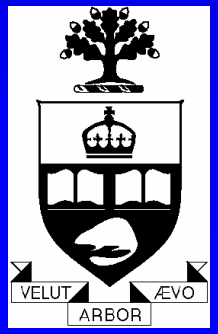
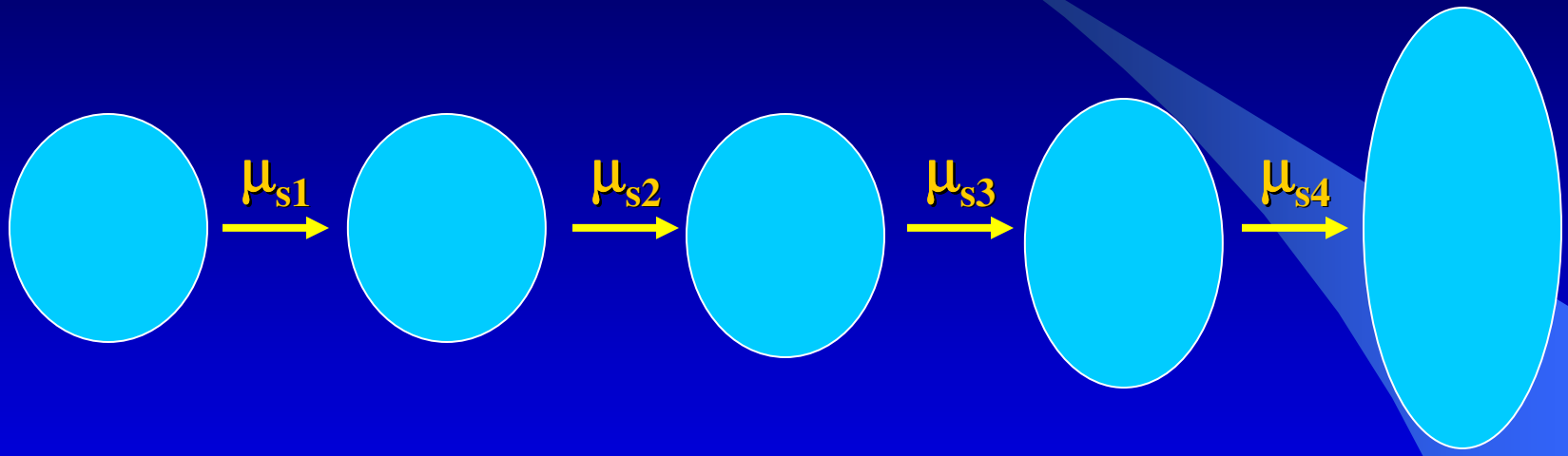
Developmental second hits and the concept of a mutation field

David Hogg
Department of Medicine
University of Toronto
Princess Margaret Hospital



Tumor progression – sequential mutations



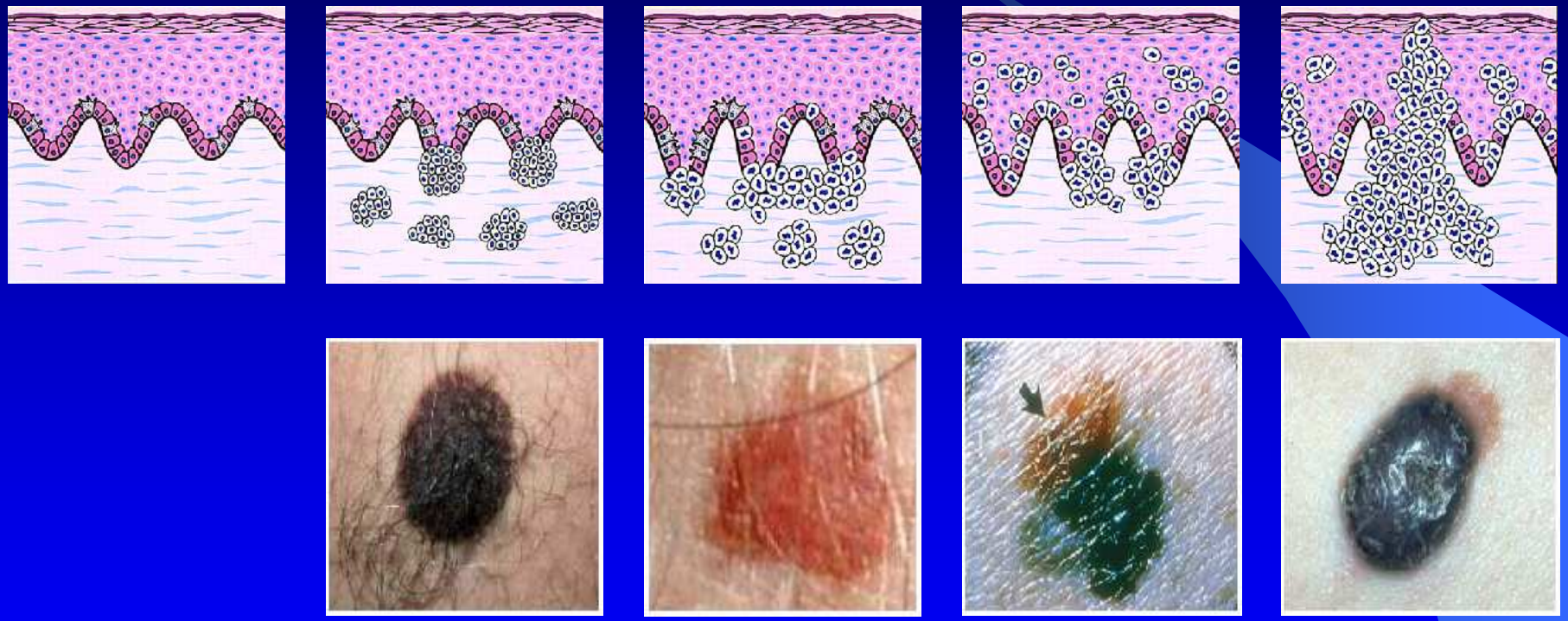
Timing of genetic changes

- There is not always a set pattern of genetic changes as a cell progresses to malignancy.
- In other words, although 5-7 changes must occur to make a cell fully malignant, the order and timing of these alteration may be flexible.

Implications of progressive genetic alterations

- Cancers are not cells with “uncontrolled growth”!
- Loss of growth control occurs very slowly over several years, in a stepwise fashion.
- Therefore, defining the genetic changes in a tumor will become as important as histological study.

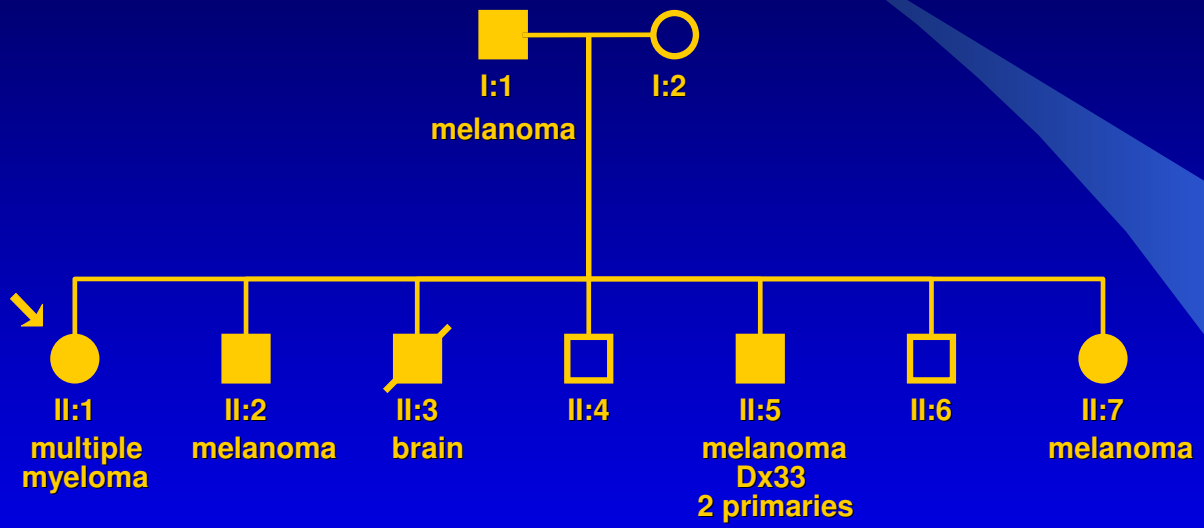
Melanoma: a multistage process



Normal Skin | Benign Nevus | Dysplastic Nevus (DN) | Radial Growth Phase melanoma | Vertical Growth Phase melanoma

Derived in part from Chin *et al*. (1998) Genes Dev. 12: 3467-3481

Melanoma family - example

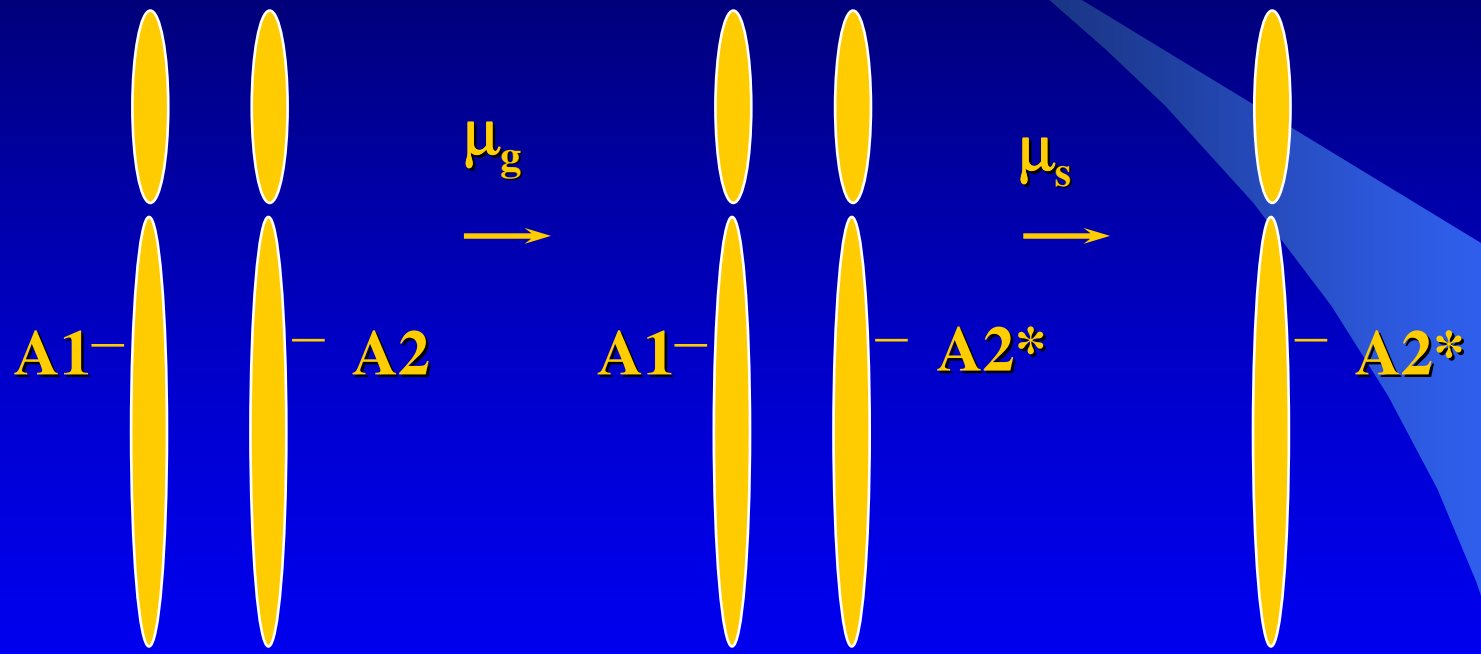


Familial Cancer - Characteristics

Compared to the corresponding sporadic cancers:

1. Age at first presentation is lower
2. Tumors may present at multiple sites in the same tissue
3. Tumors may occur in different tissues
4. There may be a family history

Loss of a functional tumor suppressor gene



Poisson distribution

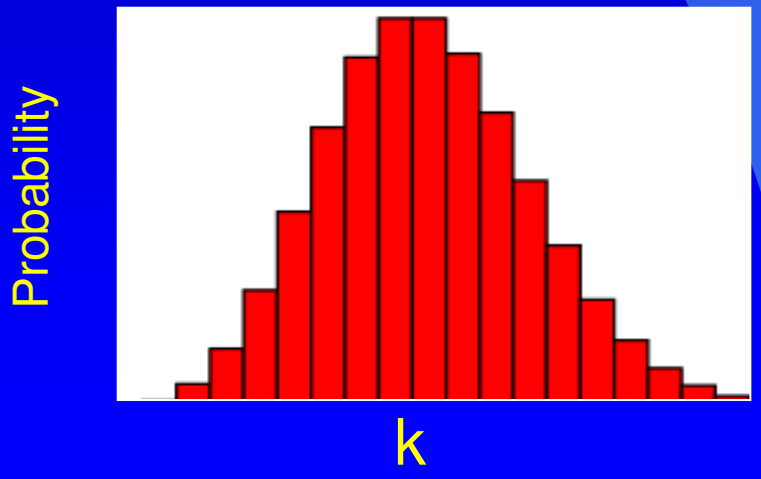
λ = Probability of an event

Probability of n events = $\frac{e^{-\lambda}\lambda^{x}}{n!}$

Familial Cancer Syndrome

1. Tumor development is a stochastic process
2. Variable number of primary lesions (k)
3. Modeled by a Poisson distribution

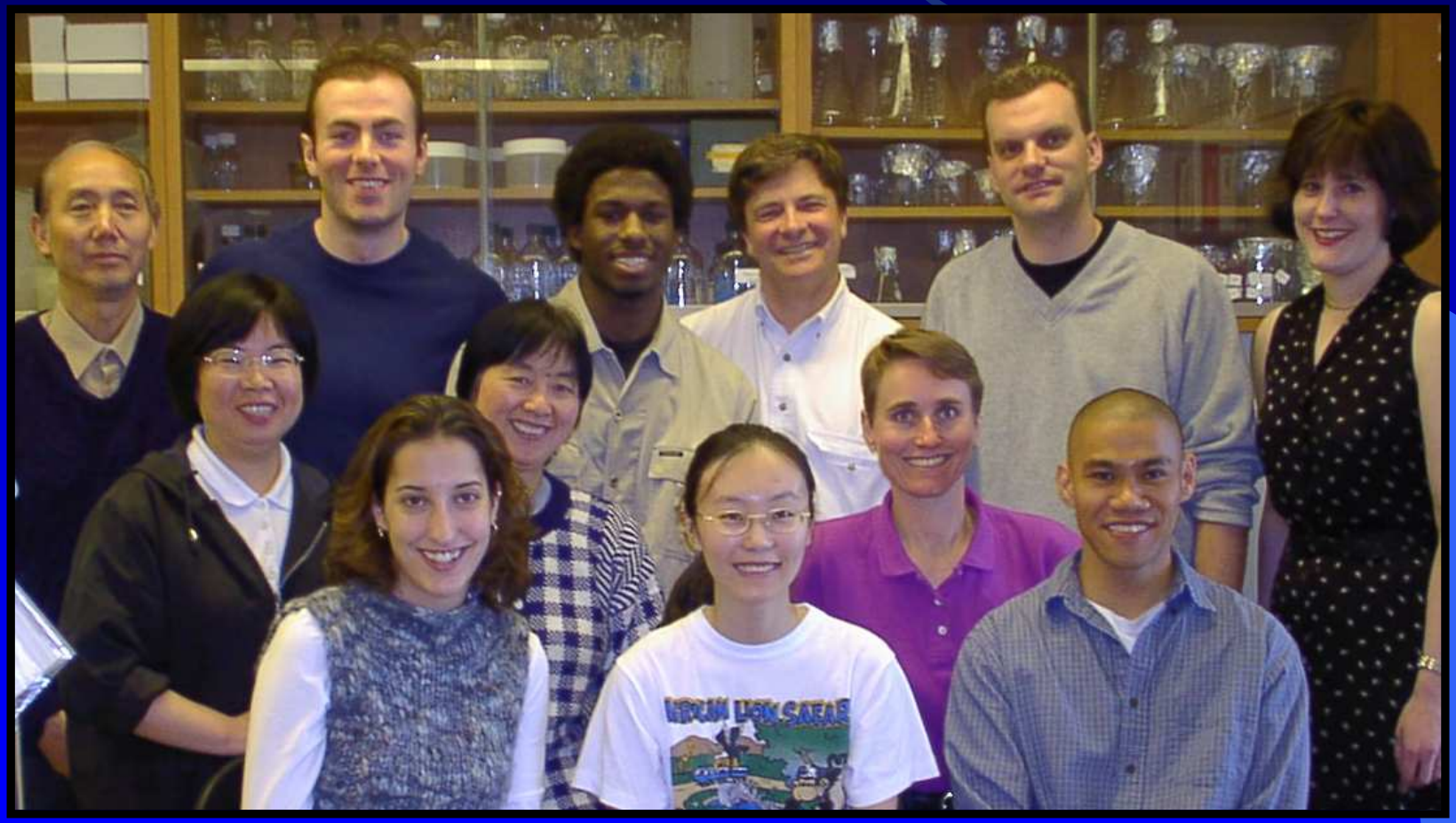
$$P(k) = \frac{e^{-\lambda}\lambda^{k}}{k!}$$



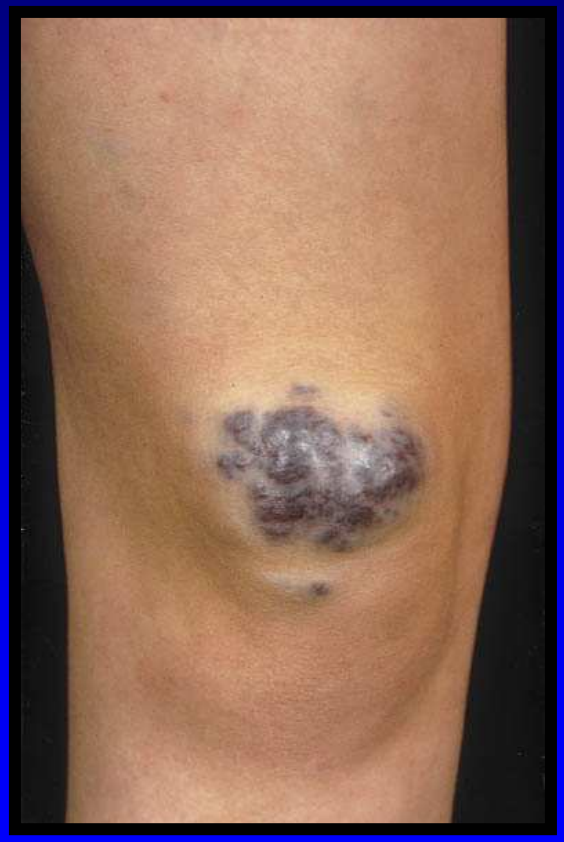
Multiple primary melanoma

- Between 2 and 8% of patients with melanoma will develop a second primary tumor
- But: the lifetime incidence of sporadic melanoma is only 1 in 80 (1.25%)
- Therefore, some factor must predispose the majority of multiple primary patients to additional tumors

Régine Mydlarski

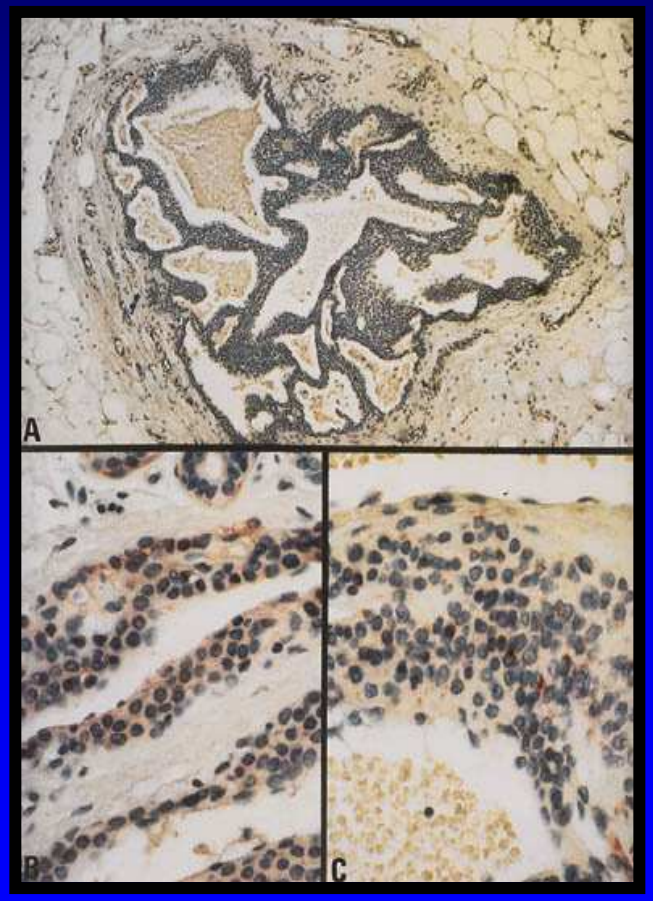


Glomangiomas



- Subtype of venous malformations
- Raised, tender, blue-red nodules
- Solitary (sporadic) or multiple (inherited) forms

Histology



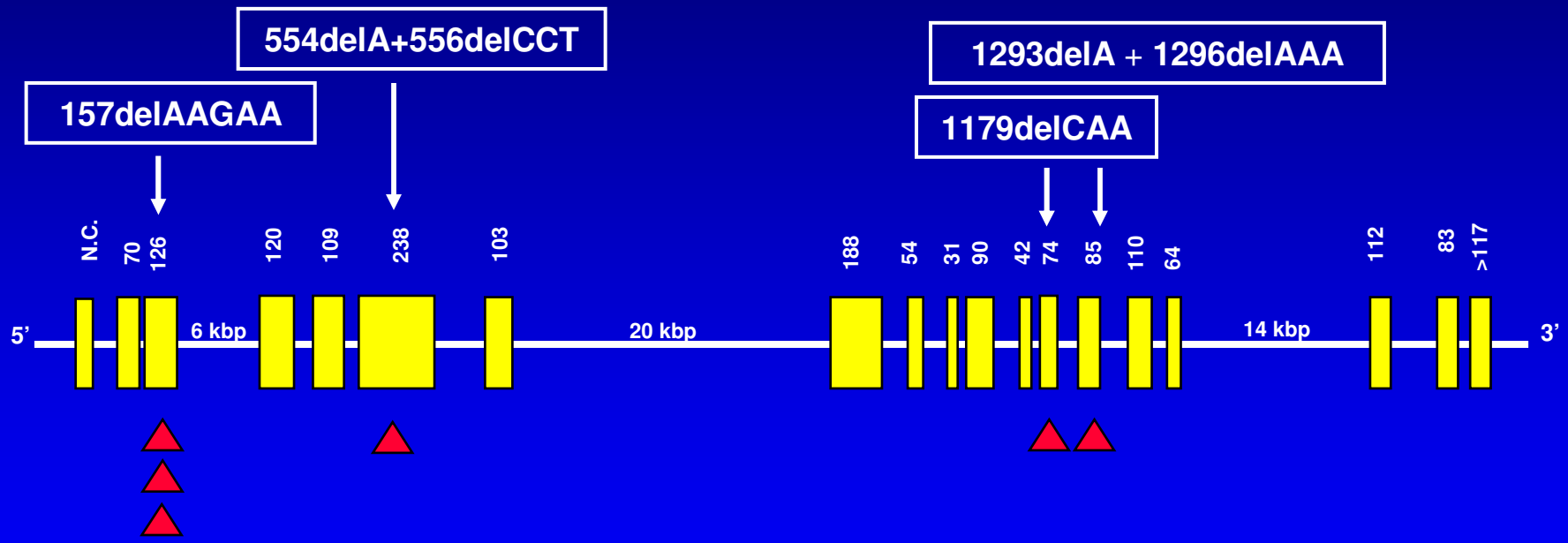
- Glomus cells
- Branching vascular channels
- Immuno
  - SMA α actin +
  - Vimentin +
  - Desmin -

Glomulin

(Brouillard *et al. Am J Hum Genet,* 2002)

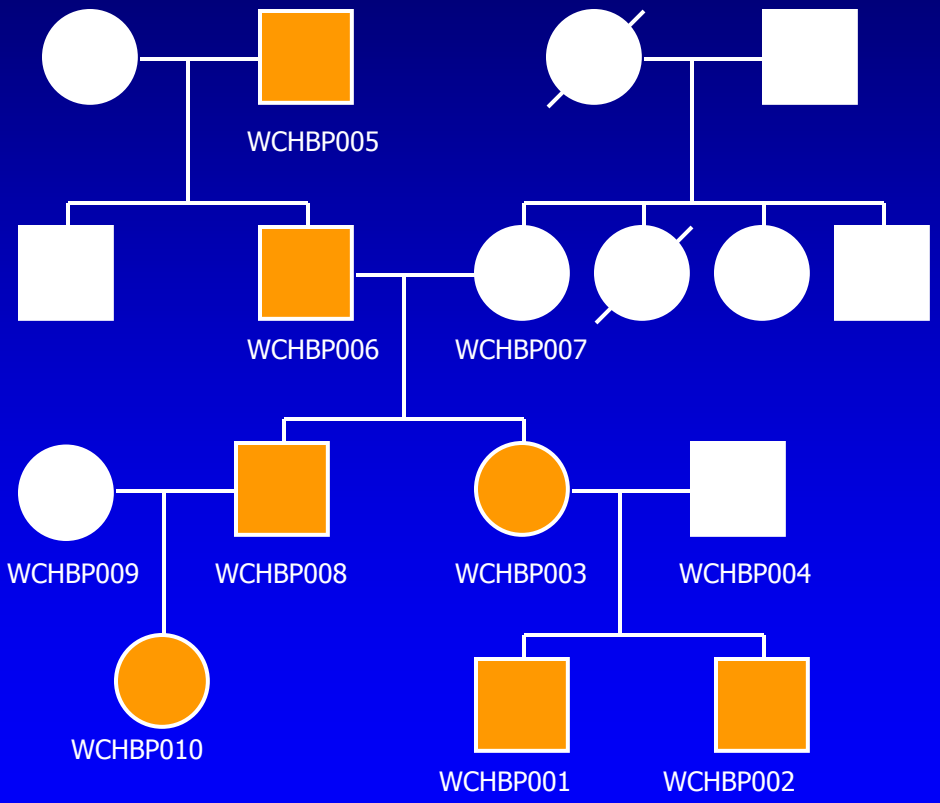
- Sequence identity to *FAP68*
- Located on chromosome *1p22*
- Function unknown
- ? Modulate signaling through TGFβ and/or HGF pathways

FAP68 (Glomulin)



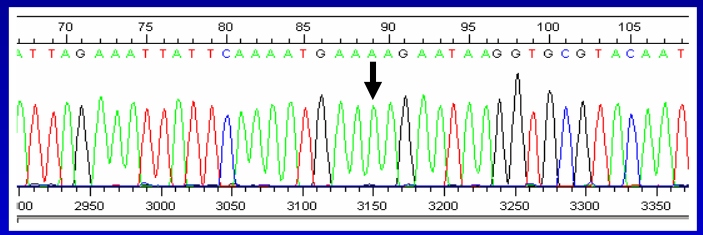
Mutational Screening

157 del AAGAA

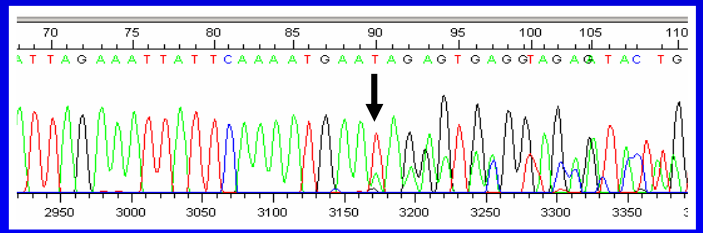


Multiple glomangiomas

Unaffected

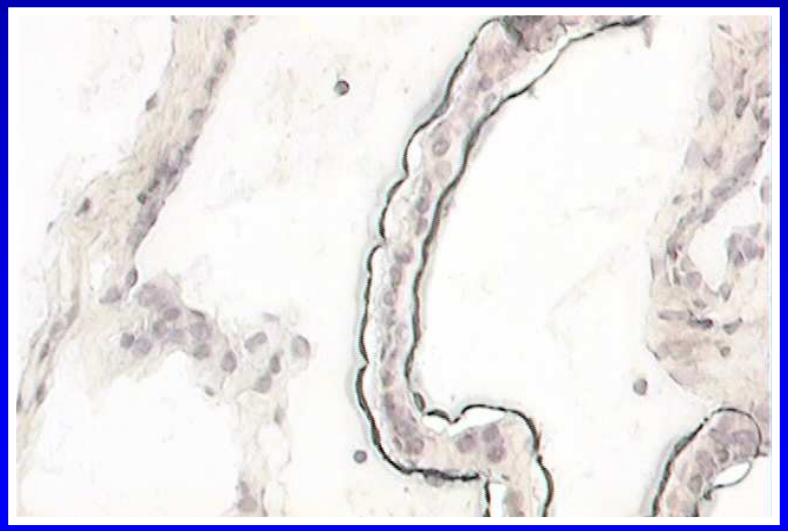


Affected

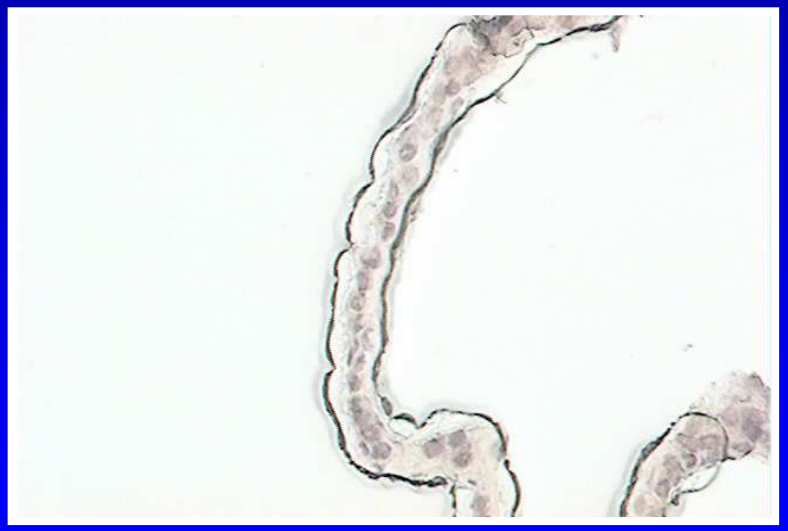


Laser Capture Microdissection

Pre-LCM



Post-LCM

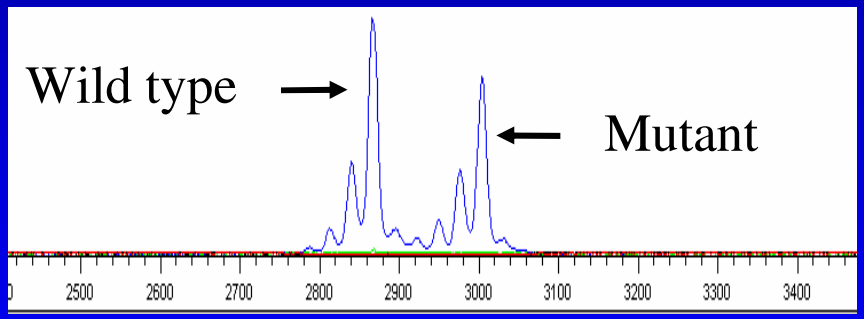


Loss of Heterozygosity

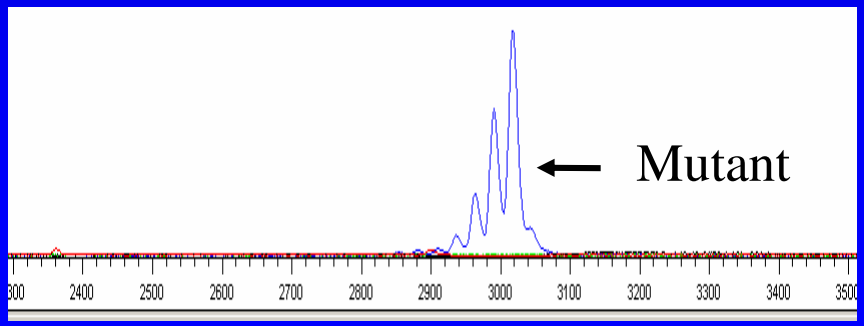
- Loss of the constitutional maternal or paternal allele of a gene

AKDE005 - D1S 2776

Genomic DNA



Tumor DNA

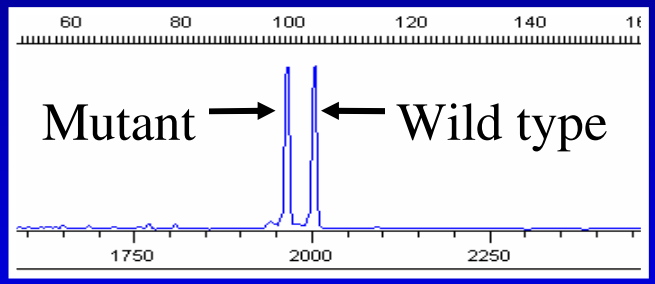


Two hits revisited...

Genomic DNA

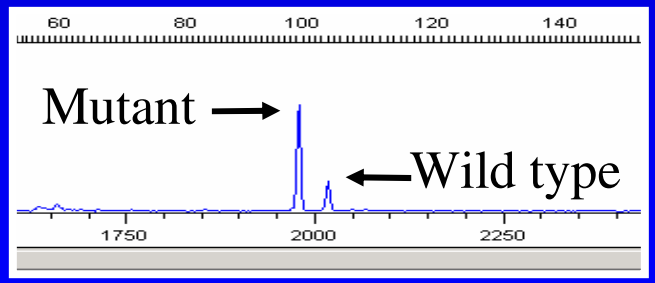
…AATGAAAAGAATAAGGT…

…AATGAA - - - - - TAAGGT…

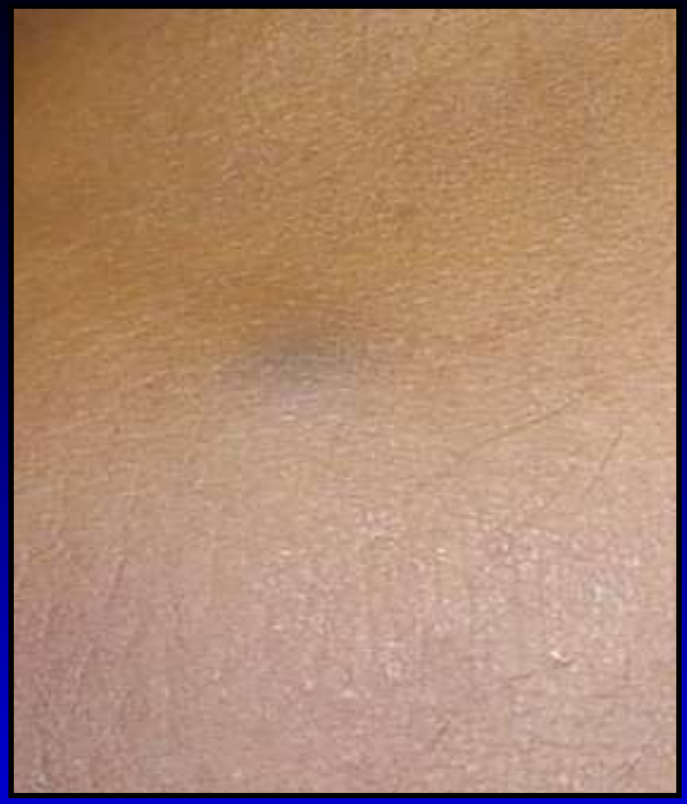


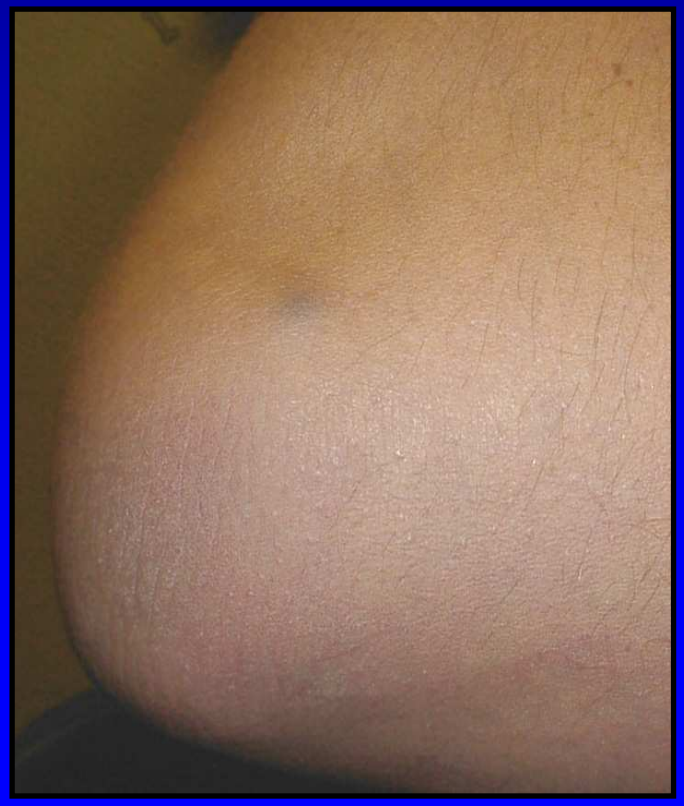
Tumor DNA

…AATGAA - - - - - TAAGGT…

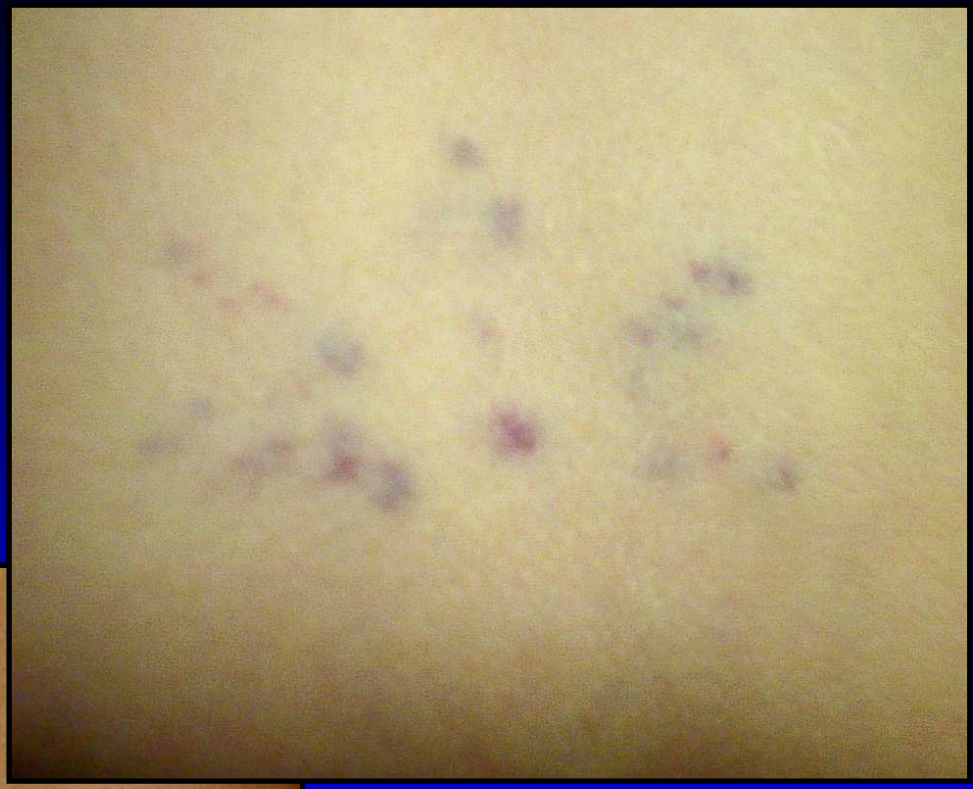


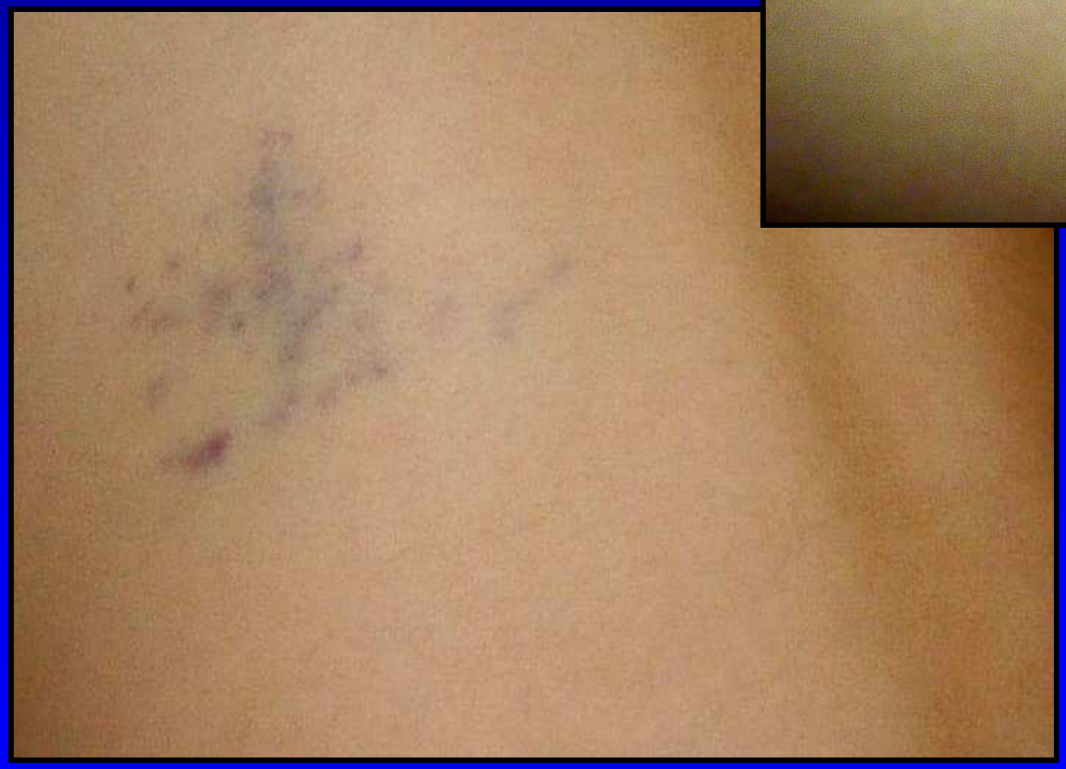
Solitary



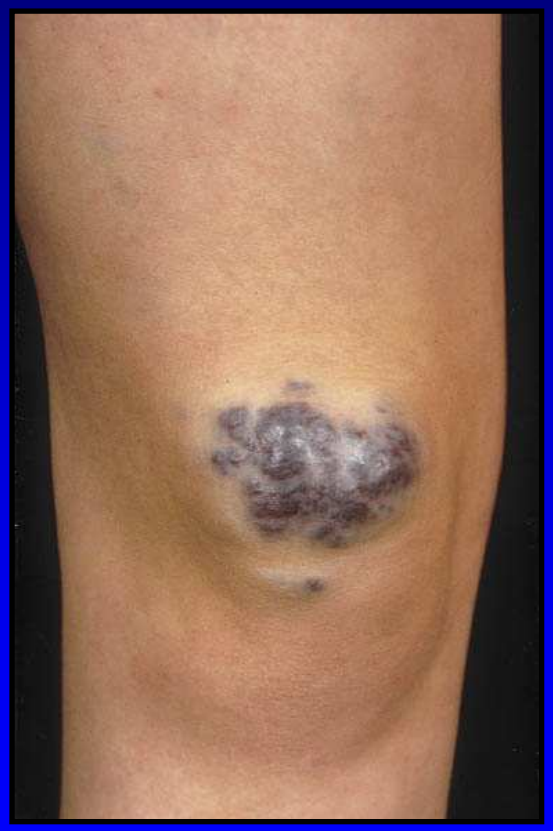


Cluster





Cluster/Segmental

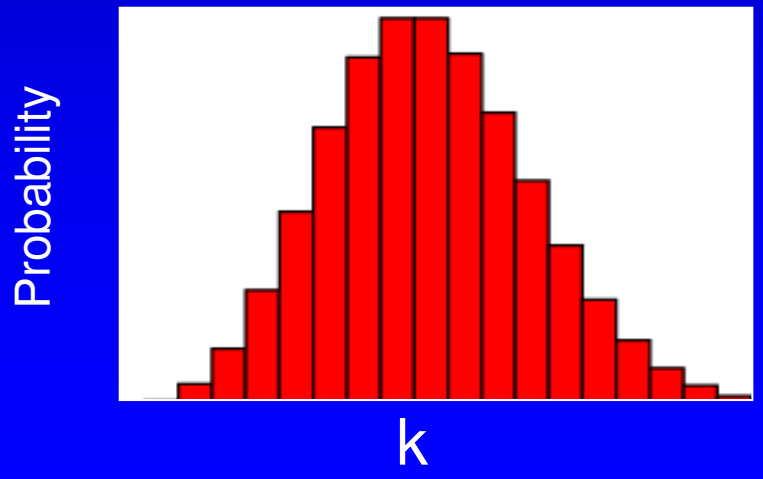


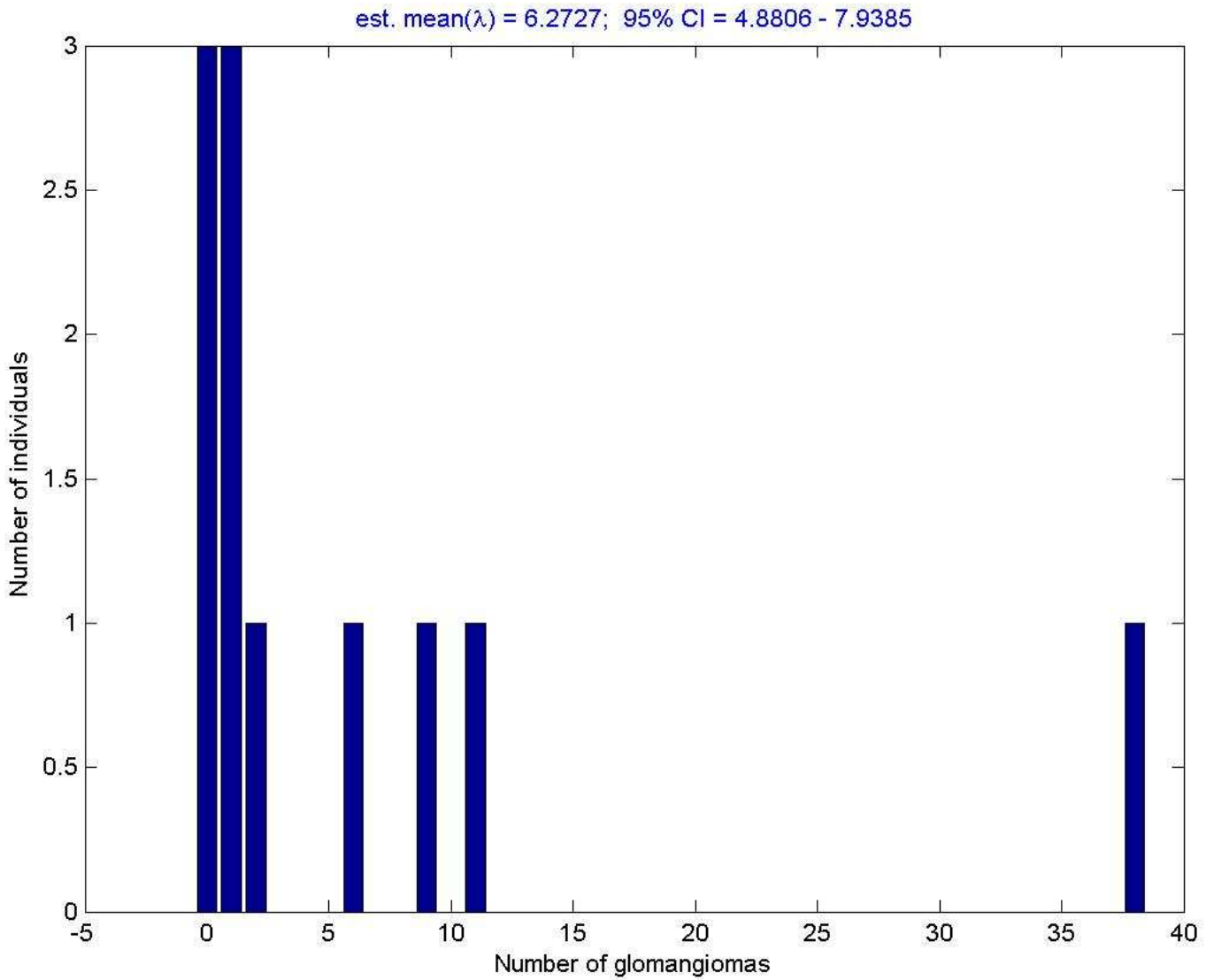
Segmental

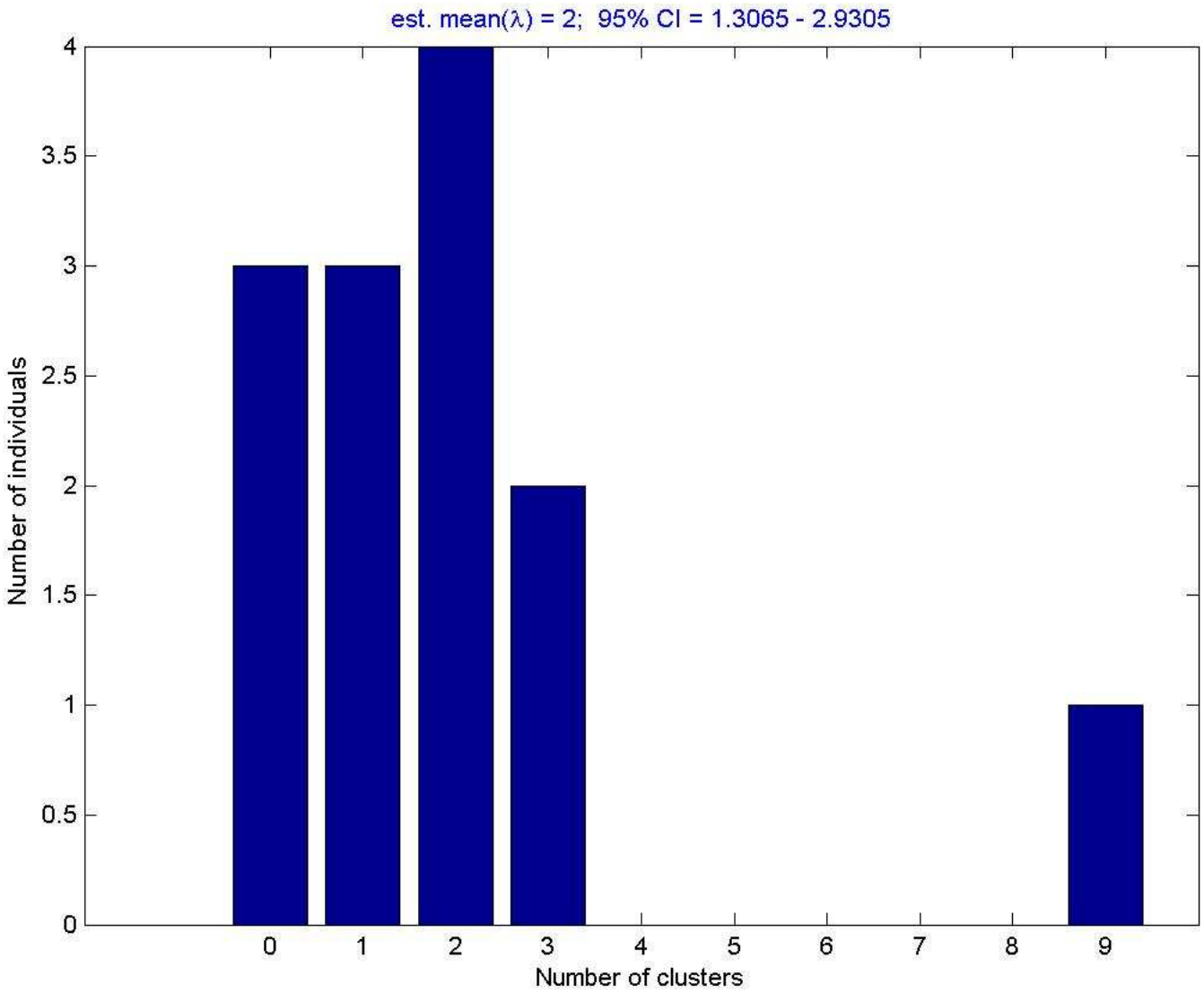
Familial Cancer Syndrome

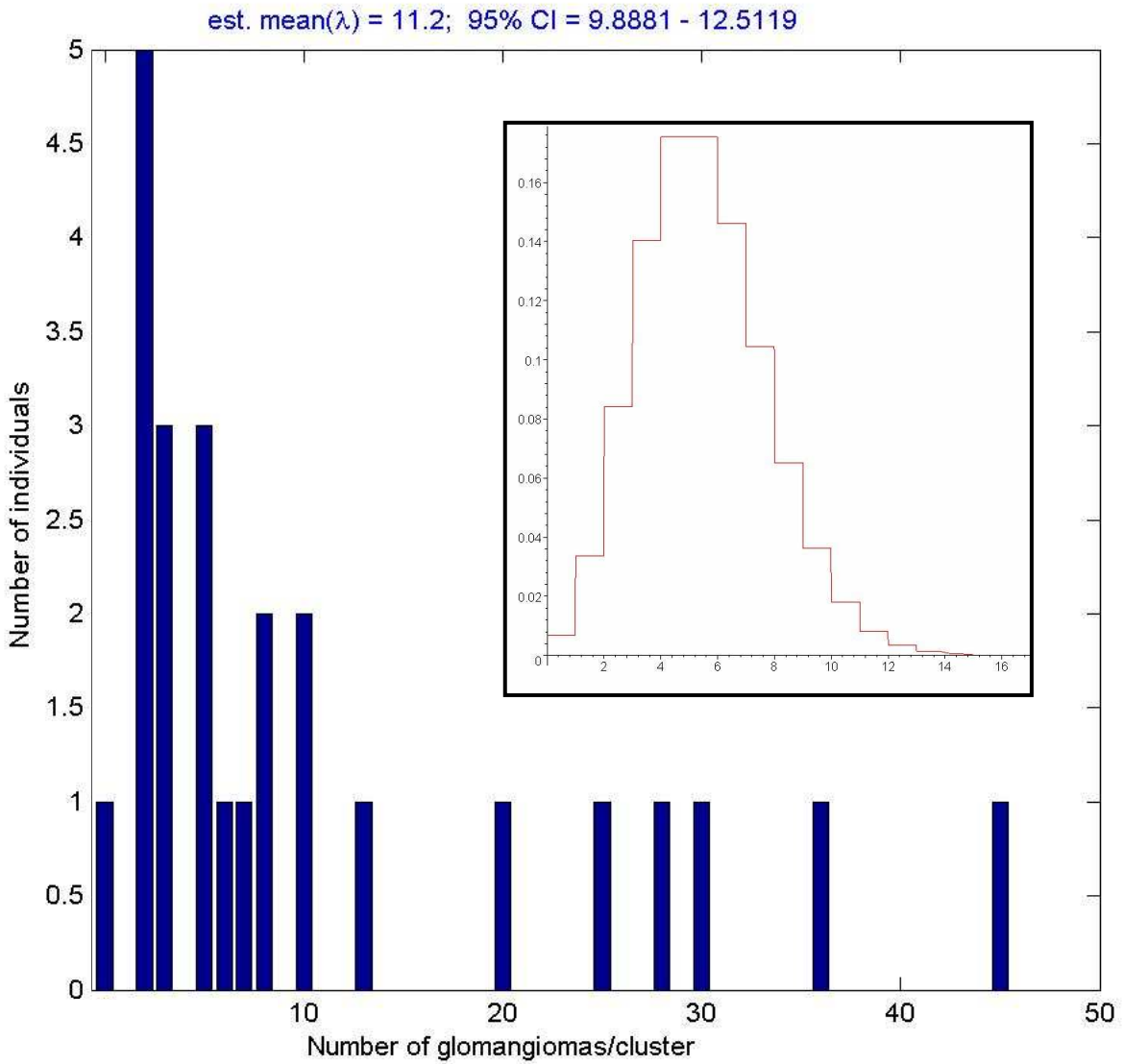
1. Tumor development is a stochastic process
2. Variable number of primary lesions
3. Modeled by a Poisson distribution

$$P(k) = \frac{e^{-\lambda}\lambda^{k}}{k!}$$

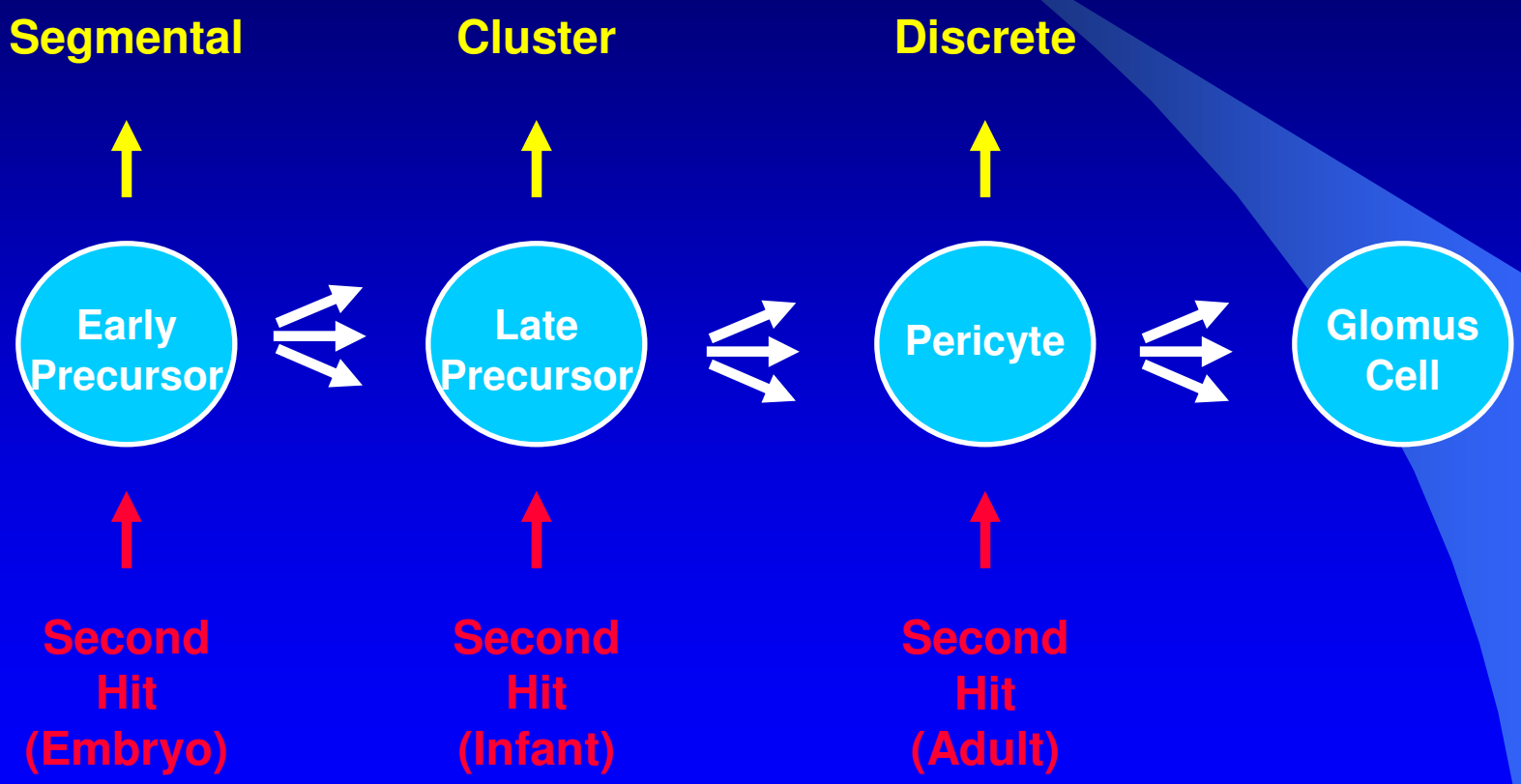



est. mean(λ) = 6.2727; 95% CI = 4.8806 - 7.9385
Number of individuals
Number of glomangiomas
0
0.5
1
1.5
2
2.5
3
-5
0
5
10
15
20
25
30
35
40


est. mean(λ) = 2; 95% CI = 1.3065 - 2.9305
Number of individuals
Number of clusters
0
0.5
1
1.5
2
2.5
3
3.5
4
0
1
2
3
4
5
6
7
8
9


est. mean(λ) = 11.2; 95% CI = 9.8881 - 12.5119
Number of individuals
Number of glomangiomas/cluster
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
10
20
30
40
50
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
2
4
6
8
10
12
14
16

Developmental Model



Goals of Mathematical Model

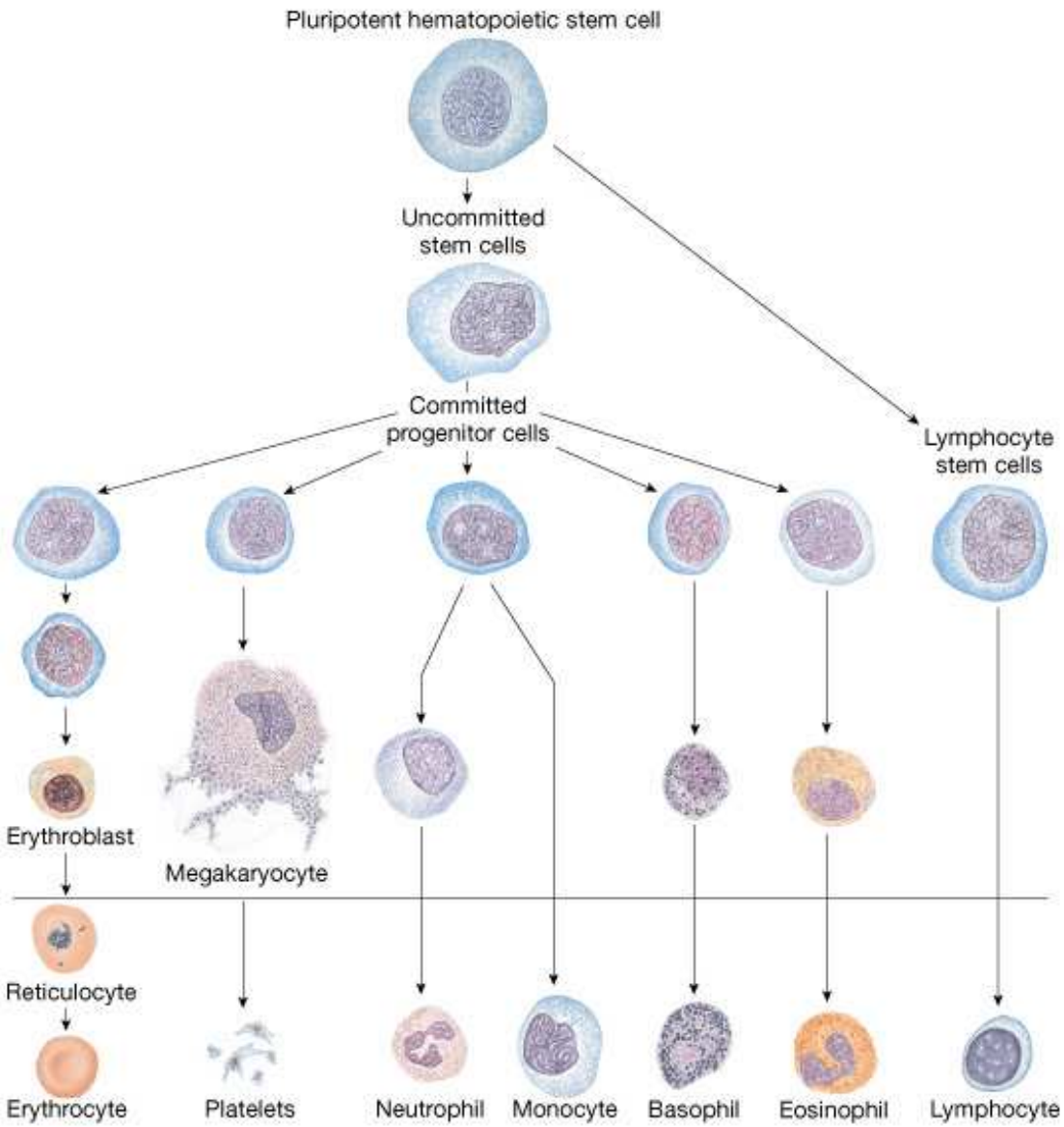
1. Explain morphological patterns of GVM
2. Account for differing frequencies of these patterns
3. Understand developmental dependence of GVM distribution

Assumption #1

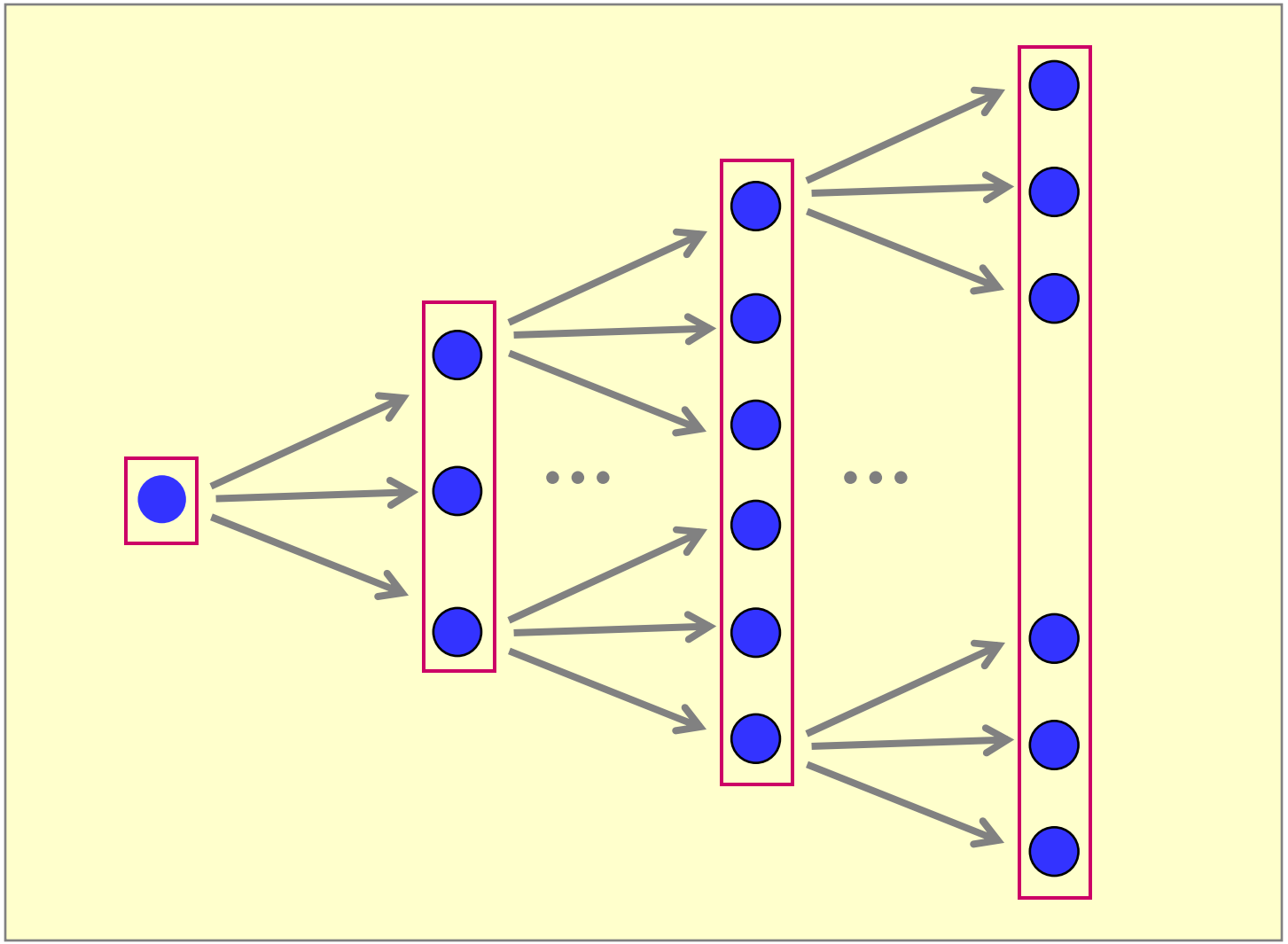
Mammalian organisms arise from a single fertilized ovum, and in the case of humans, multiply to ~10^{14} cells

Assumption #2

In the development of tissues and organs, an early stem cell is programmed to pass through a series of conceptual compartments containing progressively differentiated progenitors


Pluripotent hematopoietic stem cell
Uncommitted stem cells
Committed progenitor cells
Lymphocyte stem cells
Erythroblast
Megakaryocyte
Reticulocyte
Erythrocyte
Platelets
Neutrophil
Monocyte
Basophil
Eosinophil
Lymphocyte

Branching Hypothesis

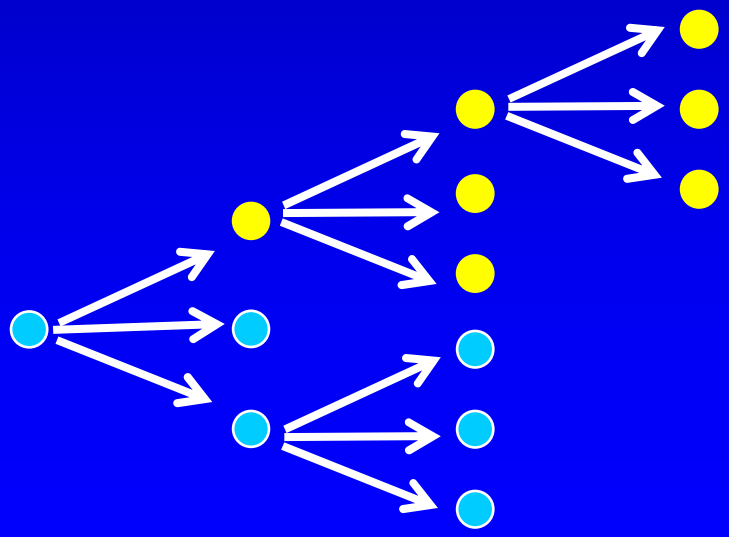


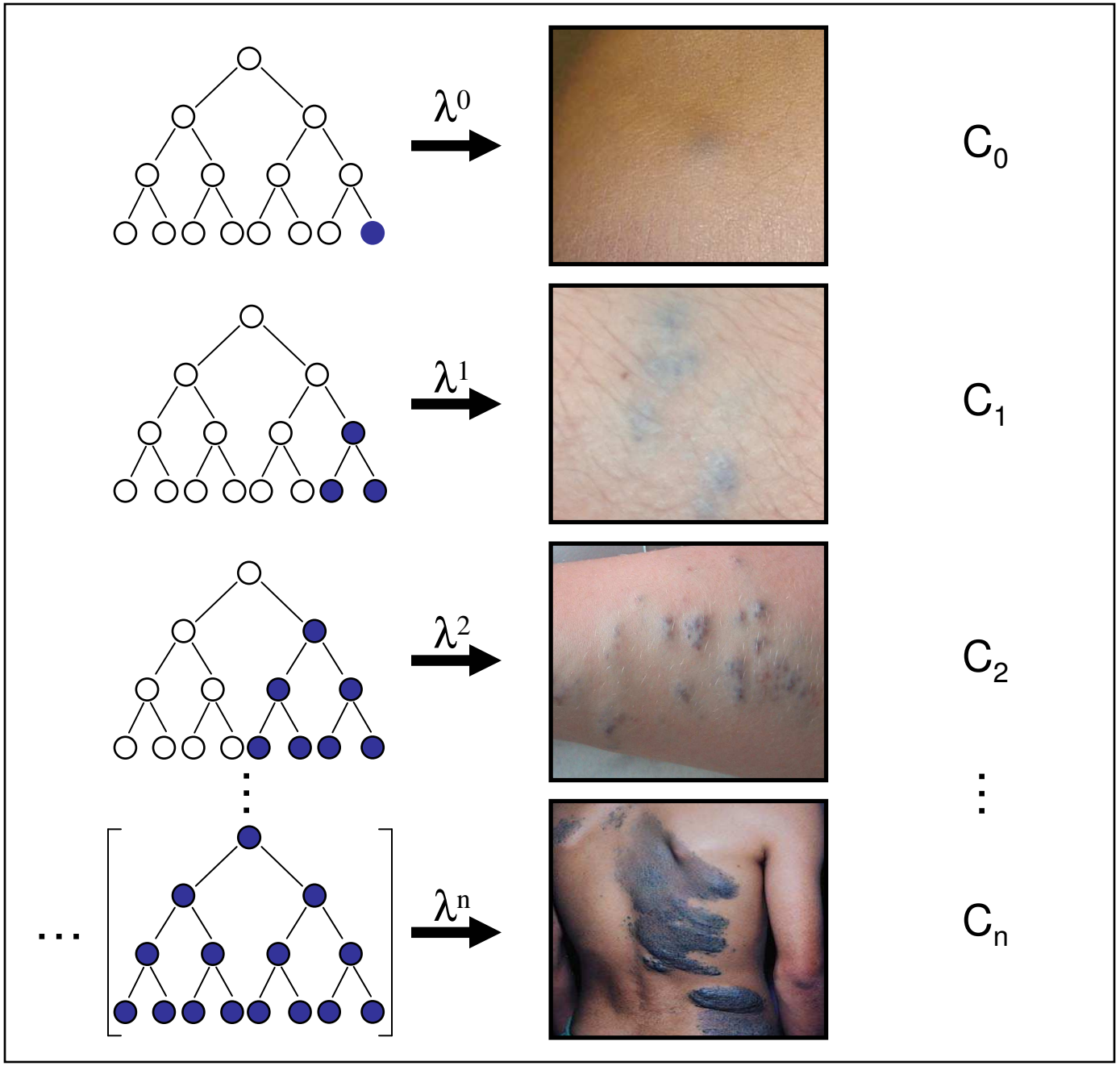
Assumption #3

The numbers of immediate progeny from parental cells are independent, identically distributed Poisson variables

Branching Hypothesis

- If a gene is mutated early in development, it will exist in its mutated form in all subsequent progeny
- As the number of progeny may be very large, the overall probability of a cell possessing this mutation is increased




λ0
λ1
λ2
λn
C0
C1
C2
Cn

Poisson distribution:

$$P(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

To derive a branching Poisson distribution, we employ the generating function:

$$G(s) = E(s^k)$$

$$G_n(s) = G\left(G\left(\ldots\left(G(s)\right)\ldots\right)\right)$$

The generating function for the Poisson distribution is:

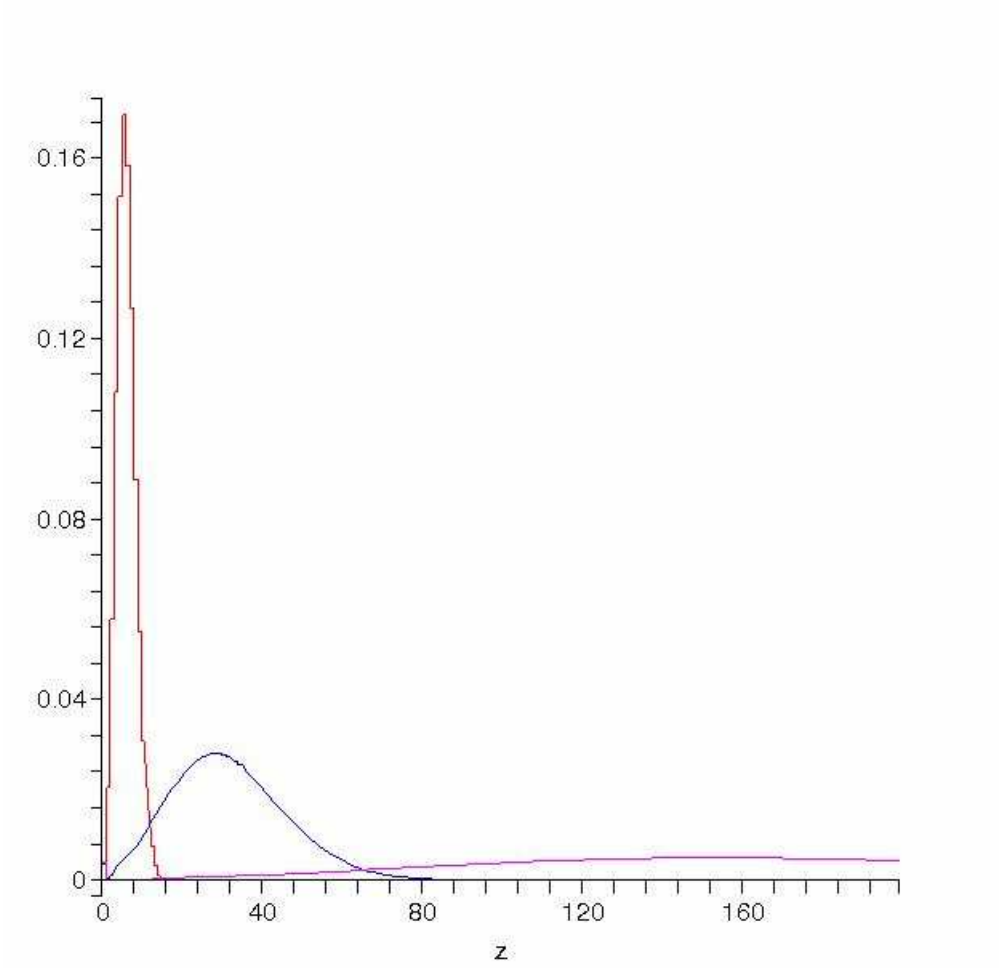
$$\begin{aligned} G(s) &= \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} s^k \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} \\ &= e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)} \end{aligned}$$

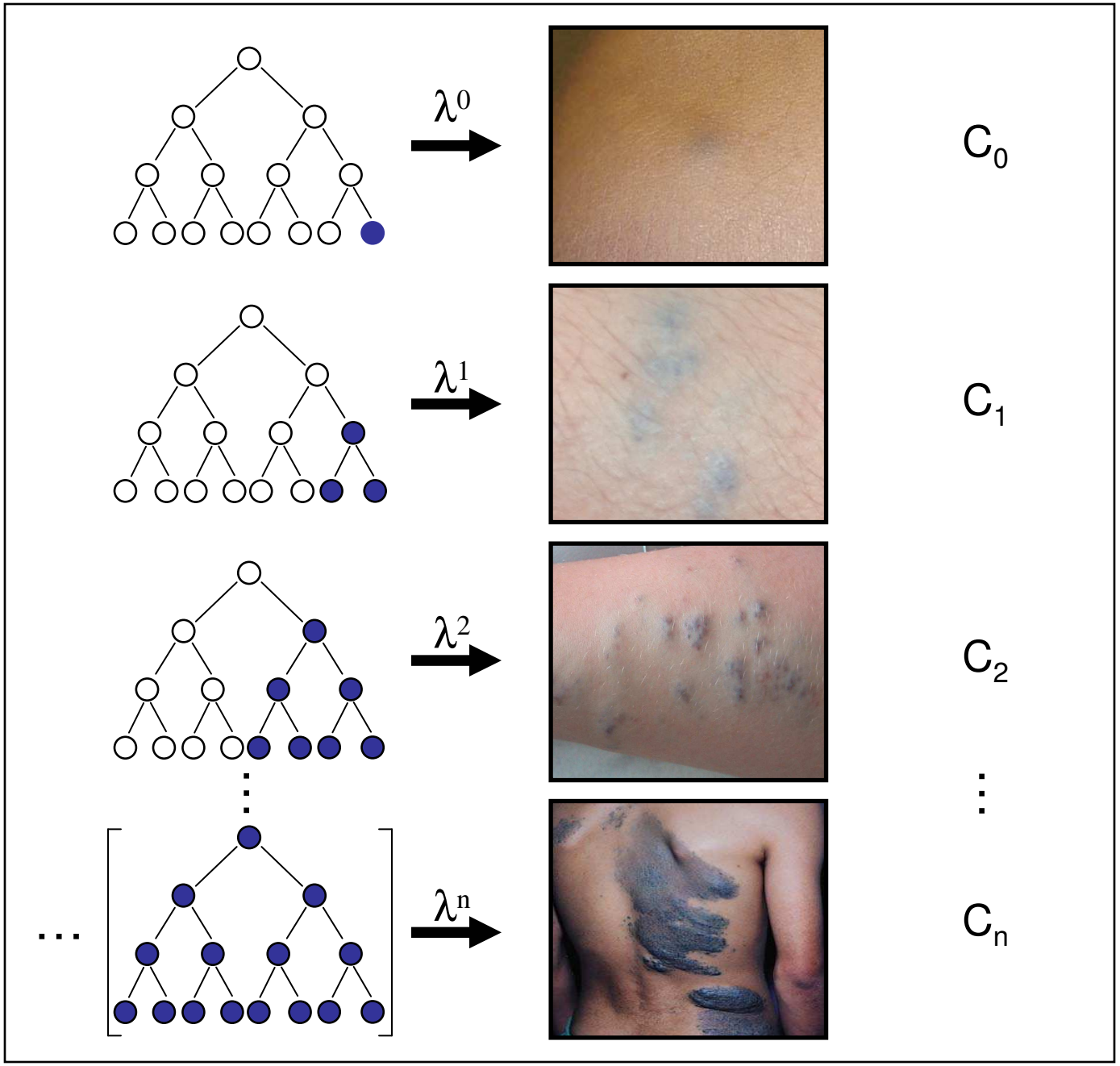
The generating function for a 2-level Poisson branching process is:

$$G_2(s) = e^{\lambda(e^{\lambda(s-1)}-1)}$$

To recover the probability of observing k progeny at n branches :

$$P_n(k) = \frac{\partial^k G_n(s)}{\partial s^k} \frac{1}{k!}\bigg|_{s=0}$$


0.16
0.12
0.08
0.04
0
0
40
80
120
160
z


λ0
λ1
λ2
λn
C0
C1
C2
Cn

Let μ be the probability of a mutation.
To scale the probability of a mutation
Occurring in each compartment,
we multiply by 1/λ:

$$P(n) = \frac{\mu}{\lambda^n}$$

We ignore μ; however, we must
Scale the sum of all P(n), since
it is > 1:

$$\sum_{n=0}^{\infty} \frac{1}{\lambda^n} = \frac{\lambda}{\lambda - 1}$$

The probability of mutation arising
in compartment n is given by:

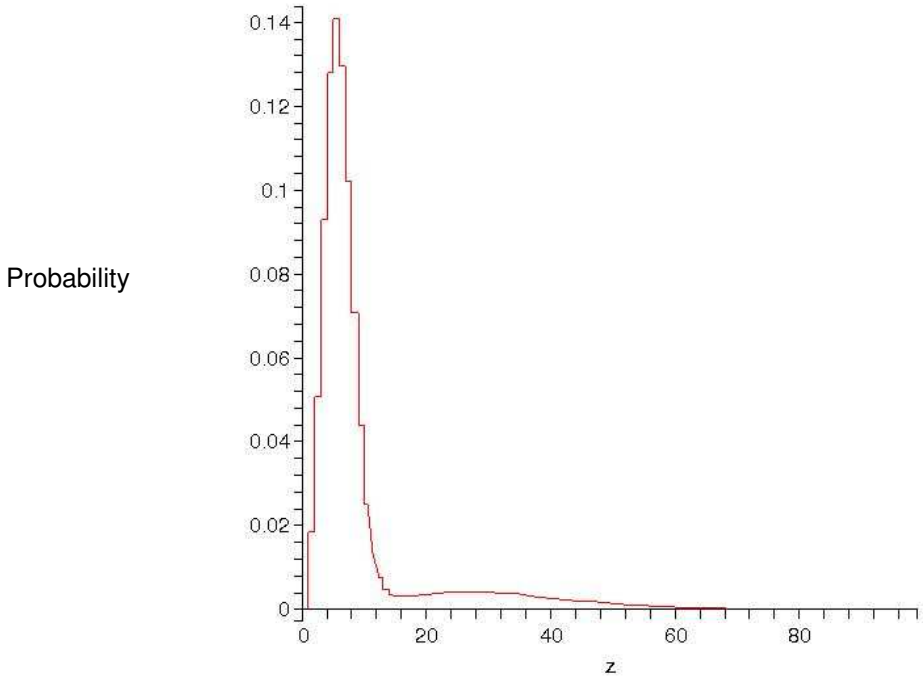
$$P_c(n) = \frac{(\lambda - 1)}{\lambda^{n+1}}$$

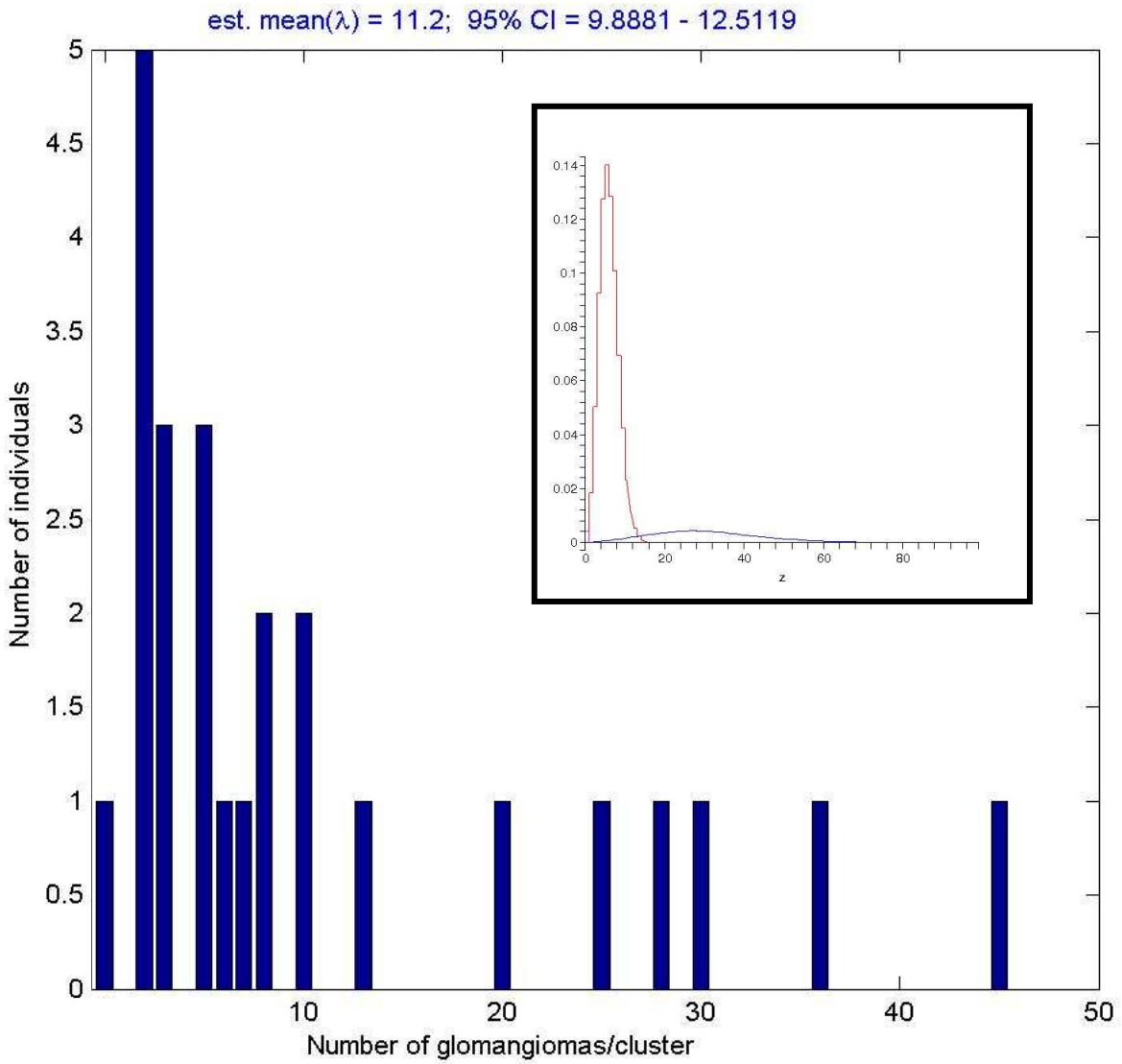
Therefore, the pmf of the combined
Poisson branching process over all
compartments becomes:

$$P(k) = \sum_{n=1}^{\infty} P_c(n) P_n(k)$$

The final curves that we construct are derived from:

$$P(k) \propto \sum_{n=1}^{3} P_c(n) P_n(k)$$




est. mean(λ) = 11.2; 95% CI = 9.8881 - 12.5119
Number of individuals
Number of glomangiomas/cluster
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
10
20
30
40
50
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
20
40
60
80
z

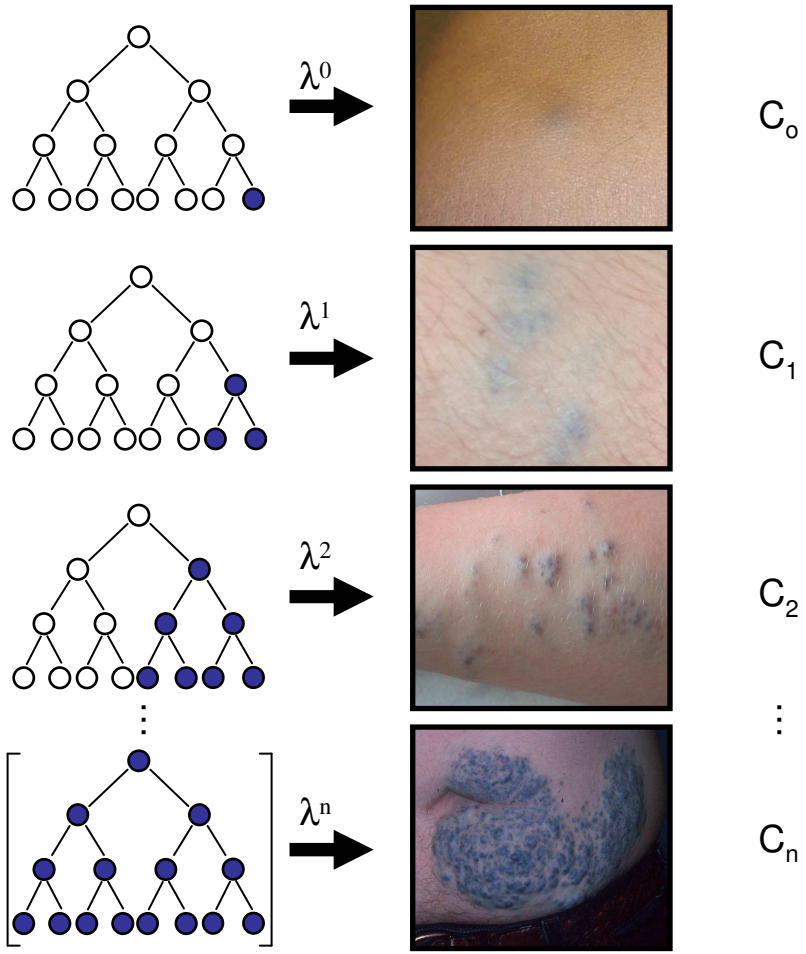
Maximum Likelihood Estimate

- The likelihood function, L, is proportional to the probability of observing the data set, θ, given the pdf, y

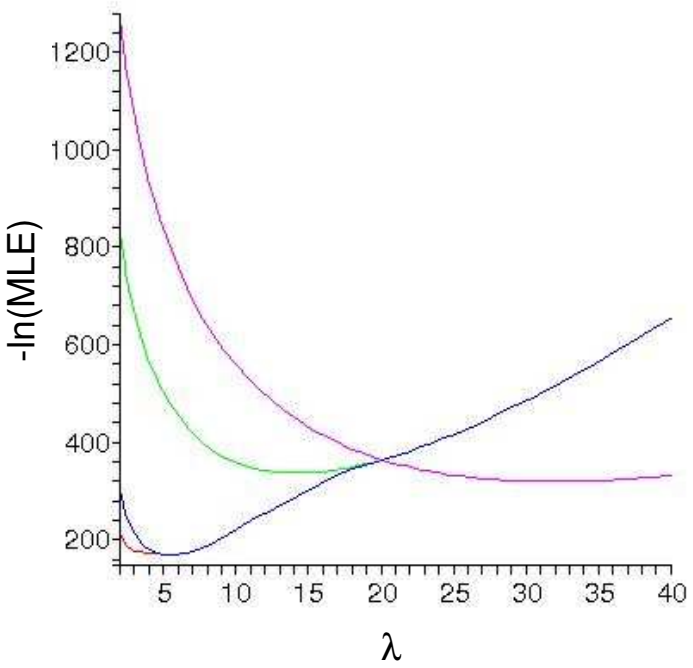
$$L(\theta|y)$$

- An MLE requires:
  - A data set – the number of GVM/cluster
  - A probability distribution – simple or branching Pd

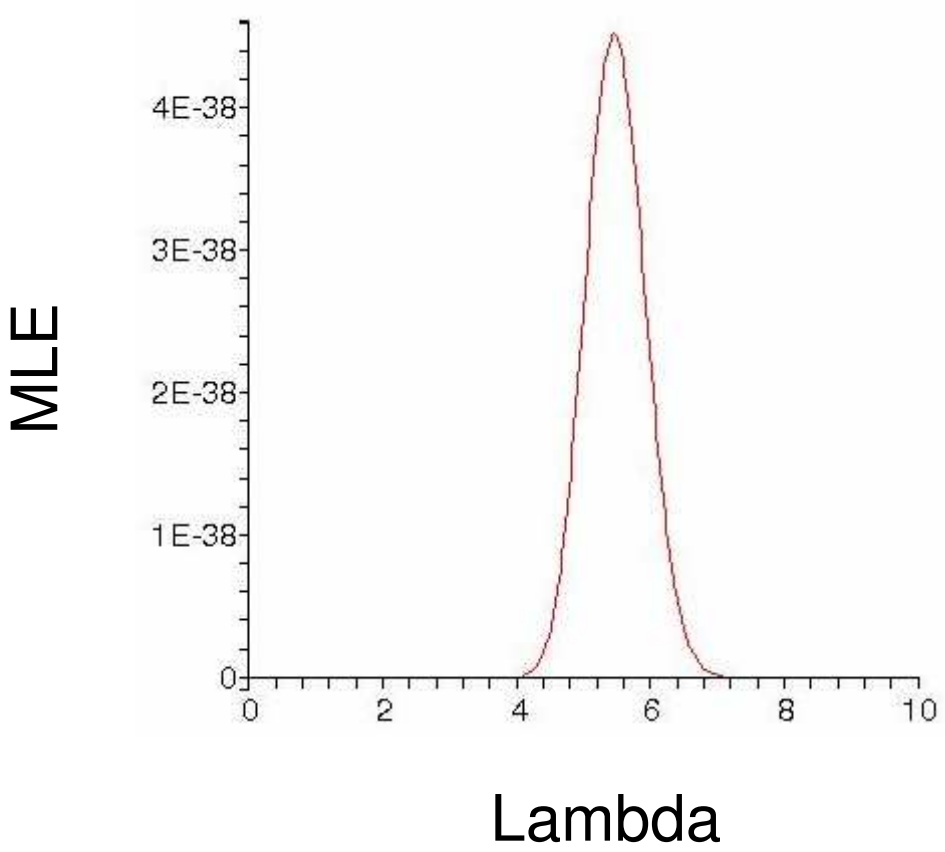
a


λ0
λ1
λ2
λn
Co
C1
C2
Cn

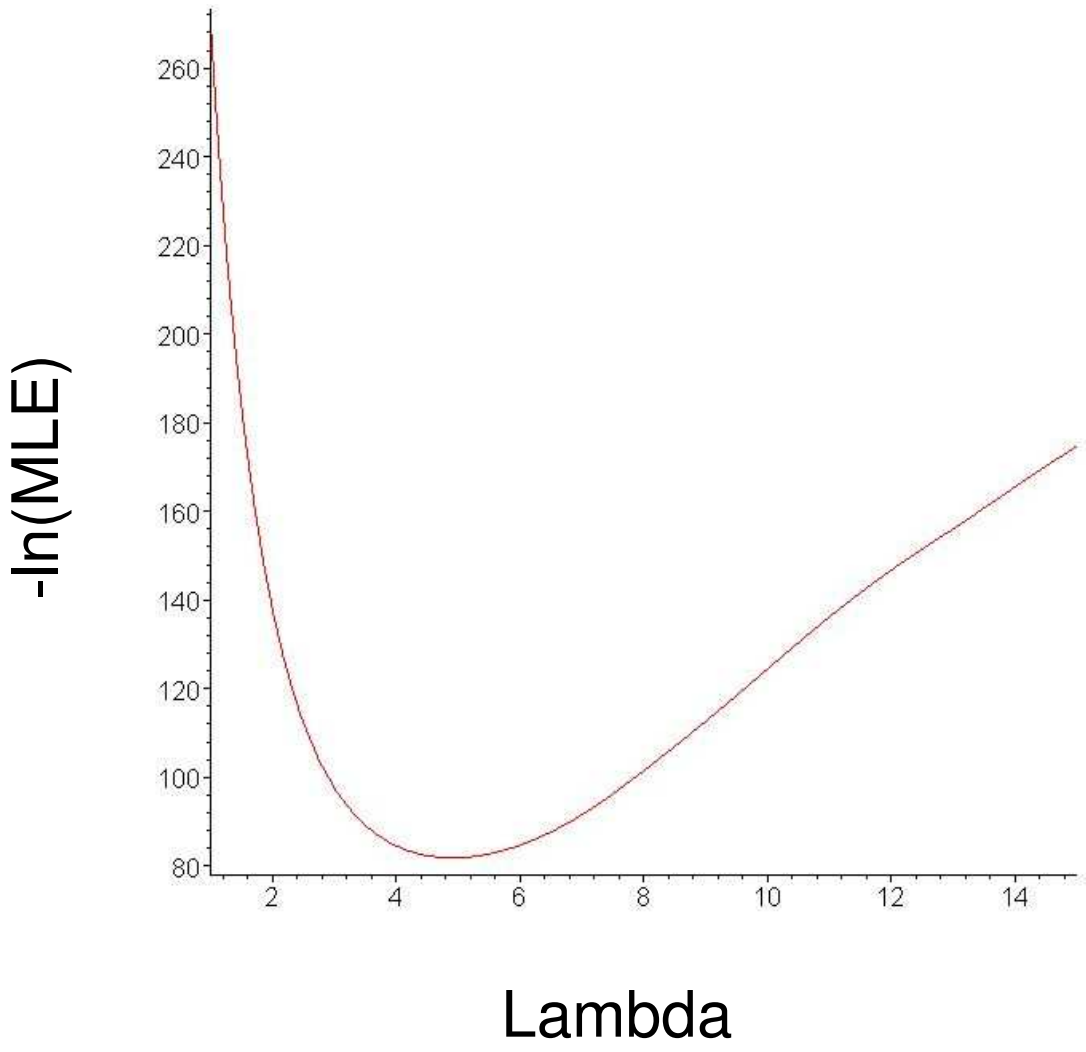
b

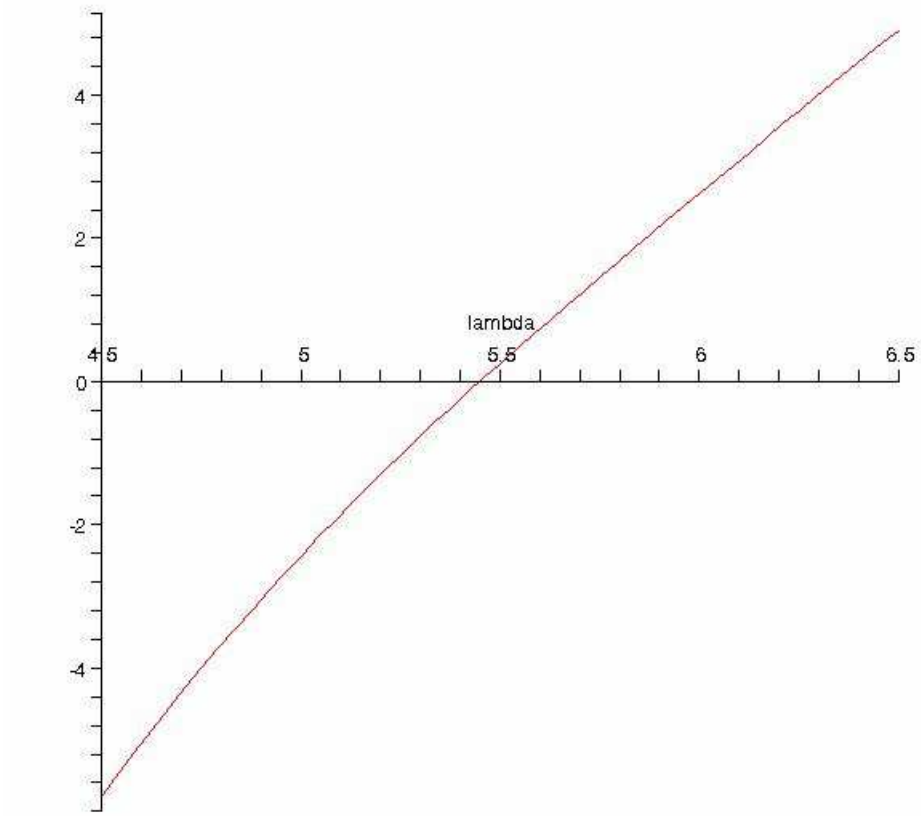

-ln(MLE)
1200
1000
800
600
400
200
5
10
15
20
25
30
35
40
λ

Maximum Likelihood Estimate

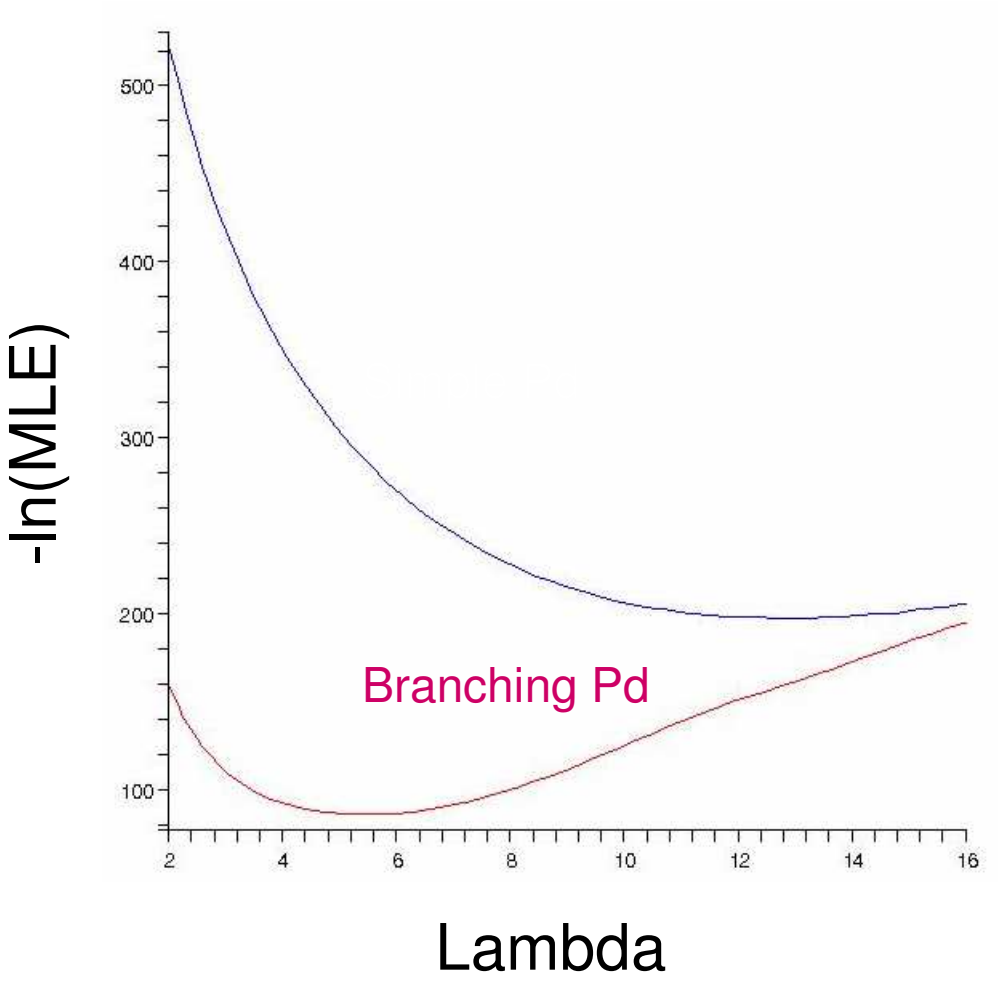


-ln(MLE)


-ln(MLE)
260
240
220
200
180
160
140
120
100
80
2
4
6
8
10
12
14
Lambda

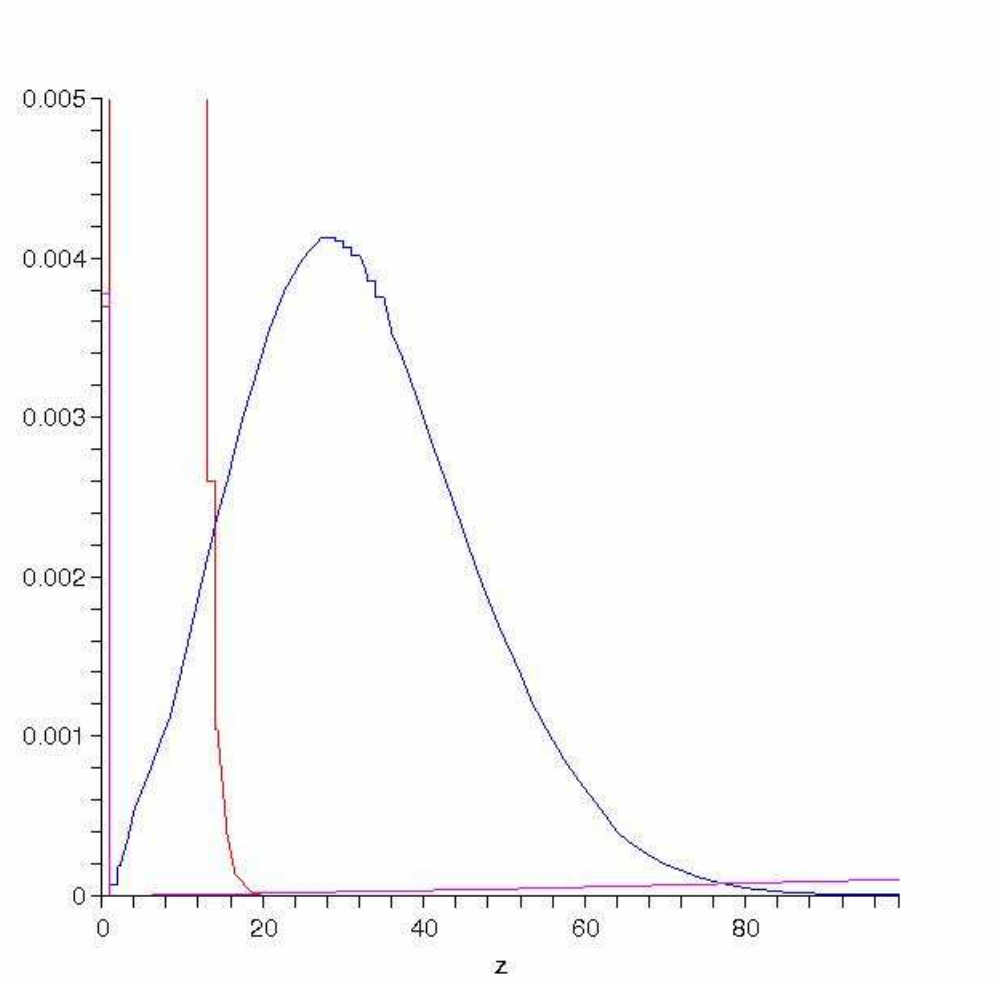

4
2
0
-2
-4
4.5
5
5.5
6
6.5
lambda

Maximum Likelihood Estimation

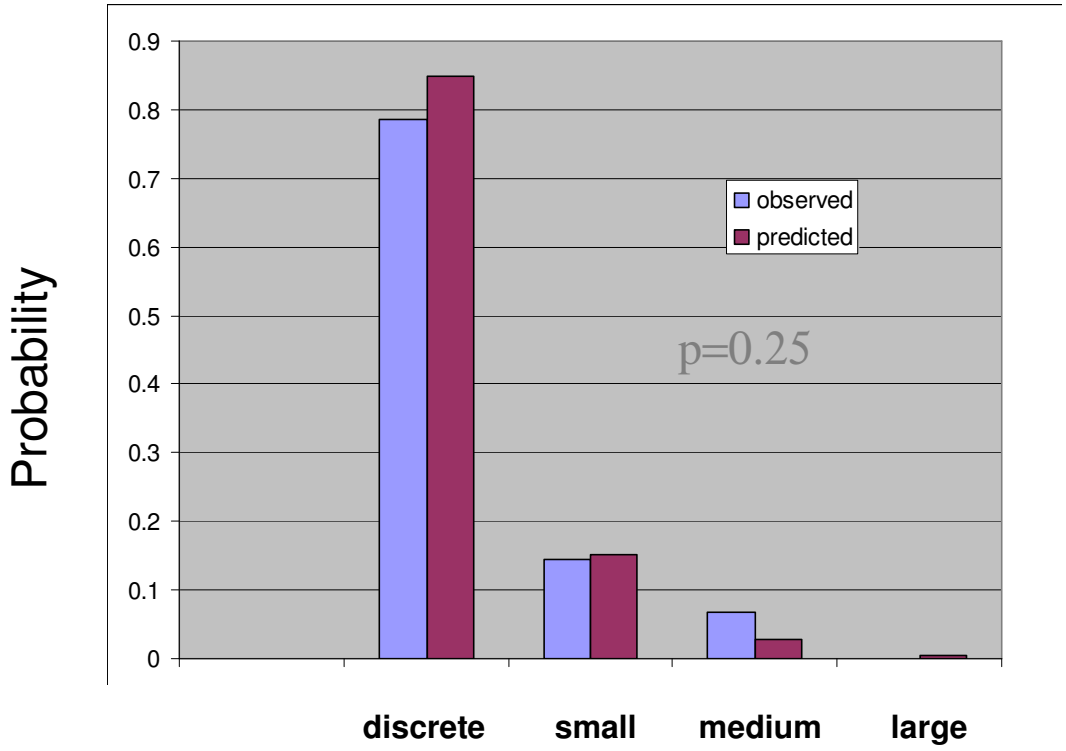


Model Predictions

1. Estimates the relative numbers of lesion types in our data set
2. Predicts that all lesions within a cluster should bear an identical second hit
3. Explain the morphology of a cluster
4. Estimate the mutation frequency of wt *FAP68* allele


0.005
0.004
0.003
0.002
0.001
0
0
20
40
60
80
z

Observed and Predicted Patterns of Hereditary GVMs



Model Predictions

1. Estimate the relative numbers of lesion types in our data set
2. All lesions within a cluster should bear an identical second hit
3. Explains morphology of a cluster
4. Estimate mutation frequency of wt *FAP68* allele

Clonality

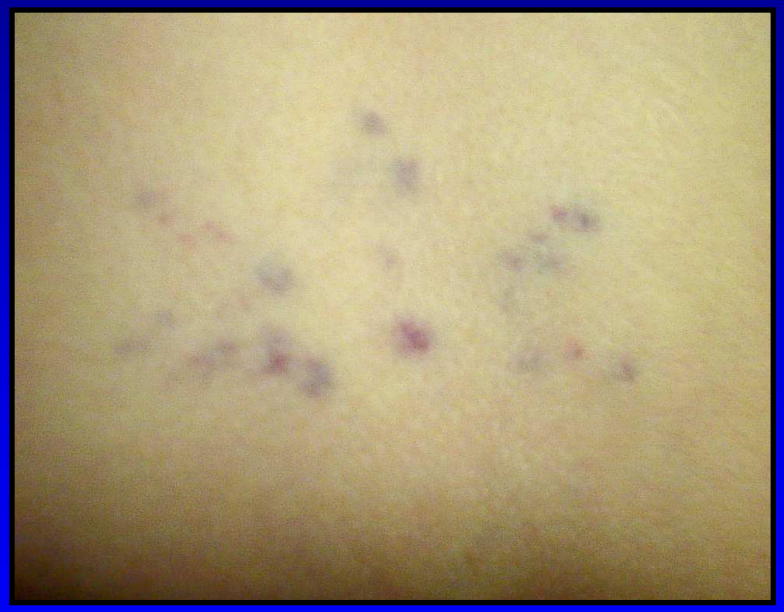
- Cluster containing 6 GVM biopsied from affected female patient
- Nonrandom XCI in GVM and NHK
- LOH absent in all 6 GVM
- No coding mutations (sequenced 19 exons)
- Methylation studies pending

Model Predictions

1. Estimate the relative numbers of lesion types in our data set
2. All lesions within a cluster should bear an identical second hit
3. Explains morphology of a cluster
4. Estimate mutation frequency of wt *FAP68* allele

Morphology of a Cluster

- Each GVM arose from the proliferation of a single GC
- 20 – 90 capillary loops/mm^2
- Small cluster size: 10 cm^2 (~$5X10^4$ cap loops)
- On average, $1/10^4$ cells bear a second hit



Model Predictions

1. Estimate the relative numbers of lesion types in our data set
2. All lesions within a cluster should bear an identical second hit
3. Explains morphology of a cluster
4. Estimate mutation frequency of wt *FAP68* allele

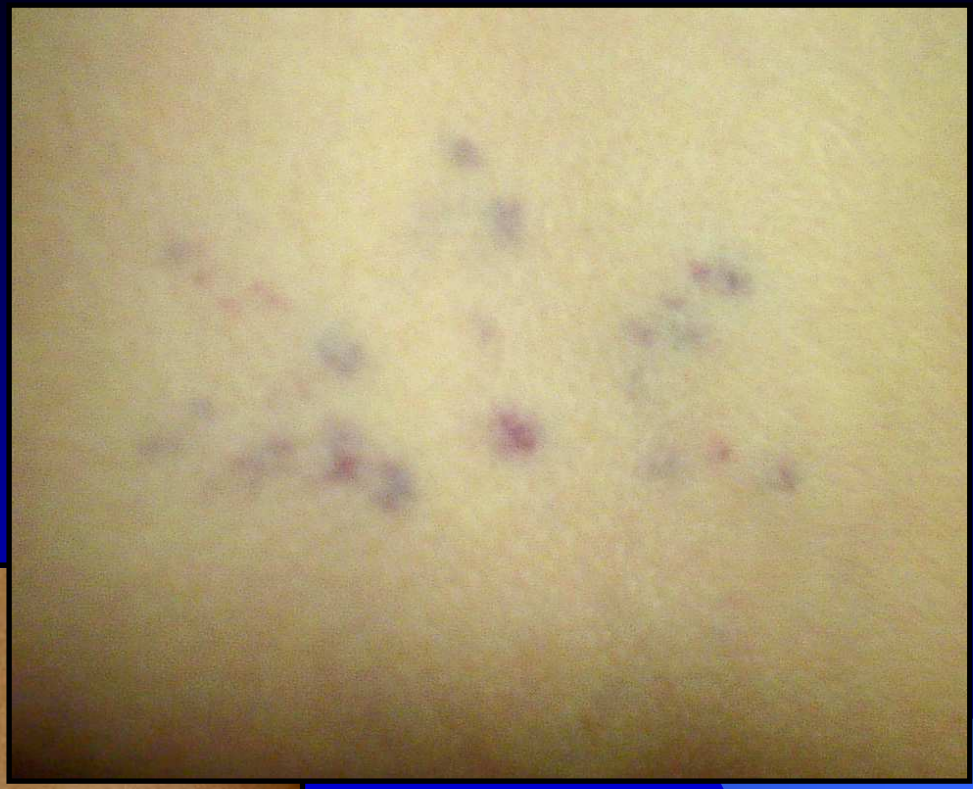
Estimate of Mutation Frequency

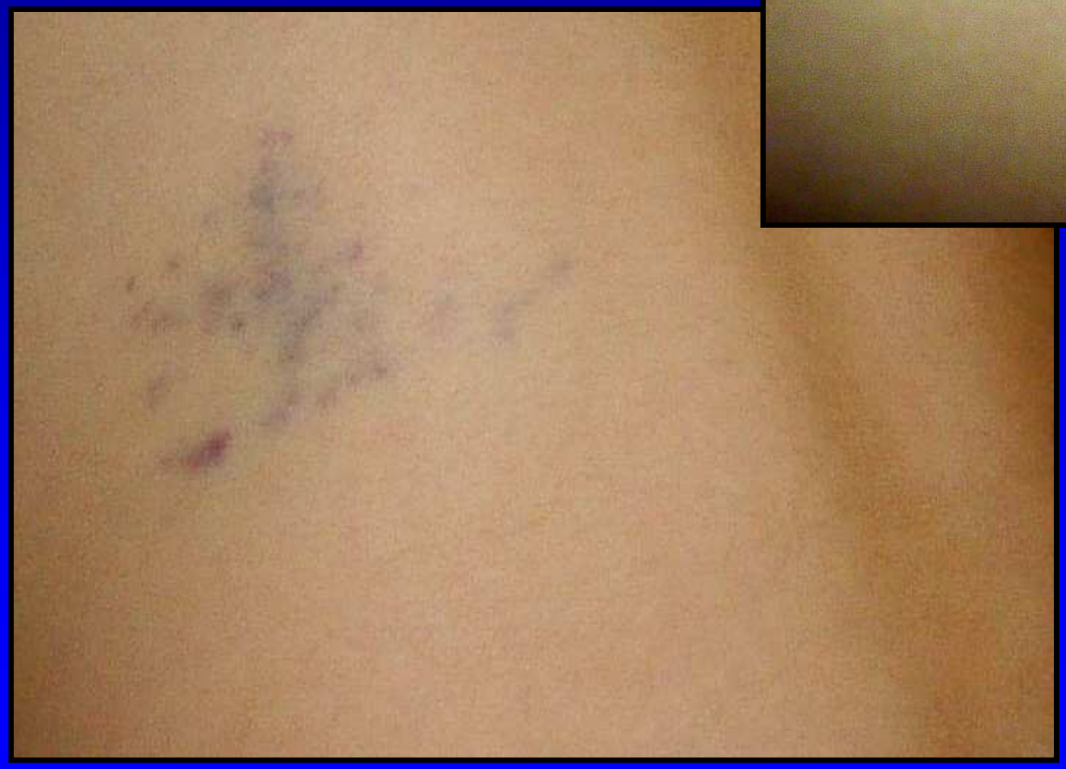
- BSA ~ 1.6 m^2 (~ $8x10^7$ capillary loops)
- Mean number of ~8 discrete lesions/patient
- Mutation rate, μ, of ~ 10^{-7}

Implications of model - I

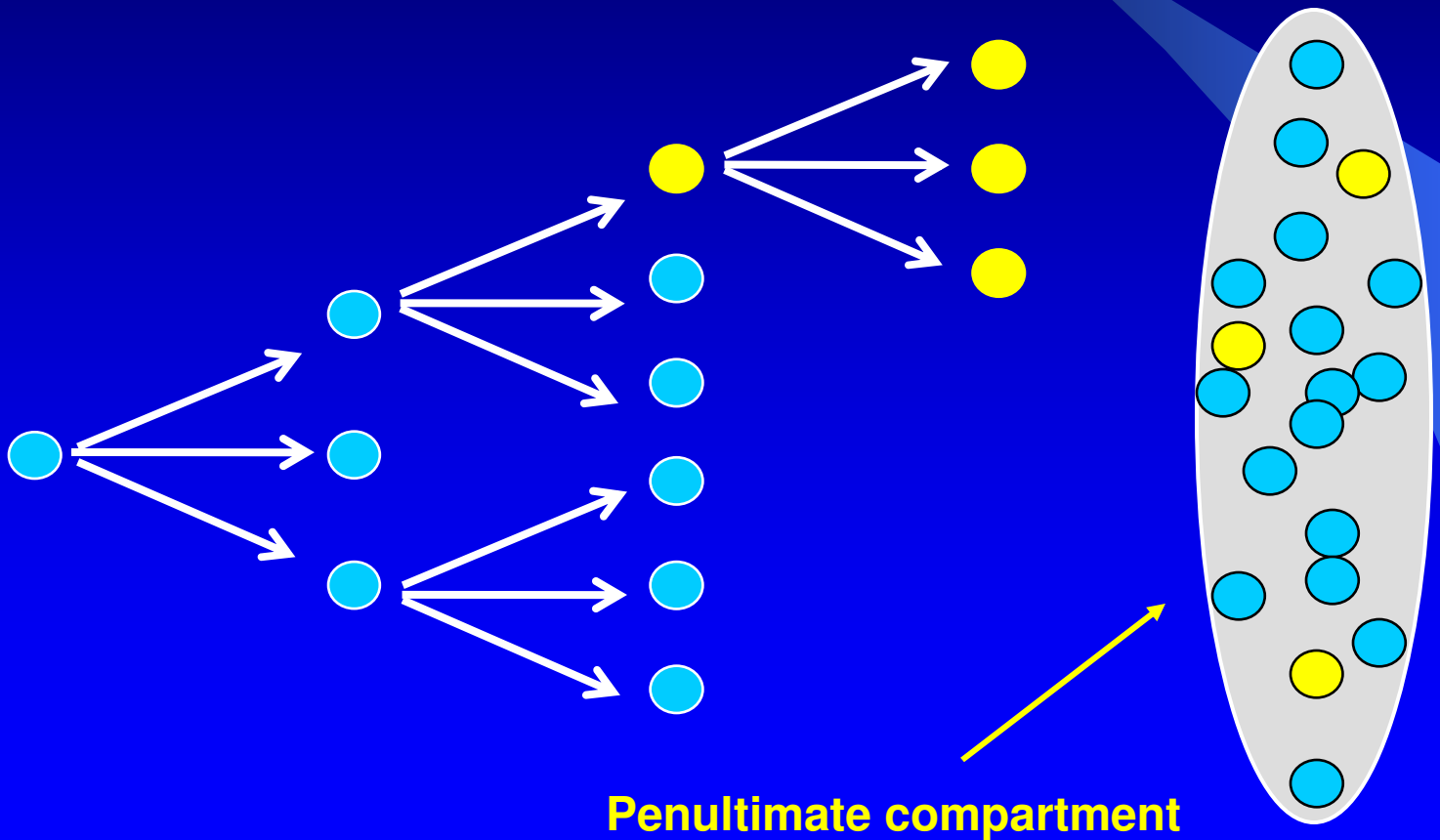
The developmental timing of a second hit is mapped onto a spatial distribution

Larger Clusters

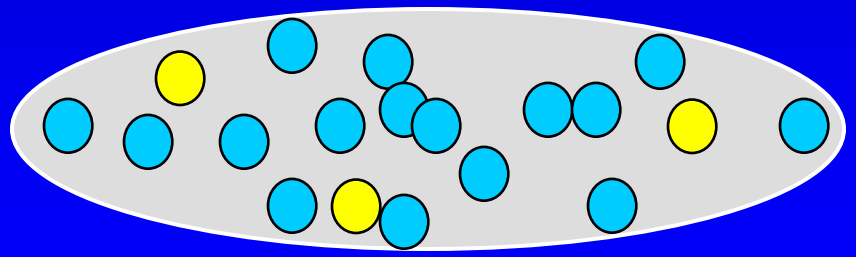




Branching Hypothesis



Sampling progenitors from a compartment



Penultimate compartment

Summary

- *FAP68* mutations co-segregate with disease
- Founder mutation: 157 delAAGAA
- Novel mutation: 1293delA + 1296delAAA
- Two hits required for the development of multiple GVM
- Development of mathematical model

Implications of Model

- Map the developmental hit to a spatial distribution
- Predict the clinical presentation when changing parameters μ and λ

In Appreciation

Dr. David Hogg
Dr. Alfons Krol
Dr. Phil Marsden
Hogg Lab Members

Canadian Dermatology Foundation