

# Michael S. Lawrence, PhD



## Lawrence Laboratory

Gabriel Golczer, PhD  
Julian Hess, BS\*  
Adam Langenbucher, BS  
Michael S. Lawrence, PhD  
Rob Morris, PhD  
Jan F. Sayilgan, BSc  
Jillian Wise, PhD  
Ben Wittner, PhD

\*Associate Computational Biologist  
based at The Broad Institute

Cancer results from alterations to DNA that lead to the activation of oncogenes or the inactivation of tumor suppressors. **The Lawrence laboratory** focuses on understanding the many ways this can happen, using computation as a powerful microscope to study the processes of DNA damage and repair, gene expression and genome replication, and cancer driver genes. Over our lifetimes, DNA slowly accumulates mutations due to environmental toxins and radiation, as well as from naturally occurring copying errors. The vast majority of mutations have little or no effect on a cell, but out of all possible mutations, a few may hit exactly the right place in the genome, where they can act as a “driver mutation,” pushing the cell toward aggressive growth and tumor formation. Sequencing the DNA in a tumor reveals not only its driver mutations, but also all the other “passenger mutations” that were present in the tumor-initiating cell. We seek insights about cancer from both driver and passenger mutations.

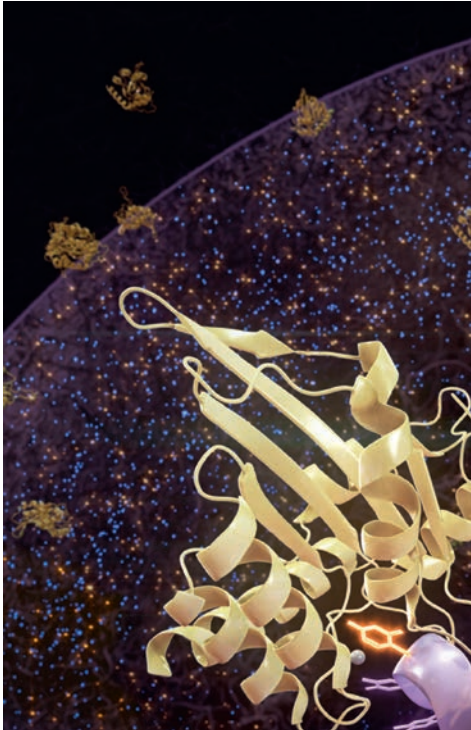
## Tumor DNA Sequencing

High-throughput DNA sequencing is a workhorse of biomedical research. There are many challenges in processing the raw DNA sequencing reads from a patient’s resected tumor or biopsy material, aligning them accurately to the reference human genome, and then scanning for loci where the tumor DNA differs from the patient’s bulk “normal” DNA (e.g. from a blood draw). Distinguishing true somatic mutations from sequencing or alignment artifacts can be tricky, especially for subclonal events present in only a fraction of tumor cells. We are refining a “panel of normals” (PoN) approach, which combats stochastic artifacts seen in the patient’s tumor sample, and not in the patient’s normal sample but widespread however in many other patients’ normal samples. We are continually discovering new artifact modes, making this a highly challenging and unpredictable area of research. Isolating true somatic mutations is crucial for downstream analyses of mutational signatures and driver events.

## Analyzing Mutational Signatures

Cancers vary over many orders of magnitude

in their total background mutation burden, ranging from very quiet tumor types such as leukemias and childhood tumors, which may have fewer than 10 somatic mutations in their exome, to carcinogen-associated tumor types such as lung cancer and melanoma, which may have over 1000. Mutations have many causes, and each mutagen can leave a telltale signature. For instance, spontaneous deamination of methylated CpG’s causes the transition mutations that dominate many tumor types. Mutagens in tobacco smoke cause G-to-T transversions. Ultraviolet radiation causes C-to-T at dipyrimidines. Agitated APOBEC enzymes cause mutations at C’s preceded by T. Loss of mismatch repair causes microsatellite instability (MSI), marked by expansion and contraction of simple-sequence repeats, as well as characteristic types of single-base changes. Tumors carrying mutations in the proofreading exonuclease domain of polymerase epsilon (POLE) tend to accrue C-to-A mutations at the trinucleotide TCT. Very rare “MSI+POLE” cancers show the highest yet known somatic mutation burdens, with upwards of 10,000 coding mutations per patient. Patients



*The mutational landscape of a cancer cell across size regimes. At the smallest scale, local DNA trinucleotide sequences (lower-left foreground) correlate with the “mutational signatures” induced by various mutagens. At the largest scale (background of image), chromatin is organized into multi-megabase domains comprising Compartment B (tightly packed, gene-poor DNA lining the nuclear periphery) and Compartment A (gene-rich open DNA in the nuclear interior). Mutations induced by APOBEC enzymes (yellow points) are distributed equally across the two compartments, but most other types of mutations (blue points) are concentrated in Compartment B. Between the large and small extremes lies the “mesoscale” regime, where genomic features like hairpin-forming ability are determined. DNA exposed in a hairpin loop is vulnerable to attack by the enzyme APOBEC3A (center), giving rise to highly recurrent passenger mutations in cancer.*

affected by MSI and/or POLE mutagenesis are known to experience better clinical outcomes, probably thanks to their high neoantigen loads which attract a powerful immune response. Our most recent research has focused on a less well-studied signal in somatic mutation datasets, mutational asymmetries between the two DNA strands. These illuminate transcriptional or “T-class” mutational patterns, associated with exposure to tobacco smoke, UV radiation, and a yet-unknown agent in liver cancer, as well as replicative or “R-class” patterns, associated with MSI, APOBEC, POLE, and a yet-unknown agent in esophageal cancer.

### **APOBEC Mutations and Mesoscale Genomic Features**

Statistical approaches for distinguishing driver mutations from passenger mutations have relied on the gold standard of recurrence across patients. Seeing exactly the same DNA base-pair mutated recurrently across patients has been taken as proof that the mutation must be under functional selection for contributing to tumor fitness. The assumption is that mutational processes, being essentially

random, are unlikely to hit the exact same base-pair over and over again. Our recent discoveries about APOBEC mutagenesis have cast doubt on this assumption. We have shown that APOBEC3A has a very strong preference for mutating cytosines presented in a short loop at the end of a strongly paired DNA hairpins. DNA hairpins occupy the genomic “mesoscale” regime, being larger than the trinucleotides that define mutation signatures, yet smaller than chromatin topological domains. Our results indicate that there are multiple routes to cancer mutational hotspots. Driver mutation hotspots in oncogenes can rise to prominence through positive selection, and are not restricted to the “favorite” sites of any particular mutagen. In contrast, special DNA sites (like hairpins) that happen to be optimal substrates for a mutagen (like APOBEC) can give rise to “passenger hotspot mutations” that owe their prevalence to substrate optimality, not to any effects on tumor fitness. These findings mean we need to be careful not to assume that all mutation hotspots are cancer drivers.

### **Selected Publications:**

- Buisson R, Langenbucher A, Bowen D, Kwan EE, Benes CH, Zou L\*, **Lawrence MS\***. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale features. *Science*. 2019 Jun 28; 364(6447):eaaw2872.
- Haradhvala NJ, Kim J, Maruvka YE, Polak P, Rosebrock D, Livitz D, Hess JM, Leshchiner I, Kamburov A, Mouw KW, **Lawrence MS**, Getz G. Distinct mutational signatures characterize concurrent loss of polymerase proofreading and mismatch repair. *Nat Commun*. 2018 May 1;9(1):1746.
- Buisson R, **Lawrence MS**, Benes C, Zou L. APOBEC3A and APOBEC3B activities render cancer cells susceptible to ATR inhibition. *Cancer Res*. 2017 Jul 11.
- Haradhvala NJ, Polak P, Stojanov P, Covington KR, Shinbrot E, Hess JM, Rheinbay E, Kim J, Maruvka YE, Braunstein LZ, Kamburov A, Hanawalt PC, Wheeler DA, Koren A, **Lawrence MS\***, Getz G\*. Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. *Cell*. 2016 Jan 28; 164(3):538-49.
- Lawrence MS\***, Stojanov P\*, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 2014 Jan 23; 505(7484):495-501.

\*Co-authors