

Uncovering the Human cell lineage tree in health and disease: The next grand scientific challenge

Ehud Shapiro

Department of Computer Science and Applied Math and
Department of Biological Chemistry
Weizmann Institute of Science

Abstract

The cell lineage tree of a person captures the history of the person's cells since conception. In computer science terms it is a rooted, labeled binary tree, where the root represents the primary fertilized egg, leaves represent extant cells, internal nodes represent past cell divisions, and vertex labels record cell types. It has approximately 100 trillion leaves and 100 trillion branches ($\approx 100,000$ bigger than the Human genome); it is unknown.

We should strive to know it, as many central questions in biology and medicine are actually specific questions about the Human cell lineage tree, in health and disease: Which cancer cells initiate relapse after chemotherapy? Which cancer cells can metastasize? Do insulin-producing beta cells renew in healthy adults? Do eggs renew in adult females? Which cells renew in healthy and in unhealthy adult brain? Knowing the Human cell lineage tree would answer all these questions and more.

Fortunately, our cell lineage tree is implicitly encoded in our cells' genomes via mutations that accumulate when body cells divide. Theoretically, it could be reconstructed with high precision by sequencing every cell in our body, at a prohibitive cost. Practically, analyzing only highly-mutable fragments of the genome is sufficient for cell lineage reconstruction. Our lab has developed a proof-of-concept method and system for cell lineage analysis from somatic mutations. The talk will describe the system and results obtained with it so far, and future plans for this project.