

Somatic Cell Genealogies and Differentiation

Olinia Geoffrey Martijn

African Institute for Mathematical Sciences

2008

Outline

The Meaning of Cell Genealogy

The meaning of stem

The meaning of Zygote, Mitotic ages and Genome

Data

Methods

Applications

Human Hair

Mouse Cancer

Future Work

Key Dates

References

What is Cell Genealogy?

- Genealogy refers to the study of family history, including the study of who the ancestors of a particular person were. Genealogy of many cells starts from the zygote. The genealogy of many cells can be divided into three sequential phenotypic phases:
 - (i) development from zygote (neogenesis)
 - (ii) a stem cell phase (stem cell latency)
 - (iii) differentiation phase
- Development and differentiation are programmed and restricted to specific times and numbers of divisions. For many cell types, development only occurs during the first few months or years after conception, and differentiation from a stem cell also typically requires a set amount of time from several days to weeks. Numbers of divisions during these phases are pre-programmed or constant regardless of adult chronological age.

What is stem?

- Stem cell definitions vary, but one precise definition is that a stem cell is a common ancestor or progenitor of a group of present day differentiated cells.
- Stem cells defined by ancestry may differ from stem cells prospectively identified by experimental manipulations.
- A stem cell must be physically alive to be prospectively identified, but common ancestors are no longer physically present.

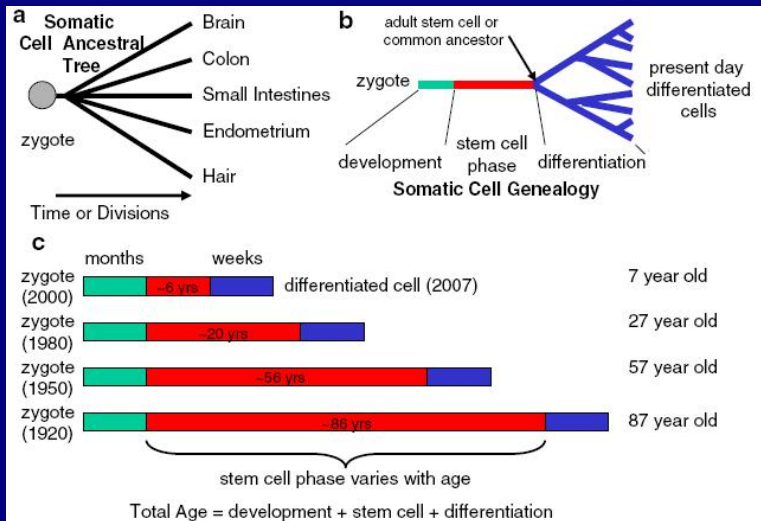


Figure: Somatic cell tree

What is Zygote, Mitotic ages and Genome?

- The Zygote refers to the cell formed by the union of a male and female sex cells.
- The total number of divisions since the zygote is known as mitotic ages.
- The Genome refers to the total amount of genetic information in the chromosomes of an organism, including its genes and DNA sequences. The genome of eukaryotes is made up of a single, haploid set of chromosomes that is contained in the nucleus of every cell and exists in two copies in the chromosomes of all cells except reproductive and red blood cells. The human genome is made up of about 35,000 genes.

Recent findings

- (i) A somatic cell tree starts from the zygote and ends with present day normal or neoplastic cells.
- (ii) In between are ancestors and dead ends, which functionally correspond to stem and nonstem cells.
- (iii) The human colon is approximately 5 ft long and composed of about 15 million clonal units called crypts. Each crypt contains about 2000 cells, which can be divided into stem and nonstem cells.
- (iv) The majority of cells are nonstem cells that differentiate and die in about a week as they migrate from the crypt base to the surface.
- (v) Multiple but an uncertain number of stem cells (from 2 to 40) reside at or near the base of the crypt. During a lifetime, one cell out of the billions of cells in the colon may transform, leading to colorectal carcinoma.

Data

Huge amount

- 10^{14} cells in the human body
- 10^9 base pairs in the human genome
- 10^{23} base pairs in the human body, only counting nuclear DNA.

Reducing data

- Focus on specific areas in the genome. eg MS

Methods

The methods can be categorised into three classes:

- Direct observation
- Invasive techniques
- Noninvasive techniques

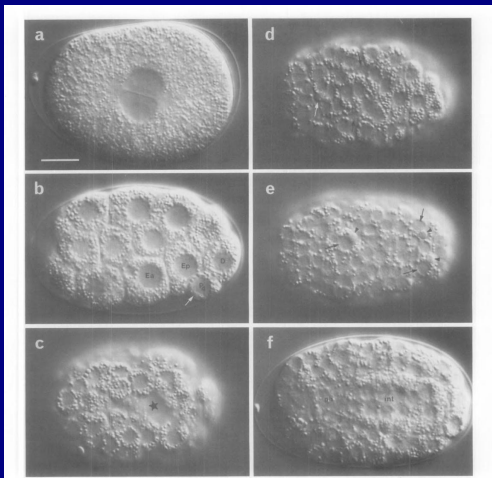
Methods

Direct Observation

- Direct observation can be used for small transparent organisms
- Sulston et al published the full cell lineage of the *C. elegans* in 1983
- Not possible for higher organisms



Sulston won the
Nobel Prize in
Medicine in 2002.



Methods

Direct Observation

- Direct observation can be used for small transparent organisms
- Sulston et al published the full cell lineage of the *C. elegans* in 1983
- Not possible for higher organisms

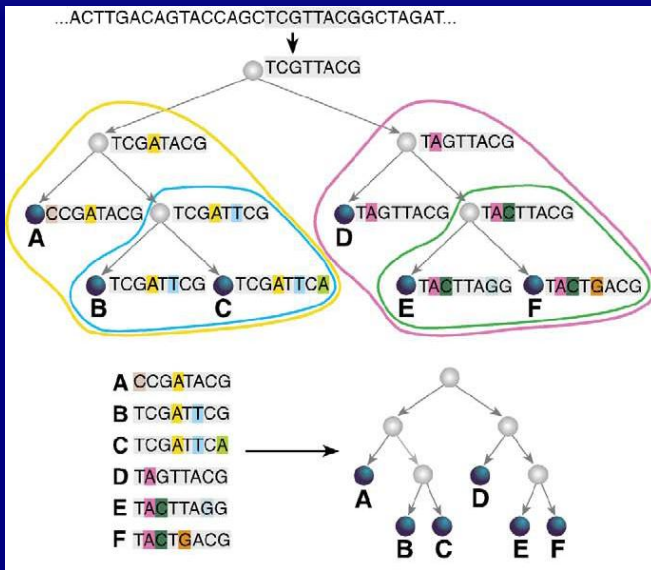
Methods

Invasive Techniques

- A single founder cell is marked with a heritable progeny. eg the injection of a tracer molecules or retroviral infection.
- This may interfere with the normal growth and function of the marked cell population.

Noninvasive Techniques

- Based upon spontaneous mutations in the founder cell.



Noninvasive techniques

Examples

- Loss or gain of large genomic fragments
- Mitochondrial DNA mutation
- Changes in the number of MS repeat units

Microsatellites

Repeated nucleotide units consisting of 1 to 6 base pairs.

GTC**AAC**AACAACAACAACAACAACAACAACGTC

Advantages of using MS

- MS slippage mutations occur during DNA replication and are coupled to cell division.
- Mutations occur at a high rate.
- MS mutations occur independently at different loci, not affecting phenotype and are unlikely to be selected against.
- High abundance in humans, mice and other organisms.

Exactly how viable is the usage of MSs to reconstruct cell lineage trees?

Theorem

Human cell lineage trees with a depth of 40 cell divisions with any topology can be reconstructed with no errors with a probability greater than 99.95%.

Uniform Model

- **all mutation events are statistically independent**
- the identifier (a vector representing MS lengths) of the root is known and used as a reference
- both daughter cells of the root lead to extant cells
- all loci have the same mutation probability
- mutations are stepwise with equal probability $+1$ and -1 .

Cell genealogy on Human Hair

- Stem cells divide to reproduce themselves and produce differentiated progeny.
- A fundamental problem in human biology has been the inability to measure how often stem cells divide. Although it is impossible to observe every division directly, one method for counting divisions is to count replication errors; the greater the number of divisions, the greater the numbers of errors. Stem cells with more divisions should produce progeny with more replication errors.
- To test this approach, epigenetic errors (methylation) in CpG-rich molecular clocks were measured from human hairs.
- Hairs exhibit growth and replacement cycles and "new" hairs physically reappear even on "old" heads.

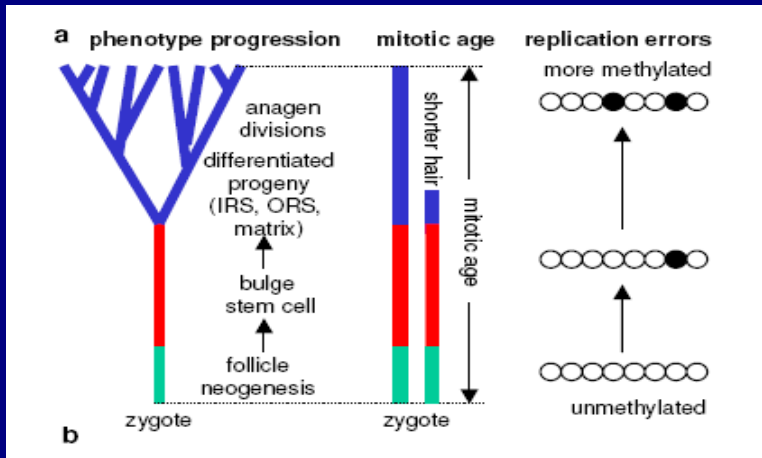


Figure: Human hair tree

Human Hair, Cont..

- Hair follicle genealogy can be divided into three phenotypic phases:
 - (i) neogenesis,
 - (ii) bulge stem cell latency
 - (iii) anagen differentiation and division.
- Mitotic age includes divisions during all three phases and may be inferred from numbers of replication errors (methylation) in sampled follicle cells.
- Short hairs should have fewer anagen divisions relative to longer hairs and therefore fewer errors (less methylation).
- Hairs from individuals of different ages have similar neogenesis and anagen intervals, and therefore any age-related increase in average methylation reflects lifelong stem cell mitotic activity.

Results

- Average hair methylation was found to be low before birth, progressively increased until two years of age, and then became constant. The lack of age-related increase in hair methylation is consistent with infrequently-dividing stem cells because the average numbers of divisions during neogenesis and anagen are likely to be similar for all hairs.

Results, Cont..

- There was no direct relationship between hair length and methylation, some long hairs having very low methylation levels and some short hairs having very high ones.
- White hairs were slightly more methylated than pigmented hairs from the same heads, possibly suggesting more lifetime divisions in white hairs. However, this difference was not statistically significant and further studies are necessary to determine whether mitotic ages differ between pigmented and white hairs.
- Relatively little is known about human hair stem cell dynamics because experimental manipulations are impractical.

Mouse Cancer, Origin and growth

- Tumor is derived from a single founder cell that has acquired a growth advantage over normal cells by a genetic modification.
- The study on Cell Lineage Analysis of a Mouse Tumor was done to reveal the lineage relations among cancer cells.
- They employed a method previously developed for reconstructing cell lineage trees from genomic variability caused by somatic mutations.
- The analysis of the reconstructed trees showed that the tumor initiated from a single **founder cell**, 5 months before diagnosis, grew in a physically coherent manner, and that the average number of the cell divisions accumulated in cancerous cells was almost twice than in the adjacent normal lung epithelial cells but slightly less than the expected figure for normal B lymphocytes.

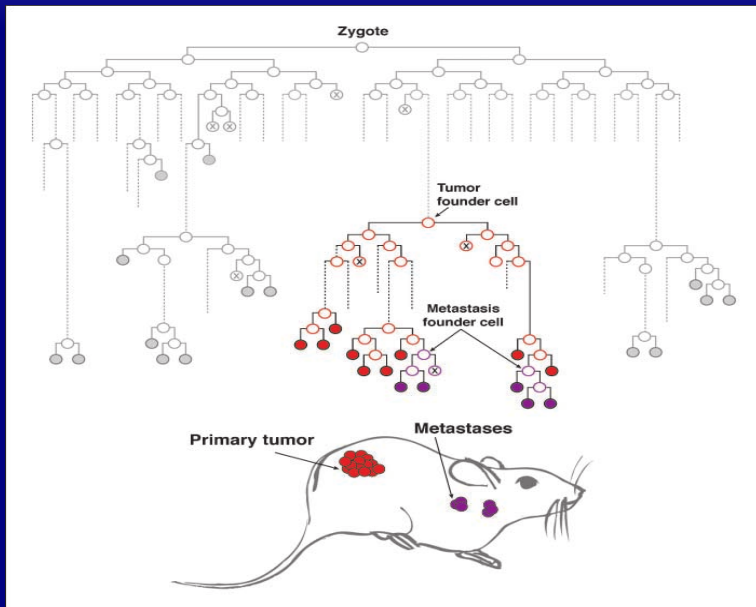


Figure: Tumor cell tree

Mouse Cancer, Cont..

- Lineage distance among cancer cells is correlated to their physical distance.
- Tumor cells share a heterozygous *TP53* mutation not shared by normal cells.
- All tumor cells have a common clonal origin.
- Cell Depth: The cell depth in cancer cells is larger than in the normal lung epitheliam cells.

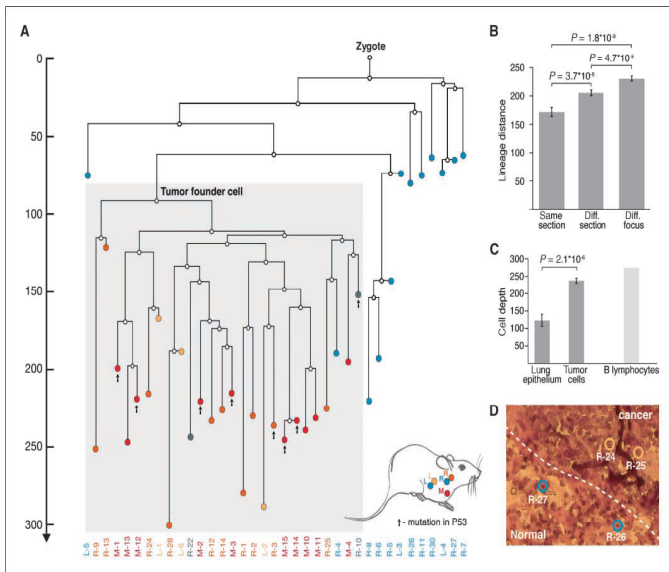


Figure: Cancer cell lineage construction

Future Work

- Human Cell Lineage Project
- Knowing cell lineage may be useful in studying cancer
- Building models and algorithms specific to cell genealogy




Key Dates

- **1983** - The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Sulston JE, et al.
- **2005** - Genomic variability within an organism exposes its cell lineage tree. Dan Frumkin, et al.
- **2008** - Cell Lineage Analysis of a Mouse Tumor. Dan Frumkin, et al.

References

1. Darryl Shibata and Simon Tavar. **Stem Cell Chronicles: Autobiographies Within Genomes**. Stem Cell Rev (2007) 3:94103.
2. Darryl Shibata. **Stem cells as common ancestors in a colorectal cancer ancestral tree**.Current option in Gastroenterology. 24:59-63, 2008.
3. Darryl Shibata and Simon Tavar. **Counting divisions in a human somatic cell tree: How, What and Why?** Cell Cycle 5:6, 610-614, 16 March 2006.
4. Dan Frumkin, Adam Wasserstrom, Shalev Itzkovitz, Tomer Stern, Alon Harmelin, Raya Eilam, Gideon Rechavi, and Ehud Shapiro. **Cell Lineage Analysis of a Mouse Tumor**. Cancer Res 2008; 68: (14). July 15, 2008.

References

-  Dan Frumkin, Adam Wasserstrom, Shai Kaplan, Uriel Feige, and Ehud Shapiro, *Genomic variability within an organism exposes its cell lineage tree*, PLoS Comput Biol **1** (2005), no. 5, e50.
-  J. E. Sulston, E. Schierenberg, J. G. White, and J. N. Thomson, *The embryonic cell lineage of the nematode caenorhabditis elegans*, Developmental Biology **100** (1983), 64–119.
-  Adam Wasserstrom, Dan Frumkin, Rivka Adar, Shalev Itzkovitz, Tomer Stern, Shai Kaplan, Gabi Shefer, Irena Shur, Lior Zangi, Yitzhak Reizel, Alon Harmelin, Yuval Dor, Nava Dekel, Yair Reisner, Dafna Benayahu, Eldad Tzahor, Eran Segal, and Ehud Shapiro, *Estimating cell depth from somatic mutations*, PLoS Comput Biol **4** (2008), no. 5, e1000058.

[2] [1] [3]