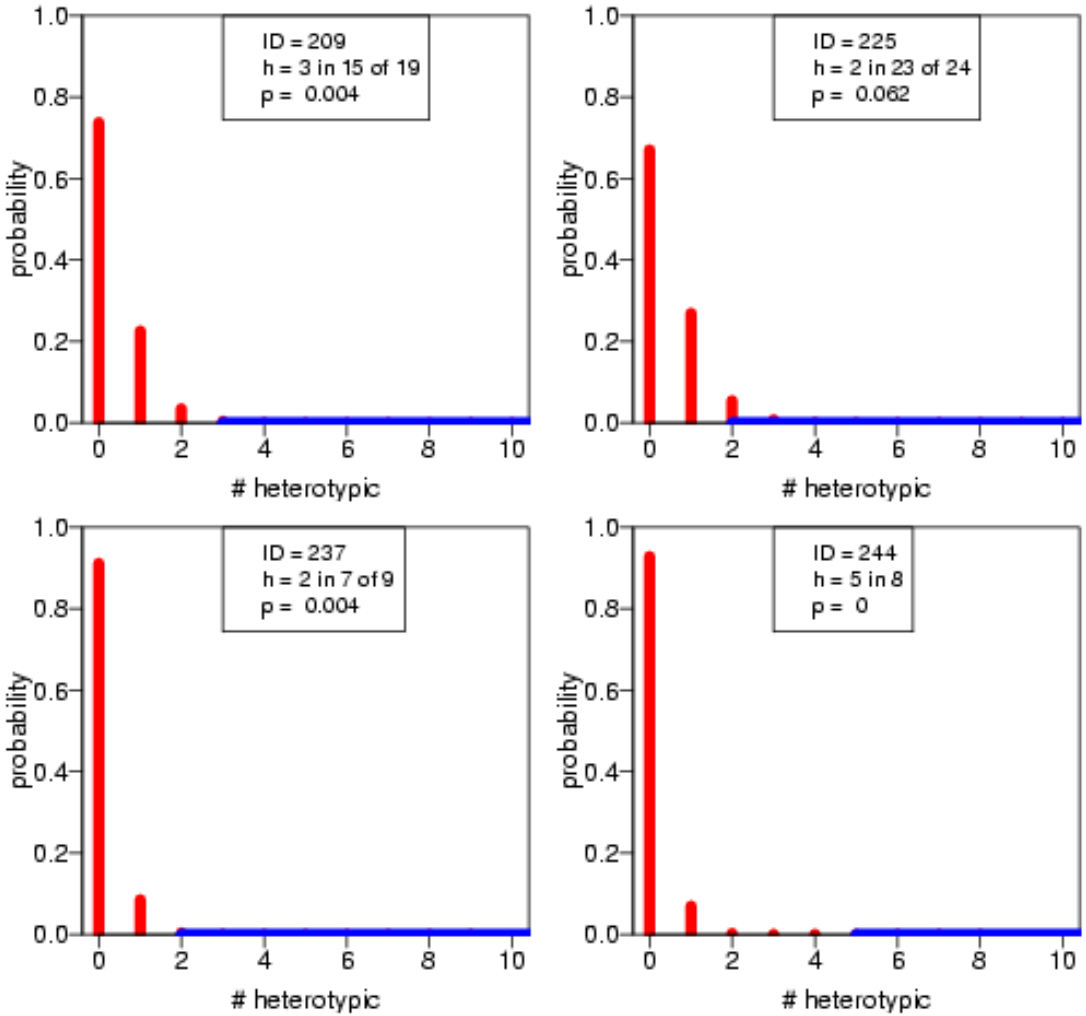
Idea: Considering low tumor multiplicity and small size, the number of collisions should be low on H_0.

Test statistic: Number of heterotypic tumors

Methods:

- Unknown, mouse specific numbers of initiated cells
- Complicated distributions induced on # collided pairs, triples, etc.
- Overdispersion
- ▸ DETAILS of stochastic geometry approxs & posterior predictive inference approach

Random collision is not plausible

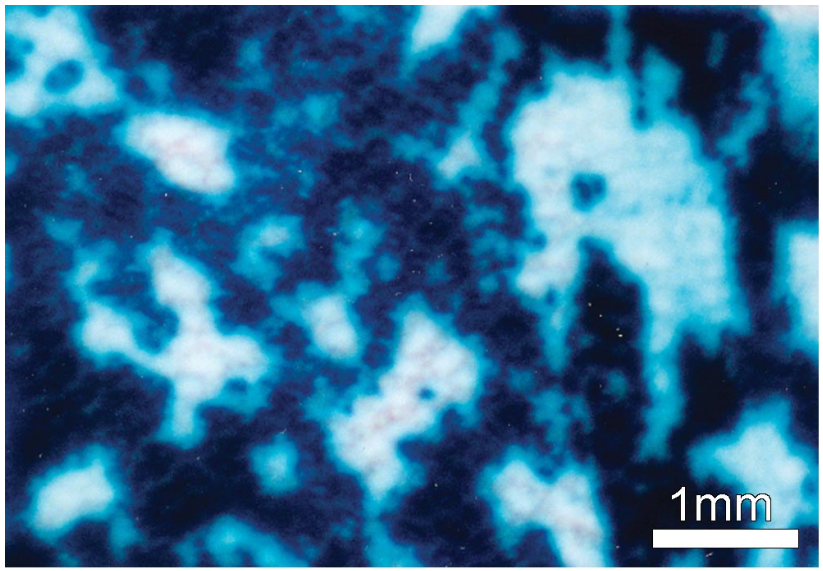


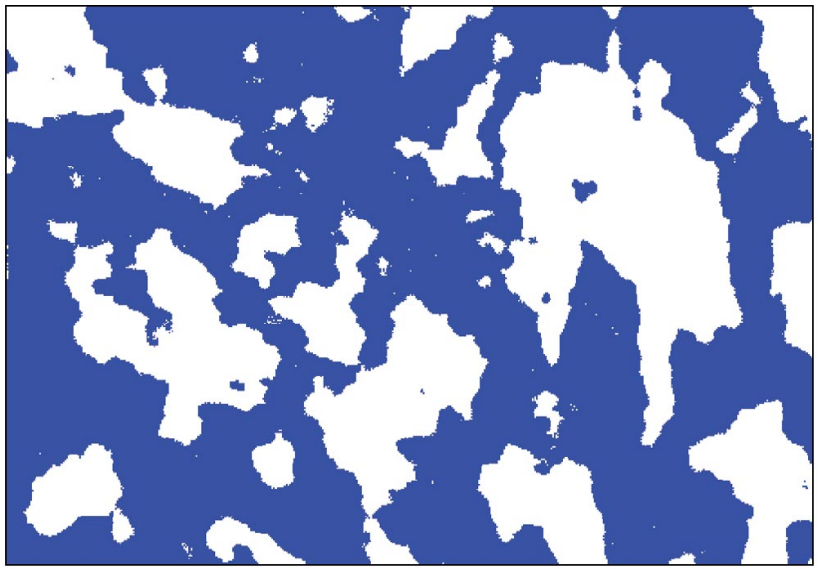
Going further

Spatial data and analysis reveal the extent of spatial interaction among clones.

Polyclonal tumors have opportunity to be heterotypic at boundaries

Intestinal epithelium adjacent to a tumor

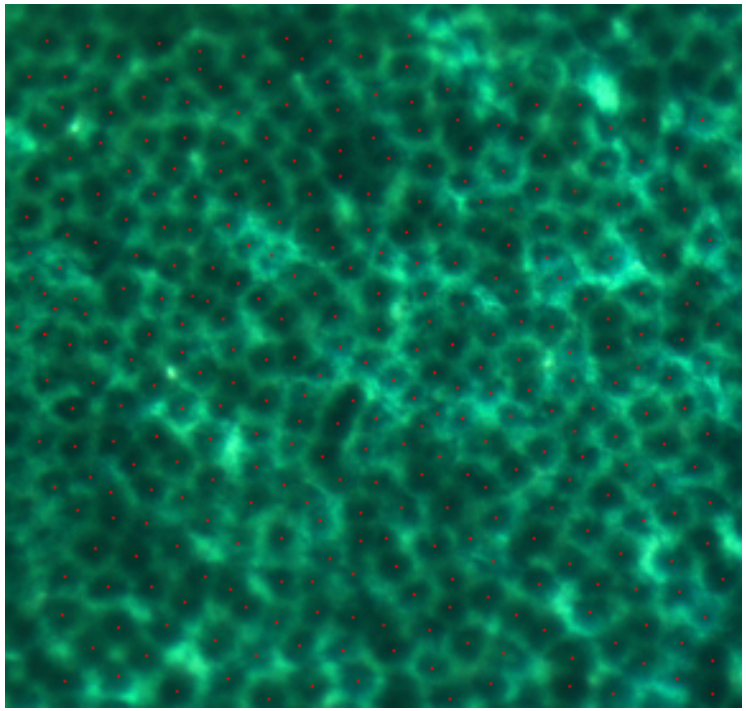




...plus more...images from regions adjacent to every tumor in 3 mice

Crypts, rather than cells, are the basic structural units

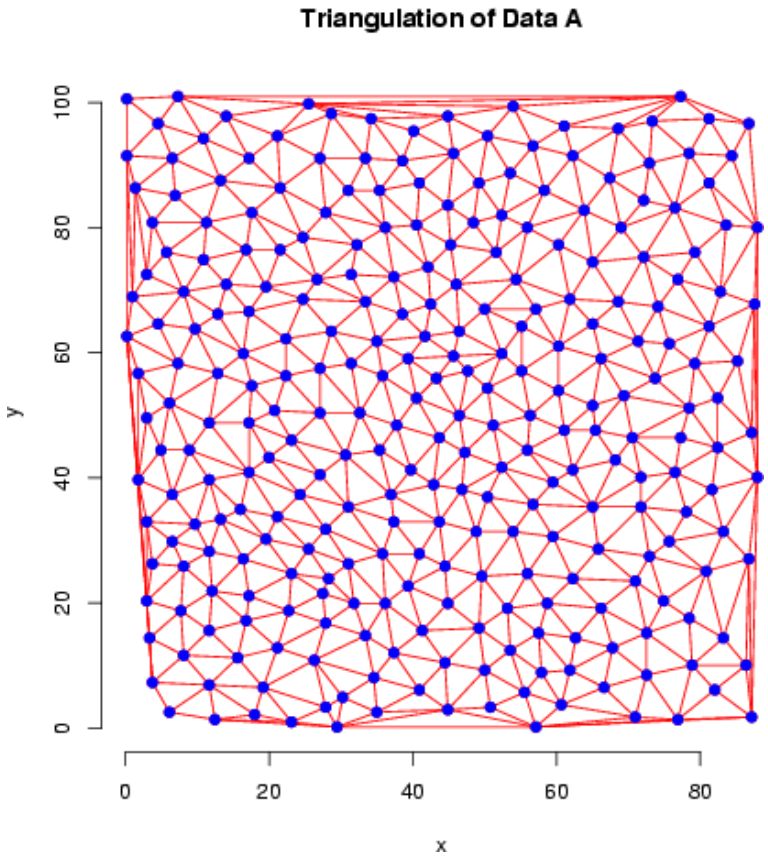
crypt an organized group of proliferating cells
- intestinal epithelium formed from $O(10^5)$ crypts
- crypts are clonal

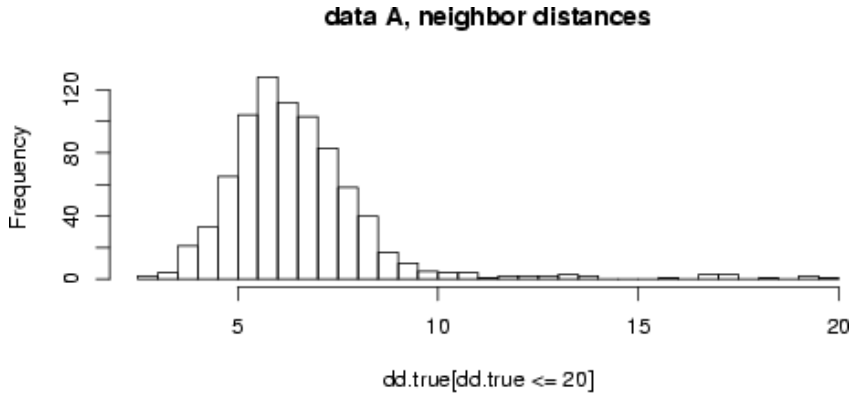


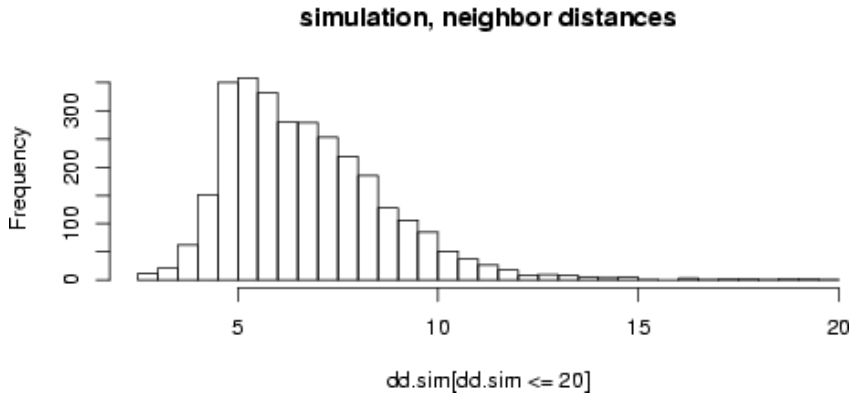
Crypts from non-chimeric mouse

Problem: Blue/white images mask crypt arrangement

Crypt arrangement data from non-chimeric mouse



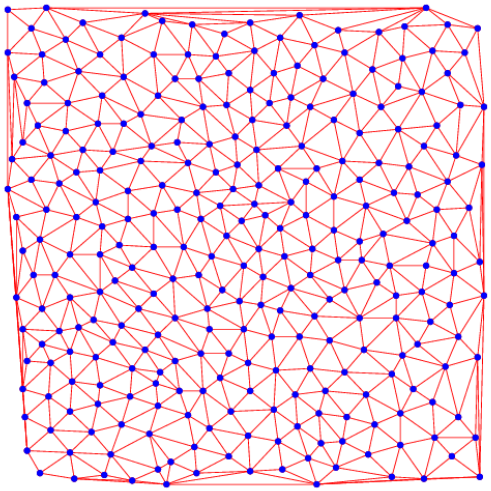


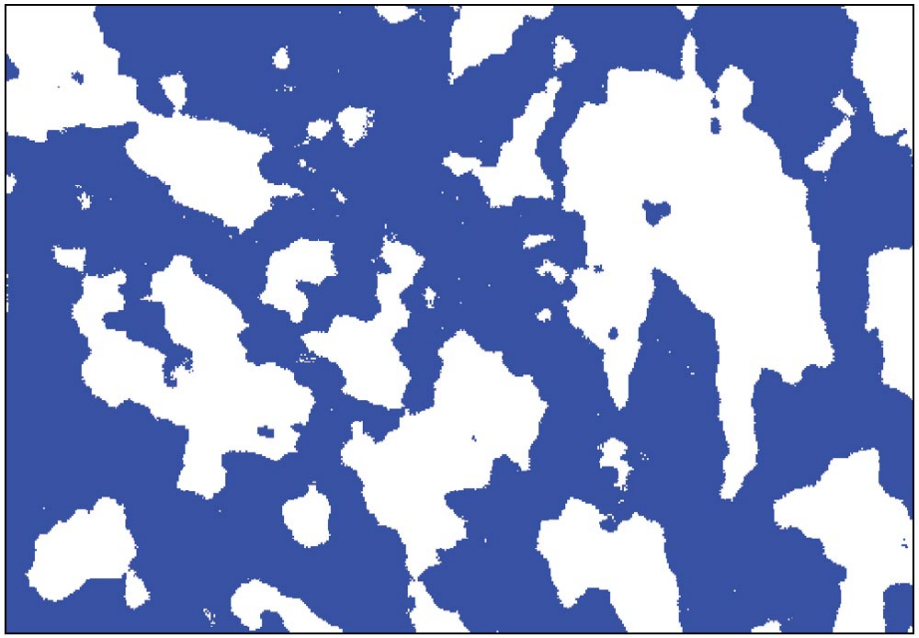


▸ On Delaunay triangulation

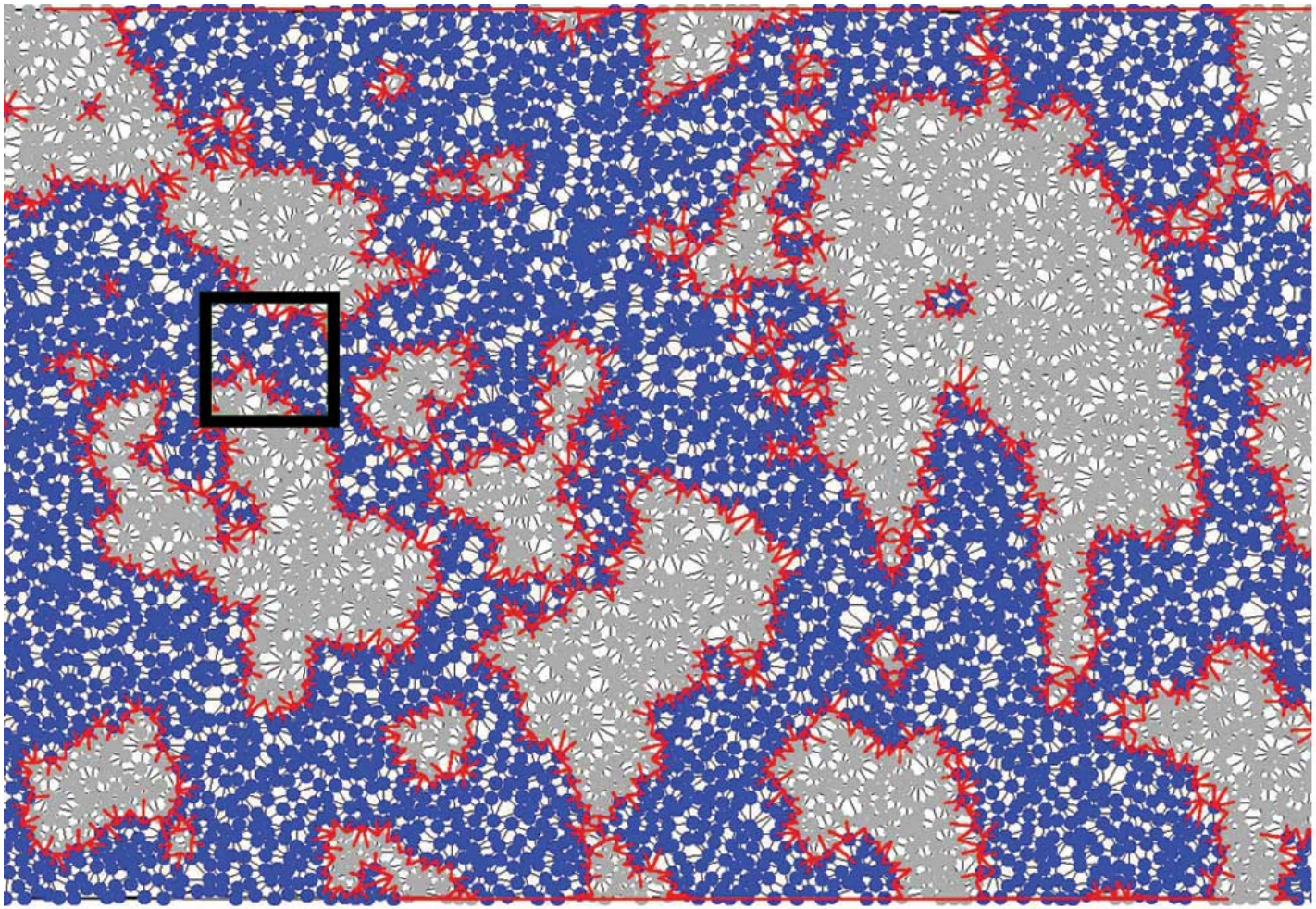
Computational inference task

Combine



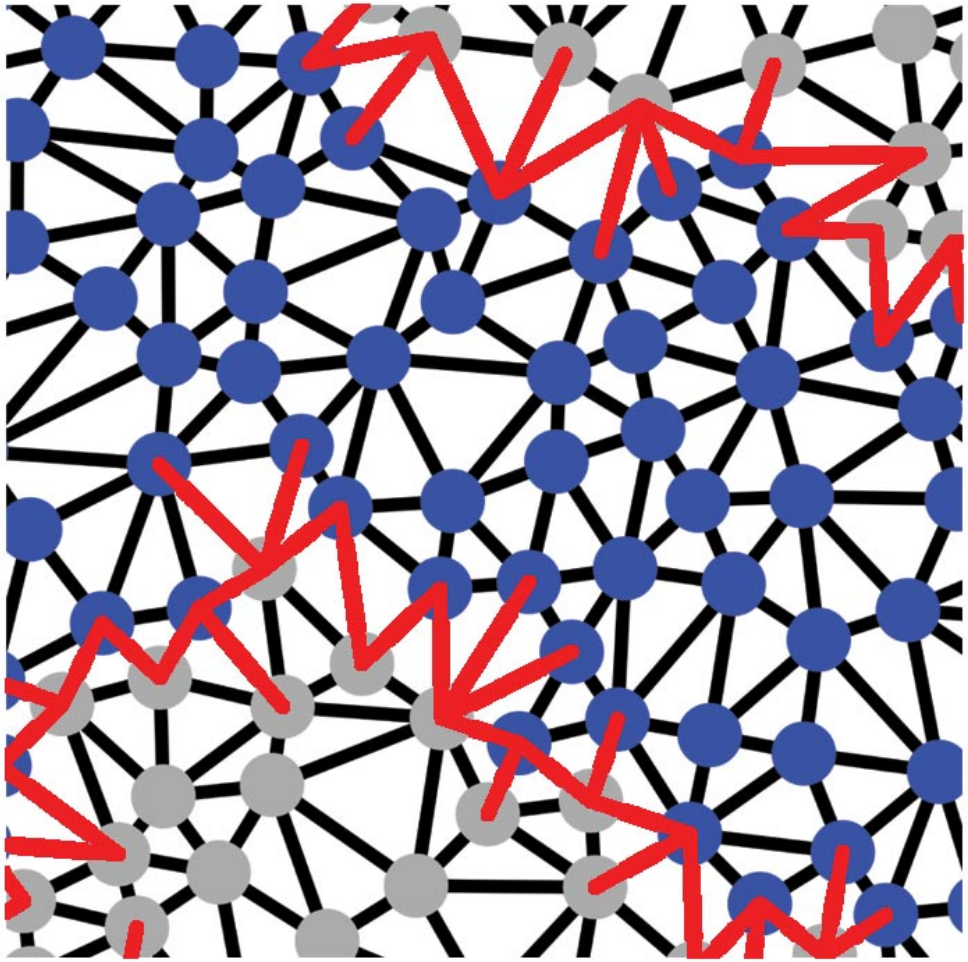


Statistical reconstruction of crypts in chimeric patches



# crypts	5642	# edges	16902	# triangles	33765
% white	39	% white	33	% white	29
		% mixed	14	% mixed	21

Blow up



Approach: Bayesian image restoration

- data I, a chimeric pattern image
- unknown crypt layout $c = \{c_i\}$, the collection of crypt centers
- crypt reconstructions $\hat{c}$ by MCMC sampling from

$$p(c|I) \propto \underbrace{p(c)}_{\textit{prior}} \underbrace{p(I|c)}_{\textit{likelihood}}$$

- estimate detection rate $P(\mathrm{HET}|C > 1)$ or $P(\mathrm{HET})$ as $f(\hat{c})$

Image prior

- point process prior for crypt centers c
- $p(c) \propto \exp\left\{-\sum_{(i,j)} h(d_{i,j})\right\}$ for a potential function h and inter-crypt distances $\{d_{i,j}\}$
- hard core model; Ripley model; (fixed n)
- Estimate prior features using crypt arrangement data.
- ▸ Details

Image likelihood

Model: $p(I|c)$

- Premise: crypts are probably pure
- Latent crypt colors (blue/white) iid Bernoulli(p)
- If crypt i is blue, each pixel in circle near c_i is white w.p. ϵ.
- If crypt i is white, each pixel in circle near c_i is blue w.p. ϵ

$$p(I|c) = \left\{ \prod_i p\epsilon^{w(i)}(1-\epsilon)^{b(i)} + (1-p)\epsilon^{b(i)}(1-\epsilon)^{w(i)} \right\} \left\{ p^B(1-p)^W \right\}$$

where $w(i)$ and $b(i)$ are numbers of white and blue pixels near i, and W and B are numbers in the intercryptal space.

Posterior sampling

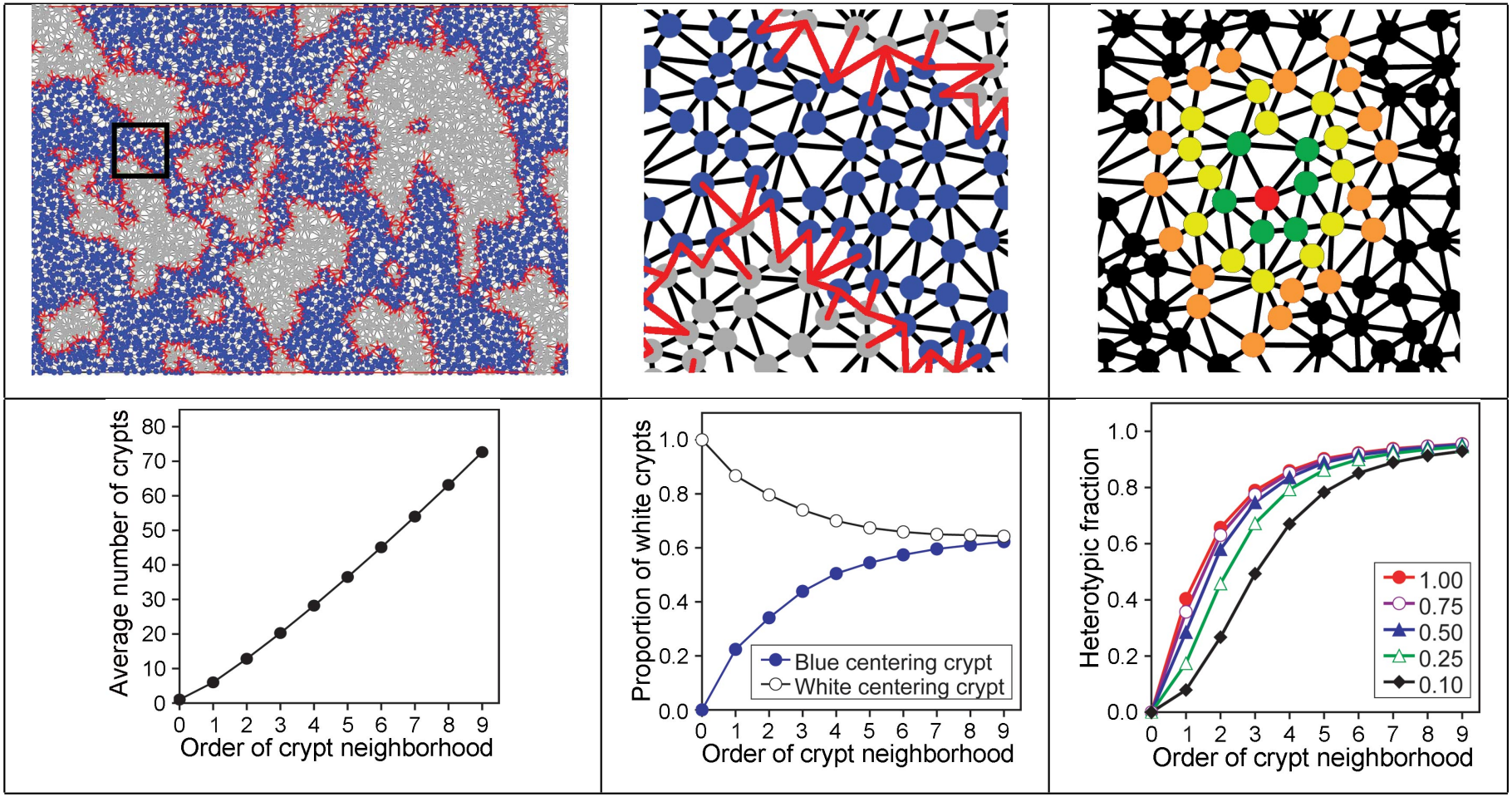
For each image:

- run Metropolis algorithm from regular hexagonal start
- pluck end state as reconstruction $\hat{c}$
- triangulate $\hat{c}$ and compute summaries $f(\hat{c})$

Fortunately:

- very low posterior variance of certain features $f(c)$
- good robustness to n

Crypt reconstruction indicates short-range interactions



Concluding remarks

- Intestinal tumors can be polyclonal.
  - Min mouse chimera; improved marker; reduced multiplicity
- Estimation of the polyclonal fraction is difficult
  - Novelli's bound
  - sensitivity to polyclonal mechanism
- Heterotypic rates are too large for random collision.
  - stochastic geometry
  - posterior predictive p-values
- Local interaction at the range of 1 to 2 neighboring crypts explains the data.
  - crypt reconstruction via Bayesian image analysis; disc model

Main references

Newton, MA (2006). On estimating the polyclonal fraction in lineage-marker studies of tumor origin. *Biostatistics*, in press.

Newton, MA, Clipson, L, Thliveris, AT, and Halberg, RB (2006). A statistical test of the hypothesis that intestinal tumors arise by random collision of initiated clones. *Biometrics*, in press.

Thliveris, AT, Halberg, RB, Clipson L, Dove, WF, Sullivan, R, Washington, MK, Stanhope, S, and Newton, MA (2005). Polyclonality of familial murine adenomas: Analysis of mouse chimeras with low tumor multiplicity suggest short-range interactions. *Proc. Natl. Acad. Sci.* USA, 102, 6960-6965.

Thanks

Appendices

Random collision hypothesis

- RC:
  - N initiated cells emerge randomly within intestine
  - Two collide if they are within δ
  - Tumors correspond to *connected* subsets of the induced graph
- Is there an argument against random collision using count and size data?
- Intuition: we see more mixed tumors than expected considering small sizes and low multiplicity

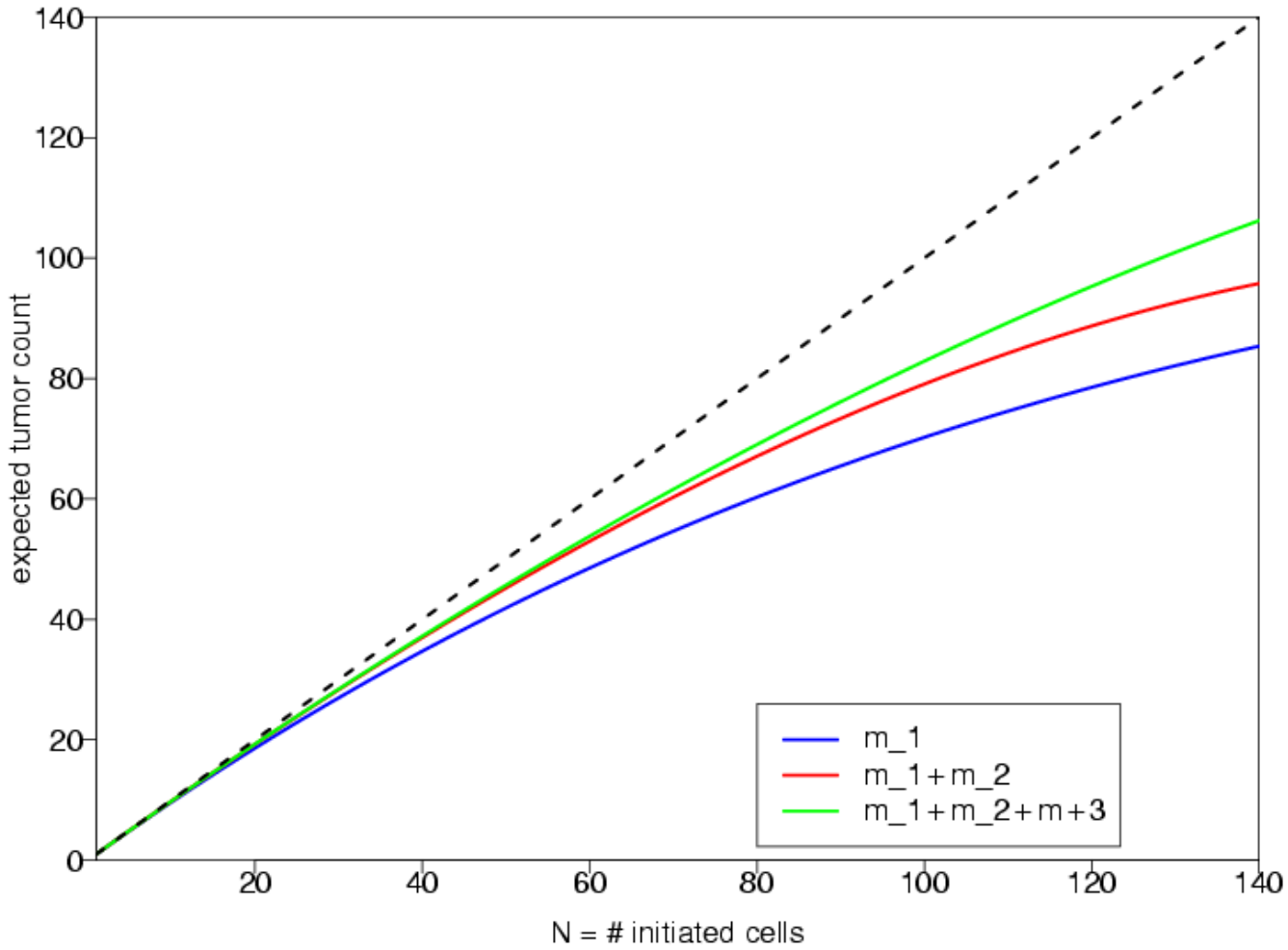
Simple collision theory

$$\begin{aligned} N &= \text{number of initiated crypts} \\ &= X_1 + 2X_2 + 3X_3 + \cdots \end{aligned}$$

where $X_j = \text{number of tumors formed from } j \text{ collided crypts}$
and thus the number of tumors is $X = \sum_j X_j$.
From Armitage (1949)

$$\begin{aligned} E(X_1) &\approx N \exp(-4\psi) \\ E(X_2) &\approx 2N \left(\psi - \frac{4\pi + 3\sqrt{3}}{\pi} \psi^2 \right) \\ E(X_3) &\approx N \left(\frac{4(2\pi + 3\sqrt{3})}{3\pi} \psi^2 \right). \end{aligned}$$

where $\psi = \pi N \delta^2 / (4A)$
Poisson approximation holds under sparse graph conditions


140
120
100
80
60
40
20
0
expected tumor count
20
40
60
80
100
120
140
N = # initiated cells
m_1
m_1+m_2
m_1+m_2+m+3

Testing random collision

- Use *size* data to set δ
- Develop Poisson/neg. binomial model for singles X_1^i, doubles X_2^i, and triples X_3^i, mouse i
- Conditional on *count* data $\{X^i = (X_1^i + X_2^i + X_3^i)\}$, simulate posterior of singles, doubles and triples
- Simulate posterior predictive of numbers sectored
  - randomly paint doubles, triples
- Compare to observed numbers sectored

Poisson/negative binomial

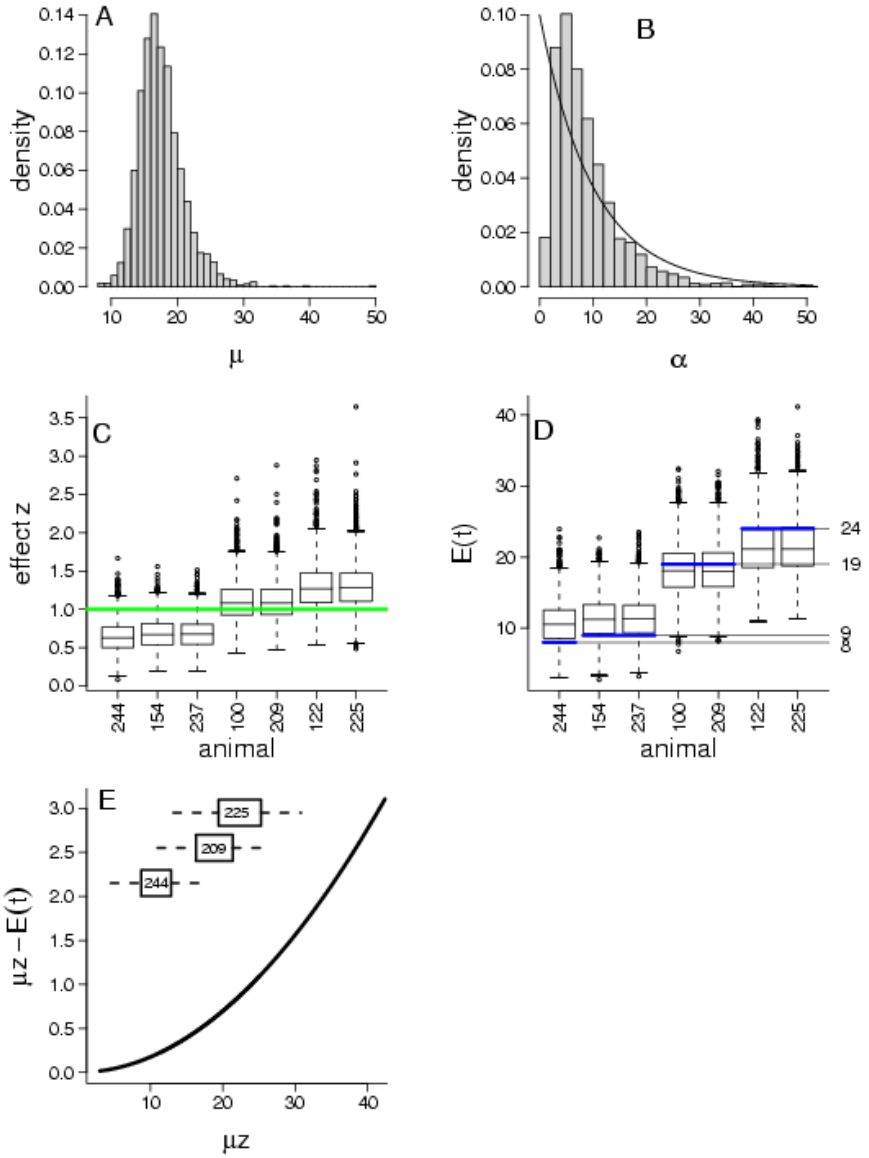
Unknowns

- μ expected number of initiated crypts per mouse
- Z_i Gamma(α, α) over-dispersion effect, mouse i
- α shape parameter
- (X_1^i, X_2^i, X_3^i) counts of singles, doubles, triples, mouse i

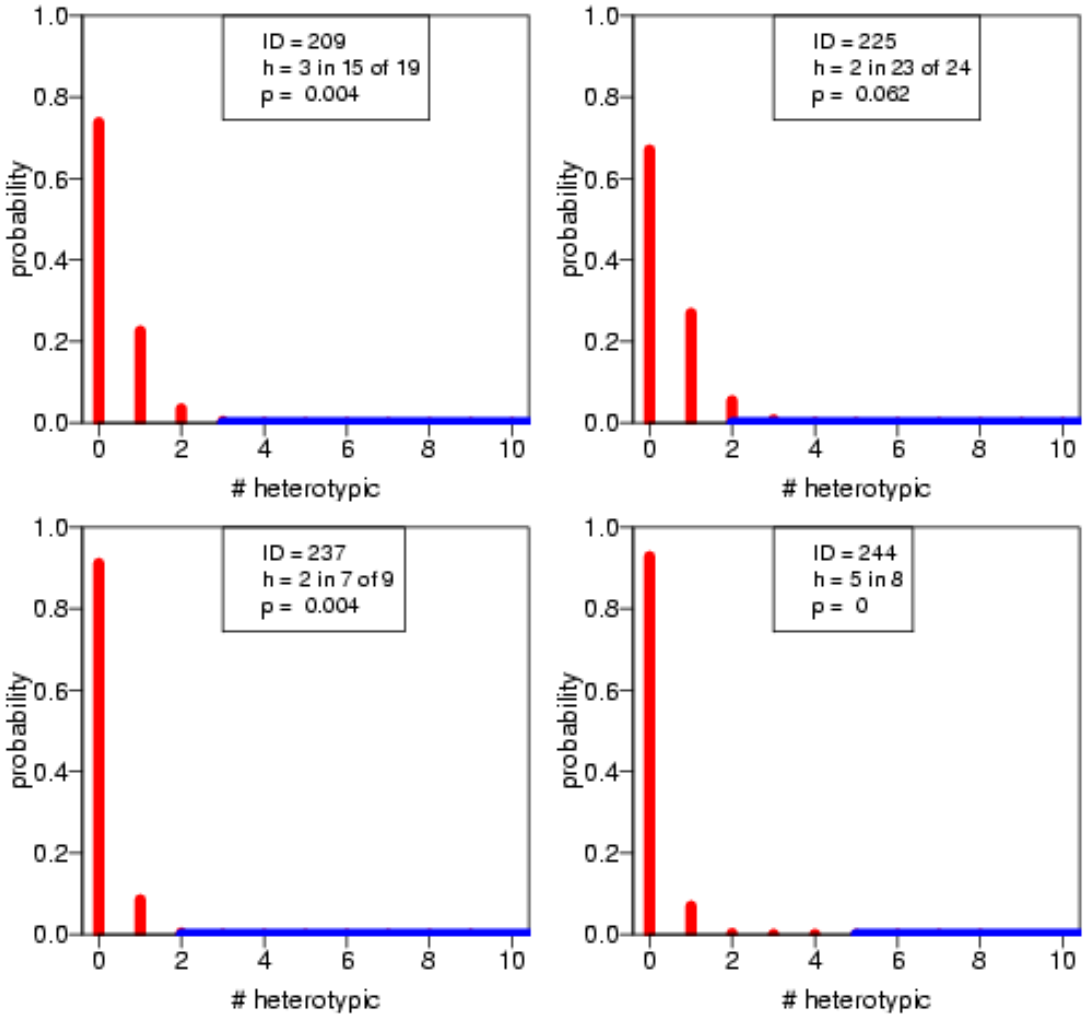
Use conditional Poisson model with N replaced by μZ_i for mouse i

Fit by MCMC

Posterior


A
density
μ
B
density
α
C
effect z
animal
244
154
237
100
209
122
225
D
E(t)
animal
24
19
8.9
E
μz − E(t)
μz
225
209
244

Posterior predictive



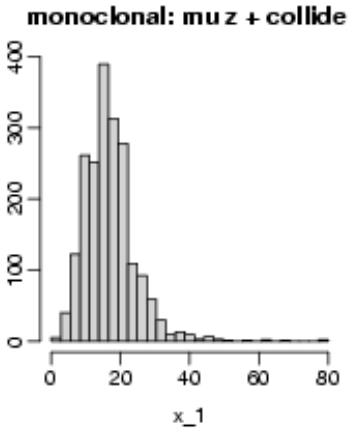
Random collision test results

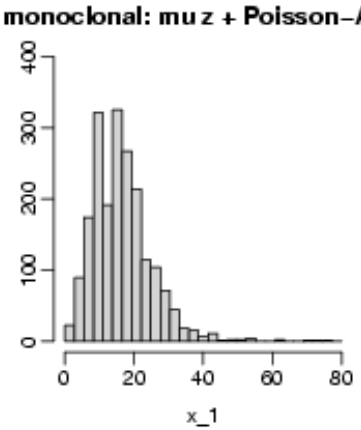
Mouse ID	%blue tissue	Tumor count					p-value
		total	white	blue	heterotypic	NA	
100	20	19	6	5	5	3	0.000
122	85	24	6	13	3	2	0.002
154	20	9	5	2	2	0	0.002
209	60	19	10	2	3	4	0.004
225	30	24	21	0	2	1	0.062
237	50	9	3	2	2	2	0.004
244	40	8	3	0	5	0	0.000

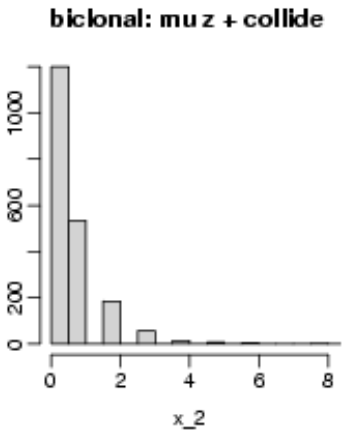
$\delta = 1.5mm$

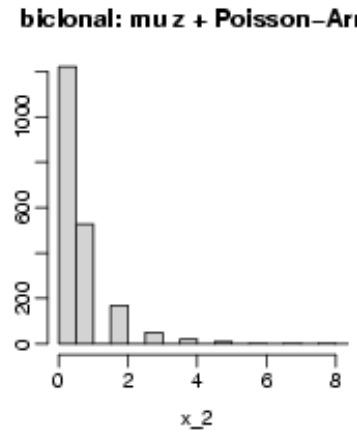
Model checking

Return to main

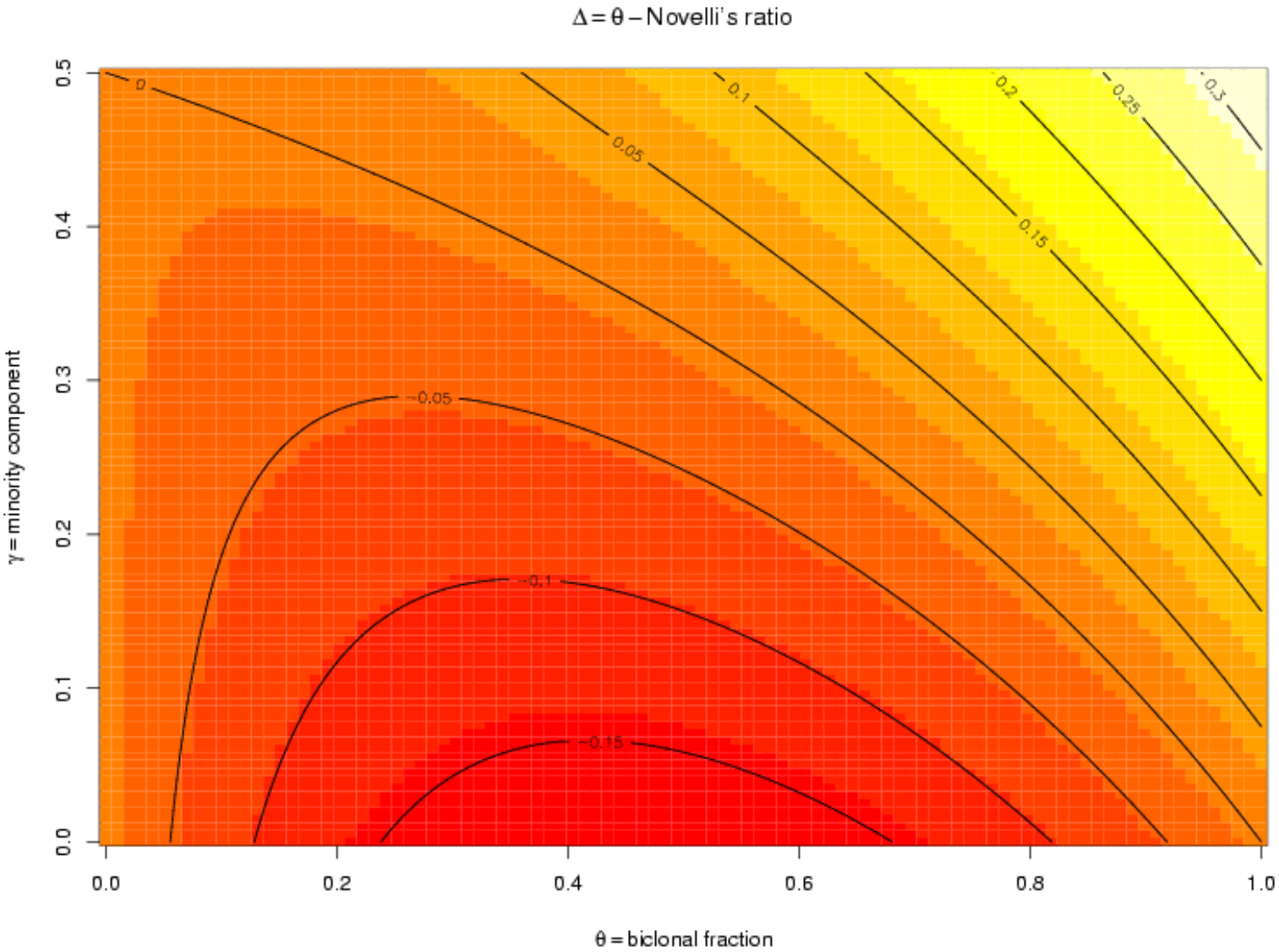




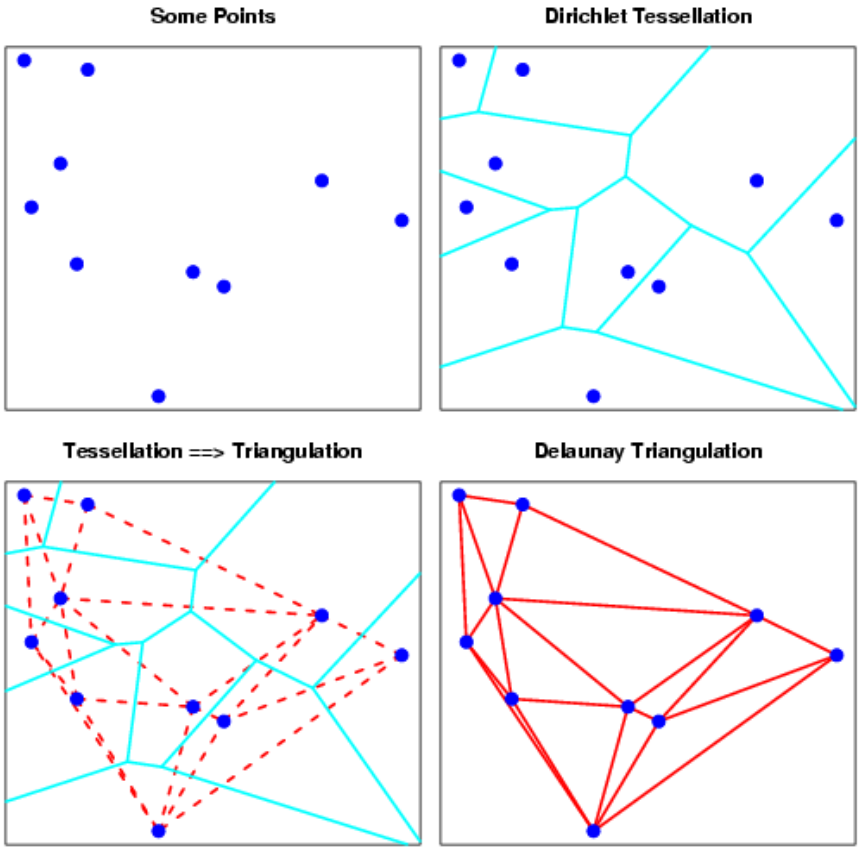




Novelli's bound compared to θ: biclonal model

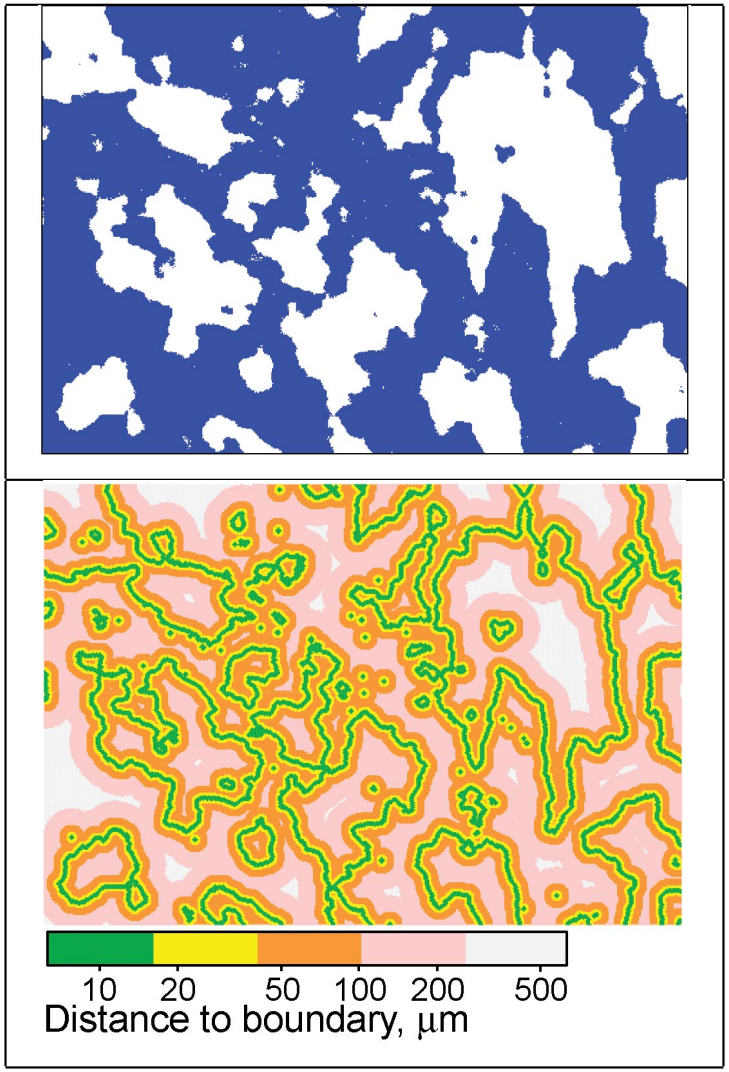


Tessellation and Triangulation

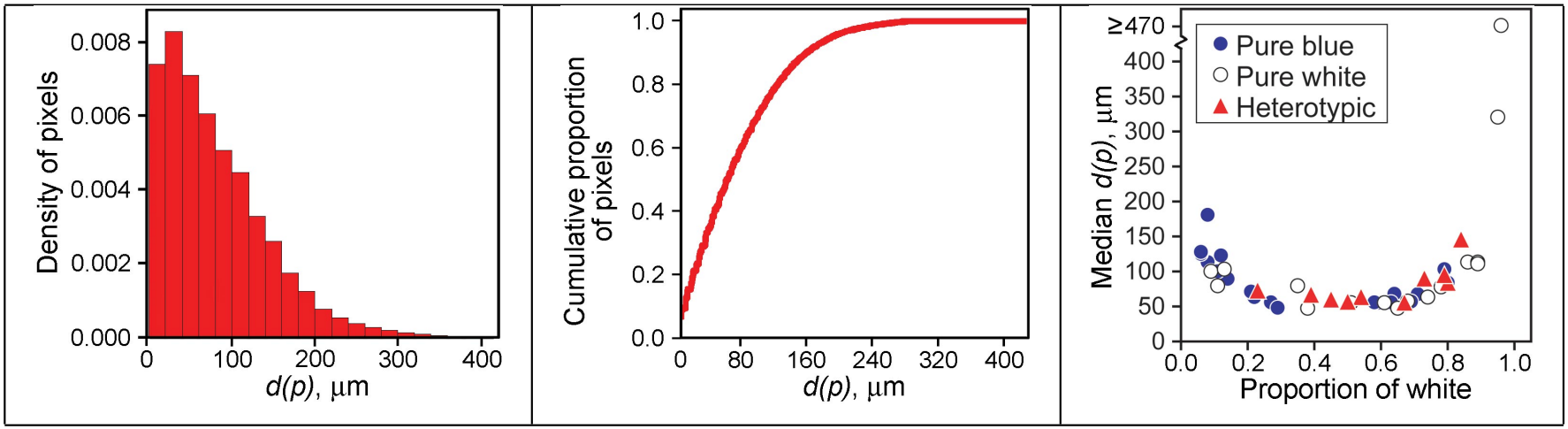


Return to crypts

A useful image transform

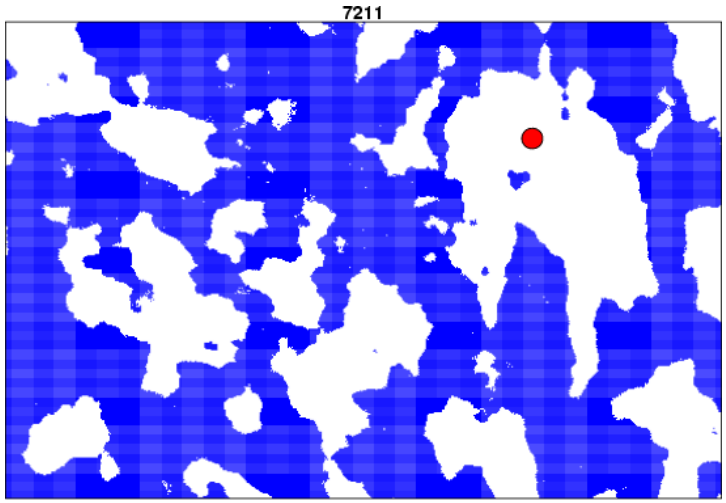


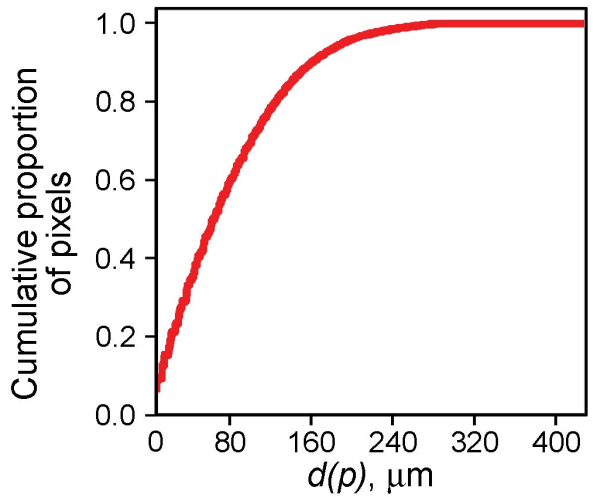
Chimeric pattern image summaries



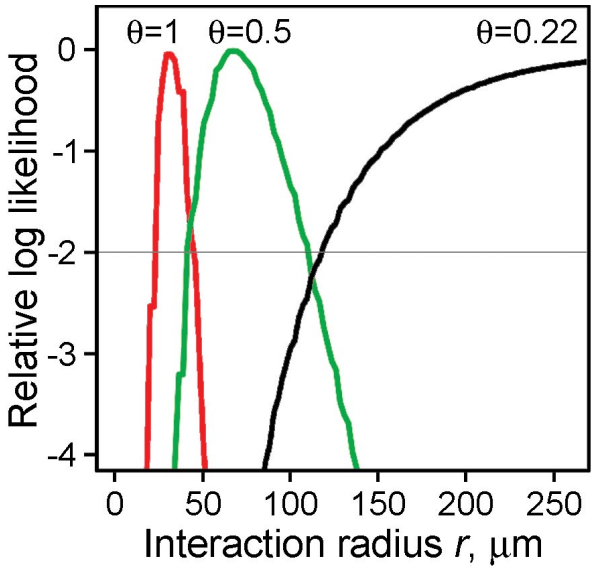
Disc model

Assume $S = \#$ heterotypic tumors $\sim$ Binomial$\{n, \theta F(r)\}$, where θ is the polyclonal fraction.





Disc model results



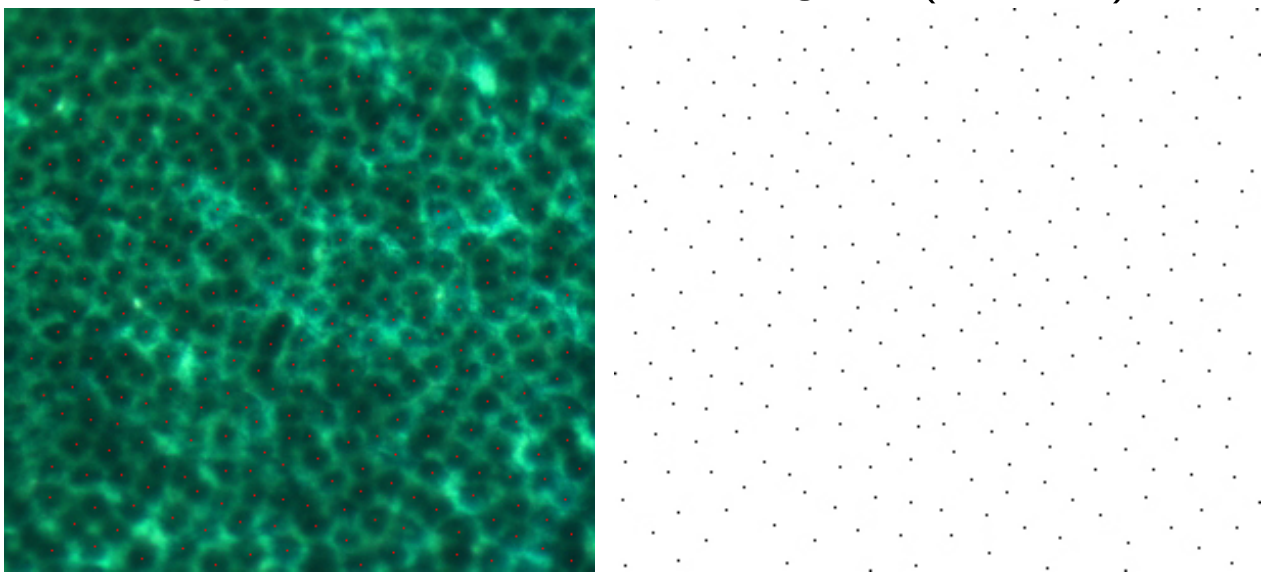
Inference - short-range interactions explain the tumor count data

Limitations - model allows elementary interactions only
- characterization is not in terms of crypt structure

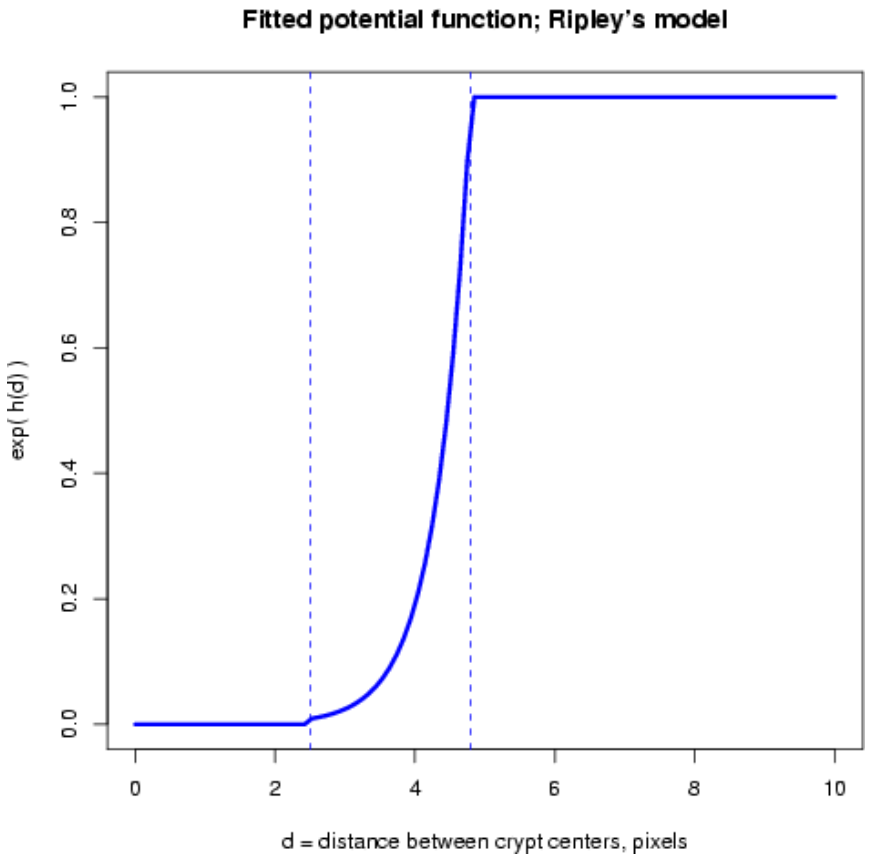
Return to crypts

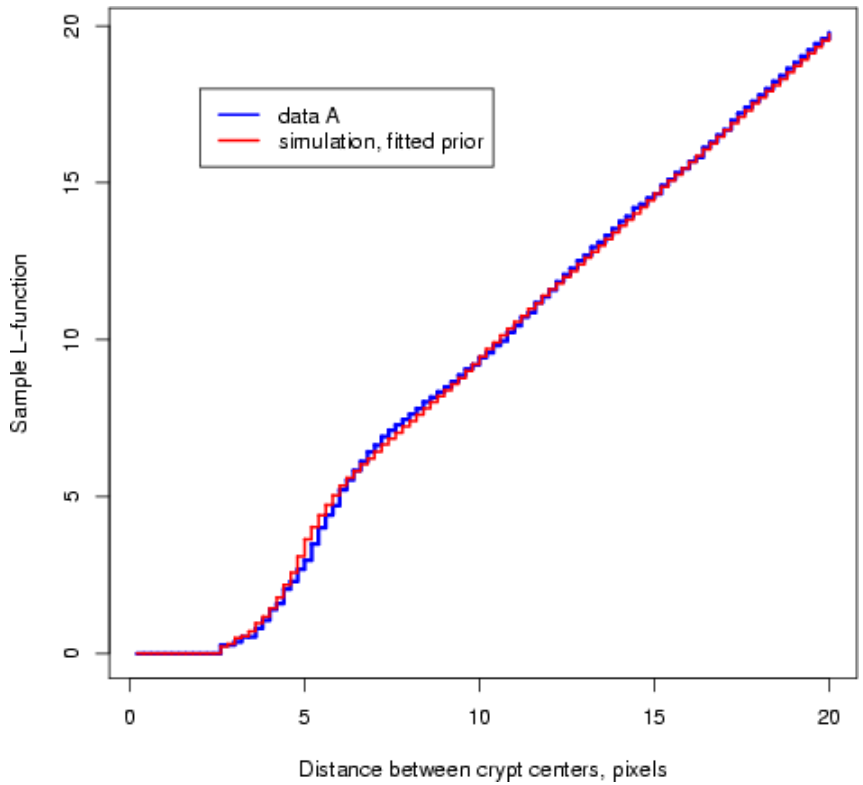
Crypt data

Crypt locations, sample region (data A)

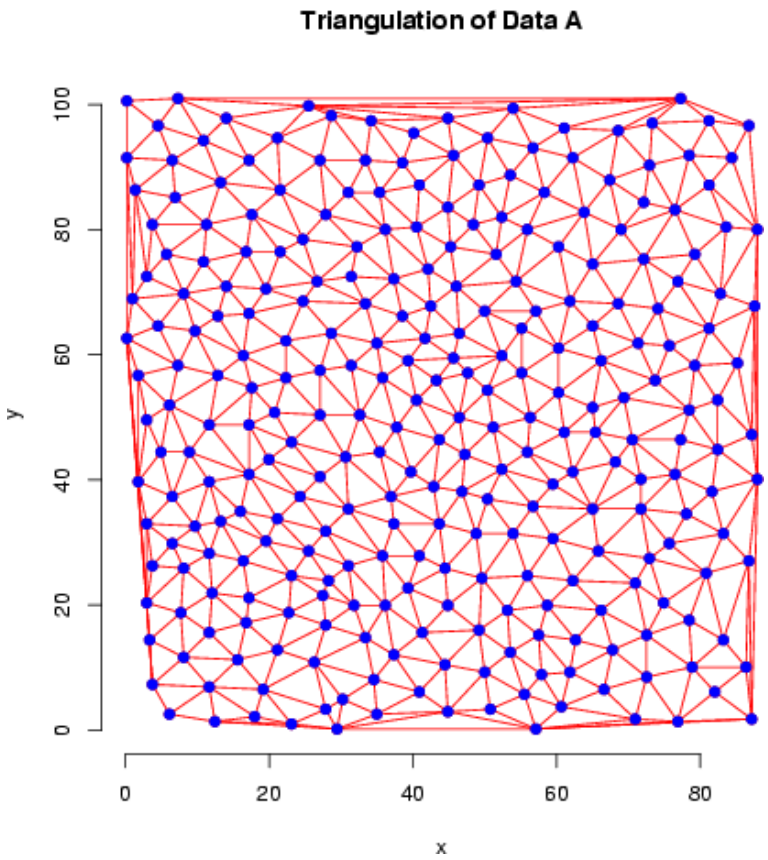


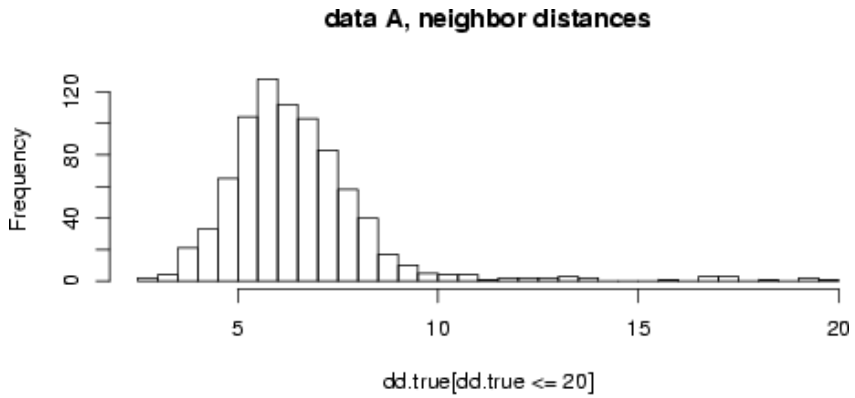
Fitting[1] and checking spatial model of crypt locations

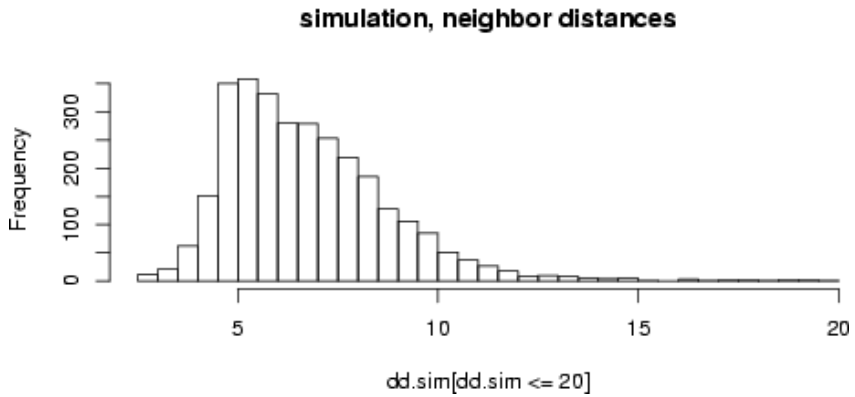




[1] maximum pseudo-likelihood







Return to main