

Stem Cell Chronicles: Autobiographies Within Genomes

Darryl Shibata · Simon Tavaré

Published online: 3 July 2007
© Humana Press Inc. 2007

Abstract Human stem cell studies are difficult because many of the powerful experimental approaches that mark and follow stem cells and their progeny are impractical. Moreover, humans are long-lived, and it would literally take a lifetime to follow stem cell fates prospectively. Considering these hurdles, an ideal method would not require prior experimental manipulations but still allow “observations” of human stem cells from birth to death. The purpose of this review is to outline how histories or fates are likely to be surreptitiously recorded within somatic cell genomes by replication errors (molecular clock hypothesis). It may be possible to reconstruct stem cell lifetimes by measuring the random somatic changes that accumulate within their genomes, or the genomes of their more-easy-to-identify progeny.

Keywords Stem cell · Molecular clock · Genealogy

Life involves the replication and transfer of information between generations. Information is stored within genomes, which also record ancestry because genomes are almost exact copies of copies. Replication errors inevitably occur, allowing for variation or evolution, and these changes may subsequently be copied and passed from cell to cell.

D. Shibata (✉)
Department of Pathology, Keck School of Medicine,
University of Southern California,
Los Angeles, CA 90033, USA
e-mail: dshibata@usc.edu

S. Tavaré
Department of Biological Sciences,
University of Southern California,
Los Angeles, CA 90089, USA

S. Tavaré
Department of Oncology, University of Cambridge,
Cambridge, UK

Ancestry is recorded by such random changes, and it is possible to infer intervals since genomes shared a common ancestor by counting differences—the greater the differences between two genomes, the greater, on average, the interval since they shared a common ancestor (“molecular clock” hypothesis [1]).

Sequences are commonly used to infer species phylogeny [1]. By analogy, it should be possible to reconstruct a somatic cell tree (Fig. 1a) because all cells within an individual are related, with the zygote as the ultimate common ancestor [2]. Somatic cell genomes represent nearly exact copies of the genome in the zygote. Somatic cells with greater mitotic ages (total numbers of divisions since the zygote) should accumulate greater numbers of replication errors. All cells within an individual have identical chronological ages (years since birth), but their mitotic ages may differ.

Stem Cell Genealogies

Before dwelling on the complexities of “molecular clocks,” it is useful to outline how stem cell biology is logically encoded by mitotic age and genealogy (the changes in phenotype between the zygote and a present day cell). The genealogy and genome of every cell starts from the zygote. The genealogy of many cells can be divided into three sequential phenotypic phases: development from the zygote, a stem cell phase, and differentiation (Fig. 1b).

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

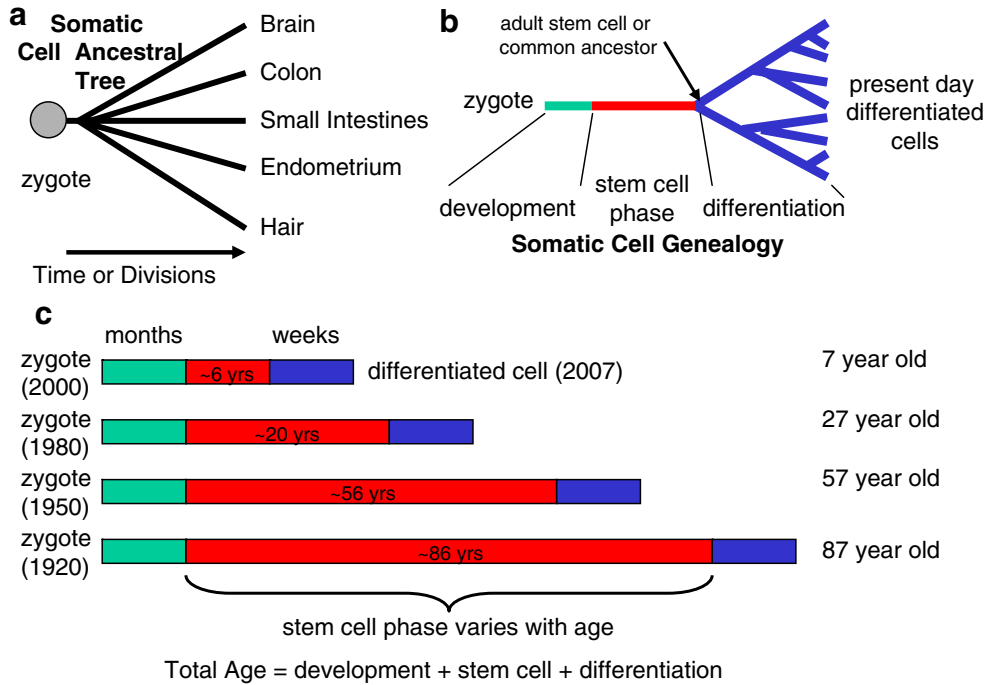


Fig. 1 Somatic cell genealogy. **a** A somatic cell ancestral tree starts at a zygote and ends with present day cells. All cells trace an origin to the zygote, which is the ultimate common ancestor. **b** The genealogy of many cells can be divided into three successive phenotypic phases: development, a stem cell phase, and differentiation. **c** Consider different aged individuals. Although development started at different

times in the past, this interval is similar for everyone. Times and numbers of divisions required for differentiation from a stem cell are also similar for everyone. Only the stem cell phase may vary between different aged individuals because stem cells can divide throughout life. Therefore, stem cell mitotic ages can be inferred by measuring the mitotic ages of differentiated cells

of adult chronological age. For example, colon morphogenesis is essentially completed by birth and differentiated colonocytes are produced daily and survive about a week [3]. By contrast, stem cell mitotic ages may vary because of their potential for limitless divisions. Therefore, changes in the mitotic ages of differentiated cells logically reflect changes in stem cell mitotic ages (Fig. 1c). The mitotic ages of easy-to-collect differentiated cells can potentially reveal how often hard-to-identify stem cells divide.

With constraints on numbers of divisions during development and differentiation, the possible geometries of somatic cell genealogies are limited (Fig. 2a). A static genealogy should be observed in non-mitotic adult tissues like the brain. Divisions are limited to the first few years of life and mitotic age or replication errors should not increase during adulthood. Mitotic tissues allow for two general genealogies. One possibility is a continuous genealogy with measurable stem cell divisions throughout life. Mitotic ages of stem cells and their differentiated cells should increase with chronological age. Another possibility is a punctuated genealogy with stem cell quiescence (infrequent division) or clonal succession (a pool of quiescent stem cells successively produce progeny [4]). In this scenario, mitotic ages of differentiated cells should not significantly increase with chronological age (Fig. 2b).

Epigenetic Molecular Clocks

Mitotic age measurements require the technical ability to count replication errors. It would be extremely difficult to measure mitotic age with mutations because DNA replication fidelity is extremely high. For example, cancer genome point mutation frequencies are about one per 100,000 to a million bases [5]. With such low mutation frequencies, one would have to sequence millions of bases to count just a few differences. It may be more practical to count somatic epigenetic changes instead of mutations because cytosine base methylation [6] in certain CpG rich regions measurably increases (from 0 to >50%) in mitotic tissues like the human colon with chronological aging [7, 8]. Like sequences, CpG methylation has a 5' to 3' order that can be measured by bisulfite sequencing, and methylation patterns are typically re-established after DNA replication (Fig. 3). Replication errors occur if the previous pattern is not copied exactly.

Unlike germline sequences that may differ between individuals, the “starting” CpG methylation state for everyone is relatively uniform because methylation is actively and passively removed or “erased” early in development before implantation [9]. Essentially everyone starts life without CpG methylation, and therefore it is possible to observe

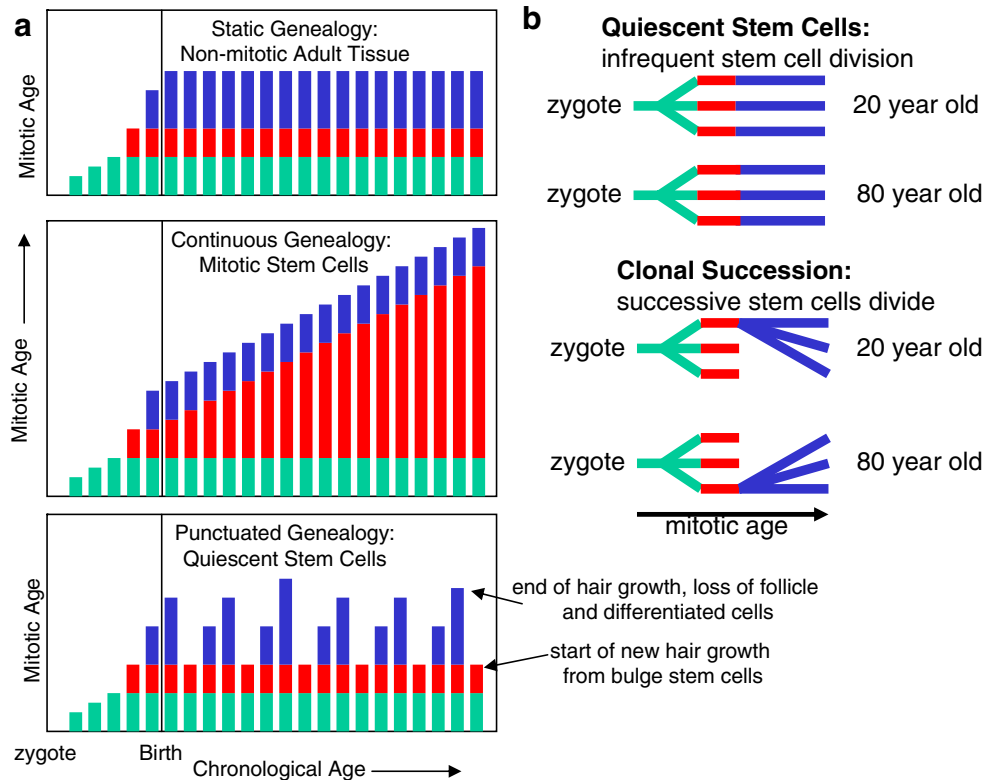


Fig. 2 Types of somatic cell genealogies. **a** A static genealogy should be observed in non-mitotic adult tissues like the brain. A continuous genealogy implies that adult stem cells divide throughout life. A punctuated genealogy implies that adult stem cell division is minimal. A punctuated hair follicle genealogy is illustrated, which shows that differentiated bulb cells divide but are lost at the end of the hair cycle when the hair falls out and the bulb physically disappears. Stem cells divide infrequently at the start of a new hair cycle, and therefore new

hairs on old or young heads have similar mitotic ages. **b** A punctuated genealogy implies relatively infrequent stem cell division. Quiescent stem cells divide infrequently but regularly contribute differentiated cells. With clonal succession [4], there is a stem cell pool—quiescent stem cells successively contribute differentiated cells. A stem cell may wait decades before it divides and produces differentiated progeny. Potentially this pool may be large enough that stem cells seldom contribute differentiated cells more than once during a lifetime

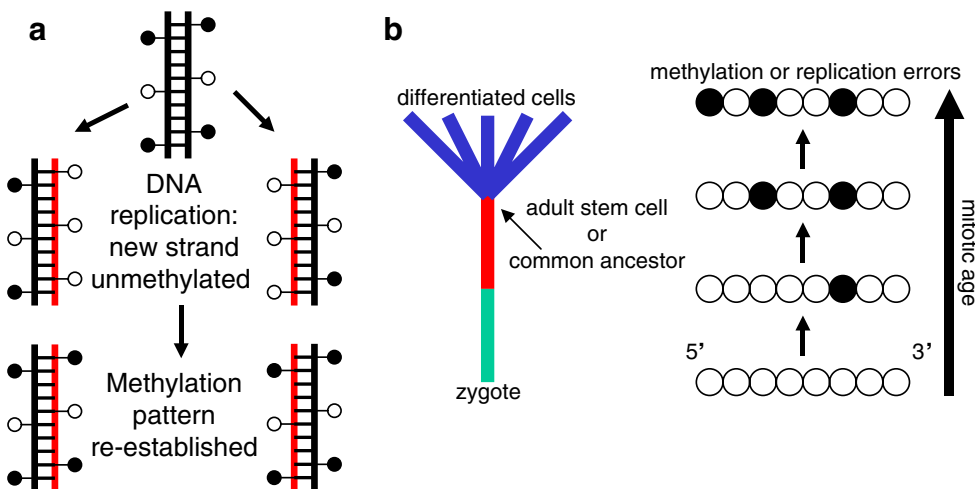


Fig. 3 An epigenetic molecular clock. **a** Methylation occurs on cytosine at CpG sites in mammals. After DNA replication, the new strand (red) lacks methylation. The methylation pattern is restored by DNA methyltransferases that covalently add methylation to hemimethylated CpG sites. Replication errors occur if the pattern is not copied exactly. Open circles represent unmethylated cytosines and black circles represent methylated cytosines at CpG dinucleotides. **b** Clock tags start unmethylated

(open circles) because methylation (black circles) is removed early in development before implantation [9]. With cell division, replication errors may randomly accumulate. Measuring methylation in differentiated cells at certain “epigenetic” molecular clocks may indicate mitotic age—the greater the numbers of divisions, the greater the numbers of replication errors. Although both methylation and demethylation may be possible, empirically net tag methylation tends to increase with age

“serial” methylation changes by sampling individuals of different chronological ages. Re-methylation can be broadly divided into programmed changes during development or differentiation, and random replication errors that may occur during any genealogical phase. About half of all human genes have CpG rich regions near or at their promoters, and extensive methylation of these CpG islands is associated with gene silencing [6, 10]. Methylation functionally involved in the control of expression is likely to occur to the same extent in a differentiated cell type, and therefore not vary with chronological aging or numbers of stem cell divisions.

In contrast to programmed methylation, random replication errors are unlikely to confer function or selection, but these “passenger” changes can passively chronicle cell fates. Replication error rates appear to differ between CpG sites because only certain CpG rich regions demonstrate measurable age-related increases in methylation in mitotic tissues. An “epigenetic molecular clock” or “tag” is a short CpG rich region whose methylation changes during the time frame of interest. Examples are illustrated in Fig. 4. Generally, a clock tag is within a gene not expressed in the cells of interest.

Epigenetic molecular clocks may be empirically identified by measuring methylation from different parts of the same organ, or from tissues of different chronological ages. If methylation is programmed during development or differentiation, then similar patterns are likely to be observed regardless of chronological age or location within an organ (Fig. 5a). In contrast, if methylation is not important, methylation patterns may predominantly reflect the accumulation of random replication errors. Numbers of replication errors should increase with mitotic age and their stochastic nature may result in different patterns between different parts of the same organ (Fig. 5).

Epigenetic Molecular Clock Data

Initial studies of human colon crypts [11] with the CSX clock tag revealed age-related increases in tag methylation consistent with random replication errors (Fig. 5b). This clock tag appears to “start” without methylation, likely

because CpG methylation is erased before implantation [9]. The error rate was estimated to be about 10^{-5} per CpG site per division [11], which is about 10,000-fold greater than the base replication error rate ($\sim 10^{-9}$ per base). This error rate is still relatively low, and measurable changes would accumulate only after years or decades. For example, average CSX tag methylation increases about 4% per decade in the colon (Fig. 6). Both methylation and demethylation replication errors appear to be possible, but net tag methylation tends to increase.

With species molecular clocks, the same sequence can be used to compare widely different species [1]. In theory, somatic cell molecular clocks may similarly accumulate replication errors regardless of cell type, allowing comparisons between different human tissues. Methylation with chronological age was measured for colon, small intestines, endometrium, brain, hair follicles, and neutrophils (Fig. 6). Consistent with an unmethylated start, CSX tag methylation is low early in life for all tissues. The brain tests whether tag methylation is a function of mitotic or chronological age because cell division is rare after childhood. Consistent with *in vitro* studies that suggest division is required for *de novo* methylation [12], the brain exhibits a static genealogy with low methylation in fetal tissues, and higher but stable adult tag methylation (unpublished data). Colon [11] and small intestinal epithelium [13] exhibit continuous genealogies with tag methylation increasing with chronological age, suggesting crypt stem cells divide throughout life. Endometrial epithelium also exhibits a continuous genealogy with increases in methylation before menopause, but methylation levels do not significantly increase after menopause when cell division largely ceases [14]. These studies are consistent with the hypothesis that average numbers of methylated sites are proportional to numbers of replication errors or mitotic ages.

Hair epithelium is mitotic but hair follicle tag methylation did not significantly increase with aging [15]. Hair biology [16] is consistent with a punctuated genealogy because the primary mitotic compartment (bulb) is physically separated from the stem cell compartment (bulge). Human hair exhibits cycles of follicle bulb growth and degeneration every few years. Bulge stem cells divide infrequently and primarily at the start of a new hair cycle to

CSX (8 CpG sites, Chr 5q34)

ggagaTTtaggaactttTgtTTTaCGCGCGttgtTttgCGTaCGggagagttgtggCGgCGattatgTagCGtgTaatgaqtgatTTgTagTTtggtT

BGN (9 CpG sites, Chr Xq28)

TTTAggagtgagtagTtgTTttCGgtTCGTCGgaTaTaTCGgaTagatagaCGtgCGgaCGgTTTATTTTtagTTTaaTtagTtagTTgCGTTtgCGTTTTTTTTTTtagtagggTtggT

SOX10 (20 CpG sites, Chr 22q13.1)

ggTaaggTaaqaaggagTagTaggaCGgCGaggCGgaCGatgaTaagttTTTCGtgTatTCGGaggTCGtTagTTaggtgTtTagCGgTtaCGaTtgggCGTtggtTTatgTCGtgCGCGTaaCGgCGTTTagTaaagTaagTCGTaCGTaaCGgTTTatgaaCGTTTatggttgggTtTaggTag

Fig. 4 Examples of epigenetic clocks or “tags.” Sequences are after bisulfite treatment that chemically converts C to T, but methyl-C is not changed. CpG sites are red and “Ts” indicate converted “Cs” at non-CpG sites. PCR primer sites are underlined. Tissue 5’ to 3’ methylation patterns can be sampled by sequencing individually cloned PCR products

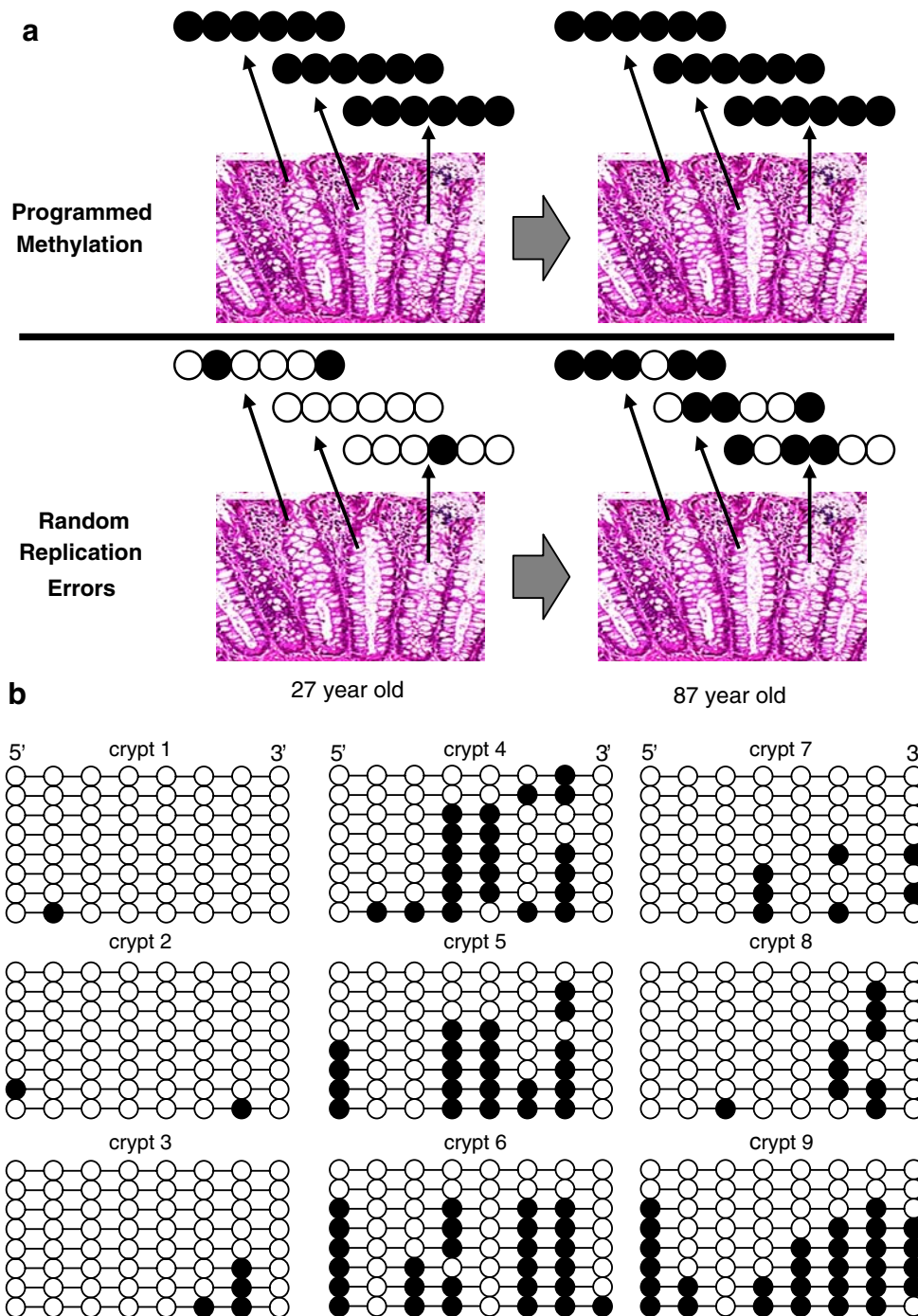


Fig. 5 Programmed methylation versus random replication errors. **a** Functional methylation likely occurs to the same extent in all cells of a given tissue (illustrated are individual colon crypts) regardless of chronological age. By contrast, methylation without functional consequences at other CpG rich regions likely represents random replication errors that occur independently in different cells. Therefore patterns may differ between cells within a given tissue. Replication

errors may increase with chronological age if tissue stem cells divide frequently. **b** CSX tag data from nine crypts from the same colon [11]. There are eight molecules arranged in a 5' to 3' horizontal order sampled from each crypt. The diverse tag patterns within and between crypts suggest methylation results from random replication errors rather than a programmed process

produce differentiated cells that migrate out of the bulge to reform a new hair follicle bulb. Differentiated follicle cells divide and accumulate replication errors, but these errors are discarded at the end of a hair cycle when the follicle

bulb disappears. In this way, hair follicles may have similar mitotic ages regardless of the age of individual because new hair follicles originate from bulge stem cells that divide only at the start each cycle. The lack of a measurable age

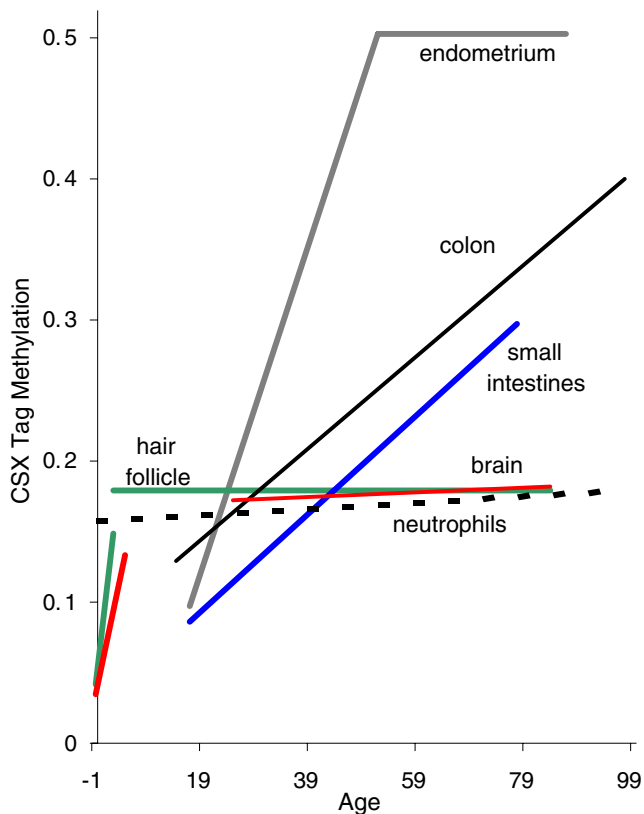


Fig. 6 Average CSX tag methylation with chronological age. Illustrated are trends for multiple human tissues. Colon (*black*) and small intestinal (*blue*) crypts appear to have continuous genealogies with tag methylation increasing throughout life. Endometrial glands (*gray*) also exhibit a continuous genealogy, but further increases in methylation are not observed after menopause when epithelial division largely ceases. The brain (*red*) exhibits the expected static genealogy with tag methylation lower in fetal and infant tissues, and higher but constant levels in adults. Hair follicle and hematopoietic divisions continue throughout life but their patterns are similar to the brain. The lack of measurable increases of hair follicle (*green*) or circulating neutrophil (*dotted*) tag methylation with aging are consistent with a punctuated genealogy, with infrequent stem cell divisions

related increase in replication errors may reflect relative bulge stem cell mitotic quiescence, or that the bulge contains a pool of stem cells that successively repopulate the bulb (clonal succession [4]).

Neutrophils are released daily from the bone marrow and survive less than a day in the blood. A punctuated genealogy might also be observed with hematopoiesis because its niche is adjacent to bone [16], which is constantly remodeled. There appears to be a stem cell pool because hematopoietic stem cells normally circulate in the blood [17]. Average neutrophil mitotic ages may not normally increase because newly formed niches may be successively colonized by relatively young stem cells, even in older individuals. Consistent with a punctuated genealogy, no significant increase in neutrophil tag methylation was observed during chronological aging (Fig. 6, unpublished data). Episodic destruction and reformation of a mitotic compartment, as with

hair and during hematopoiesis, may be a physical prerequisite for a punctuated genealogy because glands that persist exhibit continuous genealogies.

Reconstructing the ancestry encrypted within genomes requires multiple comparisons because random replication errors will produce seemingly random 5' to 3' differences (Fig. 5b). The significance of such tag patterns becomes clearer by examining multiple tissues of different ages, realizing that all genomes are copies of copies. The empirical data illustrated in Fig. 6 are consistent with the idea that numbers of genome replications and genealogy are the primary mechanisms responsible for the tag patterns found in many somatic cells. Random genome patterns can be the “words” in a replication language.

“Real” Stem Cells are “Ghosts”

Stem cell definitions vary, but one precise although retrospective definition is that a stem cell is a common ancestor or “progenitor” of a group of present day differentiated cells (Fig. 7). 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. For example, the zygote is a pluripotent stem cell, but it physically disappears after its first division. Similarly, the stem cell that divides asymmetrical to produce one stem and one non-stem cell daughter no longer exists, but the common ancestor of its differentiated progeny is the first “non-stem” cell daughter, and not the “stem” cell daughter (Fig. 7). Prospectively, this new “stem” cell daughter may function as a stem cell after transplantation, but retrospectively it is a “dead-end” because it lacks progeny.

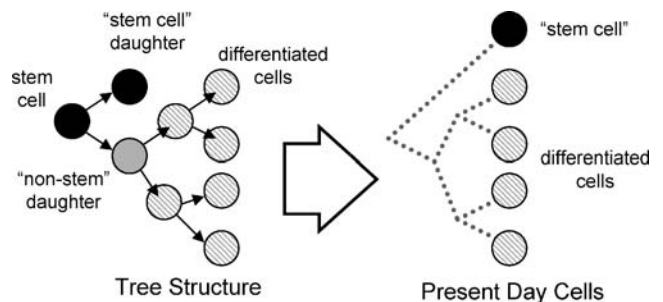


Fig. 7 The tree diagram illustrates potential differences between prospectively identified stem cells, and stem cells defined as common ancestors. A stem cell can divide asymmetrical to produce a “stem” and a “non-stem” cell daughter. Prospectively, the new “stem” cell daughter may function like a stem cell after transplantation, but retrospectively it is a “dead-end” because it has no progeny. The population of present day cells gives the impression of a stem cell hierarchy, but the present day “stem cell” is not the progenitor of the present day differentiated cells. The differentiated cells are progeny of the no longer present “non-stem” cell daughter, which is really the common ancestor or stem cell for this population

Stem cells defined by ancestry are “ghosts” because they are no longer physically present, but their progeny make them “real”. By contrast, prospective definitions of stem cells depend on the future. Experiments characterize what may happen when cells are manipulated whereas decades of normal *in vivo* divisions can be recorded within genomes. For example, genomes in a 101 year-old individual potentially chronicle the lives of long dead stem cells from over a century ago.

Niches and Stem Cell Clonal Evolution

Progeny numbers may reflect stem cell “success” (think of the zygote), but uncontrolled proliferation is undesirable. Many mammalian stem cells are thought to be maintained by stem cell niches [18] that extrinsically define and control stem cell numbers by their locations within the niche. There are typically multiple stem cells within a niche (Fig. 8). Most stem cells divide asymmetrically to produce one stem cell daughter that remains within the niche, and a non-stem cell daughter that leaves the niche and differentiates. However, symmetric divisions are also possible. To maintain a constant number of niche stem cells, one stem cell may

produce two non-stem cell daughters, balanced by another stem cell that produces two stem cell daughters.

The replacement of a population by the progeny from a single cell is called clonal evolution, a term usually applied to tumor populations [19]. The ability of stem cells occasionally to divide symmetrically also allows for the possibility that all stem cell lineages except one will eventually be lost from a niche (Fig. 8). Total stem cell numbers remain unchanged because extinction (two non-stem daughters) is balanced by expansion (two stem cell daughters). Stem cell niche clonal evolution is observed in the murine intestines because heterogeneously marked crypts become visibly homogeneous after several weeks to months, reflecting the loss with replacement of stem cell lineages. In murine intestines, symmetrical division occurs about 5% of the time [20].

Systematic visible cell fate marker experiments are not possible in humans, although heterogeneous appearing normal colon crypts tend to become homogeneous after therapeutic radiation [21]. However, it is possible to infer if human intestinal crypt stem cells divide symmetrically by measuring crypt tag diversity. Daughter cells will tend to contain nearly identical 5' to 3' patterns, but eventually differences will accumulate (Fig. 3). Whereas mitotic age is

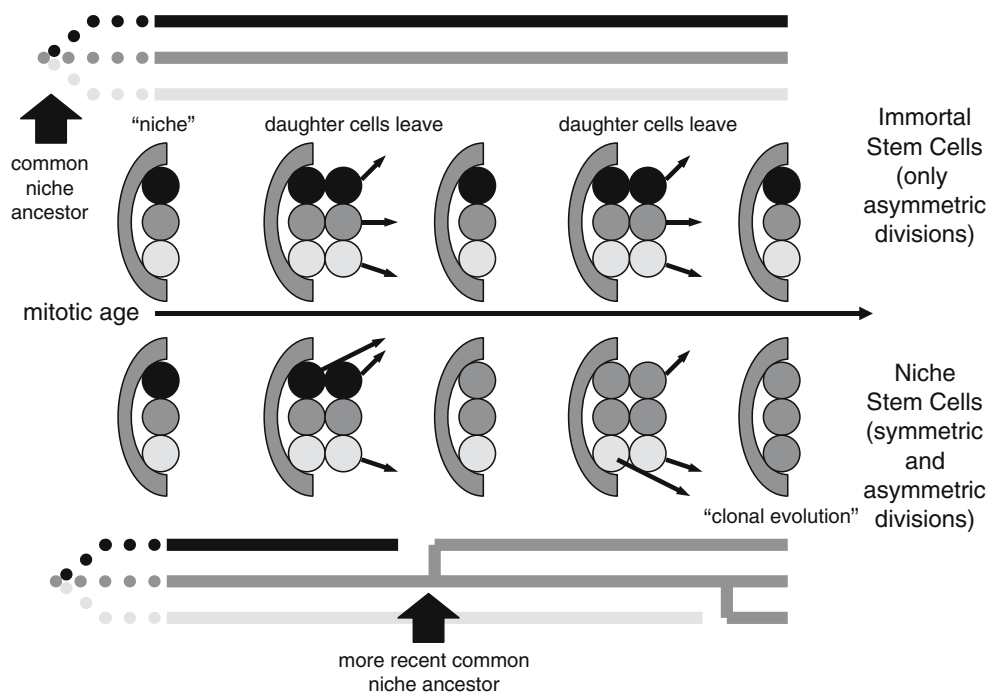


Fig. 8 A stem cell niche contains multiple stem cells. Stem cell lineages are “immortal” if they always divide asymmetrically to produce stem and non-stem daughter cells. Such lineages never become extinct. By contrast, stem cell lineages become extinct whenever symmetric division occurs because niche cell numbers are limited. Eventually all stem cell lineages except one will become extinct even if symmetrical divisions are relatively rare. This niche clonal evolution results in a “bottleneck” because the niche stem cell common ancestor progres-

sively becomes more recent. One can distinguish between immortal stem cell lineages and niche clonal evolution by measuring methylation pattern diversity. Diversity will continuously increase with aging with immortal stem cell lineages, whereas diversity will be constant with niche clonal evolution because successive “bottlenecks” limit niche diversity. Methylation tag diversity is constant in human intestinal crypts, consistent with niche clonal evolution [11] and the likely impossibility of always dividing asymmetrically

a function of numbers of replication errors, the ancestry between tags may be summarized by methylation differences between their CpG sites. This pair-wise difference (Hamming distance) is zero for identical tags, and increases on average with more distantly related tags. Diverse cell populations have many different unique tags—“old populations are diverse populations.”

Diversity increases with division because new tag patterns may arise from new replication errors. Immortal stem cells that always divide asymmetrically result in tag diversity that tends to increase continuously because errors are never lost. By contrast, symmetrical divisions reduce diversity because unique tags are potentially eliminated when a stem cell produces two non-stem daughters. Niche clonal evolution creates population “bottlenecks” because stem cells share progressively more recent common ancestor (Fig. 8). Experimentally, human colon crypt tag diversity does not increase with chronological age, consistent with stem cell niches with occasional symmetric divisions [11]. Similar to murine crypts, symmetric divisions may occur about 5% of the time in human crypts, resulting in periodic stem cell population “bottlenecks” or niche stem cell clonal evolution about every eight years [11].

The ability to reconstruct ancestry from replication errors illustrates how the life and death of stem cells may be inferred without experimental manipulations that may perturb normal behavior. Clonal evolution is not confined to tumor populations but appears to be a normal rhythm in niches, which may occur simply because it may be impossible for stem cells always to divide asymmetrically. Niches have the potential to act as evolutionary crucibles because only a limited number of progeny can remain within a niche. The dominant stem cell

clone may arise by chance (or drift), but one stem cell may acquire a mutation that confers a selective advantage over the surrounding stem cells within the niche [22]. Interestingly, certain dominant-negative APC mutations commonly found in colorectal cancers appear to enhance stem cell survival in normal appearing crypts [23]. Given the loss of most stem cell lineages during niche clonal evolution [22], one potential explanation for some cancer mutations is “contingency”—the idea that transformation of a stem cell later in life is contingent on survival during earlier rounds of niche clonal evolution. Some of the genes mutated in cancers also appear important to stem cell survival, and alterations that enhance survival during niche clonal evolution could also increase the chances of transformation later in life.

Do Older Individuals Have “Older” Cancers?

The incidence of cancer increases markedly with chronological age, and excessive cell proliferation may increase risks for cancer [24]. Cancers are characterized by uncontrolled cellular proliferation, and therefore carcinoma mitotic ages are likely to be greater than their corresponding normal epithelium. Just like normal tissues, genealogy can be extended to cancers. Starting from the zygote, the sequential phenotypic phases are development, a stem cell phase, possibly a short differentiated phase, and finally neoplasia (Fig. 9a).

Preliminary data illustrate that average colon cancer tag methylation is generally greater than corresponding normal tissue, and that tag methylation is generally greater in cancers removed from older individuals (Fig. 9b). Cancers

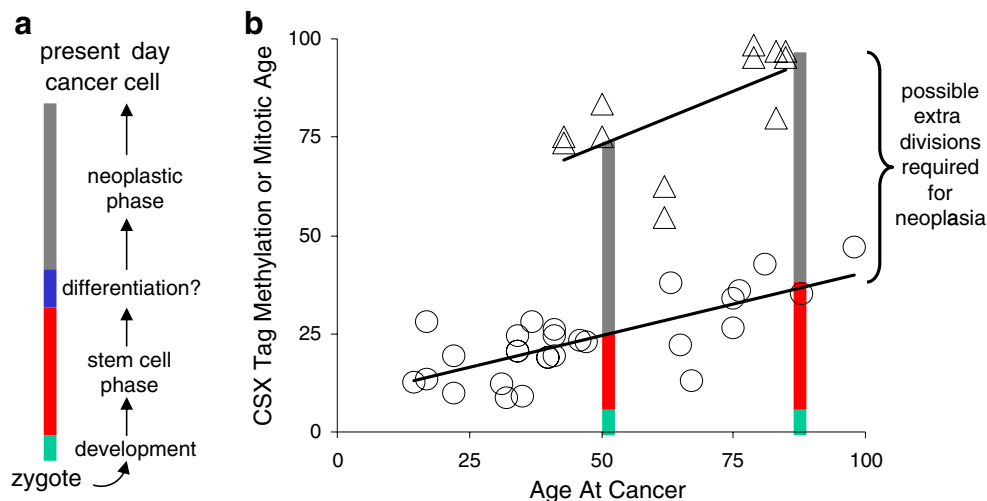


Fig. 9 Cancer genealogy. **a** A cancer genealogy is the genealogy of its normal tissue, except its termination is a neoplastic phase. The start (as always) is the zygote, followed by development, a stem cell phase, possibly a short period of differentiation, ending with neoplasia and the present day cancer cell. **b** Illustrated are unpublished data of average CSX tag methylation in normal colon crypts (circles) and colon cancers

(triangles). Cancers in older individuals should have higher mitotic ages because their normal stem cells have higher mitotic ages. The data are consistent with this model because colon cancer tag methylation increases with chronological age. Tag methylation differences between cancer and corresponding normal colon may provide information on how many extra divisions are needed for transformation and cancer growth

in older individuals may have greater mitotic ages simply because the stem cell that transforms has a greater mitotic age in older individuals. Although further studies are necessary, this type of data suggests that cancer histories are also likely to be recorded by replication errors within their genomes. If cell division increases the risk for cancer [24], potentially more colon cancers arise in older individuals because colon stem cell mitotic ages normally increase with age (Fig. 9b).

Challenges and Problems

The reconstruction of histories from sequences remains controversial because the exact manner by which changes accumulate is uncertain, and the optimal methodologies to decode sequences are uncertain [1]. The reconstruction of the past from genomes requires relatively sophisticated mathematical modeling to account for the stochastic nature of replication errors. Although conclusions based on a single locus are tenuous, the past becomes clearer when multiple genomic regions reconstruct the same ancestry [1].

Perhaps the greatest challenge is accepting that seemingly random methylation patterns (Fig. 5b) are “words” that retell stem cell ancestries. The exact mechanisms responsible for such methylation patterns are uncertain, but appropriate algorithms can reconstruct the past from random replication errors. These algorithms may be complex because methylation error rates may differ between tissues and sites within a tag. For example, for one clock tag, methylation at one CpG site appears to increase the probability of methylation at another CpG site [25].

A clock analysis is facilitated by the analysis of pure populations because genealogies likely differ between cell types. For example, bulk analysis of colon genomes is likely to yield nonsensical results because the mixtures of cells (epithelial, neutrophils, lymphocytes, smooth muscle, blood vessels) represent distinct and different mitotic histories. However, many tissues are composed of small clonal units such as colon crypts, and it is possible through microdissection or other techniques to isolate relatively pure cell populations. Although it is difficult to isolate and analyze genomes from individual cells, bisulfite sequencing of individually cloned PCR products allows for the sampling of individual tags from small clonal cell pools such as a 2,000 cell human colon crypt [11].

“Reading” the Past: The Language of Replication

Sequence comparison revolutionized species phylogeny reconstruction because it became possible to study the past from the present simply by sequencing and “reading” DNA

[1]. Similarly, it may be possible to characterize stem cells from the genomes of their progeny, without the paraphernalia traditionally associated with stem cell studies. Read one way, the 5′ to 3′ order of bases in a human genome retells the emergence of modern humans “out-of-Africa” about 50,000 years ago. Read another way, the 5′ to 3′ order of CpG methylation in a somatic cell genome potentially retells an emergence “out-of-embryo” decades ago.

The basic premise of translating molecular clocks from species to somatic cells is that even somatic cell genomes are not “created” but represent copies of copies. Evolution or change is possible because genomes are almost perfect rather than exact copies of prior genomes. Although certain methylation patterns may be spontaneously created or programmed, other patterns may simply represent outcomes of random replication errors, which potentially “chronicle” past divisions. The current approach with epigenetic molecular clocks may not be optimal, but the general strategy of using genomes to reconstruct ancestry can overcome many problems that currently hinder systematic studies of human stem cell biology. What yet unread stories await within our cells?

References

1. Bromham, L., & Penny, D. (2003). The modern molecular clock. *Nature Reviews Genetics*, 4, 216–224.
2. Shibata, D., & Tavaré, S. (2006). Counting divisions in a human somatic cell tree: how, what and why? *Cell Cycle*, 5, 610–614.
3. Potten, C. S., & Loeffler, M. (1990). Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development*, 110, 1001–1020.
4. Kay, H. E. (1965). How many cell-generations? *Lancet*, 15, 418–419.
5. Wang, T. L., Rago, C., Silliman, N., Ptak, J., Markowitz, S., Willson, J. K., et al. (2002). Prevalence of somatic alterations in the colorectal cancer cell genome. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 3076–3080.
6. Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development*, 16, 6–21.
7. Ahuja, N., Li, Q., Mohan, A. L., Baylin, S. B., & Issa, J. P. (1998). Aging and DNA methylation in colorectal mucosa and cancer. *Cancer Research*, 58, 5489–5494.
8. Issa, J. P. (2000). CpG-island methylation in aging and cancer. *Current Topics in Microbiology and Immunology*, 249, 101–118.
9. Morgan, H. D., Santos, F., Green, K., Dean, W., & Reik W. (2005). Epigenetic reprogramming in mammals. *Human Molecular Genetics*, 14, R47–R58.
10. Jones, P. A., & Laird, P. W. (1999). Cancer epigenetics comes of age. *Nature Genetics*, 21, 163–167.
11. Yatabe, Y., Tavaré, S., & Shibata, D. (2001). Investigating stem cells in human colon by using methylation patterns. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 10839–10844.
12. Velicescu, M., Weisenberger, D. J., Gonzales, F. A., Tsai, Y. C., Nguyen, C. T., & Jones, P. A. (2002). Cell division is required for de novo methylation of CpG islands in bladder cancer cells. *Cancer Research*, 62, 2378–2384.

13. Kim, J. Y., Siegmund, K. D., Tavaré, S., & Shibata, D. (2005). Age-related human small intestine methylation: evidence for stem cell niches. *BMC Medical*, 3, 10.
14. Kim, J. Y., Tavaré, S., & Shibata, D. (2005). Counting human somatic cell replications: methylation mirrors endometrial stem cell divisions. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 17739–17744.
15. Kim, J. Y., Tavaré, S., & Shibata, D. (2006). Human hair genealogies and stem cell latency. *BMC Biology*, 4, 2.
16. Fuchs, E., Tumber, T., & Guasch, G. (2004). Socializing with the neighbors: stem cells and their niche. *Cell*, 116, 769–778.
17. Wright, D. E., Wagers, A. J., Gulati, A. P., Johnson, F. L., & Weissman, I. L. (2001). Physiological migration of hematopoietic stem and progenitor cells. *Science*, 294, 1933–1936.
18. Spradling, A., Drummond-Barbosa, D., & Kai, T. (2001). Stem cells find their niche. *Nature*, 414, 98–104.
19. Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science*, 194, 23–28.
20. Marshman, E., Booth, C., & Potten, C. S. (2002). The intestinal epithelial stem cell. *Bioessays*, 24, 91–98.
21. Campbell, F., Williams, G. T., Appleton, M. A., Dixon, M. F., Harris, M., & Williams, E. D. (1996). Post-irradiation somatic mutation and clonal stabilisation time in the human colon. *Gut*, 39, 569–573.
22. Calabrese, P., Tavaré, S., & Shibata, D. (2004). Pre-tumor progression: clonal evolution of human stem cell populations. *American Journal of Pathology*, 164, 1337–1346.
23. Kim, K. M., Calabrese, P., Tavaré, S., & Shibata, D. (2004). Enhanced stem cell survival in familial adenomatous polyposis. *American Journal of Pathology*, 164, 1369–1377.
24. Preston-Martin, S., Pike, M. C., Ross, R. K., Jones, P. A., & Henderson, B. E. (1990). Increased cell division as a cause of human cancer. *Cancer Research*, 50, 7415–7421.
25. Nicolas, P., Kim, K. M., Shibata, D., & Tavaré, S. (2007). The stem cell population of the human colon crypt: analysis via methylation patterns. *PLoS Computational Biology*, 3, e28.