[Cell Cycle 8:14, 2187-2193; 15 July 2009]; ©2009 Landes Bioscience

Review

Many colorectal cancers are "flat" clonal expansions

Kimberly D. Siegmund,[1] Paul Marjoram,[1] Simon Tavaré[2] and Darryl Shibata[3,*]

[1]Department of Preventive Medicine; University of Southern California Keck School of Medicine; Los Angeles, CA USA; [2]Department of Biological Sciences; University of Southern California; Los Angeles, CA USA; and Department of Oncology; University of Cambridge; Cambridge, UK; [3]Department of Pathology; University of Southern California Keck School of Medicine; Los Angeles, CA USA

Key words: tumor progression, clonal evolution, cancer stem cells, methylation, gompertzian

Population geneticists can reconstruct the ancestries of macroscopic populations from polymorphisms in present day individuals. For example, the migration "out of Africa" is recorded in human genome variation in different parts of the world. Here we apply this approach to human colorectal cancer cell populations and polymorphic passenger methylation patterns. By sampling molecular variation from different parts of the same cancer, it should be possible to infer how individual tumors grow because recent clonal expansions should be less diverse than older expansions. Average diversity was different between cancers implying that some cancers are older clonal expansions than others. For individual cancers, methylation pattern diversity was relatively uniform throughout the tumor (right versus left side, superficial versus invasive), which is more consistent with a single, uniform or "flat" clonal expansion than with stepwise sequential progression. Many colorectal cancers appear to invade and expand early, but subsequently stall. Epiallele diversity within individual small cancer gland fragments was high and more consistent with frequent rather than extremely rare cancer stem cells (CSCs). These studies suggest that many human colorectal cancers are relatively old uniform clonal expansions, that cancer cell populations contain frequent long-lived CSC lineages, and that some passenger methylation patterns record somatic cell ancestry.

©2009 Landes Bioscience. Do not distribute.

A deficit in cancer research is the inability to observe serially exactly how human tumors grow. Cancers are thought to progress through clonal evolution, or the sequential emergence of progressively more "fit" neoplastic populations.[1] While tumor progression may occur through successive phenotypic stages, the full malignant phenotype of a cancer may also be present at the time of transformation.[2] There are many ways progression may occur, and a priori it is difficult to determine exactly how and when an individual human tumor arises. All models of human tumor progression are based on inference because the growth of individual human cancers cannot be directly observed.

Simplistically, because larger, widespread tumors must start as smaller, localized tumors, size and stage should correlate with tumor age. Depending on preconceived notions, the history or "age" of a tumor may be judged based on its current appearance. However, exceptions are common. Small primary cancers may have already metastasized,[3] and advanced prostate cancer is not prevented by annual screening.[4] In lieu of direct serial observations, it appears possible to reconstruct how individual cancers grow from the genomic diversity in present day human cancer cell populations.[5] In this approach, the growth of individual human cancers is also inferred, but this inference is based on the comparison of genomes directly sampled from different tumor regions. Genome comparisons are often used to reconstruct ancestral aspects of macroscopic populations, such as human migration histories. Although polymorphisms appear to be randomly scattered among individuals, patterns emerge when genomes are annotated with their geographic locations. African populations tend to have higher genome diversity than European or Asian populations, consistent with the dispersal of humans "out of Africa".[6] The same approaches should also apply to tumor cell populations and their genomes. Genomes are historical documents because they are almost perfect copies of prior copies. The greater the number of replications since a common ancestor, on average the greater the number of differences (a "molecular clock" hypothesis[7]). Starting from a single transformed cell, its progeny and their genomes will become more abundant and physically more separated and on average more polymorphic during dispersal. A visible tumor contains billions of cancer cells (~one billion per cm^3). If growth occurs through sequential clonal expansions, newer tumor populations should be less polymorphic than older populations. By contrast, if the full malignant potential is present at the time of transformation, then all parts of the cancer are essentially created simultaneously and should have similar ages (Fig. 1A).

Using Space to Calibrate Time

Our approach to maximize the historical information in cancer genomes adds the dimension of space.[5] During clonal expansion, cancer cells and their genomes increase, on average, both their physical and genetic distances (Fig. 2A). (Genetic distance is

*Correspondence to: Darryl Shibata; 1441 Eastlake Avenue; Norris Cancer Center; NOR 6410; Los Angeles, CA 90033 USA; Email dshibata@usc.edu

Submitted: 05/20/09; Accepted: 06/01/09

Previously published online as a *Cell Cycle* E-publication:
http://www.landesbioscience.com/journals/cc/article/9151

the number of pairwise differences or replication errors between genomes). Because cancer dimensions (length, width and height) are created by cell division, physical distance becomes a surrogate for genetic distance—cells from opposite sides of the same cancer were likely last related around the time of transformation. Therefore, the approximate mitotic age of a cancer (numbers of divisions since transformation) can be inferred by measuring numbers of pairwise differences between genomes from opposite sides of the cancer.

Solid tumors like colorectal adenocarcinomas facilitate a space-time analysis because tumor cells are physically partitioned into smaller glandular units, which can be isolated relatively intact from fresh tumors (Fig. 2B). Each gland can be considered a subclone because cells from adjacent glands generally do not mix. Neoplastic glands, similar to colon crypts, are likely created by fission,[8] in which a single gland divides to produce two glands (Fig. 2C). In this manner, adenocarcinomas grow by cell division and then gland fission when the number of cells within a gland exceeds a threshold. Moreover, glands may remain static in size if cell division is balanced by cell death. Two daughter glands may remain adjacent but drift apart by independently acquiring, in situ, different combinations of somatic replication errors. Experimentally, sampling glands from opposite sides of the cancer allows for genome comparisons between glands (to infer cancer mitotic age) and within glands (to infer gland mitotic age). Depending on how tumors grow, the mitotic age of a cancer gland may or may not be similar to the mitotic age of its cancer.

©2009 Landes
Do not distribute

The Geometry of Cancer Space-Time

A cancer ancestral tree starts at the first transformed cell and ends at present day cancer cells. The tree for a one cm^3 tumor has a billion branch tips that eventually converge or coalesce back to the first transformed cell. There are two basic trees (Fig. 1B) that coalesce early (palm-tree) or late (star-shaped). A star phylogeny has the interesting property that mitotic ages are similar regardless of physical location.

Cancer ancestral tree shape can be anticipated by observations that tumor growth tends to follow a rapid initial exponential increase in size followed by a progressive decline in growth rates ("Gompertzian kinetics").[9] Visible tumors (>1 cm^3) grow slowly, and often surgical removal of colorectal cancers may be delayed without apparent clinical consequences.[10] About 30 cell doublings can produce a billion cells. Rapid initial growth is consistent with a star-shaped tree because most tumor cells and their lineages are created early in progression during the exponential growth phase (Fig. 2D). After exponential growth, if cell division is balanced by cell death, genetic distances between cells will increase while their physical distances are static (Fig. 2C).

Cancer ancestral trees are geometrically oriented within a tumor because branch tips terminate at present day cancer cells. Therefore, it should be possible to distinguish between palm-tree or star-shaped phylogenies by sampling genomes from different parts of the same cancer and comparing the physical distances between genomes with their genetic distances (Fig. 1C). With a star phylogeny, genome diversity is similar regardless of location (an isotropic or "flat" distribution), whereas with a palm-tree phylogeny, genomes from opposite cancer sides will be more diverse than those from the same side (Fig. 1B).

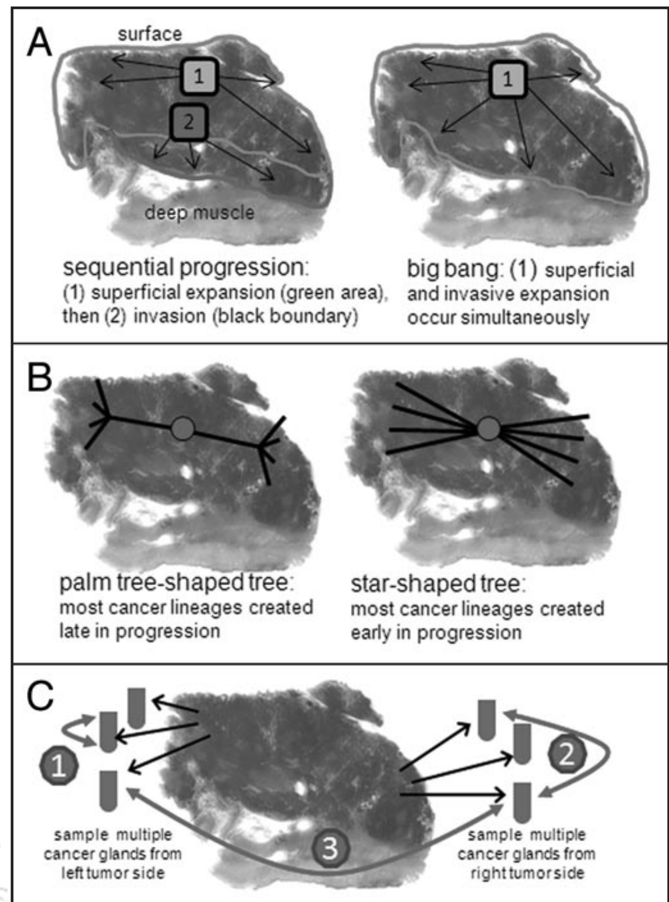

Figure 1. A microscope slide of a human colorectal cancer. How this cancer was created: (A) Sequential stepwise progression where the superficial tumor ("1") is created after transformation, and the deeper portion ("2") arises later from a new subclone with the ability for invasion. The invasive portion is younger and should have less genomic diversity relative to the older, superficial region. Alternatively, the entire tumor may arise through a single uniform "flat" clonal expansion after transformation. No further alterations are required for invasion and therefore superficial and invasive portions have similar mitotic ages. (B) Two basic phylogenies. A star-shaped tree has the special property that all parts of a tumor have similar mitotic ages because they are created early in progression (also see Fig. 2D). A palm tree-shaped tree has more local structure because most cancer cell lineages are created locally and later in progression. (C) The experimental approach samples genomes from multiple colorectal cancer glands isolated from opposite sides (right or left) of the same cancer. In this manner, one can sample the tips of cancer ancestral trees, which are physically oriented within the tumor from growth. Genomes from opposite cancer sides likely shared a common ancestor around the time of transformation, and are potentially the least related genomes in a cancer. Genome comparisons are possible, (1) within glands (to estimate gland mitotic age), (2) between glands from the same side (to estimate the age of that side), or (3) between glands from opposite sides (to estimate the mitotic age of the cancer).

Measuring Cancer Genome Diversity

Genome comparisons traditionally analyze 5' to 3' base order, but the rarity of somatic mutations in colorectal cancers hampers

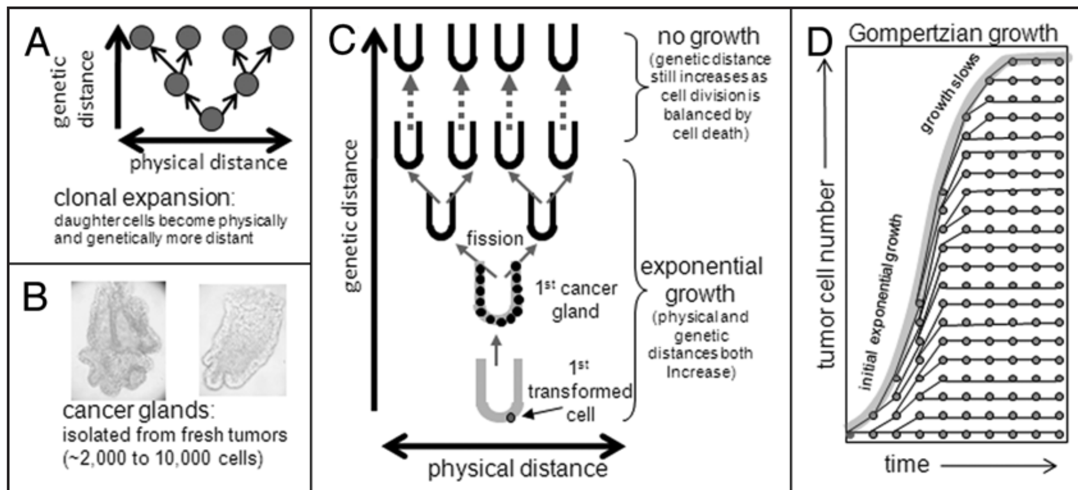

Figure 2. Physical and genetic distances are linked. (A) As the first transformed cell (red dot) undergoes clonal expansion, both physical and genetic distances increase because cell duplication creates both the physical dimensions (height, length, width) and replication errors (genetic distance). (B) Examples of gland fragments isolated from a colorectal adenocarcinoma. These glands contain about 2,000 to 10,000 cells and are relatively free of stromal contamination (>95% cancer cells). Glands enable adjacent cells to remain in close proximity during progression. Comparisons of genomes within a gland allow inferences on gland mitotic ages whereas comparisons between glands infer mitotic ages of different parts of the cancer. (C) In adenocarcinomas, cancer cells are physically segregated into smaller subclones or glands that prevent mixing. During growth, cancer glands may divide by fission to form new daughter glands. After growth slows and physical distances are relatively stable, genetic distances between cells and glands can continue to increase if cell division is balanced by cell death. (D) Gompertzian growth is characterized by early exponential growth that slows with time—most visible tumors grow slowly. Cell doubling during the initial early exponential growth creates cancer cell lineages. Because these lineages are created early in progression, ancestry is a star-shaped tree.

©2009 Landes Bioscience

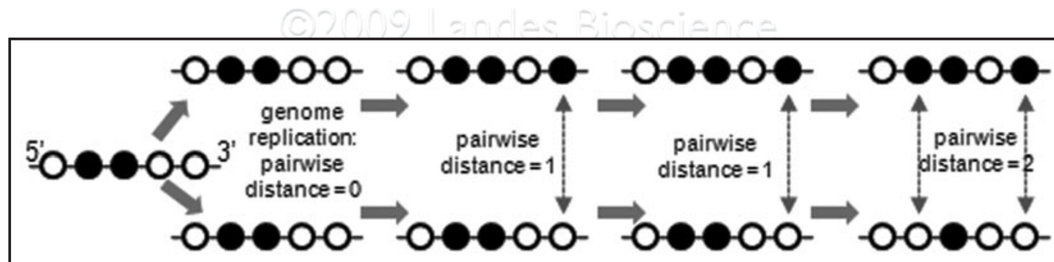

Figure 3. Methylation patterns are usually copied by DNA methyltransferases after DNA replication. However, after lineage duplication, replication errors (de novo methylation or demethylation) may occur independently in each lineage. The greater the number of replications since the common ancestor, on average the greater the number of pairwise differences. In this manner, the drift of CpG methylation can be used, like sequences, to reconstruct somatic cell ancestry (circles represent CpG sites, filled are methylated and open are unmethylated).

this approach. Cancer genome sequencing projects have detected less than one clonal somatic mutation per 100,000 bases.[11] Therefore, one would have to sequence millions of bases to find a handful of somatic mutations.

An alternative approach measures passenger DNA methylation pattern diversity.[12] Genomes contain both genetic and epigenetic information. DNA methylation, like bases, has a 5' to 3' order at CpG sites but is a binary code (methylated or unmethylated). Methylation patterns are usually duplicated after DNA replication by DNA methyltransferases, but epigenetic replication error rates appear to be greater than somatic mutation rates at certain "epigenetic somatic cell clocks". These somatic cell clocks are short stretches of CpG rich sequences (8 to 14 CpG sites) in intergenetic or non-promoter regions in which methylation has no obvious function or selective value. Such "passenger" methylation is therefore free to drift and become polymorphic between cancer cells. As with DNA sequences, differences between two methylation patterns can be measured as a pairwise (or Hamming) distance (Fig. 3). For a population of epialleles, greater average pairwise distances imply greater pattern diversity and a longer interval since a common ancestor.

DNA methylation patterns can be read by bisulfite sequencing,[13] and typically methylation patterns in both normal and tumor tissues are polymorphic. For example, using 454-pyrosequencing to sequence hundreds of cancer epialleles, nearly every possible methylation pattern was found in human breast cancers.[14] Considering that a single epigenome is present in the first transformed cell, high methylation pattern diversity in present day cancer cells implies lack of selection (i.e., passenger changes), low replication fidelity, many long-lived cancer cell lineages, and a high cancer mitotic age.

Colon Cancers are Relatively "Flat" Clonal Expansions

The data from 12 human colorectal cancers[5] are discussed in this section. Although methylation patterns were sampled from two different loci, for simplicity we discuss the data from a single

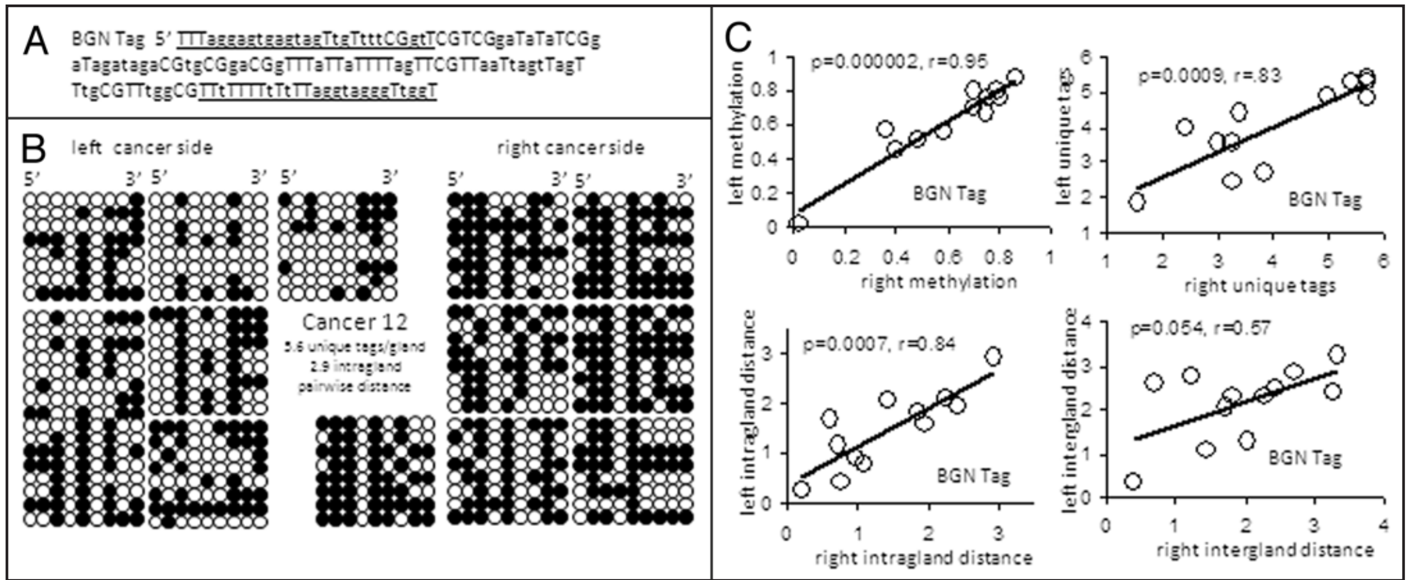

Figure 4. Cancer DNA methylation patterns. (A) The BGN locus on the X-chromosome is in body of the BGN gene and contains 8 or 9 CpG sites (bisulfite converted sequence is illustrated with PCR primer sites underlined). Potentially only a single epiallele is present per cell when cancers from male patients are sampled. (B) BGN methylation patterns sampled from the left and right sides of a cancer. Eight epialleles are sampled from each of seven glands from each cancer side (14 total glands and 112 total epialleles). (C) DNA methylation patterns from opposite cancer sides are similar because average parameter values (percent methylation, numbers of unique patterns per gland, intragland pairwise distances, and intergland pairwise distances) from right versus left sides of 12 colorectal cancers correlate.

locus. This locus has eight or nine CpG sites in the non-promoter region of the BGN gene on the X-chromosome (Fig. 4A). BGN is expressed in connective tissue and methylation is unlikely to confer selection in colorectal cancer cells. To facilitate analysis, colorectal cancers from male patients were examined because only a single epiallele may be present per cell. Aneuploidy may complicate interpretation because a cancer cell may contain multiple BGN epialleles, but each BGN epiallele is a copy of the single original epiallele. Whereas two autosomal epialleles may begin to drift apart from birth, aneuploidy and multiple X-chromosome BGN epialleles per cell likely occurs closer to the time of transformation. Chromosomal instability is characterized by both copy gains and losses,[15] and losses of duplicate BGN epialleles would tend minimize intracell diversity because of a bottleneck effect.

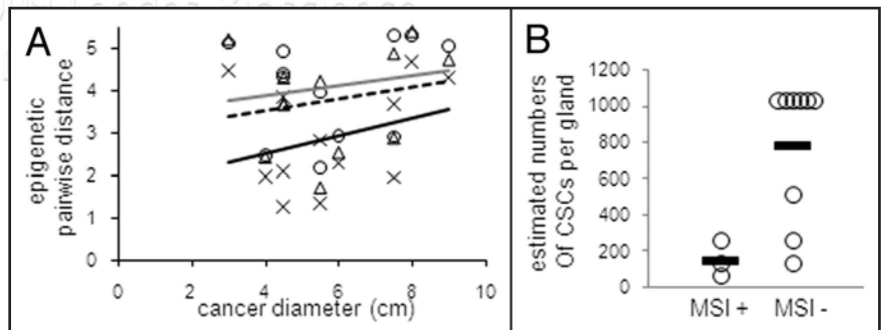

Figure 5. Colorectal cancer diversity. (A) Average BGN pairwise distances for 12 colorectal cancers with respect to cancer diameter. For each cancer, average pairwise distances within glands ("X"), between glands from the same side (triangle), and from opposite sides (circle) were similar, implying that mitotic ages within individual cancers are relatively uniform in its different parts. This relatively homogeneous topographical diversity is more consistent with a single, "flat" clonal expansion rather than stepwise sequential progression. Diversity was different between cancers, implying different cancers have different mitotic ages. There was a trend for greater diversity in larger cancers, but this trend was not significant. (B) MSI^+ colorectal cancers tended to have lower estimated numbers of CSCs per gland compared to MSI^- cancers.

Experimentally 14 cancer gland fragments (each about 2,000 to 10,000 cells) were sampled from opposite sides of each human colorectal cancer (7 from the left side and 7 from the right side, Fig. 1C), and eight epialleles from each gland were measured after bisulfite sequencing of cloned PCR products (112 epialleles per cancer). An example of the polymorphic cancer passenger methylation patterns is illustrated in Figure 4B. Such seemingly random patterns may arise from the independent accumulation of random replication errors (methylation and demethylation) in multiple long-lived cancer lineages in the many divisions after transformation. The methylation patterns can be compared within glands and between glands to infer relative mitotic ages. Average parameter values are used because with the stochastic nature of replication errors, single cancer epialleles are relatively uninformative. For the 12 cancers, average pairwise distances were different between cancers (Fig. 5A), consistent with the idea that cancers are removed from individuals at different times after transformation—cancers with greater mitotic ages should have greater epiallele diversity. Larger cancers were generally more diverse populations relative to smaller cancers, but this trend was not significant because some small cancers were also diverse populations.

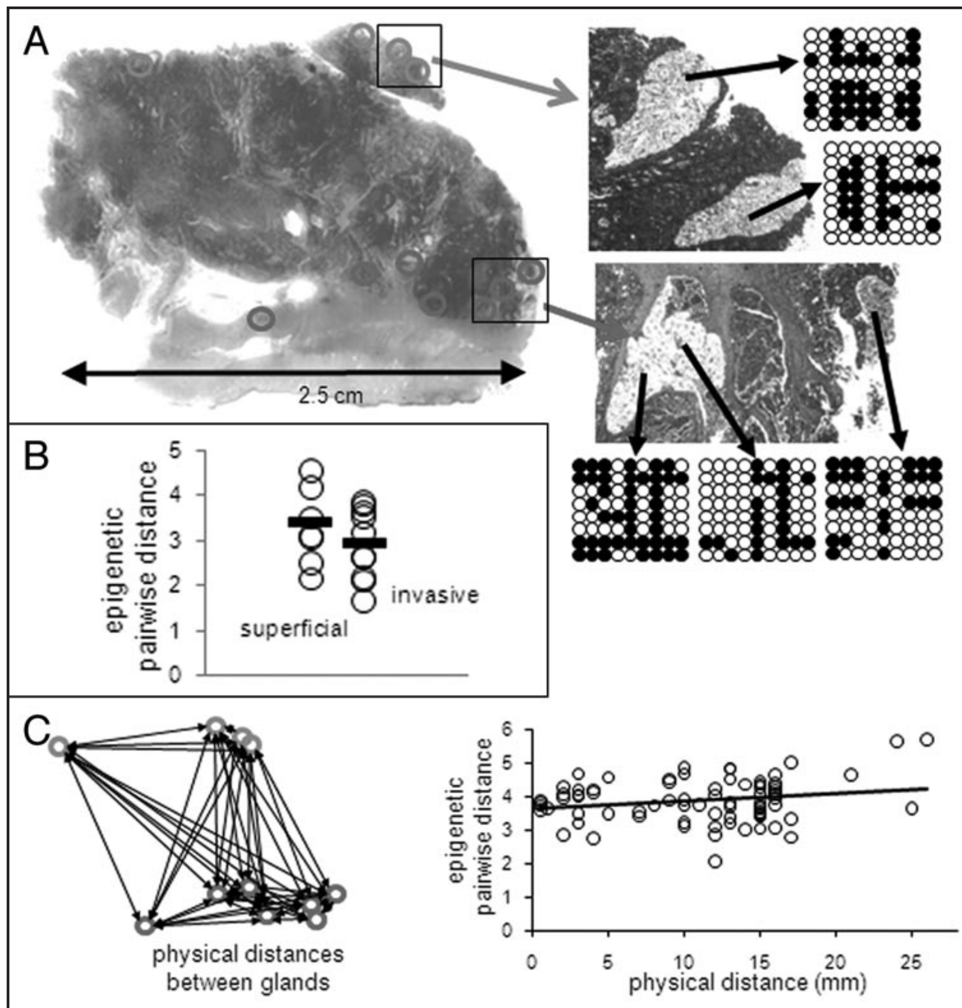

Figure 6. LCM sampling of cancer glands. (A) LCM allows sampling of glands defined by location and phenotype. The locations of the LCM isolated glands are indicated by circles (top are superficial glands and bottom are invasive glands). This is the same cancer as Figure 4. The BGN methylation patterns from the LCM isolated glands (images after LCM) are also diverse. (B) Average BGN epiallele diversity within superficial glands was not significantly different from invasive glands, indicating similar mitotic ages. (C) Physical distances can be measured between LCM isolated glands and then plotted with respect to the average epigenetic pairwise distances between glands. There is a slight but non-significant increase in epigenetic pairwise distance with physical distance. Adjacent superficial or invasive glands are almost as related as more physically distant glands.

Although methylation pattern diversity was similar throughout individual cancers, a problem with this analysis is its relatively poor spatial resolution because glands are sampled from 1 cm^3 tumor fragments. Moreover, it is uncertain whether the glands originate from invasive or superficial parts of the tumor. To improve resolution, we also performed laser capture microdissection (LCM) on microscopic slides, which allows sampling of glands precisely defined by location and phenotype (Fig. 6). The relative mitotic ages of each LCM isolated gland can be estimated by the pairwise distances between its eight sampled epialleles. The edge of an expanding population should be younger and less diverse than its stationary parts. Invasive glands should be less diverse than superficial glands if invasive glands are actively growing or if invasion occurred much later in progression. However, there were no significant differences in average pairwise distances between superficial and invasive glands,[5] implying that invasive and superficial glands have similar mitotic ages (Fig. 6B). Although the invasive portion of a cancer may appear to be actively "invading," methylation pattern diversities are more consistent with a history of very early invasion but subsequent stalling. That the invasive portion is not actively growing may help account for clinical observations that delays in the diagnosis and treatment of colorectal cancer do not appear to adversely affect survival.[10]

Epigenetic pairwise distances can be directly compared with physical distances by analyzing multiple glands on a single microscope slide. Are adjacent glands more related than widely spaced glands? With a palm-tree phylogeny, adjacent glands should be more related than distant glands because growth occurs locally. With a star phylogeny, adjacent glands should be as unrelated as distant glands because growth occurs uniformly (Fig. 1B). In general, epialleles differences between glands increased with the physical distances between glands, but this trend was slight and not statistically significant (Fig. 6C). Adjacent glands are almost as unrelated as widely spaced glands, indicating a star-shaped phylogeny.

Theory predicts that genomic diversity in opposite sides of a cancer will be similar with either star or palm-tree shaped clonal expansions (Fig. 1B). Consistent with this expectation, methylation pattern parameters (average percent methylation, average numbers of unique patterns per gland, and average diversity or pairwise distances) correlated between right and left tumor sides of the 12 cancers (Fig. 4C). Glands from the left and right sides of a cancer had similar diversities (average pairwise distances) or mitotic ages, consistent with a single uniform or "flat" clonal expansion. More consistent with a star phylogeny, BGN pairwise distances within glands, between glands sampled from the same side, and between glands from opposite cancer sides were similar for individual cancers (Fig. 5A).

The data are consistent with the idea that many colorectal cancers are single, relatively uniform "flat" clonal expansions because all parts of individual cancers appear to have similar

diversity or mitotic ages. This conclusion does not rule out the possibility that invasion occurs after superficial growth, but any such stepwise progression appears to occur very early after transformation such that the ages of both the invasive and superficial cancer regions are indistinguishable. A single clonal expansion is also consistent with Gompertzian growth because rapid initial exponential growth would tend to homogenize the ages of different parts of a tumor when they are essentially created at the same time (Fig. 2D).

Estimating Cancer Stem Cell Numbers

Cancer trees estimate cancer stem cell (CSC) numbers because a somatic cell ancestral tree is essentially a stem cell tree.[12] An ancestral tree contains three cell types: a common ancestor (first transformed cell), present day cells, and ancestral cells. Ancestral cells can be subdivided into dead-ends (no present day progeny) and ancestors (with present day progeny). Functionally, stem cells are ancestors and non-stem cells are dead-ends (Fig. 7A). CSCs are long-lived lineages in a cancer cell ancestral tree.

CSC frequencies can be estimated from cancer gland epiallele diversity. If CSCs are rare (i.e., a single CSC or less per cancer gland), there are few long-lived lineages (palm-tree shaped tree) and cancer gland diversity will be limited. If there are multiple CSCs per gland (star-shaped tree), cancer gland diversity will be higher. Simulations with different numbers of CSCs per cancer gland indicate that the measured epiallele diversity is too high for very rare CSCs.[5]

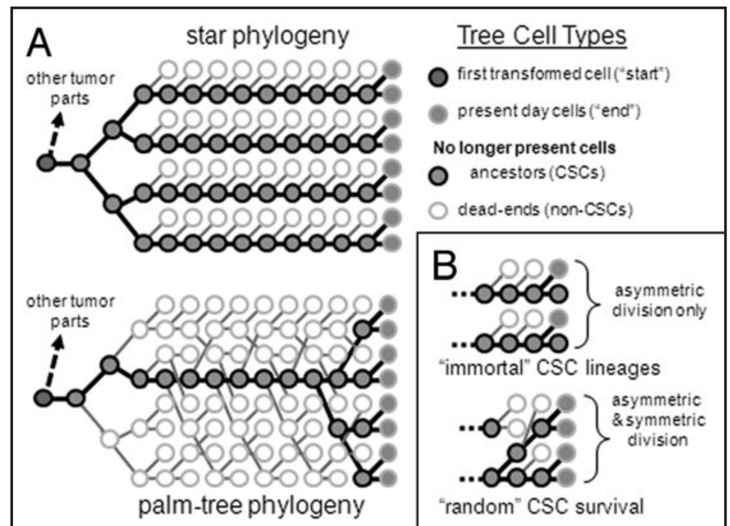

Figure 7. A cancer somatic cell ancestral tree. (A) An ancestral tree has three types—a start or ultimate common ancestor (first transformed cell), present day cells, and no longer present cells. No longer present cells can be divided into ancestors (with present day progeny) and dead-ends (no present day progeny). CSCs are ancestors and dead-ends are non-CSCs. (B) CSC survival may be deterministic ("immortal" with asymmetrical division only) or probabilistic ("random" with asymmetric and symmetric divisions). The BGN epiallele diversity was too high for very rare CSCs and more consistent with multiple CSCs per cancer gland (4 to 16 deterministic CSCs or 128 to 1,024 probabilistic CSCs).

©2009 Landes Bioscience. Do not distribute.

An alternative possibility is an equal survival probability for every cancer cell (i.e., every cell is a "CSC"). However, cancer gland epiallele diversity was too low for this model. Simulations with intermediate numbers of CSCs per gland and a stem cell hierarchy with constant numbers of CSCs were more consistent with the measured cancer gland diversities, with estimated numbers of long-lived CSC lineages between 4 to 1,000 CSCs per 8,000 cell gland (0.05 to 12% CSCs). The lower numerical CSC estimates are consistent with simulations of "immortal" CSC lineages, and the higher CSC estimates are with simulations of random CSC survival (Fig. 7B). Either with immortal CSCs or with random CSC survival, the basic stem cell hierarchy of normal colorectal crypts appears to be maintained after transformation—some cancer cell progeny appear to have limited potential for further division or survival. Frequent CSCs contribute to high cancer passenger methylation pattern diversity because replications errors accumulate independently in each CSC lineage.

CSC Numbers and Survival

The CSC concept is controversial and CSCs are difficult to define.[16] CSC frequencies estimated with epiallele diversity are much higher than estimated by xenotransplantation studies, which detected fewer than one CSCs per 50,000 unfractionated colorectal cancer cells.[17,18] Although the etiologies behind these different CSC estimates are uncertain, transplantation can be more successful when immunologic differences are minimized between the tumor and its host.[19] Whereas xenotransplantation prospectively measures the ability for human tumorigenesis in a murine immunedeficient environment (or "tumor initiating cells"[20]), genome diversity retrospectively infers the behaviors of CSCs during in vivo human tumorigenesis.

A conundrum of cancer biology is whether clinical outcomes benefit from rare or more frequent CSC. If CSCs are rare, therapy may be ineffective if only non-CSCs are eliminated. By contrast, therapy may be more effective when CSCs are rarer because relatively fewer cells must be eliminated. Our CSC estimates[5] revealed that colorectal cancers with microsatellite instability (MSI^+) tended to have relatively fewer CSCs per gland compared to MSI^- cancers (Fig. 5B). Although the numbers of examined cancers are small, this association suggests that cancers with less frequent CSCs may be easier to treat because better survival is observed for patients with MSI^+ colorectal cancers.[21]

Converting Epigenetic Distance into Chronological Time

Average pairwise distances between BGN epialleles (8 or 9 CpG sites) from opposite sides of the same cancer ranged from 0.4 to 4.1 differences among the 12 human colorectal cancers.[5] The average pairwise distance between sides could be translated into a cancer chronological age if the methylation error rate is known. Unfortunately, this error rate is uncertain, and the mitotic or chronological ages of most human cancers are unknown. Although the age of any single human cancer is unknown, the average human colorectal cancer is unlikely to be removed immediately after transformation or remain undetected for many years. Therefore, it may be possible to calibrate an epigenetic somatic cell clock after many

human cancers are examined—on average a cohort of cancers is likely to have a chronological age of a few years. In our study,[5] the average chronological age of our 12 colorectal cancers was about 21 months (range of 8 to 37 months) with an error rate of 3 x 10^{-4} changes (methylation or demethylation) per division, assuming one division per day. Corresponding estimated cancer mitotic ages were between 250 and 1,130 divisions since transformation. (The estimated mitotic age of the cancer in Fig. 4 was 1,100 divisions since transformation).

Mitotic Age versus Cancer Stage

In general there is no precise relationship between tumor size and stage, although higher stage is associated with larger cancers.[3] Metastases may be present with small primary tumors. Similarly, there were no significant relationships between estimated cancer mitotic age and cancer size, although larger tumors tended to have greater mitotic ages (Fig. 5A). After an initial rapid clonal expansion, smaller cancers may have higher mitotic ages if detection or removal is delayed.

Summary

A priori it is difficult to determine how an individual human tumor arises because progression may occur through different pathways. Instead of inferring which textbook model applies to an individual tumor, it may be possible to use molecular phylogeny to directly read the past written in the genomes of their cells. Molecular phylogeny is complex because quantitative approaches are needed to infer histories encoded by random replication errors. Our analysis of polymorphic passenger methylation patterns from different parts of the same tumor suggests that growth more often occurs through a single uniform clonal expansion rather than stepwise sequential clonal evolution (Fig. 1). A single clonal expansion is a "default" pathway because further alterations are not required after transformation. Most mutations in a metastases were also present in its primary,[22] suggesting that few additional mutations are acquired after transformation. A clonal expansion with uniform mitotic ages is also consistent with the uncontrolled proliferation hallmark of cancer.[23] Cancer cells lack the mechanisms that allow normal cells to sense environmental signals and control cell division. Such uncontrolled proliferation implies monotonic division regardless of location or direction, and would tend to homogenize tumor mitotic age.

©2009 Landes Bioscience. Do not distribute.

The relatively old or "stable" cancer populations inferred from the diverse passenger methylation patterns may appear to be incompatible with the reality that cancers kill through continued growth and invasion. More recent growth can occur through stepwise progression with a new subclonal expansion (Fig. 1A), which may be missed if the tumor is not extensively sampled. However, cell numbers can still increase slowly in older cancer populations, which may eventually kill the patient. The inference that most colorectal cancers are relatively old or static cell populations may be biased because all 12 examined cancers were removed surgically with the intent to cure. Growth patterns may differ in more advanced cancers or at the time of death.

The current analysis of somatic methylation patterns is sparse compared to population genetic studies of macroscopic populations. Whereas our analysis compared ~20 CpG sites, the geographic distributions of European populations could be inferred when ~500,000 SNPs were measured.[24] Population geneticists typically compare multiple loci because changes likely accumulate differently along a chromosome.[6] More extensive sampling and analysis may reveal greater cancer subpopulation structure and it is likely that human tumors grow with a myriad of patterns. Given the impracticality of serial observations or time travel, reconstructing histories from genome variations in different parts of the same cancer may provide a practical systematic approach to learn how individual human tumors actually grow.

References

1. Nowell PC. The clonal evolution of tumor cell populations. Science 1976; 194:23-8.
2. Bernards R, Weinberg RA. A progression puzzle. Nature 2002; 418:823.
3. Yasuda A, Sato K, Kitano S. Histopathologic characteristics of advanced colorectal cancer smaller than 2 cm in size. Colorectal Disease 1999; 1:19-22.
4. Andriole GL, Grubb RL, 3rd, Buys SS, Chia D, Church TR, Fouad MN, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med 2009; 360:1310-9.
5. Siegmund KD, Marjoram P, Woo YJ, Tavaré S, Shibata D. Inferring clonal expansion and cancer stem cell dynamics from DNA methylation patterns in colorectal cancers. Proc Natl Acad Sci USA 2009; 106:4828-33.
6. Kidd KK, Pakstis AJ, Speed WC, Kidd JR. Understanding human DNA sequence variation. J Hered 2004; 95:406-20.
7. Bromham L, Penny D. The modern molecular clock. Nat Rev Genet 2003; 4:216-24.
8. Garcia SB, Park HS, Novelli M, Wright NA. Field cancerization, clonality and epithelial stem cells: the spread of mutated clones in epithelial sheets. J Pathol 1999; 187:61-81.
9. Norton LA. Gompertzian model of human breast cancer growth. Cancer Res 1988; 48:7067-71.
10. Ramos M, Esteva M, Cabeza E, Campillo C, Llobera J, Aguiló A. Relationship of diagnostic and therapeutic delay with survival in colorectal cancer: a review. Eur J Cancer 2007; 43:2467-78.
11. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. Science 2007; 318:1108-13.
12. Shibata D, Tavaré S. Counting divisions in a human somatic cell tree: how, what and why? Cell Cycle 2006; 5:610-4.
13. Clark SJ, Harrison J, Paul CL, Frommer M. High sensitivity mapping of methylated cytosines. Nucleic Acids Res 1994; 22:2990-7.
14. Korshunova Y, Maloney RK, Lakey N, Citek RW, Bacher B, Budiman A, et al. Massively parallel bisulphite pyrosequencing reveals the molecular complexity of breast cancer-associated cytosine-methylation patterns obtained from tissue and serum DNA. Genome Res 2008; 18:19-29.
15. Lengauer C, Kinzler KW, Vogelstein B. CIN Genetic instability in colorectal cancers. Nature 1997; 386:623-7.
16. Hill RP, Perris R. "Destemming" cancer stem cells. J Natl Cancer Inst 2007; 99:1435-40.
17. O'Brien CA, Pollett A, Gallinger S, Dick JE. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature 2007; 445:106-10.
18. Ricci-Vitiani L, Lombardi DG, Pilozzi E, Biffoni M, Todaro M, Peschle C, et al. Identification and expansion of human colon-cancer-initiating cells. Nature 2007; 445:11-5.
19. Kelly PN, Dakic A, Adams JM, Nutt SL, Strasser A. Tumor growth need not be driven by rare cancer stem cells. Science 2007; 317:337.
20. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer and cancer stem cells. Nature 2001; 414:105-11.
21. Sinicrope FA, Rego RL, Foster N, Sargent DJ, Windschitl HE, Burgart LJ, et al. Microsatellite instability accounts for tumor site-related differences in clinicopathologic variables and prognosis in human colon cancers. Am J Gastroenterol 2006; 101:2818-25.
22. Jones S, Chen WD, Parmigiani G, Diehl F, Beerenwinkel N, Antal T, et al. Comparative lesion sequencing provides insights into tumor evolution. Proc Natl Acad Sci USA 2008; 105:4283-8.
23. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000; 100:57-70.
24. Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, et al. Genes mirror geography within Europe. Nature 2008; 456:98-101.