Introduction to Cancer Bioinformatics and cancer biology

Anthony Gitter

Cancer Bioinformatics (BMI 826/CS 838)

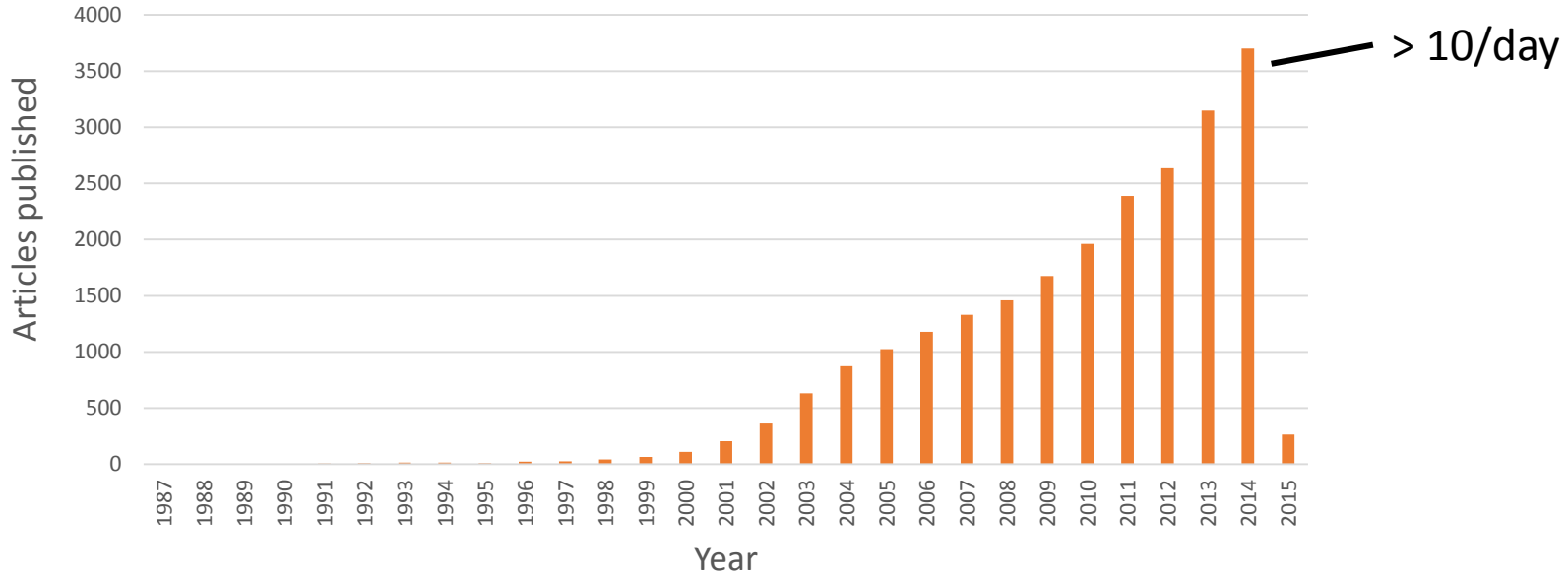
January 20, 2015

Why cancer bioinformatics?

- Devastating disease, no cure on the horizon
- Major focus of large-scale genomics efforts
  - Abundant data, interpretation is the challenge
  - Problem is complex data not “big data”
- Real impact on basic and translational research
  - Heavy computational components to high-profile biology papers
  - Purely computational/statistical papers can have impact
    - January 2, 2015 in *Science* Tomasetti2015

Caveats in cancer bioinformatics

- 27173 “cancer bioinformatics” papers in PubMed



- Easy to make predictions, hard to support them
  - Many genes have some relation to cancer

Introductions

- Professor Anthony Gitter
  - Assistant professor in Biostatistics and Medical Informatics
  - Affiliate faculty in Computer Sciences
  - Investigator in Morgridge Institute
- Class introductions
  - Home department
  - Background (BMI/CS 576?)
  - Research
  - Interest in Cancer Bioinformatics

Course overview

- ***Interactive*** discussions of research papers
- Grades
  - Presentation and discussion of research papers: 60%
  - Project: 40%
- No textbook
- Office hours by appointment
- Review syllabus at https://www.biostat.wisc.edu/~gitter/BMI826-S15

Presentations

- Journal club/reading group style presentations
  - Minimal slides
  - Figures or equations from paper
- May require reading some referenced papers
- Your thoughts on areas for improvement or future work
- Schedule online later today

Project

- Experience making real predictions with real data
- Groups of 2-3 students
- Ideas will be presented in class
- Tentative schedule
  - 2/24: Ideas presented
  - 3/10: Proposals due
  - 4/9: Status report due
  - 5/7: In class presentations
  - 5/11: Reports due

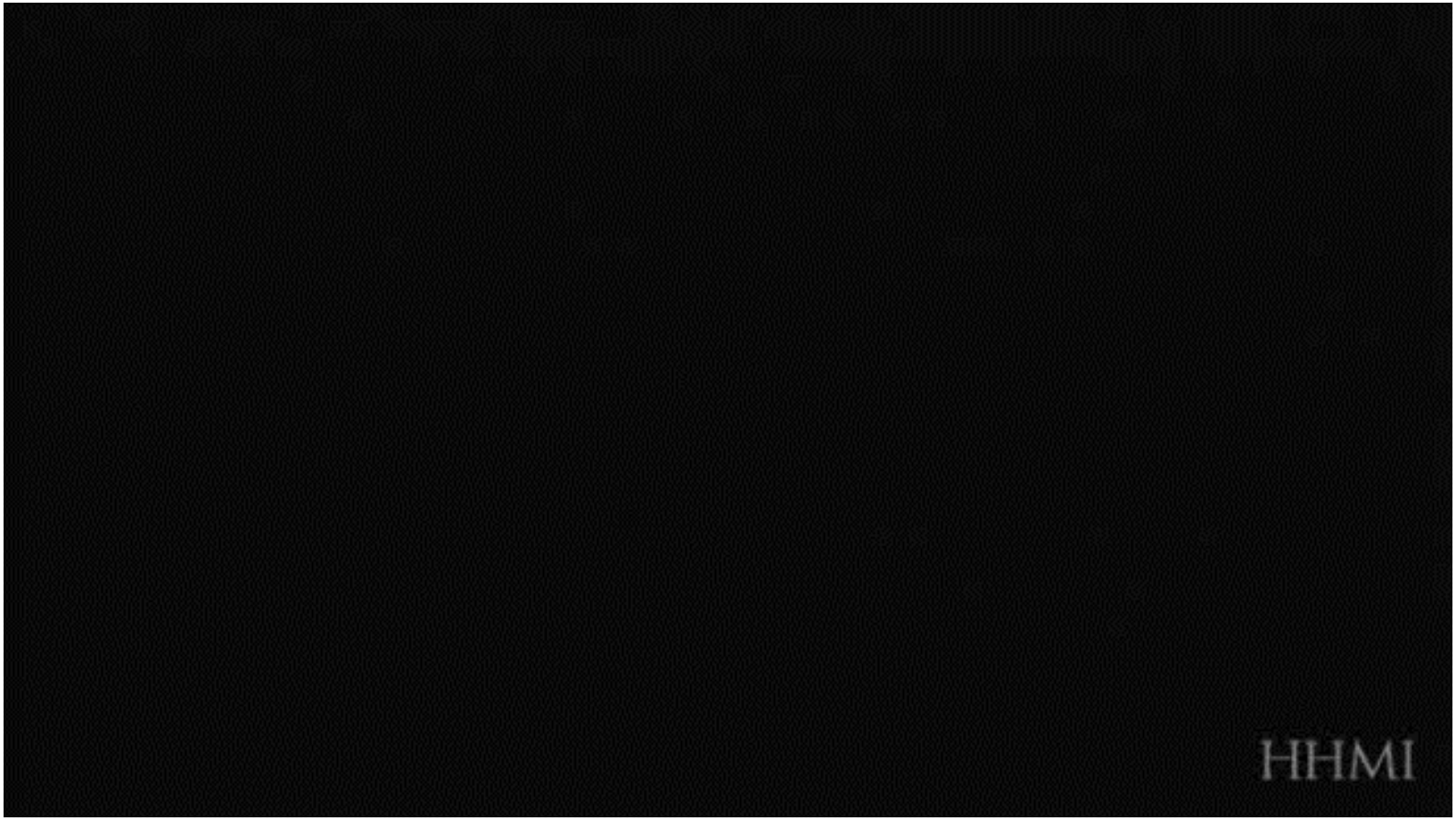
What is cancer?

- Normal cells acquire deficiencies
  - Grow and divide more than their peer normal cells
  - Overcome body's defense mechanisms
- Over time, become increasingly abnormal, invasive, and detrimental

Causes of cancer

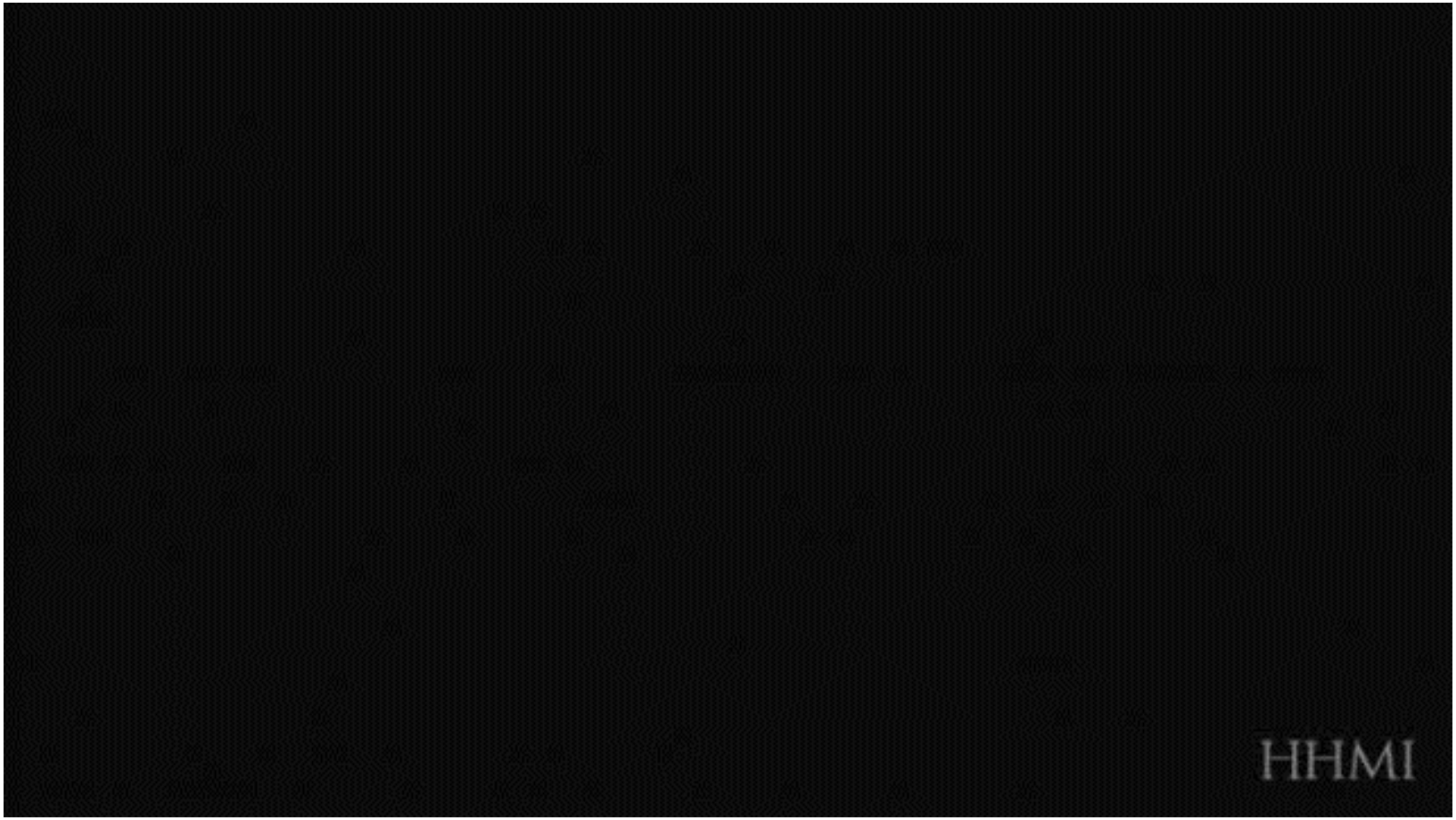
- Germline genetics
  - Predisposition due to inherited DNA
  - E.g. BRCA1/BRCA2 mutations increase risk of breast and ovarian cancer
- Environmental factors
  - Carcinogens: smoke, asbestos, UV radiation
  - Viruses
- Somatic alterations
  - E.g. non-inherited DNA mutations
  - Recent estimates (Tomasetti2015) that 65% of differences in cancer risk explained by errors in DNA replication (stem cell division)

DNA damage



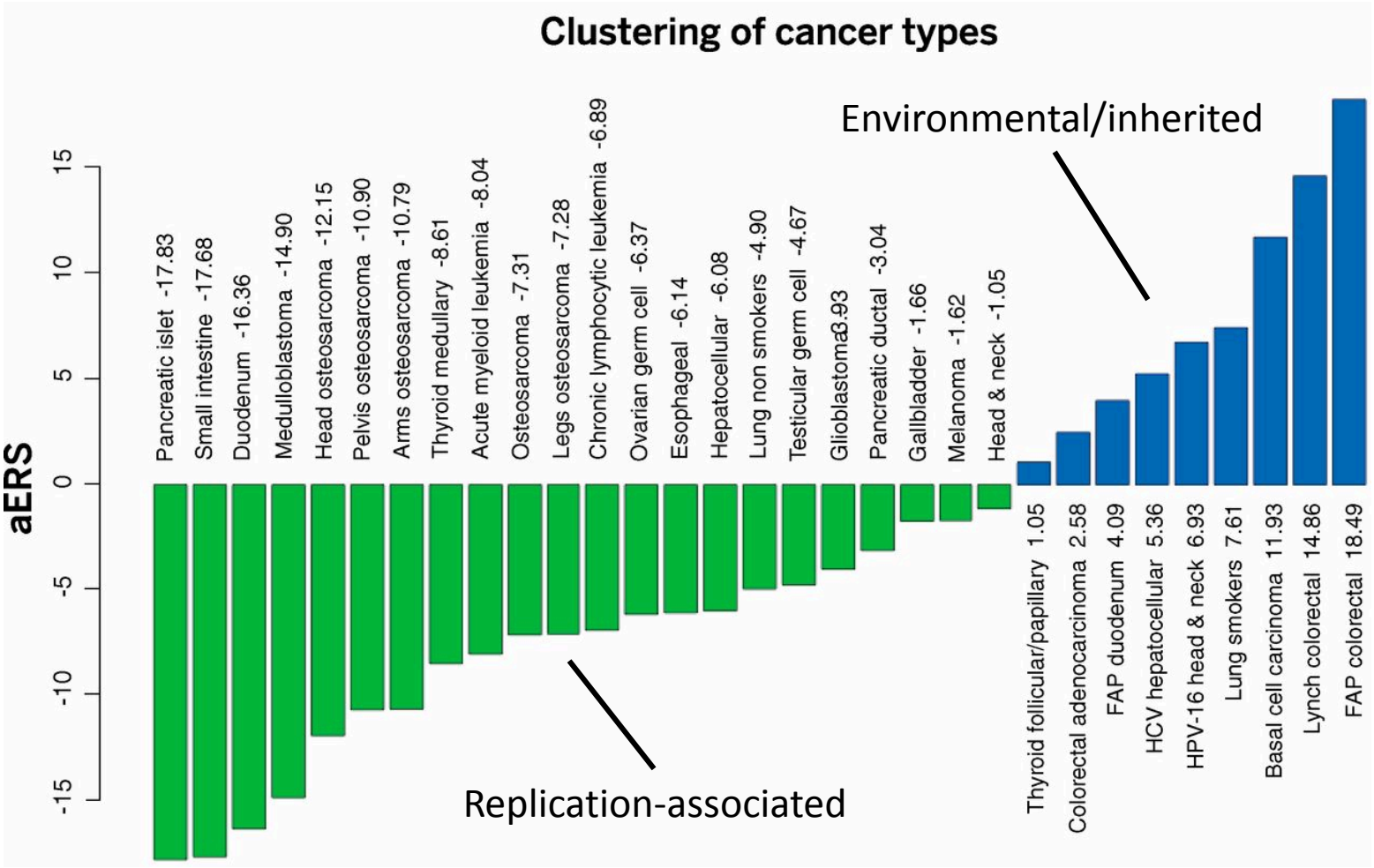
HHMI BioInteractive

DNA replication



HHMI BioInteractive

Variation in cancer risk among tissues can be explained by the number of stem cell divisions



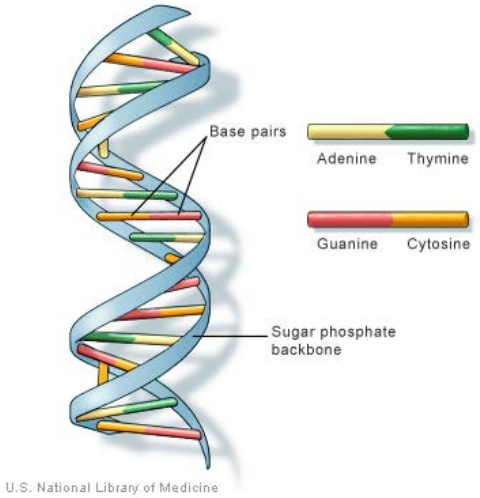
Tomasetti2015

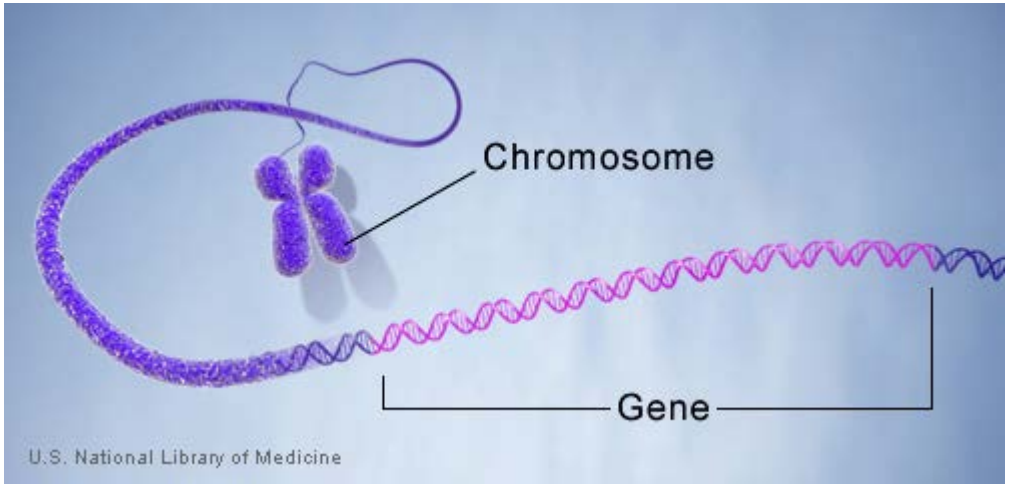
Causes of cancer continued

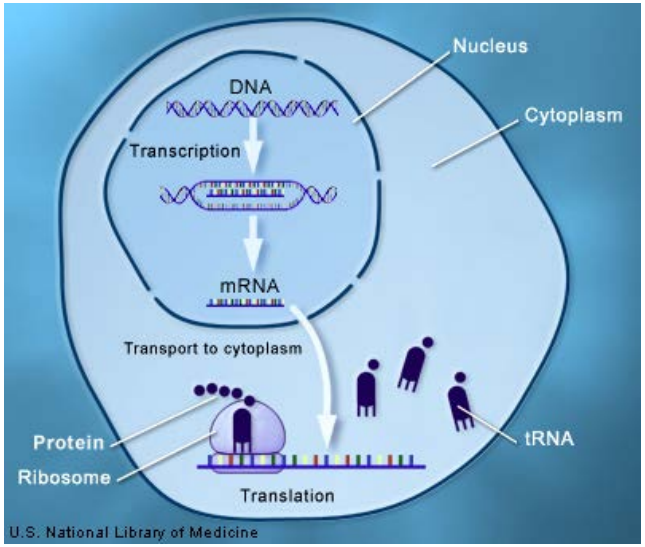
- Causes include:
  - Germline genetics
  - Environmental factors
  - Somatic alterations

- Specific causes have the same net effect
- Confer a growth advantage upon the mutated or altered cells

Basic units of a genome and cell

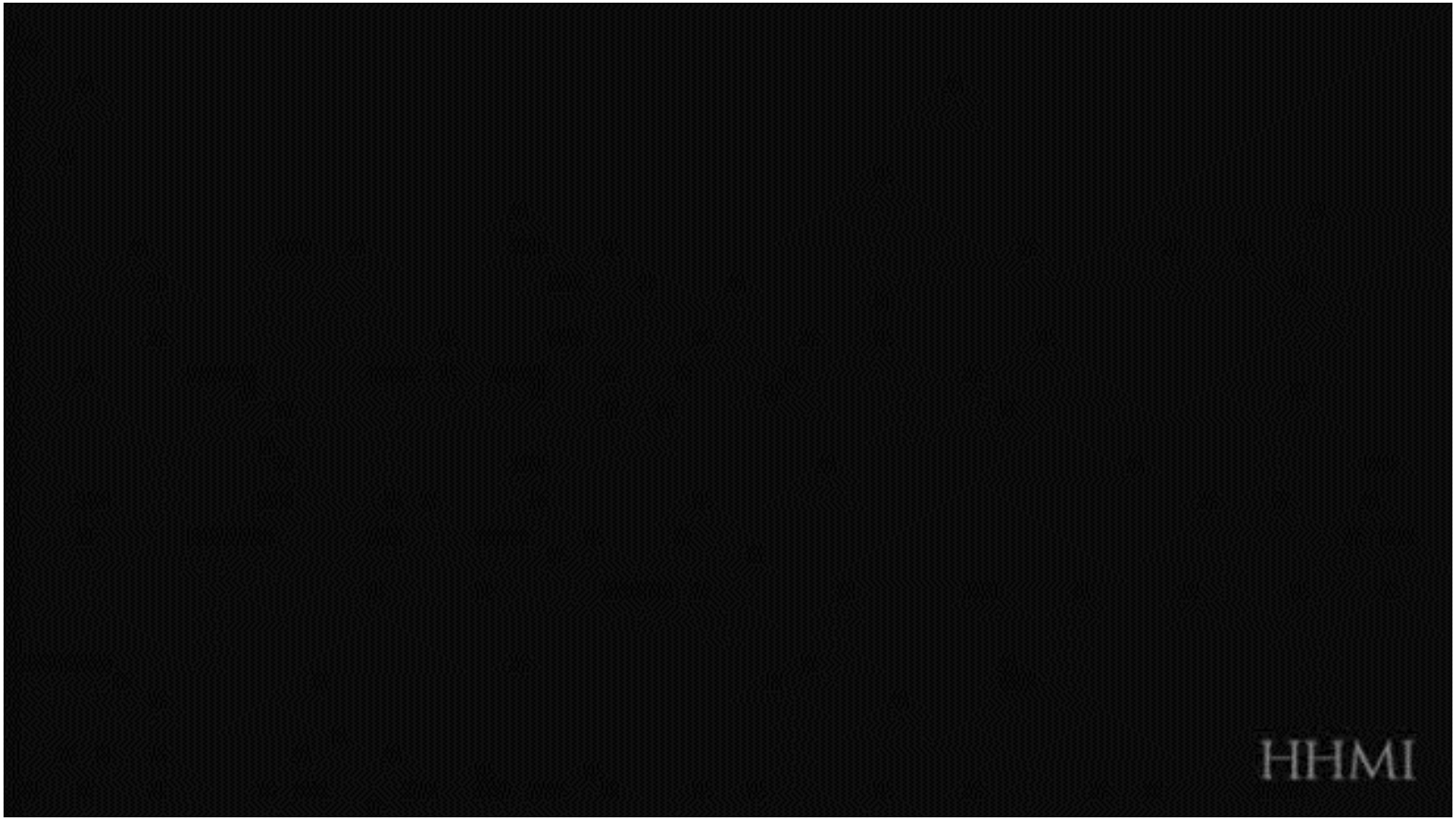






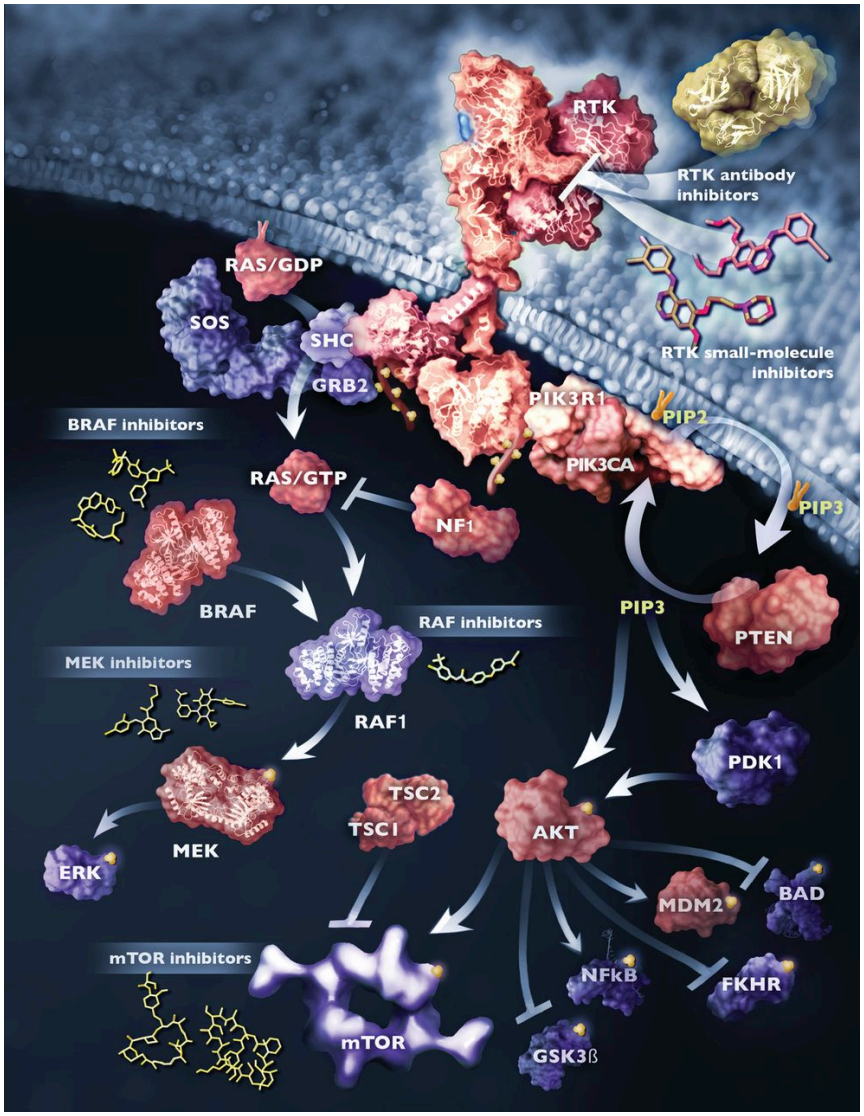
[NLM Handbook]

Transcription



HHMI BioInteractive

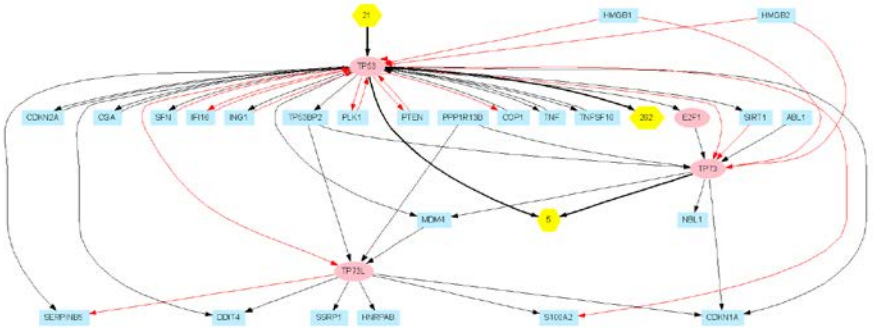
Signaling



Vogelstein2013

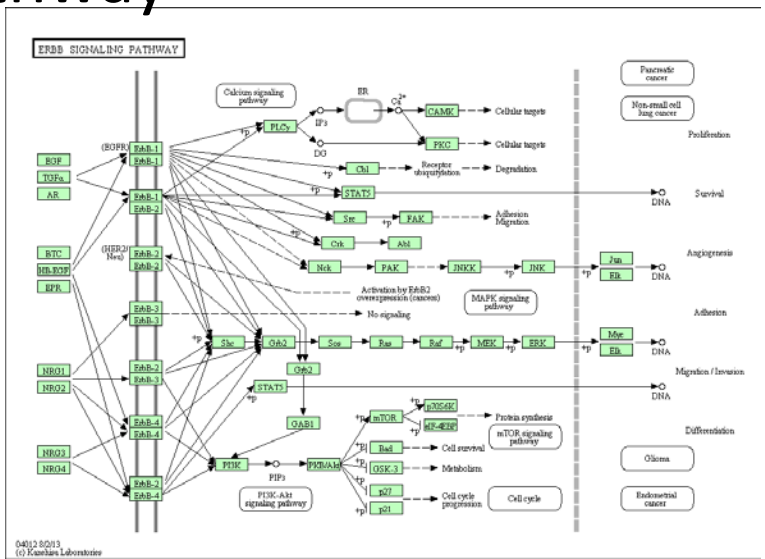
Abstracting transcriptional and signaling networks

- Transcriptional regulation



TRED

- Signaling pathway

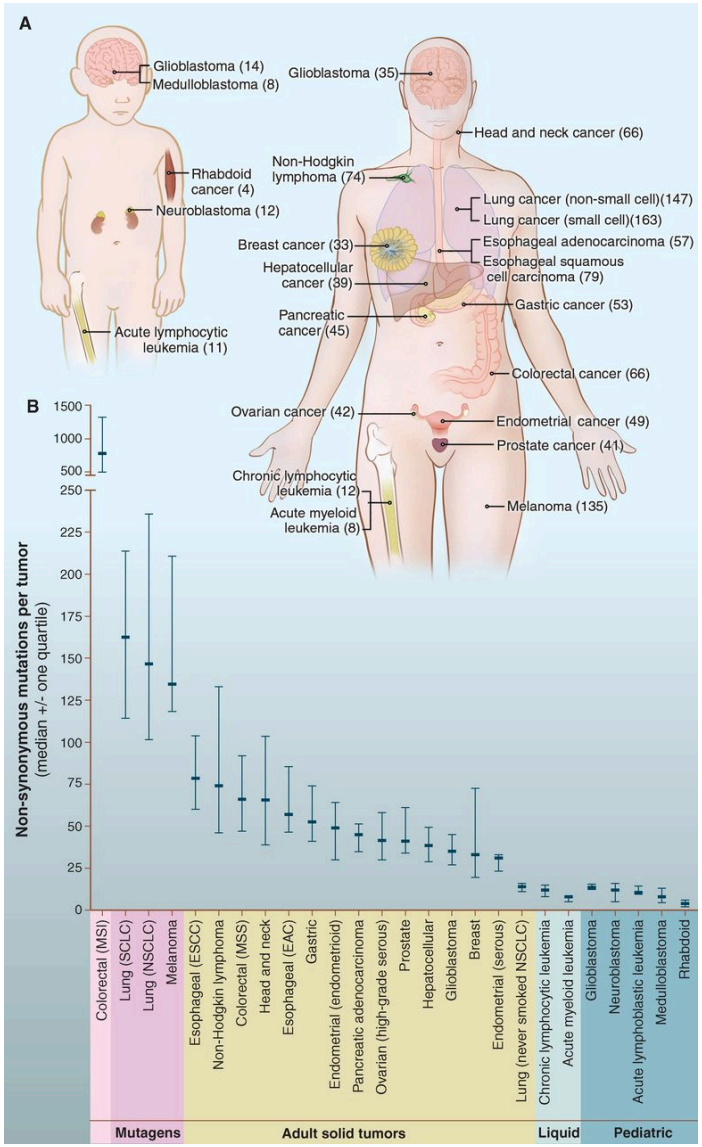


KEGG

Types of genomic abnormalities

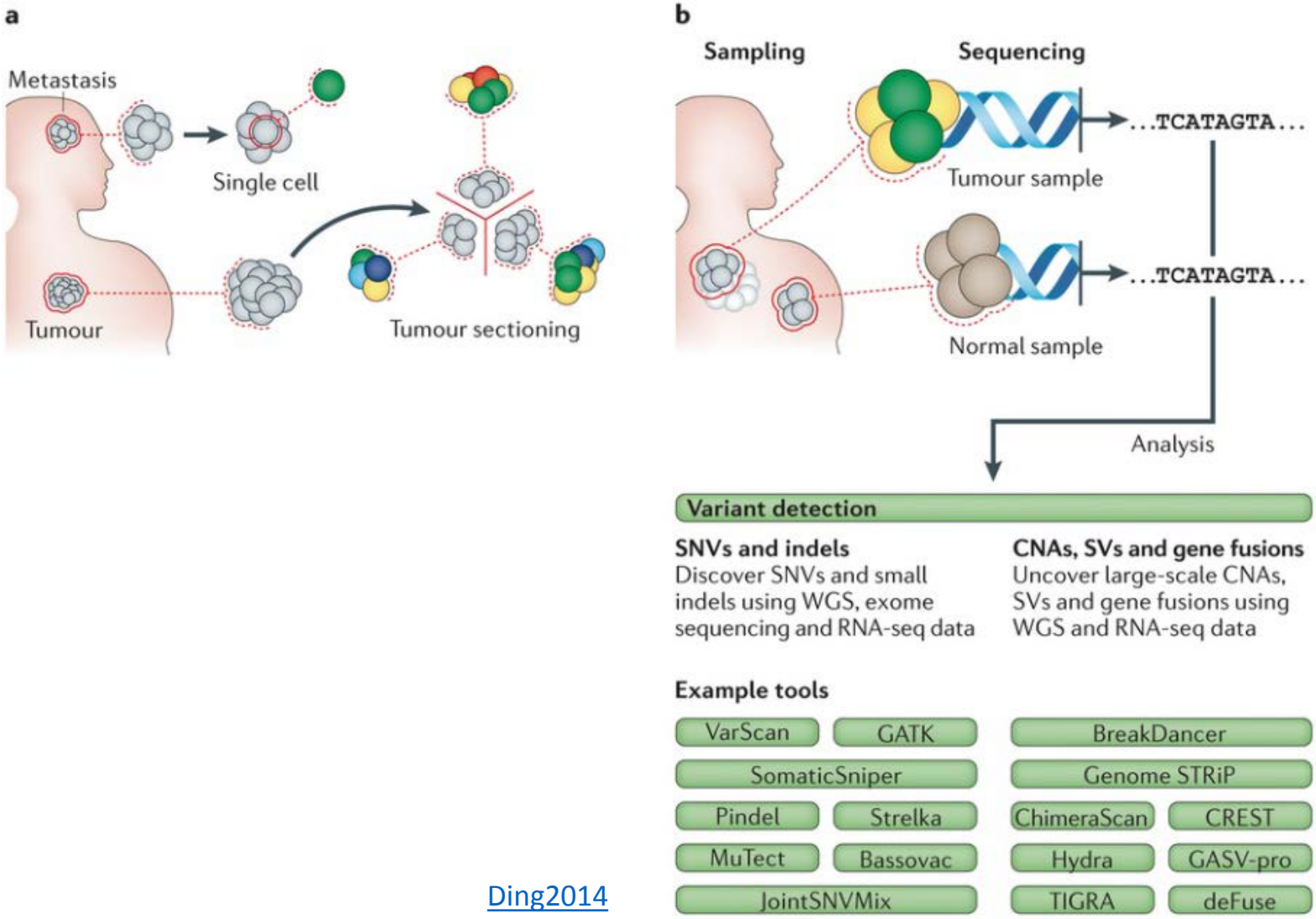
- DNA mutations
  - **Silent**: do not change amino acid
  - **Missense**: modified codon creates new amino acid
  - **Nonsense**: premature stop codon, truncates protein
  - **Insertion/deletion** (indel)
- Copy number changes
  - **Amplifications**
  - **Deletions**
- Translocations (gene fusions)

Number of somatic mutations



[Vogelstein2013]

Detecting genomic alterations

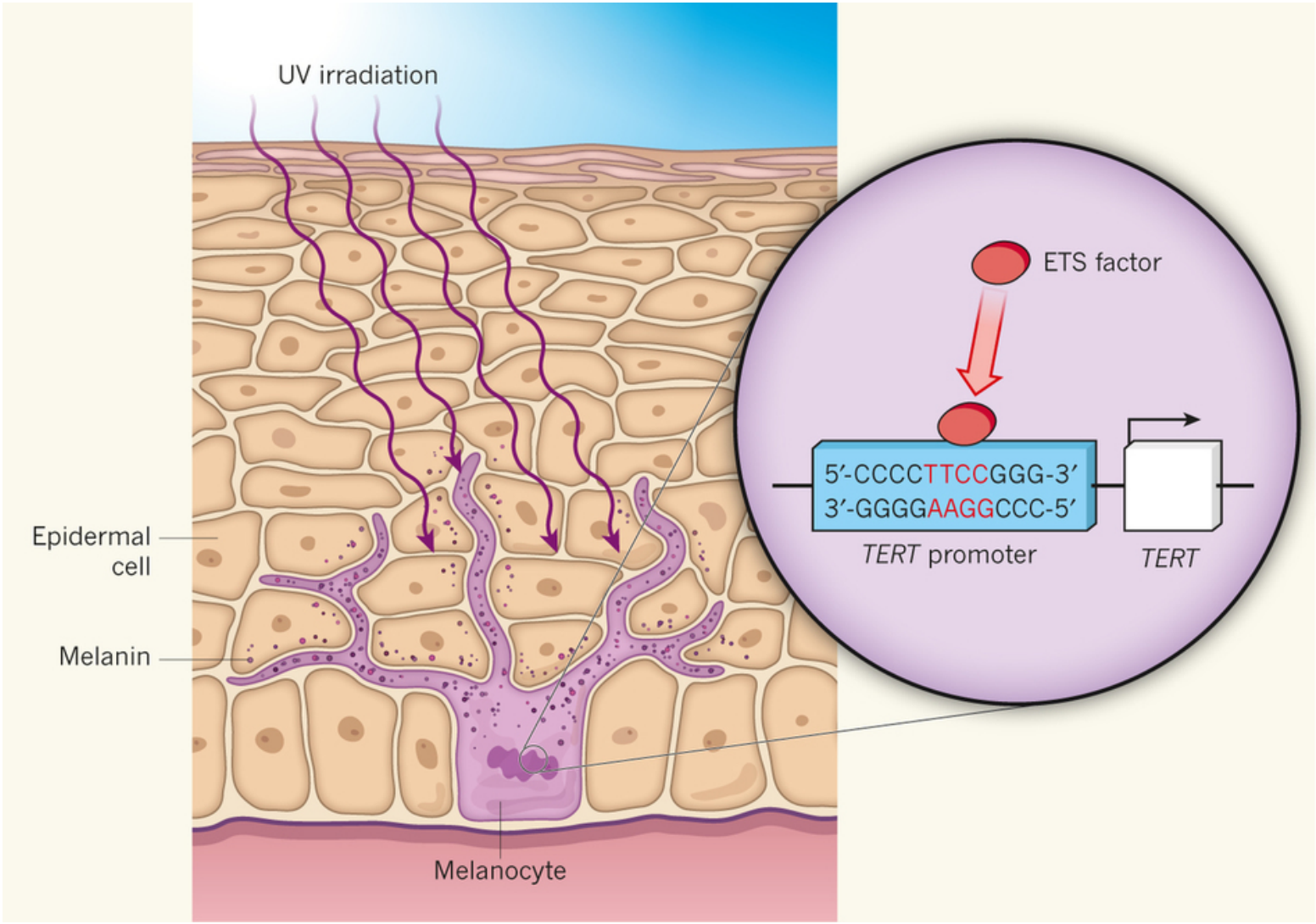


Ding2014

Other types of perturbations

- Not all mutations directly affect coding sequence of a gene
  - Splicing
  - Transcription
- Chain reactions along pathways and networks
- **Epigenetic**: Changes in gene expression or cellular phenotype caused by mechanisms other than changes in the DNA sequence (Vogelstein2013)
  - DNA methylation

Promoter mutations



Patton2013

Cancer progression

- Mutations accumulate over time
  - More environmental exposures
  - More DNA replication cycles

- Benign tumor -> malignant tumor -> metastasis
- Cancer stages (Cancer Institute):
  - **Stage 0**: Local tumor in tissue of origin
  - **Stage 1**: Invades neighboring tissue
  - **Stage 2/3**: Regional spread, lymph nodes
  - **Stage 4**: Distant spread

Barriers to treating cancer

- Distinguishing root cause
- Several types of heterogeneity
- Redundancy of signaling pathways, resiliency to treatment
- Metastasis

Driver vs. passenger

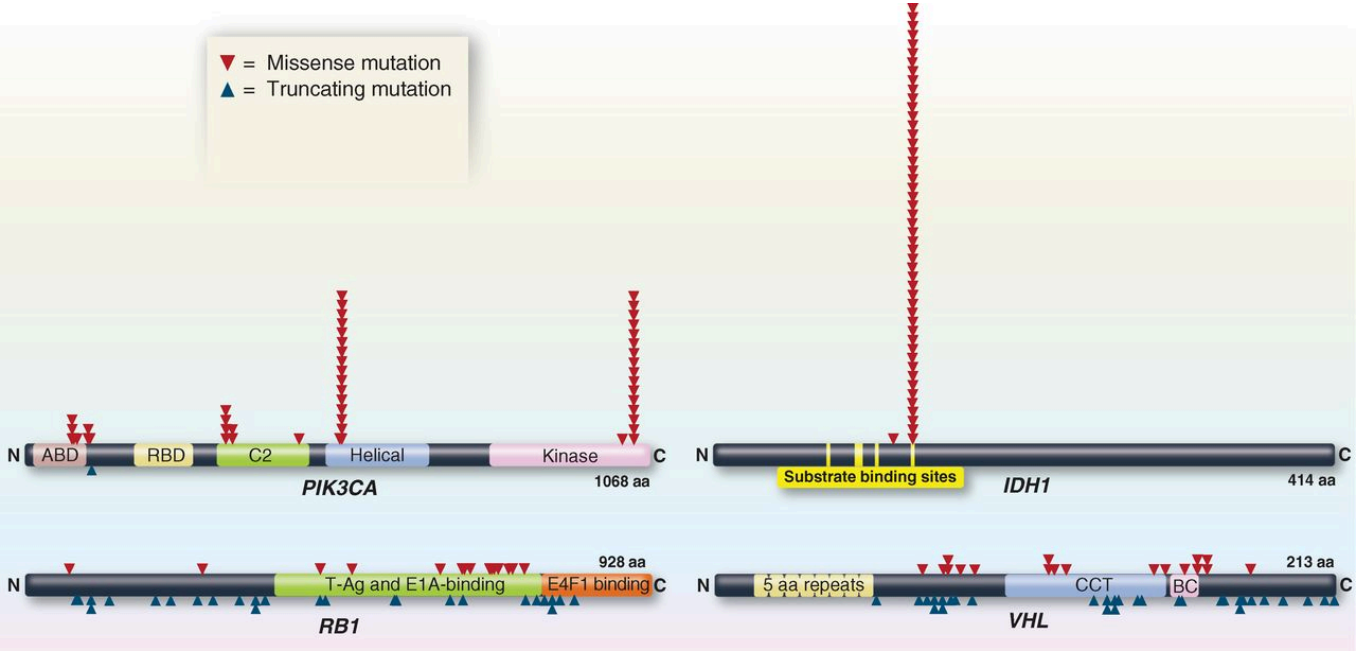
- Cancer cells disable DNA protection mechanisms
  - Acquire many more mutations
  - Most of them are not causal
- **Driver**: Confers selective growth advantage
- **Passenger**: No impact on growth
- Recent estimates ~100s of drivers across cancers
  - 138 (Vogelstein2013)
  - 291 (Tamborero2013)

Distinguishing driver vs. passenger

- Strategies for identifying the driver mutations (Ding2014)
  - Recurrence and frequency assessment
  - Variant effect prediction
  - Pathway or network analysis

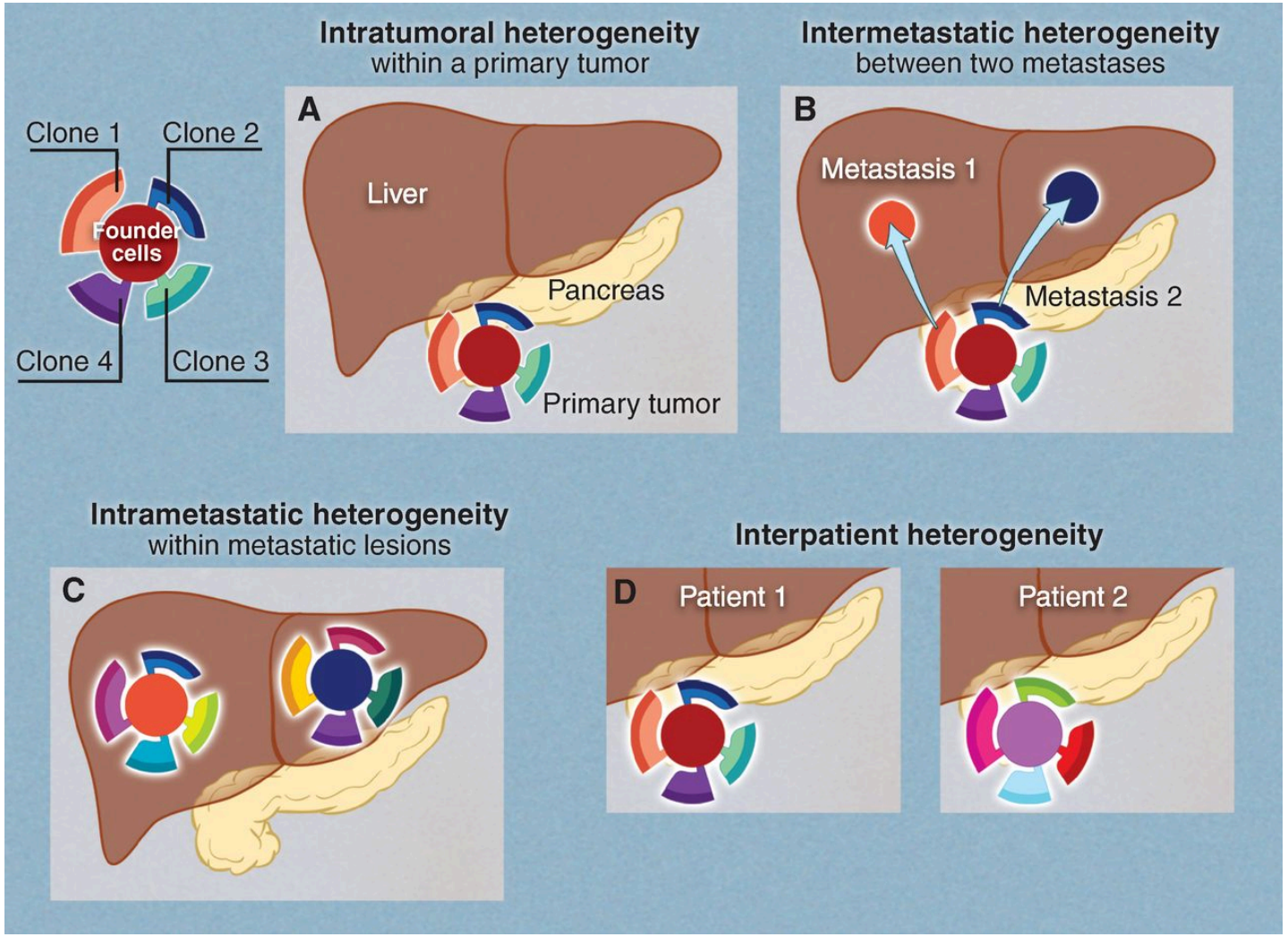
Oncogenes vs. tumor suppressors

- **Oncogene**: mutation activates, increases selective growth advantage
- **Tumor suppressor:** mutation inactivates (often truncates), increases selective growth advantage



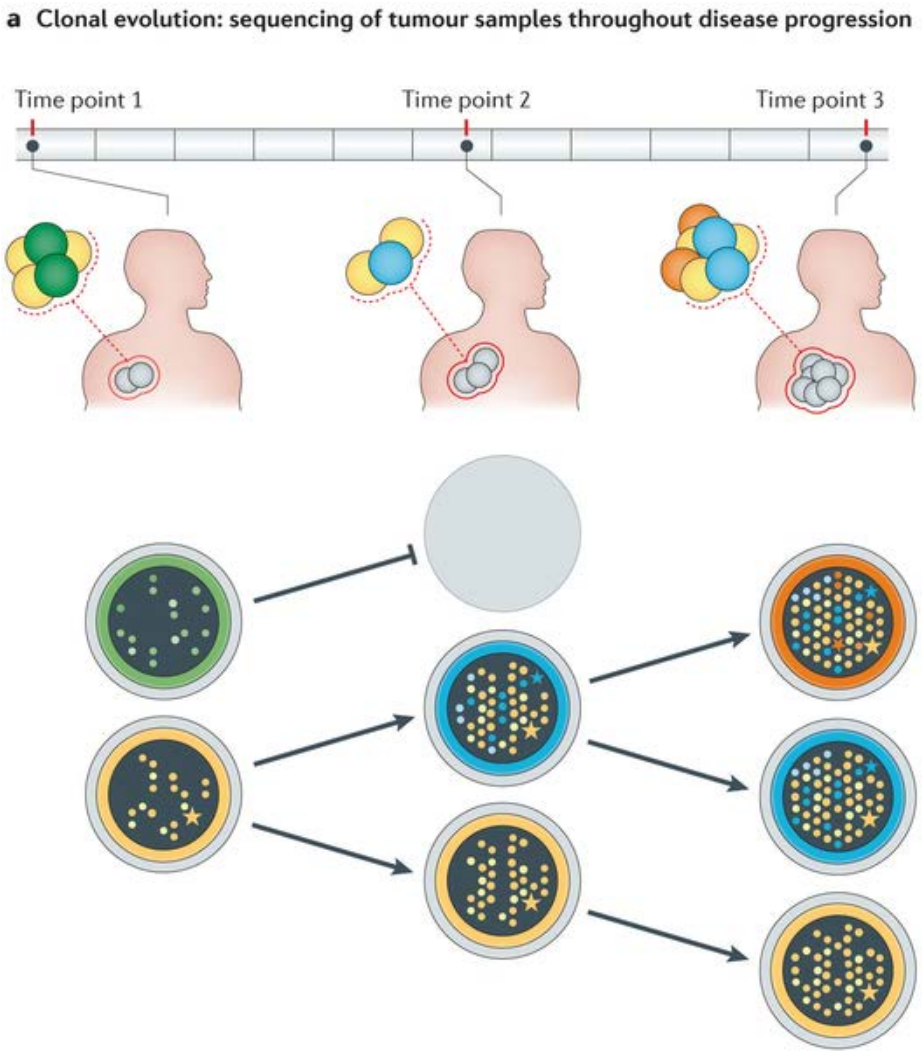
Vogelstein2013

Types of heterogeneity



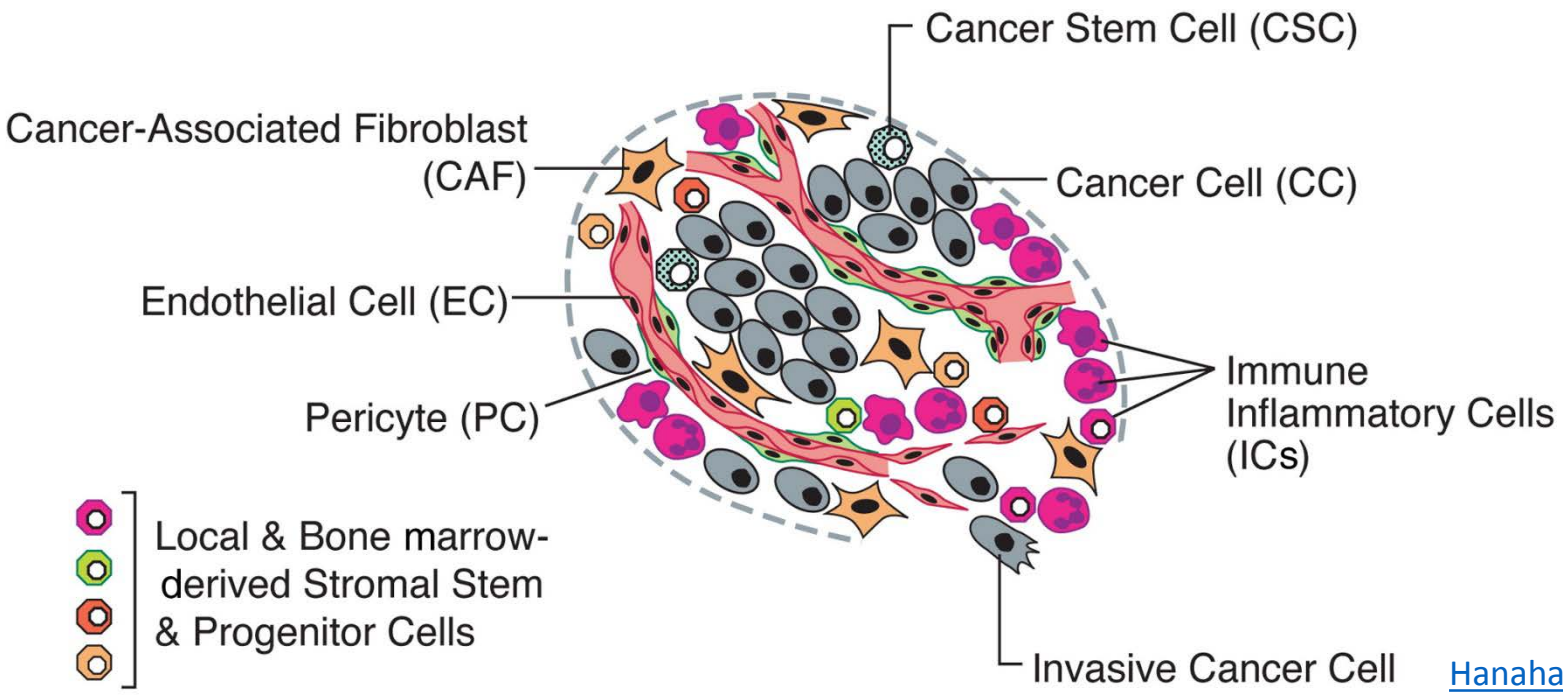
Vogelstein2013

Tumor evolution



Ding2014

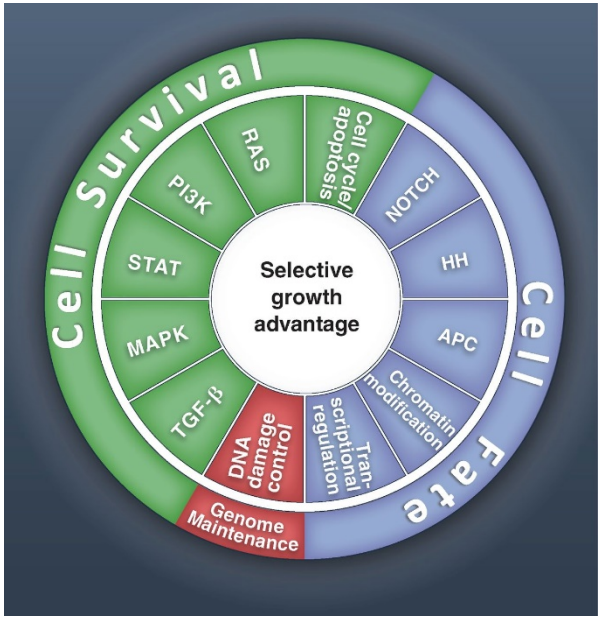
Tumors are mixtures of cell types



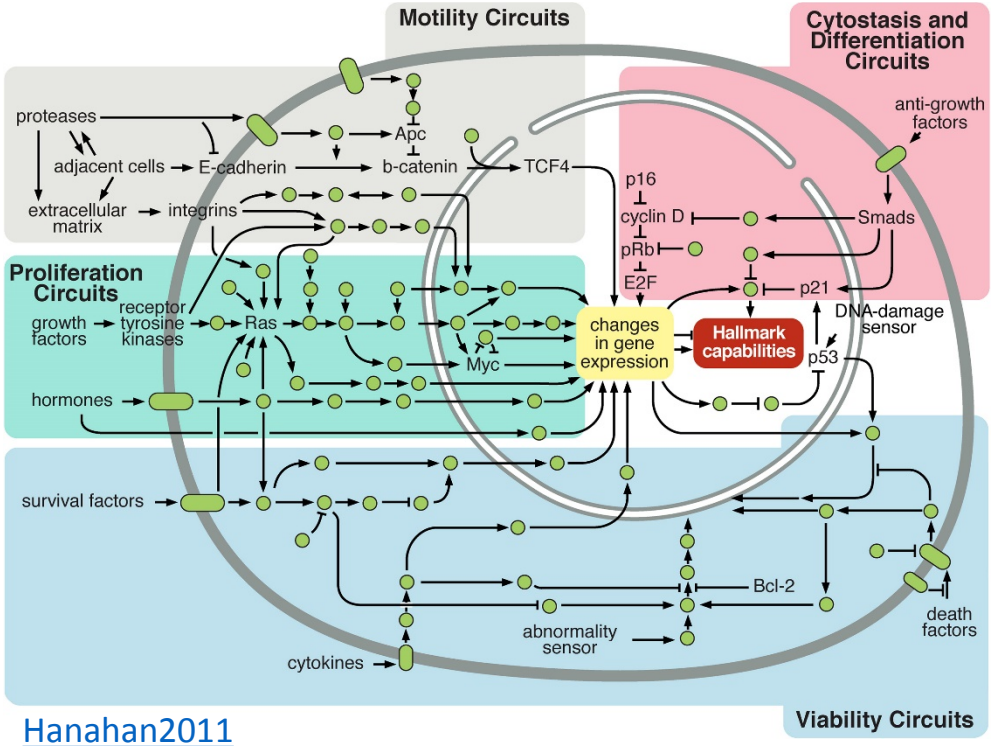
Hanahan2011

Convergence of driver events

- Amid the complexity and heterogeneity, there is some order
- Finite number of major pathways that are affected by drivers



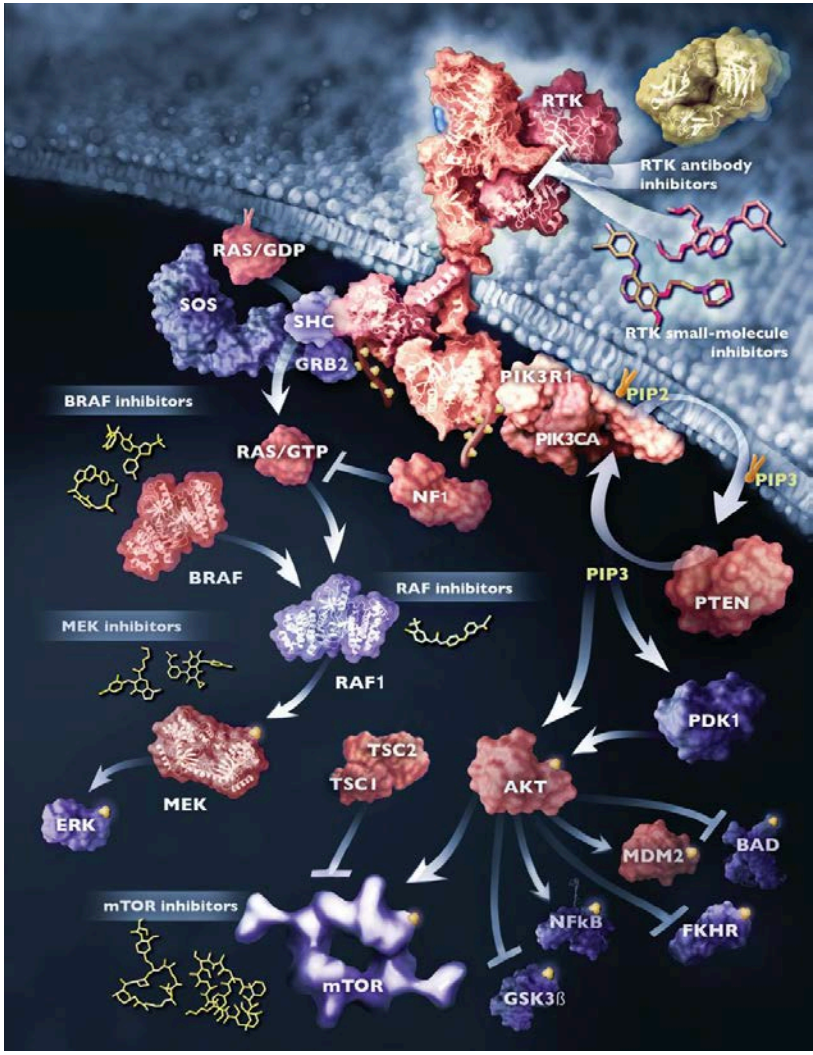
Vogelstein2013



Hanahan2011

Similar pathway effects

- Tumor 1: EGFR receptor mutation makes it hypersensitive
- Tumor 2: KRAS hyperactive
- Tumor 3: NF1 inactivated and no longer modulates KRAS
- Tumor 4: BRAF over responsive to KRAS signals



Vogelstein2013

References and glossary

- Material in these slides based upon
  - Hanahan2011
  - Vogelstein2013
  - Ding2014

- Vogelstein2013 contains an excellent glossary of cancer terms