nature biotechnology

Login

Search this journal go Advanced search

High order chromatin architecture shapes the landscape of chromosomal alterations in cancer

Geoff Fudenberg[1], Gad Getz[2], Matthew Meyerson[2–5] & Leonid A Mirny[2,6,7]

[1]Harvard University, Program in Biophysics, Boston, Massachusetts, USA. [2]The Broad Institute of the Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts, USA. [3]Harvard Medical School, Boston, Massachusetts, USA. [4]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [5]Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [6]Harvard-MIT, Division of Health Sciences and Technology, Cambridge, Massachusetts, USA. [7]Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. Correspondence should be addressed to L.M. (leonid@mit.edu).

Received 9 September; accepted 21 October; published online 20 November 2011; doi:10.1038/nbt.2049

Cristina Sisu

Journal Club

8th December 2011

Hypothesis

- 3D chromatin organization and spatial co-localization influences the set of somatic copy-number alterations in cancer.

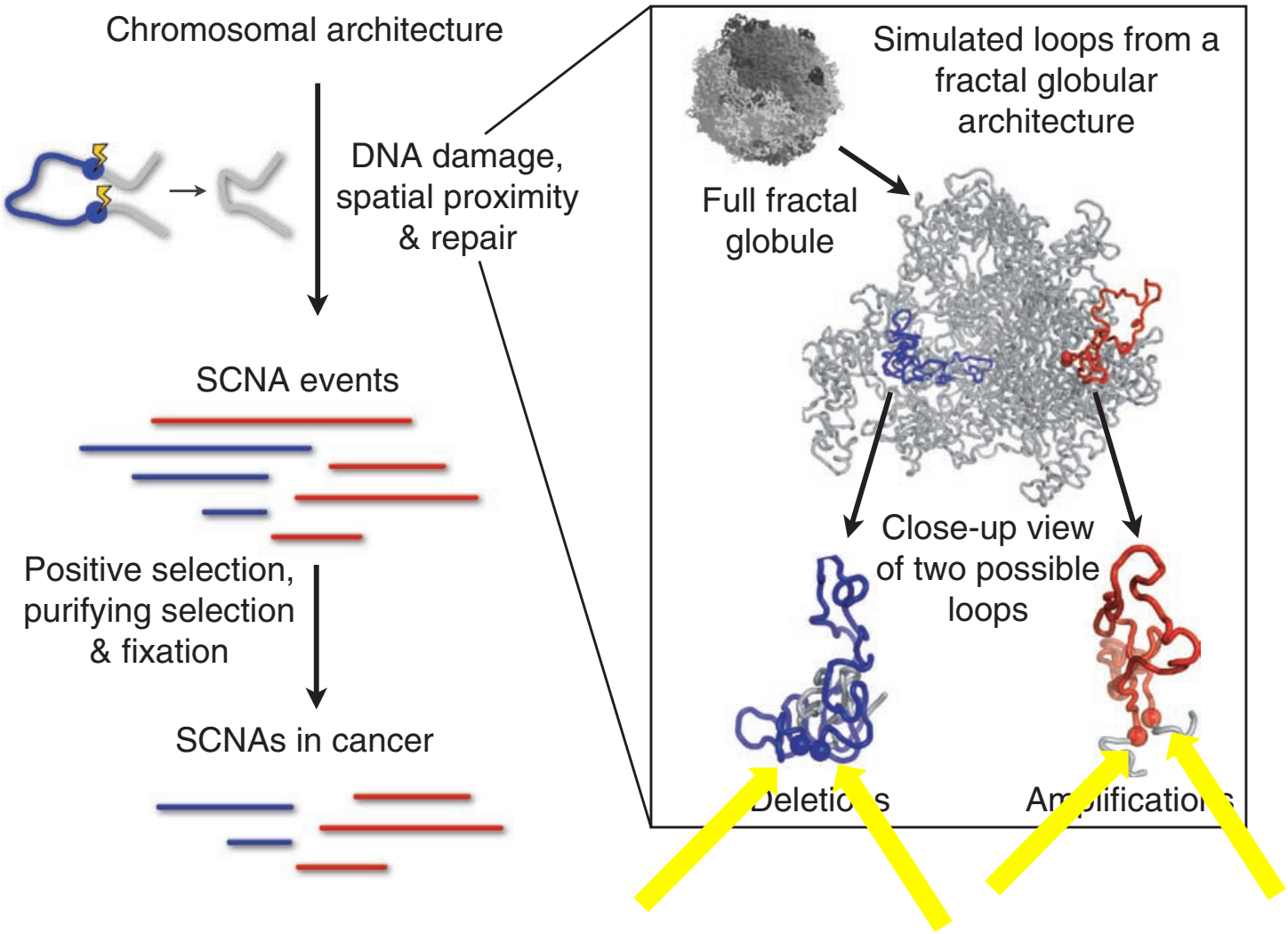
Question

- Are distant genomic loci that are brought together by the 3D chromatin architecture during interphase more likely to undergo structural alterations and become end points for amplification and deletions observed in cancers?

Background

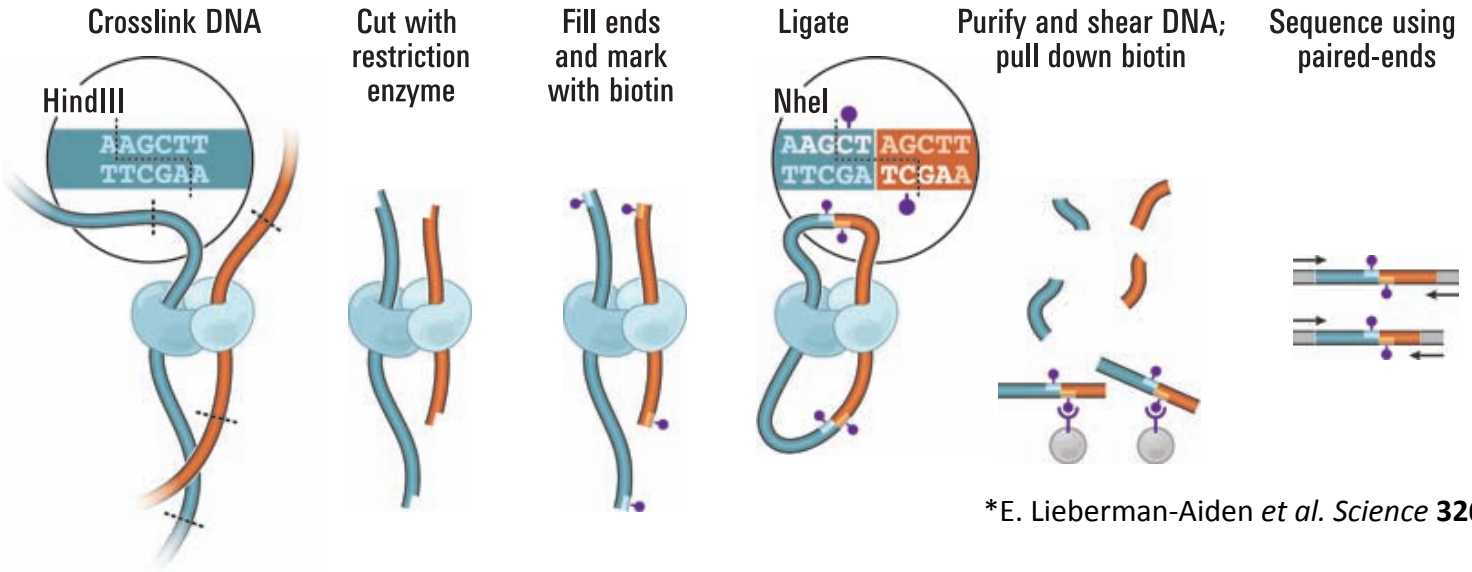
- SCNA:
  - **S**omatic **C**opy **N**umber **A**lteration(s)
  - a sequence that is found at different copy numbers in an individual's germ-line DNA and in the DNA of a clonal sub-population of cells
  - somatic changes in the number of copies of a DNA sequence that arise during the process of cancer development
  - the most common genomic alterations in cancer
  - focal SCNA have led to the identification of cancer-causing genes and aided the design of potential therapeutic strategies.
  - challenge:
    - discrimination of SCNAs from all CNVs

Cancer – An Evolutionary Process



Methods (1)

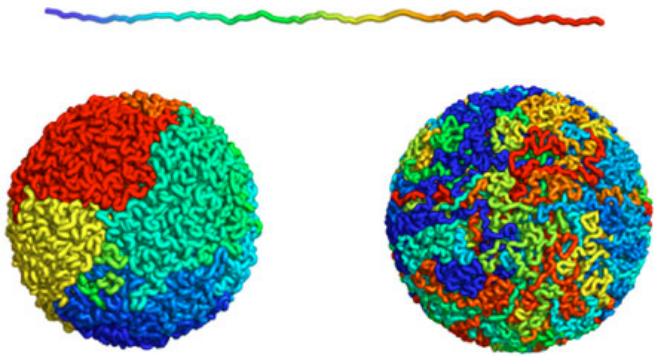
- 39,568 SCNAs:
  - 26,022 amplifications and 13,546 deletions
  - Intra-arm position (not in near telomeric or centromeric regions)
  - L > 1Mb
- HiC*:
  - Experimental method for high-throughput chromosome conformation capture
  - Capable of identifying long-range interactions in an unbiased genome-wide fashion



*E. Lieberman-Aiden *et al. Science* **326**, 289 (2009)

Methods (2)

- Fractal globule model (FG*):
  - A compact polymer state that emerges during polymer condensation as a result of topological constraints which prevent one region of the chain from passing across another one



Fractal Globule Equilibrium Globule

*L.A. Mirny, *Chrom. Res*. **19**, 37 (2011)

- Bayesian information criterion (BIC*)
  - Criterion for model selection given a number of models
  - Penalizes the models based on their complexity – number of parameters

$$\mathrm{BIC} \equiv -2 \ln \mathcal{L}_{\max} + k \ln N$$

L_{max} : maximum likelyhood achievable by the model
k = number of parameters in the model
N = number of data points in the model

*G. Schwarz *Annals of Statistics* **6,** 461 (1978)

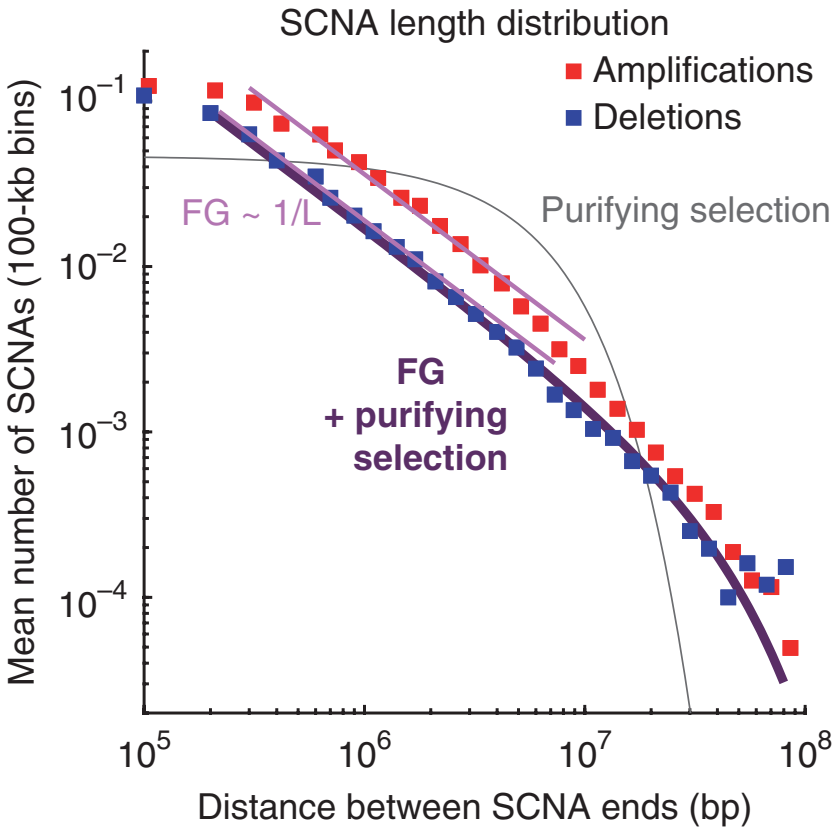
  - Adapted for SCNA:

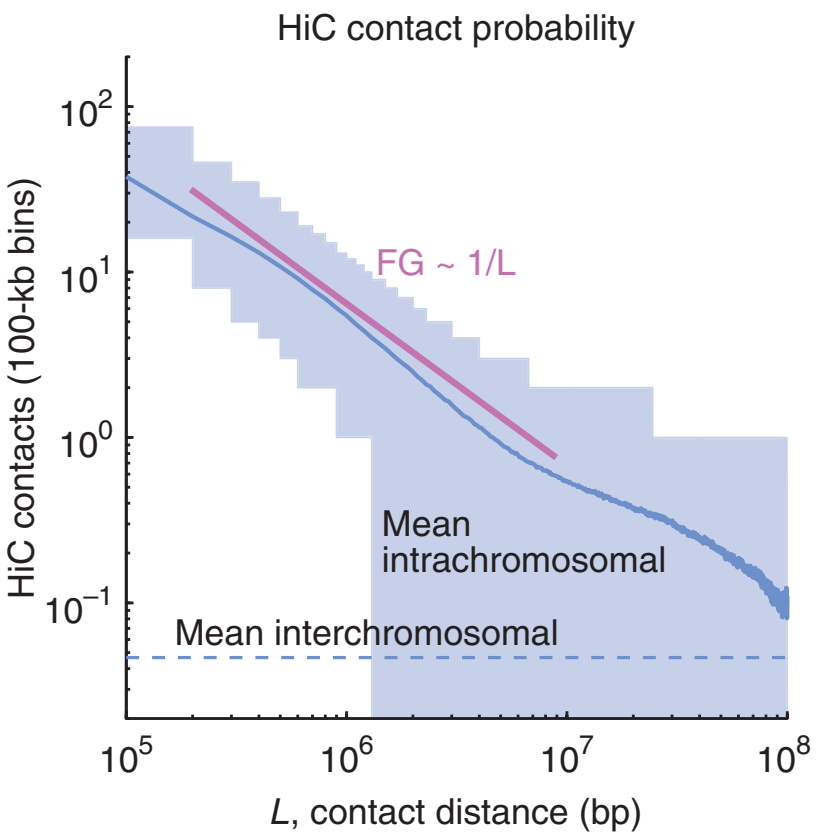
$$BIC = \log L(\mathrm{SCNA} \mid \mathrm{Model}) - \frac{1}{2} k \log(n)$$

L: length of SCNA
k: number of parameters
n: number of SCNA events

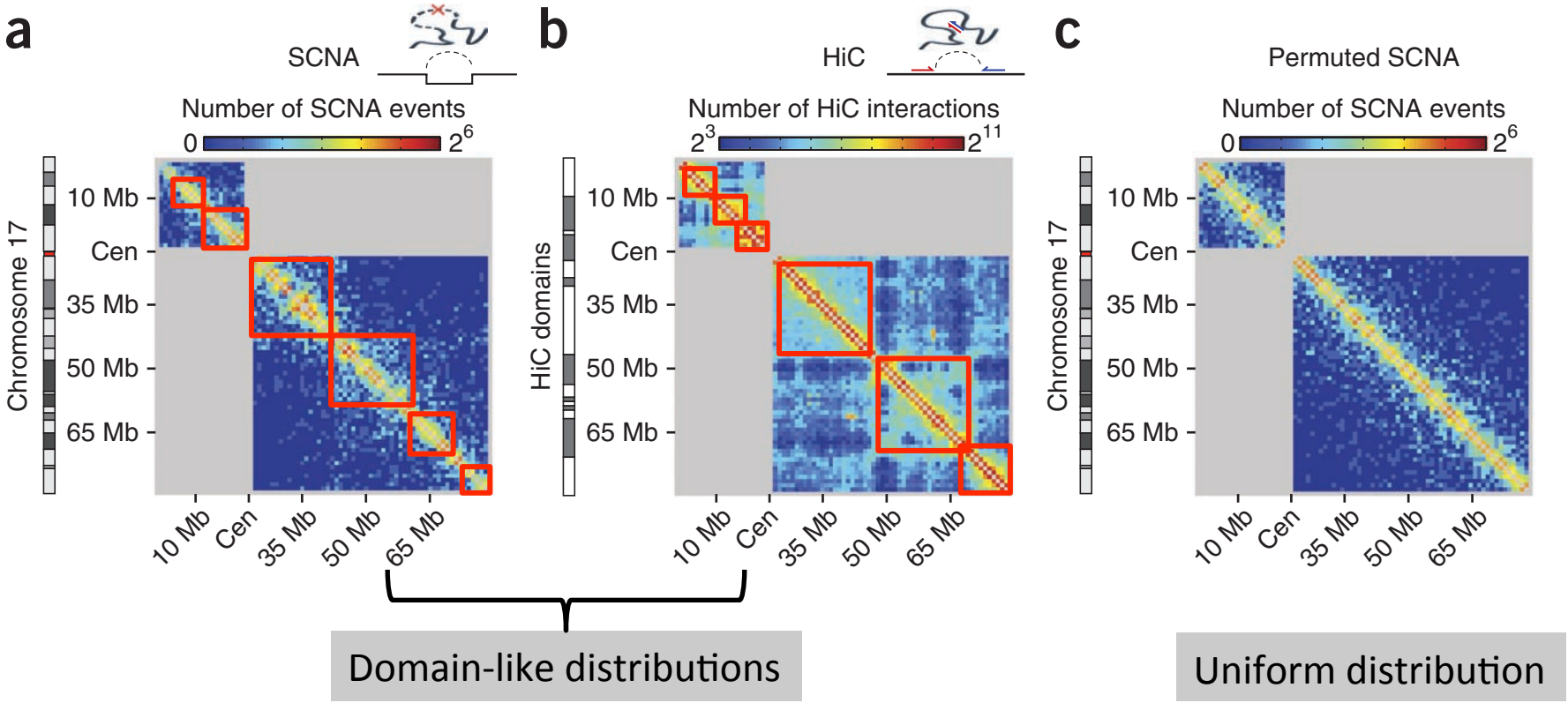
    - High value for BIC are preferred

Patterns of Chromatin Structure in the SCNA 'Landscape'





SNCA ~ 3D Structure



- GM06990 cell line
- Bin size $1Mb^2$
- Pearson's correlation coefficient r=0.55, P<0.001
  - => use Poisson likelihood for rare events analysis

Mutational and Evolutionary Models of SCNA

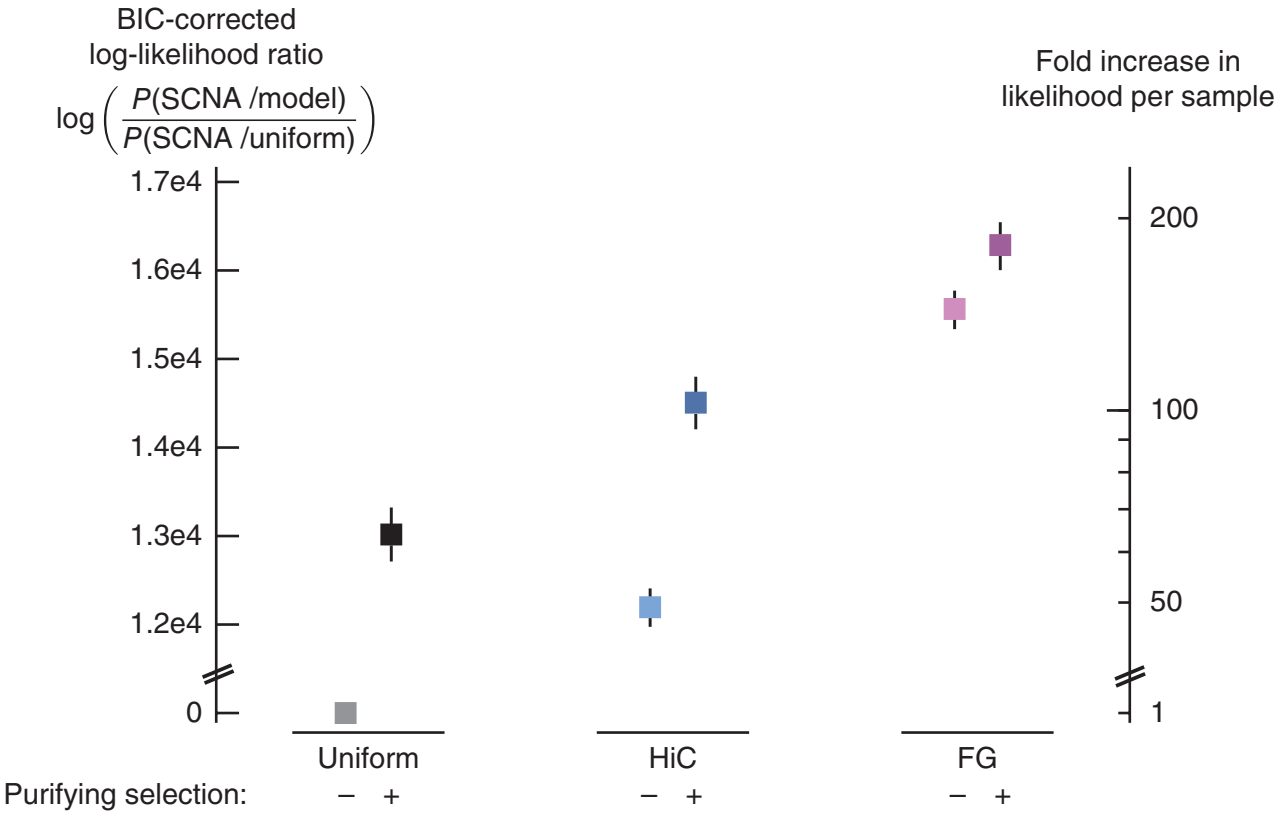
- Probability of observing SCNA starting at position *"i"* & ending at position *"j"*

$$P_{ij} = \mu_{ij} \cdot \pi(L)$$

  - μ_{ij} = mutation probability - SCNA to occur in a single cell
  - $\pi(L)$ = probability to have this mutations fixed in the population of cancer cells
  - L = i-j is the SCNA length

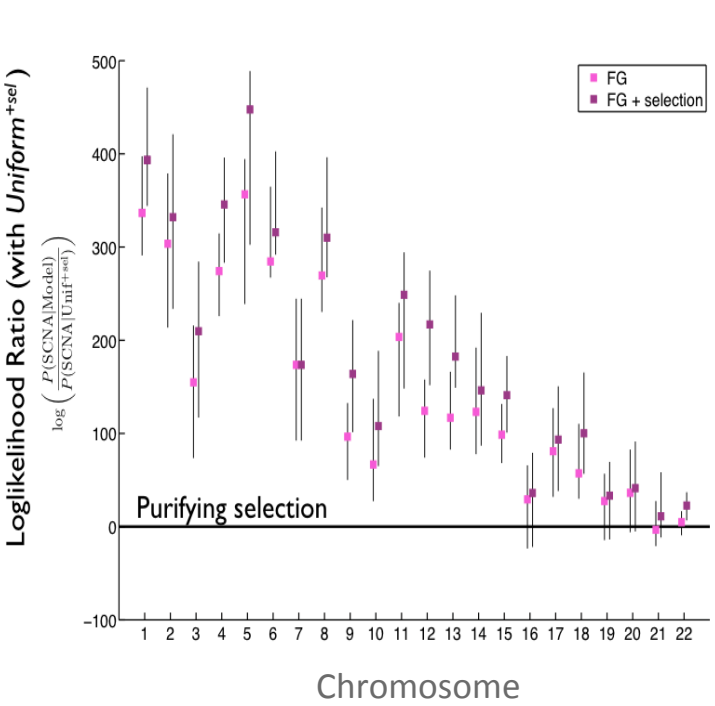
- Mutation probability depends on the model of SCNA apparition:
  - Uniform model - the two ends of the SCNA are selected randomly
  - HiC model – the probability of SCNA depends on the probability of a 3D contact between the 2 points given the HiC data
  - FG model – the probability of SCNA depends on the probability of a 3D contact between the 2 points given the fractal globule model

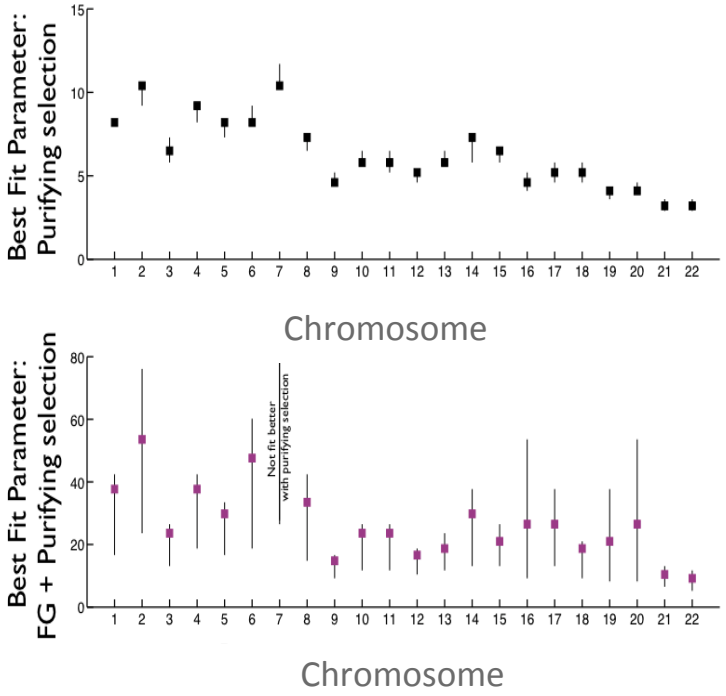
Model Selection – Results (1)

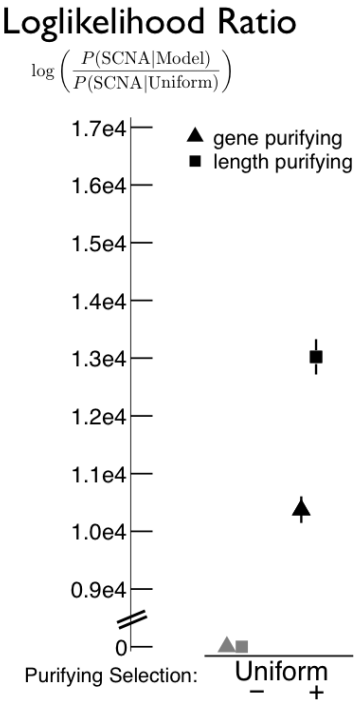


- HiC & FG perform better than random selection
- FG coupled with purifying selection fits best the SCNA data

Model Selection – Results (2)



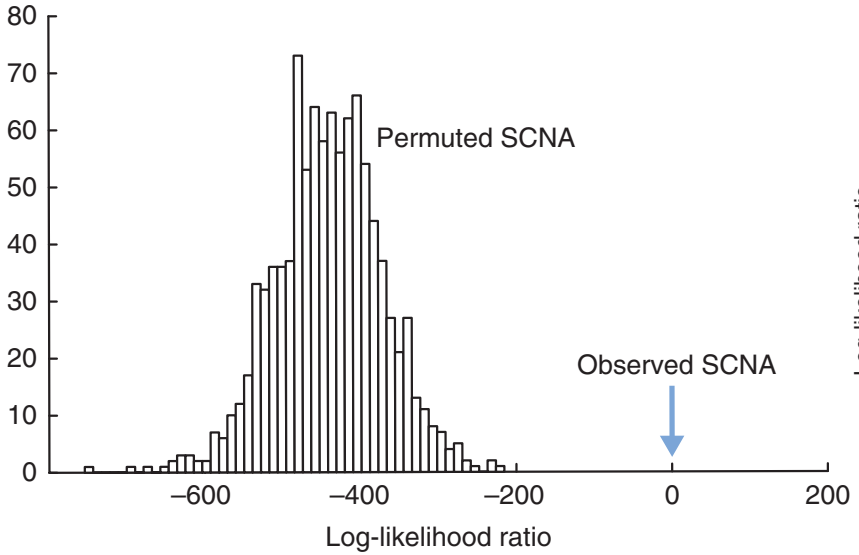


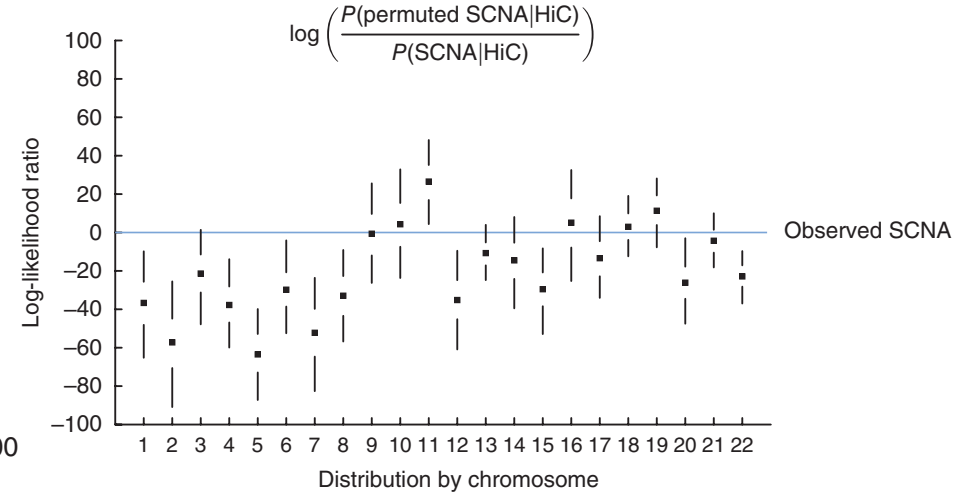


- For long chromosomes FG is better then purifying selection alone
- Short gene-rich chromosomes are subjected to stronger purifying selection,
  - *BUT: purifying selection alone is more depends stronger than SCNA length rather than the number of gene affected by SCNA !!*

Back to HiC

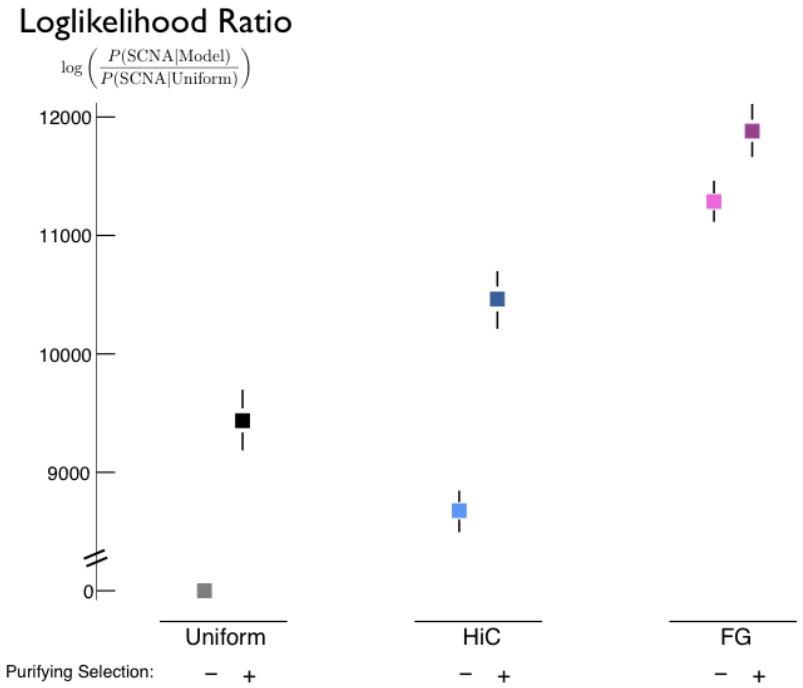
- SCNA landscape reflects chromatin structure in HiC
- Test using permutation analysis
  - randomly relocate the SCNA start points
  - keep *L* the same

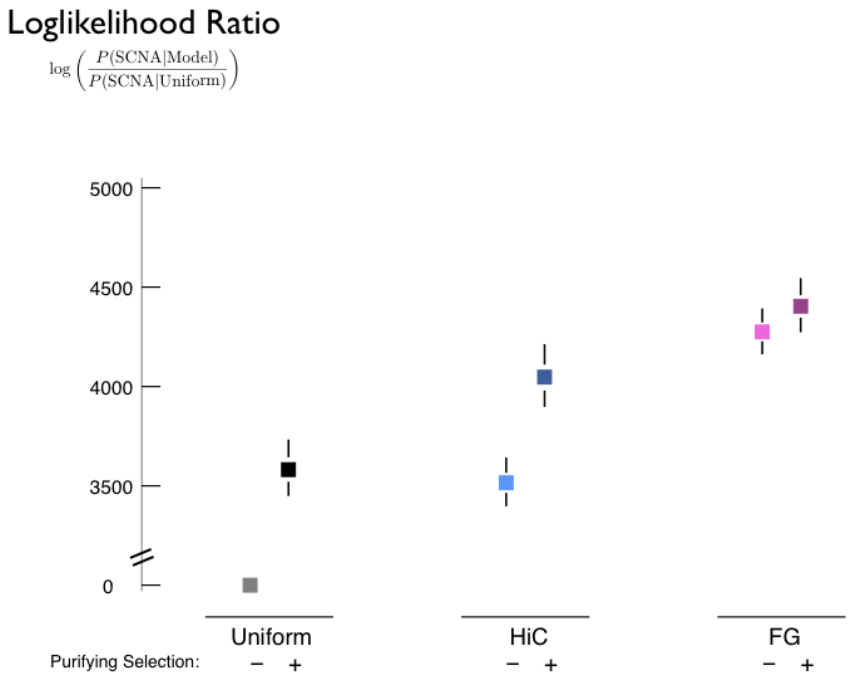




- HiC fits better the SCNA data than the random selection

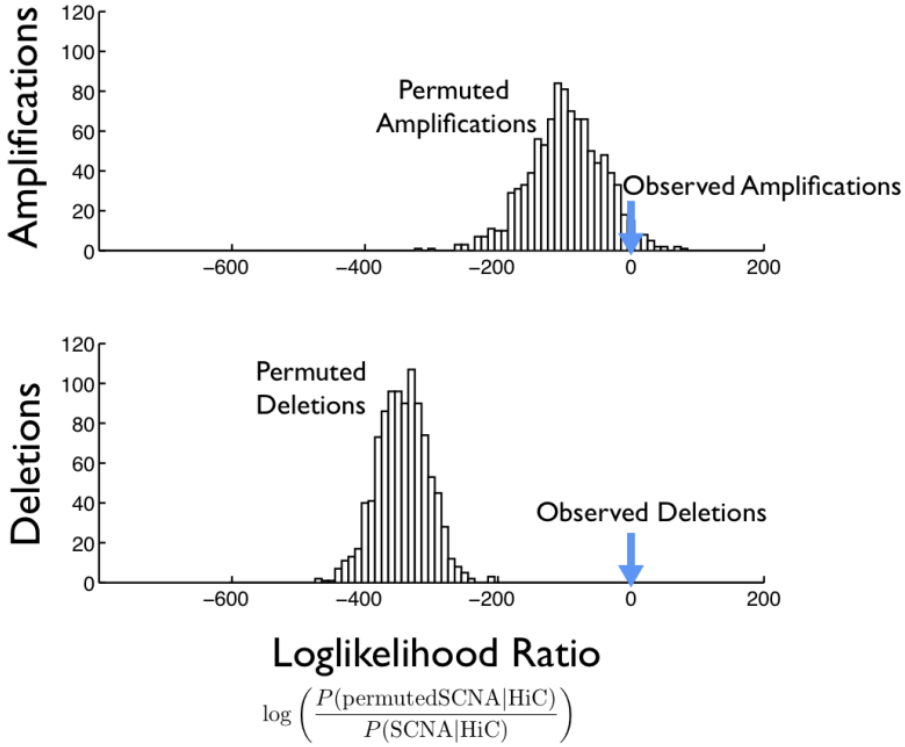
Amplifications vs Deletions

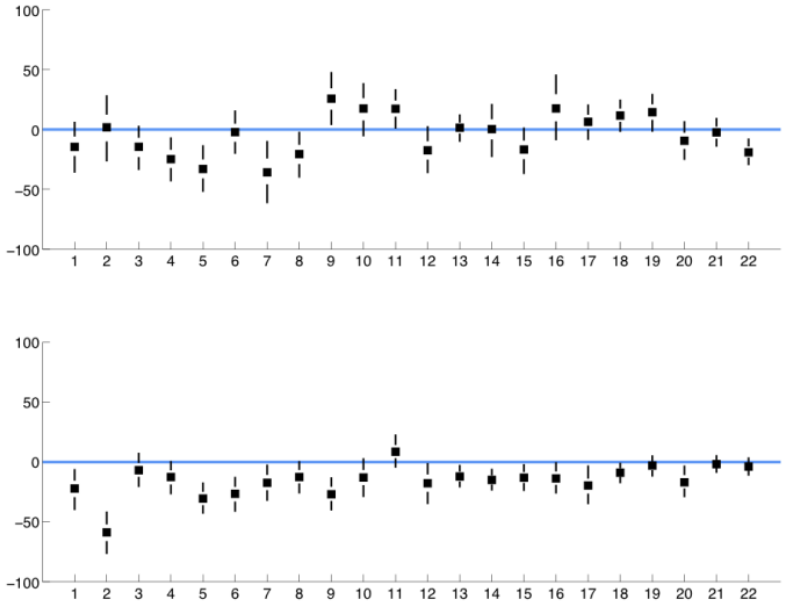




- Differences in strength of selection on genomic alterations:
  - Loss of a loop could easily lead to a deletion
  - Amplifications occur through more complex processes (may require assistance form homologous or non-homologous chromosomes not related to intra-chromosomal spatial proximity during interphase)

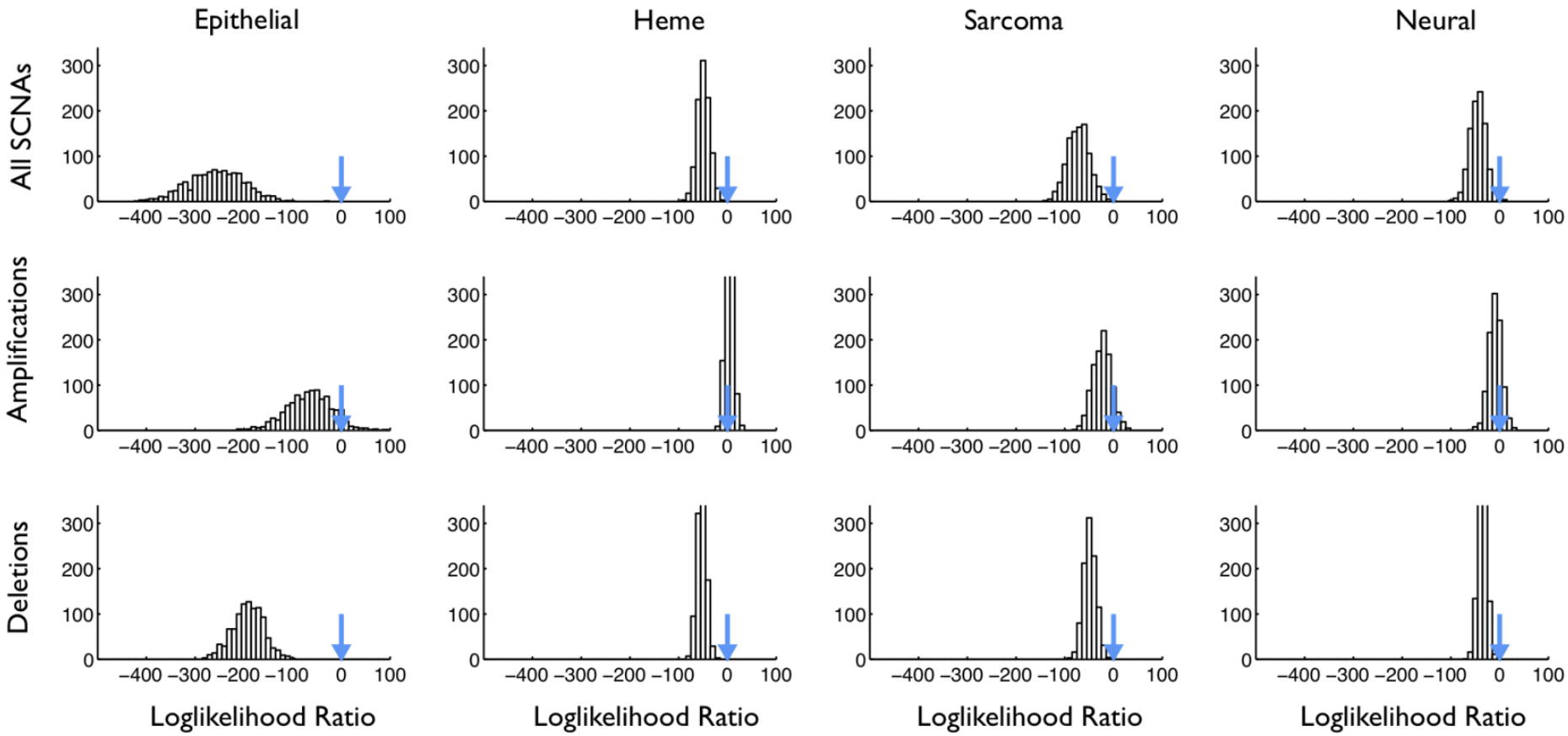
Permutation Analysis





Distribution of Loglikelihood Ratios by chromosome

Permutation per Cell Type (Cancer Lineage)



Deletions are more significant in cancers than amplifications.

Conclusions

- The probability of a 3D contact between 2 loci based on FG model explains best the length distribution of SCNAs.
- There is significant connection between megabase-level, position specific information, 3D chromatin structure observed in HiC and SCNA.
- SCNA data reflect mutational mechanisms and purifying selection in addition to commonly considered positive selection.
- Experimental evidence support physical proximity as a key factor of genomic rearrangements.
- Cell type specific experimental 3D contacts (HiC) fit the observed distribution of SCNAs better than the fractal globule model.