Cancer as an evolutionary process

Niko Beerenwinkel
ETH Zurich

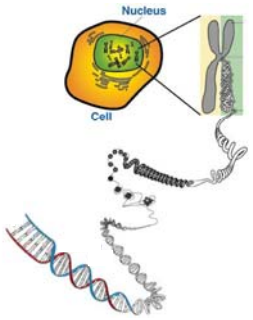
Outline

- Introduction
- Cancer initiation
- Cancer progression
- Genetic heterogeneity and treatment
- Dependencies among clonal mutations

Introduction

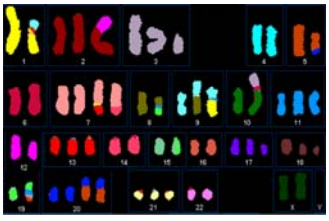
Cancer is a genetic disease

- Cancer cells harbor various types of genetic alteration, including
  - point mutations
  - insertions
  - deletions
  - chromosome rearrangements
  - mitotic recombination
  - loss or gain of whole chromosome arms



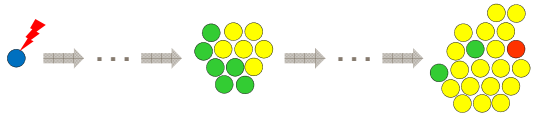
Most cancer cells are *aneuploid*

normal karyotype



karyotype of a colon cancer cell

Cancer progression



origin adenoma carcinoma

Genetic progression (accumulating mutations)

Cancer develops in multiple steps

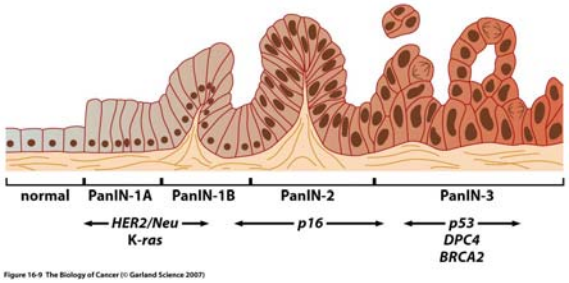
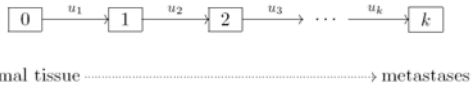


Figure 16-9 The Biology of Cancer (© Garland Science 2007)

Multistage theory

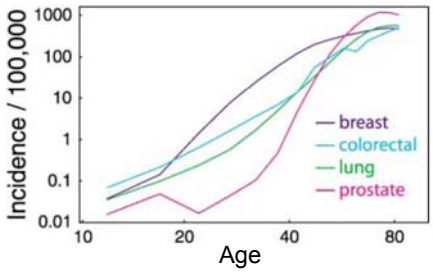


- u_j = u small $\Rightarrow$ Prob(one step by time t) = $1 - e^{-ut} \approx ut$.
- Then the incidence is $I_k(t) = (ut)^{k-1}u = u^k t^{k-1}$ and

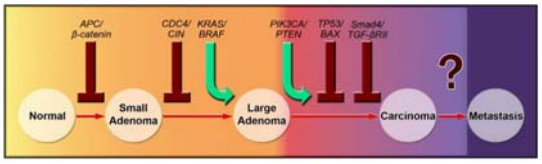
$$\log I_k = k \log u + (k-1) \log t$$

Armitage & Doll 1954, Nordling 1953

Cancer incidence data

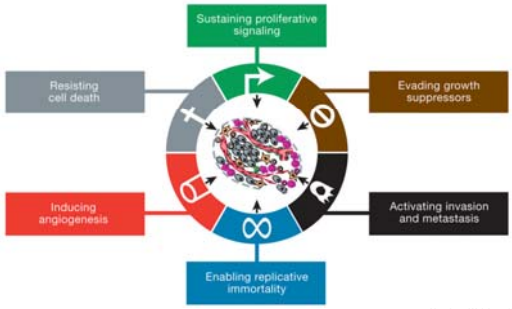


Vogelgram: Linear genetic progression



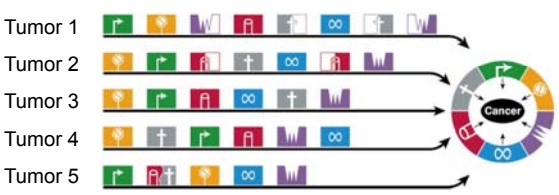
Vogelstein et al. 1988, Jones et al. 2008

The hallmarks of cancer



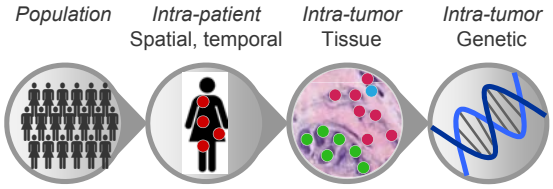
Hanahan, Weinberg (2001, 2011)

Hallmarks occur in different orders



Hanahan, Weinberg (2001, 2011)

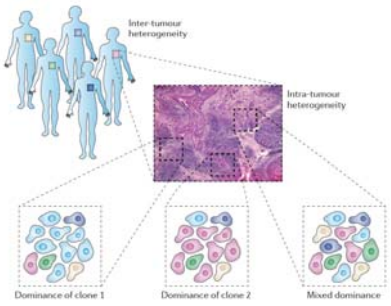
Cancer is heterogeneous at multiple levels



Florian Markowetz

13

Inter-patient and intra-patient diversity

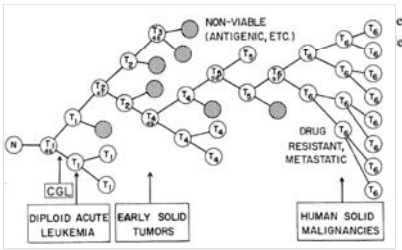


Marusyk et al (2012)

14

Cancer as an evolutionary process (Nowell, Science 1976)

The Clonal Evolution of Tumor Cell Populations

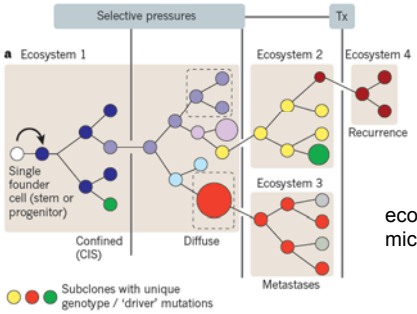


ermits stepwise selection
erlies tumor progression.

Peter C. Nowell

15

Cancer ecology and evolution

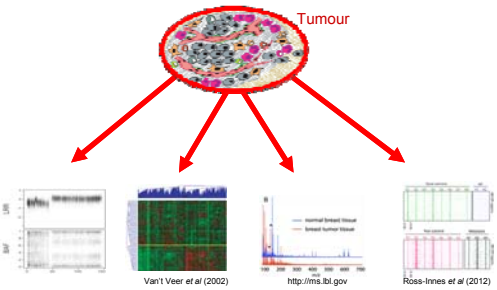


ecosystem = tissue microenvironment

Greaves and Maley (2012)

16

Molecular profiling of tumors



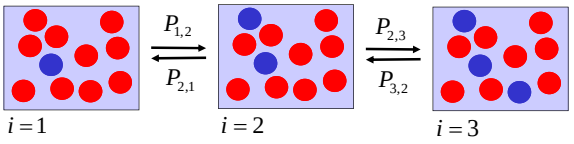
DNA RNA Protein ChIP

17

Cancer initiation

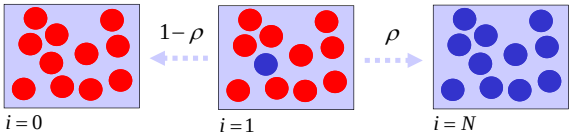
18

Finite populations: The Moran process



$$P_{i,i+1} = \frac{i}{N}\left(1-\frac{i}{N}\right)$$
$$P_{i,i-1} = \left(1-\frac{i}{N}\right)\frac{i}{N}$$
$$P_{i,i} = \left(\frac{i}{N}\right)^2 + \left(1-\frac{i}{N}\right)^2$$

Key results

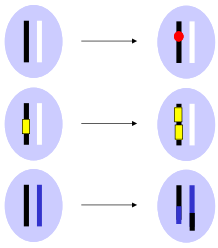


- Fixation probability of a single allele with selective advantage r: $\rho = \frac{1-1/r}{1-1/r^N}$
- Mean waiting time: $\tau \propto N^2$
- Rate of loss of diversity: $2/N^2$

Oncogenes

- Oncogenes increase fitness, if one allele is mutated or inappropriately expressed. They are activated by:

1. a specific point mutation
2. a gene amplification
3. or chromosomal fusion



Nowak (2006)

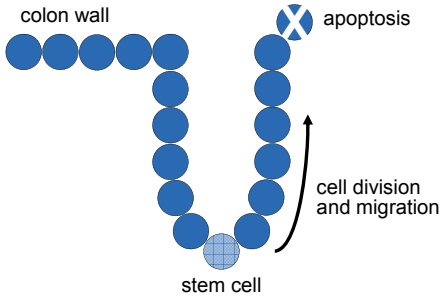
Fixation of oncogene mutations

- The probability that a mutant with selective advantage r has been fixed in a (small) population of size N by time t is

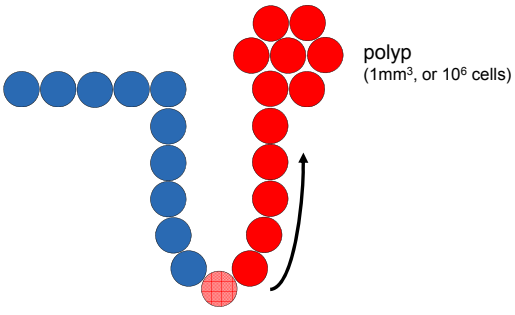
$$P(t) = 1 - e^{-Nu\rho t}$$

- Observation: Large compartments accelerate the accumulation of advantageous mutations, small compartments slow it down.
- Most tissues with high cell turnover are organized in many small compartments.
- For example, colonic crypts…

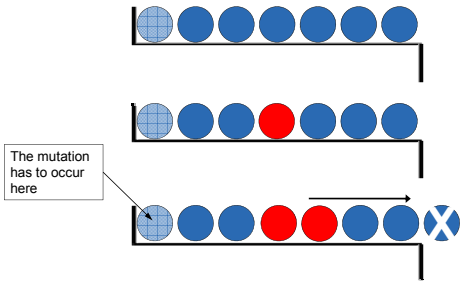
The colon is organized into 10^7 crypts, each consisting of 1,000 to 4,000 cells.



Colon cancer arises in a crypt



The linear process of cancer



$\rho = 1/N$ and $P(t) = 1 - e^{-ut}$, independent of r

Nowak (2006)

Tumor suppressor genes (TSGs)

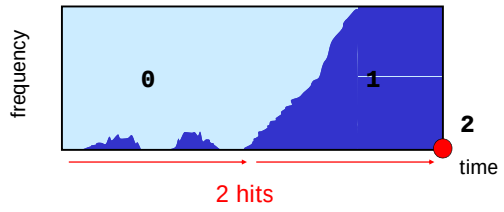
- Somatic mutations in TSGs are recessive: inactivation of one allele is (nearly) neutral, while inactivating the second allele confers a fitness advantage. TSGs are inactivated by
  1. two point mutations

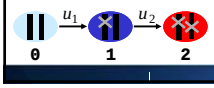
or

  2. one point mutation followed by loss of heterozygosity (LOH).

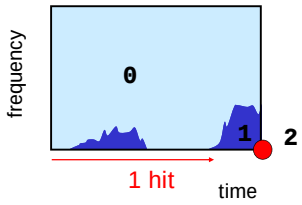
Nowak (2006)

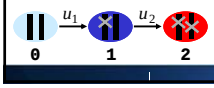
TSG inactivation in small populations: two hits



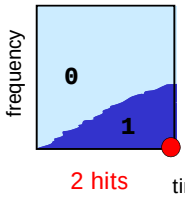


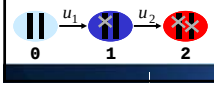
TSG inactivation in intermediate populations: one hit ("stochastic tunneling")



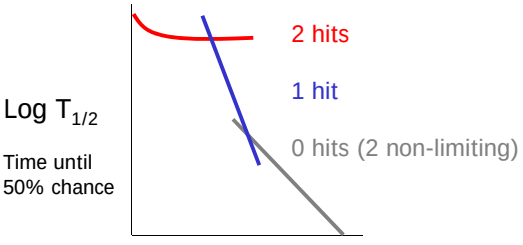


TSG inactivation in large populations: two hits (not rate limiting)



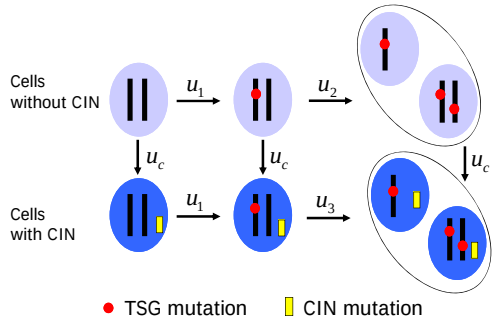


Summary: three dynamic laws for TSG inactivation



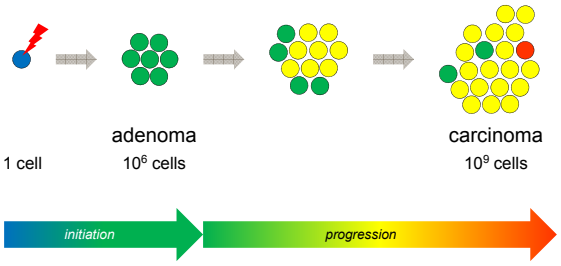
Nowak (2006)

TSG inactivation and genetic instability

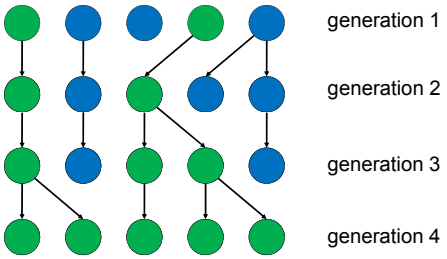


Cancer progression

Genetic progression of cancer

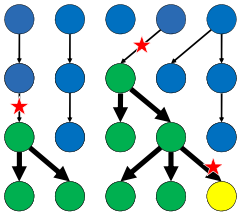


The Wright-Fisher process



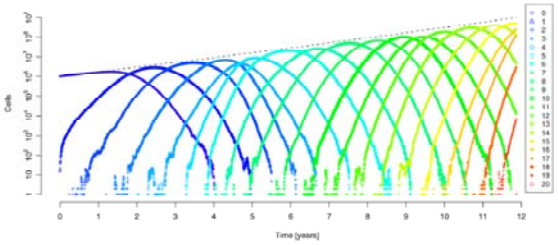
Wright-Fisher process, extensions

- Multinomial sampling in discrete generations
- Four model parameters:
  - N — population size
  - d — number of driver genes
  - u — mutation rate
  - s — selective advantage



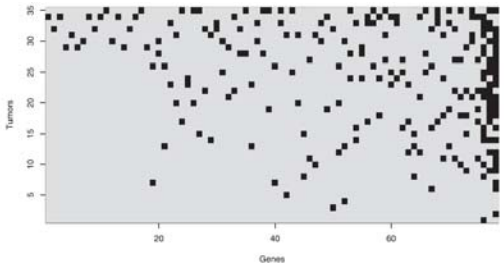
$$X(t+1) \mid X(t) \sim \text{Mult}(N(t), \theta_0(t;u,s,d), \ldots, \theta_k(t;u,s,d))$$

Evolutionary dynamics of cancer progression



Beerenwinkel et al (PLoS Comput Biol 2007)

The waiting time to cancer



- How long does it take until the first cell with any 20 out of 100 mutations occurs?

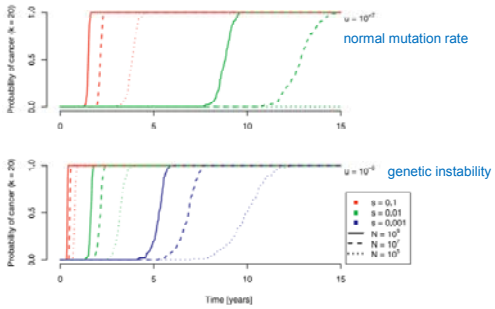
The speed of adaptation

- The time until the first cell with *k* mutations appears is

$$\tau_k \approx \frac{k \log^2 [s/(ud)]}{2s \log N}$$

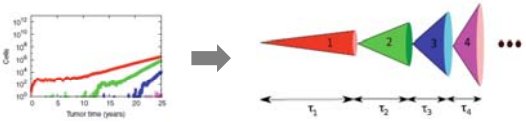
- Thus, we waiting time to cancer is dominated by the selective advantage *s*.

The probability of developing cancer

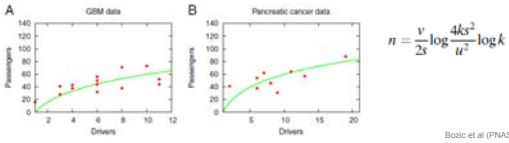


Accumulation of driver and passenger mutations

- Branching process model:



- Data:



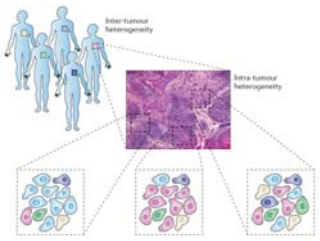
$$n = \frac{v}{2s} \log \frac{4ks^2}{u^2} \log k$$

Bozic et al (PNAS 2010)

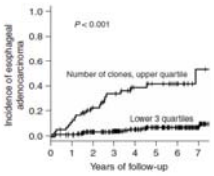
Intra-tumor genetic diversity and treatment

Intra-tumor diversity itself has prognostic value

Diagnostics



Prognostics

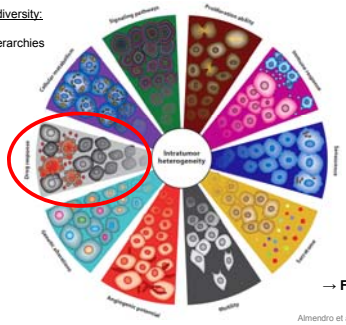


→ **Personalized medicine**

Marusyk et al (Nat Rev Cancer 2012), Maley et al (Nat Genet 2006)

Causes and consequences of tumor diversity

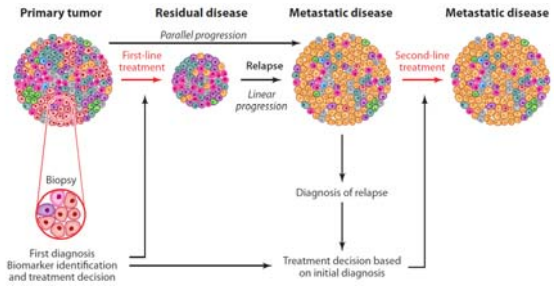
Sources of tumor diversity:
- Genetic variation
- Differentiation hierarchies
- Environment
- Stochasticity



→ **Functional variation**

Almendro et al (Annu Rev Pathol Mech Dis 2013)

Intra-tumor diversity affects the value of biomarkers

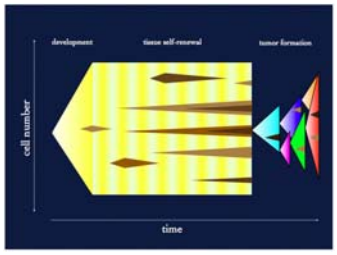


Almendro et al (Annu Rev Pathol Mech Dis 2013)

Most mutations originate prior to tumor initiation

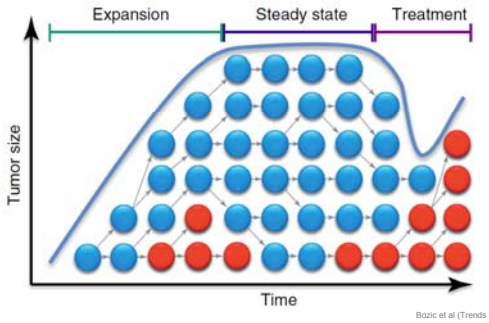
$$\mathrm{E}[N_{\text{passenger}}] = S\mu t$$

where *S* is the number of bases, μ is the per-base mutation probability, and *t* the number of times the tissue has self-renewed



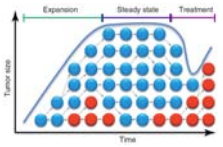
Tomasetti et al (PNAS 2013)

Resistance development



Bozic et al (Trends Mol Med 2012)

The probability of treatment success



$$P(\text{eradication}) = P_1 P_2 P_3$$

Prob. of no resistance mutation arising during **expansion**

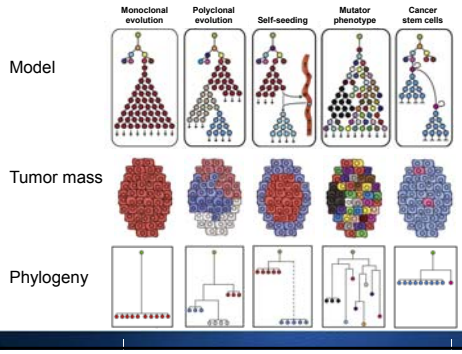
Prob. of no resistance mutation arising during **steady state**

Prob. of no resistance mutation arising during **treatment**

(density-dependent branching process model)

Bozic et al (Trends Mol Med 2012)

Competing tumor progression hypotheses



Navin et al (Mol Oncol 2010)

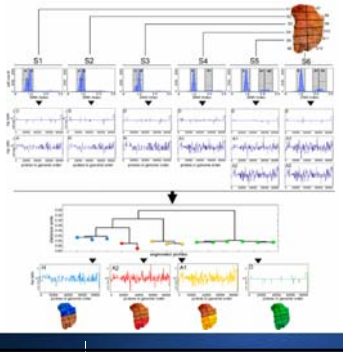
Strategies for intra-tumor sampling

- Deconvolution of pooled sample
- Laser capture microdissection
- Cell sorting
- Single-cell analysis

- Tree reconstruction methods:
  - Rearrangement phylogeny (Greenman et al., 2011)
  - TuMult (Letouze et a., 2010)
  - MEDICC (Schwarz et al.)

→ Roland Schwarz' talk!

Sector Ploidy Profiling



Correlation among copy number porfiles

NJ tree

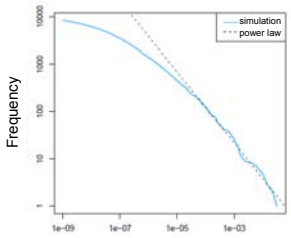
Navin et al (Genome Res 2010)

Estimating intra-tumor diversity from next-generation sequencing data

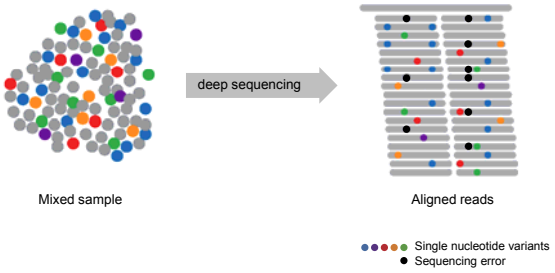
Wright-Fisher model: Expected mutant distribution

- Include neutral mutations (passengers)
- Trace individual genotypes

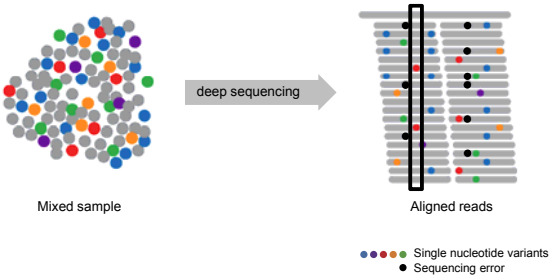
→ Most genetic variants occur at low frequencies



Next-generation ultra-deep sequencing



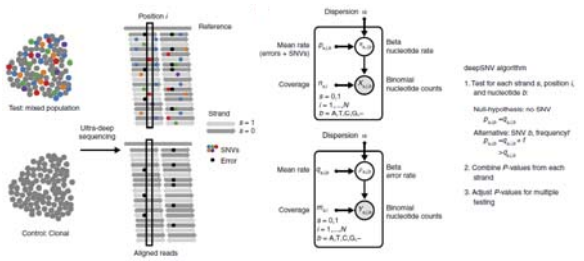
Calling single-nucleotide variants (SNVs)



Challenges in NGS-based diversity estimation

1. Most SNVs are expected to occur at low frequencies
2. Sample processing and sequencing errors are not uniform
3. Need to test many positions

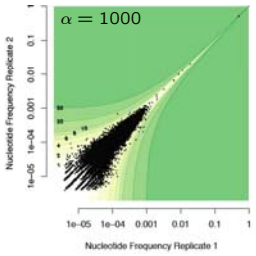
deepSNV: Likelihood ratio test



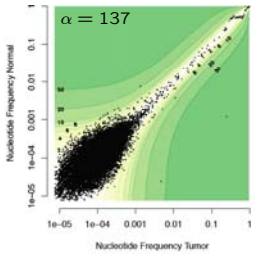
$$-2 \log \frac{L(X_{s,i,b}, Y_{s,i,b} \mid H_0)}{L(X_{s,i,b}, Y_{s,i,b} \mid H_1)} \sim \chi^2_1$$

Overdispersion

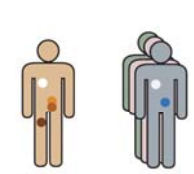
phiX replicates



tumor vs. normal

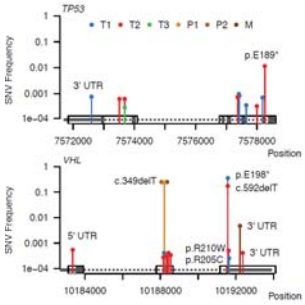


Application: Renal cell carcinoma



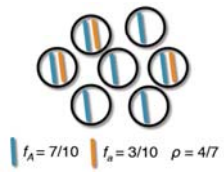
1x multiple lesions
Primary 1
Primary 2
Metastasis

3x tumor-normal:
Tumor 1
Tumor 2
Tumor 3

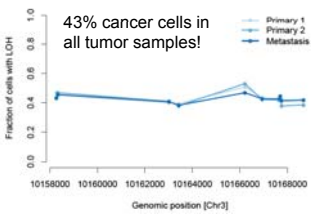


Loss of heterozygosity

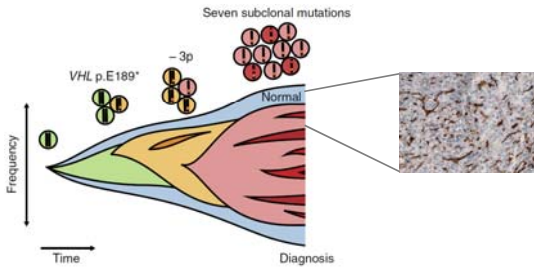
- 14 hemizygous SNPs
- Allele frequency ratio $r = f_A / f_a \approx 2$ on chr. 3



$$\rho = \frac{r/r_0 - 1}{r/r_0 + n - 1}$$

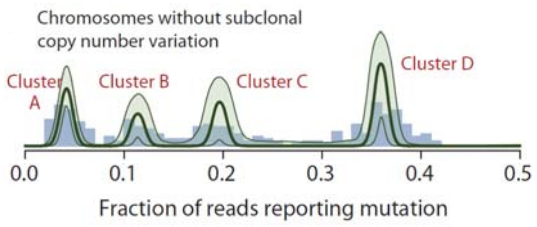


Evolutionary history



A breast tumor sequenced at 188-fold coverage

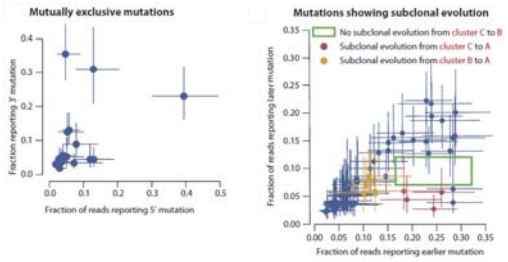
- Frequency distribution of mutations:



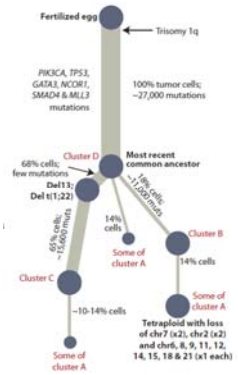
Nik-Zainal et al (Cell 2012)

Phasing pairs of subclonal mutations

- Independent vs. subclonal evolution:



Nik-Zainal et al (Cell 2012)



Nik-Zainal et al (Cell 2012)

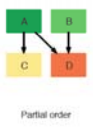
Dependencies among clonal mutations (order constraints)

Accumulation of driver mutations

- Cancer mutations accumulate in a series of clonal expansions

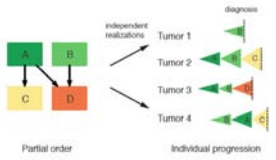
… but not always in the same linear order.

Partial order

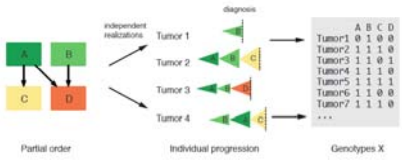


Partial order

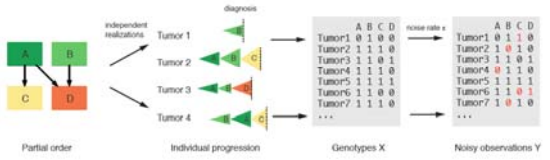
Individual progression



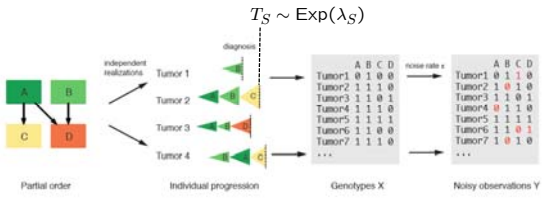
Genotypes at diagnosis



Observed genotypes



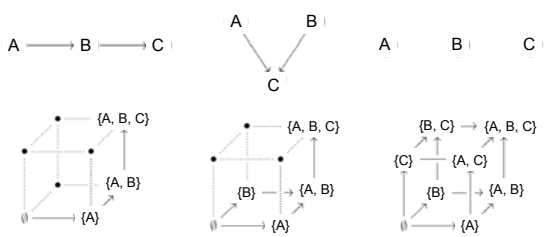
Waiting time distribution



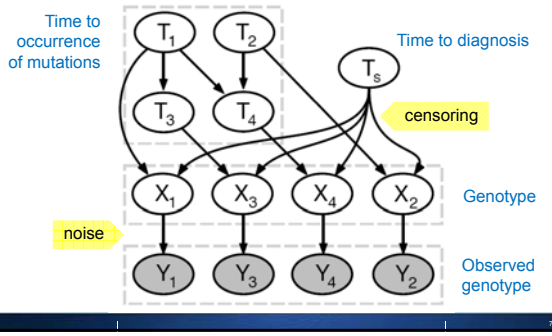
$$
\begin{aligned}
T_A &\sim \text{Exp}(\lambda_A) \\
T_B &\sim \text{Exp}(\lambda_B) \\
T_C &\sim T_A + \text{Exp}(\lambda_C) \\
T_D &\sim \max(T_A, T_B) + \text{Exp}(\lambda_D)
\end{aligned}
$$

Gerstung et al (PLOS ONE 2012)

Posets define the geometry of genotype space



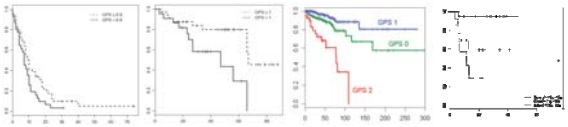
Hidden conjunctive Bayesian network (H-CBN)



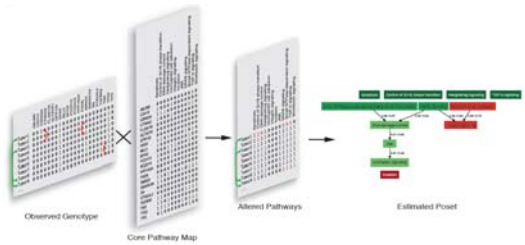
The genetic progression score (GPS) is a prognostic biomarker

$$\mathrm{GPS} = \mathrm{E}\left[\max\{T_j \mid X_j = 1\}\right]$$

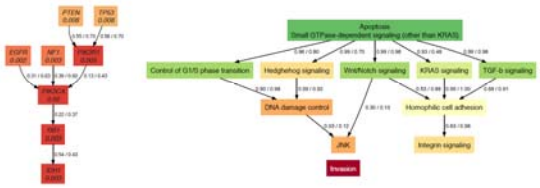
Glioblastoma | Prostate cancer | Meningioma | Oral cancer



Progression on the level of genes and pathways



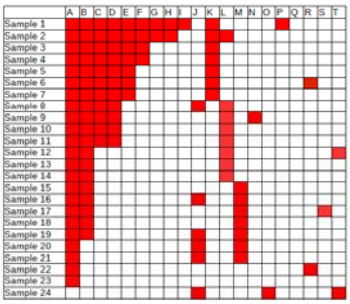
Glioblastoma multiforme



genes pathways

Drivers and passengers

- Frequency
- Mutual exclusivity
- Prior knowledge
- Dependency



SUMMARY

- Tumor development is a stochastic evolutionary process.
- Intra-tumor heterogeneity is abundant and associated with failure of targeted treatment.
- Tumor diversity can be detected and quantified by deep-coverage next-generation sequencing.
- Intra-tumor phylogeny is an open problem, characterized by specific features of cancer evolution and by experimental techniques.

Acknowledgements

- Jonas Behr
- **Simona Constantinescu**
- **Ariane Hofmann**
- Janos Kelemen
- Pejman Mohammadi
- Hesam Montazeri
- **Thomas Sakoparnig**
- Fabian Schmich
- David Seifert
- Juliane Siebourg
- **Ewa Szczurek**
- Armin Töpfer

Collaborators:
Tibor Antal (Edinburgh)
Christian Beisel (D-BSSE, ETH Zurich)
Nick Eriksson (23andMe)
Moritz Gerstung (Sanger Institute)
Manoj Mahimkar (Tata Memorial Center, Mumbai)
Holger Moch (University Hospital Zurich)
Martin Nowak (Harvard)
Jörg Rahnenführer (TU Dortmund)
Alejandro A. Schäffer (NCBI, NIH)
Arne Traulsen (MPI Evolutionary Biology)
Bert Vogelstein (Johns Hopkins)

Funding:
SNF, SHCS, SystemsX.ch,
Swiss Cancer League, SIB

www.cbg.ethz.ch